Investigating Research Integrity

Proceedings of the First ORI Research Conference on Research Integrity

Investigating Research Integrity

Proceedings of the
First ORI Research Conference on Research Integrity

editors

Nicholas H. Steneck, Ph.D.

Mary D. Scheetz, Ph.D.

[HHS Logo]

2002

publication information page

Preface

Over the past twenty years, a consensus has developed that integrity is vitally important to the health of federally funded research and that the key stakeholders–individual scientists, research institutions, associations and societies, government sponsors, and the general public–all play important roles in fostering research integrity. However, there is little consensus about the importance of and a lack of empirical scientific evidence on specific problems than can and do undermine integrity in research. Even those of us who are experienced in research integrity issues have in the past based too much of our thinking on personal experience, personal and philosophical biases, individual case exposes, and the public, political, and media response thereto. Accordingly, to advance to the next level in our understanding, it is time for new approaches to the study and discussion of research integrity.

Since its establishment in 1992, the Office of Research Integrity (ORI) has conducted a number of studies on research misconduct and research integrity, some of which are ongoing. The goal of these studies has been to develop a knowledge base for addressing important research integrity issues, including: the impact of misconduct allegations on exonerated scientists, the experience of whistleblowers in the aftermath of making allegations, the content of research guidelines adopted by medical schools, and the the incidence of research misconduct. Over time, it became apparent to ORI that a more comprehensive, coordinated effort in collaboration with extramural research scholars was needed to develop baseline knowledge for understanding research integrity issues. This recognition led to the development of the first Research Conference on Research Integrity in November 2000 and the revised papers published in this volume. ORI has also begun, with support from the National Institute of Neurological Disorders and Stroke, a related "Research Program on Research Integrity."

In the background report that begins this volume, *Assessing the Integrity of Publicly Funded Research*, Dr. Nicholas Steneck (ORI's consultant for the November 2000 conference and the related research program) has summarized the state of the empirical research on research integrity. This report provided important background information for participants at ORI's Research Conference on Research Integrity and for scholars and others in the research community generally.

The research conference background report and the conference papers published in this volume will hopefully provide an important catalyst for identifying important problems and for improving our understanding of research integrity issues. Although research integrity has been a high profile topic for some twenty years and some important preliminary studies have been conducted, the publications in this volume, while contributing valuable information, make clear how little we really know about many key issues, such as: how often research misconduct occurs, what situations tend to encourage or prevent it, how human subjects are best protected, how often conflicts of interest occur in research and how they affect the integrity of the research, how common questionable research practices are and what harm they cause to the research process, how students and research trainees learn the ethics of science, and what career pressures or other factors influence their ability and desire to follow the most honorable scientific practices.

These unanswered questions provide a significant opportunity for the Public Health Service and the research community to build a knowledge base for examining research integrity through further research. Research will permit us to understand in a more thorough and genuine way the influence

that research integrity issues have on the careers of scientists, the operation of research laboratories, the generation of accurate and useful research results, and the confidence of the public and political community in the research enterprise. It will also provide a science base for making important decisions—by government, by research institutions, by the community of scientists, and ultimately by the general public—in response to future research integrity issues and concerns that will inevitably arise.

Chris B. Pascal, J.D., Director
Office of Research Integrity

Introduction

Researchers and research institutions are universally committed to maintaining high standards for integrity in research. Precisely what this commitment entails, however, and whether it is being fulfilled are questions that have not been subject to rigorous critical investigation. What is "research integrity"? Can it be assessed? Do current research practices meet the high standards individuals and institutions say they embrace? How are standards for *best practices* transmitted? Are current approaches to fostering integrity appropriate and effective? Are all segments of the research community appropriately contributing to the promotion of high standards for integrity in research? Many individuals have provided answer to these questions, based on personal experience and anecdotal evidence. Few scholarly studies have been undertaken to confirm or refute what is commonly believed to be true about research integrity but is seldom demonstrated.

The papers published in this volume were originally presented at the first ORI Research Conference on Research Integrity in Bethesda, Maryland, on November 19-20, 2000, and subsequently reviewed and edited for publication. Abstracts for other papers and posters presented at the conference but not published in this volume can be accessed at *http://ori.dhhs.gov*. Together, this work represents the first comprehensive effort by a group of scholars to take a broad but critical look at evidence underlying our assumptions about integrity in publicly funded research.

The organization of the Proceedings reflects the collective interests and judgments of the scholars who responded to the call for abstracts for the Conference. Roughly half of the papers focused on factors that influence attitudes toward integrity and actual research practices. These factors are explored in these papers from the perspective of students and mentors, institutions and professions, medical practice and clinical research, conflict of interest, and, the most-studied subcategory of integrity, research misconduct. A second group of papers looked specifically at the way research integrity is taught, either across institutions or in one institution or course. Finally, a significant number of scholars tackled important methodological issues, looking at specific ways to detect misconduct, publication practices, and different theoretical perspectives.

To speed dissemination and to facilitate access, all of the papers published in this volume have previously been made available on the web. This limited-edition, bound copy is intended to create a more permanent archive of the first Research Conference on Research Integrity. As this volume goes to press, the call for abstracts for the second Research Conference on Research Integrity is being transmitted to continue the work begun in November 2000.

Nicholas H. Steneck, Ph.D.
Department of History, University of Michigan
Office of Research Integrity, DHHS

Mary D. Scheetz, Ph.D.
Office of Research Integrity, DHHS

Contents

Assessing the Integrity of Publicly Funded Research

Nicholas H. Steneck, Department of History, University of Michigan, USA

Keywords ***Accuracy, Authorship, Bias, Conflict of interest, Duplicate publication, Misconduct, Other research practices, Peer review, Research on research integrity, Self correction***

Since the early 1980s, when research integrity became a major national concern as a consequence of reports of misconduct in research, several thousand publications have in one way or another reported on, analyzed, and/or expressed opinions about the integrity of publicly funded research. Despite widespread interest in research integrity, however, the integrity of researchers has not been subject to the same critical study as other professionals. The research articles listed at the end of this paper account for no more than 3-4% of the total literature on research integrity.

The lack of research on research integrity presents a significant problem for government, research institutions, and professional societies. If integrity is defined as being honest in your dealings with others, there is ample evidence to suggest that from time to time publicly funded research falls short of this mark. As the articles summarized in this Paper confirm, researchers do commit misconduct; research results are inappropriately influenced by bias, conflicts of interest, and just plain carelessness; and researchers allow personal ambitions and biases to get in the way of the supposed objectivity of the research process. Publicly funded research does not always achieve the high standards that researchers, research institutions, and professional societies commonly set for themselves. This much is known.

In contrast, too little is known about the causes and significance of, or remedies for, research practices that fall short of the ideals set for the responsible practice of research.

- Is research misconduct rare or are the cases reported simply the tip of some unmeasured iceberg?
- Are there accepted norms or standards for research and, if so, how are they set, learned, and monitored?
- Are the regulations that currently govern publicly supported research sufficient and well enough enforced?
- Which practices that seem to fall short of accepted standards matter most from the standpoint of protecting the public's investment in research?
- Are there ways to foster integrity and thereby to prevent misconduct?
- Do research ethics courses make any difference?
- What influence does the research climate have on research integrity?

Each of these questions has at one time or another been raised and answered in the literature on research integrity. Few of the answers given have been based on critical understandings of research

---

The information and views presented in this report are those of the author and do not reflect the official views or policies of the Office of Research Integrity or the co-sponsoring organizations.

Corresponding author: Nicholas H. Steneck, Department of History, University of Michigan, Ann Arbor, MI 48109-1003, 734-647-4868 (voice), 734-647-4881 (fax), nsteneck@umich.edu.

as a profession, largely, as noted, because research as a profession has not be the subject of careful observation and controlled study.

The remainder of this Paper presents a brief analysis and summary of the research literature on research integrity.

- Section one presents an overview of what is known about the frequency of research misconduct (FFP).
- Section two discusses the complex and growing literature on research practices that seemingly compromise professional standards but may not constitute outright misconduct.
- Section three surveys the research that has been done on approaches to providing instruction on the responsible conduct of research (RCR).
- Section four explains how the literature cited in this Paper was selected, some of its characteristics, and the limitations of this analysis.

The bibliography at the end provides a complete list of references cited in the Paper, a summary of the RRI literature sorted by topics, and a comprehensive listing, sorted by first author, of the RRI literature with abstracts.

Throughout this Paper, I have used the terms "research misconduct," "scientific misconduct," or simply "misconduct" to refer to the three behaviors outlined in the common government definition of research misconduct, namely fabrication, falsification,and plagiarism (FFP) in proposing, conducting or reporting the results of research. While none of these behaviors is self-explanatory, the crucial element in each is a deliberate intent to deceive or mislead. Deliberate deception is clearly not consistent with good research practice and is generally agreed to constitute misconduct.

A second term used throughout this report, "integrity," is more difficult to define. Integrity is a measure of wholeness or completeness. When applied to professional behavior, it is essentially a measure of the degree to which someone's (or some institution's) actions accord with ideal or expected behavior. However, the ideals or expected behaviors for professional conduct are complex, not always well defined, and subject to change or reinterpretation. I have, therefore, adopted a fairly inclusive definition of integrity and assumed that it can be thought of as a measure of the degree to which researchers adhere to the rules or laws, regulations, guidelines, and commonly accepted professional codes and norms of their respective research areas.

Finally, a note of caution needs to be added. This survey of the RRI literature is of necessity selective and evolving. It places more emphasis on the biomedical sciences than the physical or social sciences. It does not do justice to the rich literature on peer review. It almost certainly has missed important articles that need to be included in the RRI literature. As a result, it will almost certainly be updated, and therefore comments and additions are welcomed.

Misconduct

Opinion about the extent of misconduct (FFP) in publicly funded research is sharply divided. In public testimony and editorials, researchers have commonly argued that research misconduct is rare. Support for this position is based on the fact that the documented cases of misconduct are few in number in comparison with the total number of individuals engaged in research. Approximately 200 cases of misconduct have been confirmed by the federal government over the last decade. Dividing cases by total researchers, this works out to a rate of about 1 in 10,000 over 20 years, assuming approximately 2,000,000 active researchers, or 1 in 100,000 per year. Critics of the way publicly funded research is conducted and administered counter that the reported cases represent the tip of a larger but uncharted iceberg. Support for this view is based in part on documented and presumed examples of the reluctance of researchers and research institutions to pursue cases of misconduct (for early warnings about possible larger numbers, see: 1, 2). Which, if either, opinion is correct remains to be determined.

Direct evidence

Research undertaken to clarify the extent of scientific misconduct suggests that it may be more common than the 1 in 10,000 or lower estimates. Evidence for this position comes from three direct approaches to measurement:

- It is reasonable to presume, based on research in other fields, that confirmed cases underestimate actual cases (3). Further research is needed to determine whether under-reporting in research is trivial or significant.

Year Author	Population Place	Sample Size	Responses (%)	Mis-conduct	FFP
1976 St. James-Roberts	Readers, *New Scientist* England	??	199 (?)	92%	?
1987 Tagney	Phys, biol, behav, & soc. scientists Major research university, US	1100	245 (22%)	–	32%
1992 Kalichman	Biomedical trainees UC San Diego, US	2010	549 (27%)	36%	–
1993 Swazey	Chem., civil eng., microbiol., sociol. US survey, faculty/graduate	4000	--/-- (72/59%)	44/50%	6/9%
1993 Hals	PIs, biomedical sciences Health Region IV, Norway	159	119 (70%)	27%	–
1995 Bekkelund	Biomedical researchers Norway, random survey	274	215 (80%)	22%	3%
1996 Eastwood	Post-doctoral training fellows US, random national survey	1005	324 (33%)	58%	3-12%

Table 1. Surveys of the Level of Misconduct in Research

- Surveys of knowledge of misconduct consistently report knowledge rates above 1% (Table 1). Reported knowledge of misconduct remains above 1% (1 in 100, or 100 times higher than the 1 in 10,000 estimate) even when researchers are asked about their own research group and when misconduct is specifically limited to FFP. One survey specifically asked researchers whether the misconduct they were aware of was public knowledge. Of the roughly one-in-four researchers who were aware of misconduct (27%), 47% said that the cases were not public knowledge (4).
- Audits of research procedures and results have turned up "significant problems" or "major deviations" a levels that range at and above the 10% level (5-8). These results do not correlate directly with FFP, since they do not take into account whether discrepancies result from deliberate actions.

The results of surveys, audits, and estimates of the rate of under-reporting raise two important issues for further consideration. First, however the results of surveys and audits are ultimately interpreted or clarified, there remains the troubling discrepancy between public statements about how "rare" misconduct in research supposedly is and the more private belief on the part of many researchers that it is in fact fairly common. How can these two views be reconciled?

Second, whatever the actual rate of misconduct, it is not so much the rate as the significance of the misconduct that matters most. Summarizing the results of scientific data audits of the Cancer and Leukemia Group B's clinical trials, Weiss et al. conclude that "scientific improprieties have occurred very rarely..." (8, p. 459). "Very rarely, in this case, is based on a quantitative estimate of 0.28% (p. 462)–28 cases of misconduct for every 10,000 clinical researchers or one case for every 357 clinical researchers. On what basis can this rate be judged as either "rare" or "significant"? Clearly, understanding the importance of misconduct in research requires not only better estimates of numbers but also of significance. How much does a case of misconduct in research actually cost the public in terms of wasted research dollars, of deceptive findings that mislead other researchers until the misconduct is discovered, and perhaps of negative impacts on patient health?

Indirect evidence

Gathering information on the likely prevalence of misconduct in research can be approached indirectly. For example, many studies have documented that cheating is common in the educational system at all levels and in all programs. The rates vary from well above 50% for high school and college undergraduates (9-12) to levels between 10% and 30% for professional students (13-20). One survey specifically asked whether misconduct at this level was indicative of future performance. Of 246 faculty and administrators responding, 216 (86%) felt that it was so indicative (14, p. 34). If this estimate of the relationship between student conduct and later professional conduct is true, it would support the contention that the prevalence of misconduct in research may be higher than the small number of confirmed cases suggest.

The prevalence of a willingness to engage in misconduct has been documented into graduate and post-doctoral research education. Kalichman's and Eastwood's surveys report that significant numbers of students (above 10%, except for fabricating data) will omit or change evidence and add honorary authors if it will help get papers published or grants funded (Table 2) (21, 22). Students who are in the beginning stages of becoming researchers clearly feel that career pressures may make it necessary to engage in practices that they also know are wrong.

That significant numbers of beginning researchers may in fact do what they say they will do has been confirmed in a series of audits of the research publications listed on residency fellowship applications. These audits report significant numbers (15% and higher) of misrepresentations, from seemingly trivial offenses such as inflating author rank to listing articles "in press" when they were not, listing papers in journals that do not exist, and listing bogus articles in real publications (Table 3) (23-27). Similar practices are generally counted as FFP when they occur in research grant applications or resumes submitted for promotion.

One final piece of indirect evidence that should be noted is the confirmed reluctance of researchers to report suspected misconduct.

- As noted above, Hals reported that roughly one-in-four researchers (27%) who knew of misconduct, said that the cases they knew of were not public knowledge, which could mean they were not reported (4).
- In Tagney's survey conducted at one research institution, roughly half of those who reported suspecting misconduct took no action (28).
- Korenman's study of the attitudes of researchers and institutional representatives toward misconduct found that researchers were more likely to favor informing

Action	**1992 Kalichman**	**1996 Eastwood**
Past misconduct (yes/no?)	15.1%	12%
Future misconduct (yes/no?)	14.8%	
...modify data for paper	7.3%	15%
...modify data for a grant application	13.5%	--
...fabricate date for a paper or grant application	1.3%	< 2%
...select or omit data for paper or grant application	14.2%	27%
...list an undeserving author	--	41%

Table 2. Self-reported attitudes toward misconduct

Author ***Specialty***	**1995 Sekas** ***Gastro-enterology***	**1996 Gurudevan** ***Emergency Medicine***	**1997 Panicek** ***Radiology***	**1998 Bilge** ***Pediatrics***	**1999 Dale** ***Orthopaedic Medicine***
Total applications	236	350	201	404	213
...with citations	53 (22%)	113 (32%)	87 (43%)	147 (36%)	64 (30%)
...misrepresented	16 (30%)	23 (20%)	14 (16%)	29 (20%)	11 (17%)
Total citations	--	276	261	410	76
...misrepresented	--	44 (16%)	39 (15%)	41 (10%)	14 (18%)
Research experience	138 (59%)	--	--	--	--
...not confirmed	47 (34%)	--	--	--	--

Table 3. Misrepresentation in medical resident training program applications

colleagues whereas institutional representatives favored reporting to supervisors and deans (29).

These findings confirm the suspicions of the "tip-of-the iceberg" school, which argues that reported cases are not an accurate measure of actual levels of misconduct. No controlled studies of under-reporting have been undertaken to assess the rate of under-reporting, making it difficult to conclude whether it is significant.

Cheating or misconduct on the path toward becoming a researcher does not, of course, demonstrate that misconduct continues once students become researchers. Under-reporting may not seriously compromise estimates of the amount of misconduct. Reasons can be given to suggest that some of the estimates of misconduct given in the various surveys reported above may be too high as well as reasons to suggest that they may be too low. The differences between the "rare" and "tip-of-the-iceberg" schools can therefore not be resolved easily. What is important to note, however, is that in seeking to refine understandings and resolve the differences between the two schools, the range of uncertainty that exists is significant. In terms of decimal points, the range is not a matter of one or two orders of magnitude but closer to four or five orders of magnitude, varying from 1 in 100,000 or less to 1 in 100 or more. And this, in turn, makes it difficult, if not impossible, to estimate the public costs of misconduct when determining what policies are needed to protect the public's investment in research.

Other Research Practices

Over the past twenty years or longer, the discussion of "research integrity" has focused primarily on "research misconduct," based on widespread agreement that misconduct (FFP) is wrong or fraudulent. While it is true that research misconduct clearly can undermine the integrity of publicly supported research and therefore needs to be taken seriously, so can other research practices, such as sloppy research, inappropriate bias, conflict of interest, or poor mentoring.

The existence of other research practices that can compromise integrity has been recognized by the research community, but there has been no agreement on how to respond to them or how seriously they should be taken. In its 1992 report, Responsible Science, the NAS/NAE/IOM Panel on *Scientific Responsibility and the Conduct of Research* specifically set out a separate category of research behavior called "Questionable Research Practices." The Panel recognized that such practices "...violate traditional values of the research enterprise and ... may be detrimental to the research process," but it was not willing to include them under "misconduct." It did concede, however, that since "...the relationship between these two categories is not well understood ... [i]t may be difficult to tell, initially, whether alleged misconduct constitutes misconduct in science or a questionable research practice" (30, pp. 5-6, 29).

Whether or not "other questionable practices" constitute misconduct is irrelevant for the purposes of this Report. What is relevant is the fact that any practice that deviates significantly from the "rules, regulations, guidelines, and commonly accepted professional codes or norms for the responsible conduct of research" (the definition for integrity given in the Introduction) can compromise and currently are compromising the integrity of publicly funded research. However, until more is known about these practices, it will be difficult to suggest how seriously they need to be taken.

The remainder of this section summarizes some of the research on other practices that can compromise the integrity of research. The summary is intended to be more illustrative than exhaustive. Some aspects of research practice, such as authorship and peer review, have been the subject of intense study and hundreds of publications, thanks in large part to the Congresses on Biomedical Peer Review organized by JAMA editor, Drummond Rennie (31). Exhaustive coverage is therefore not possible. Rather, the goal of this section is to focus on some areas of potential concern and illustrate some of the findings that have emerged.

Accuracy

Accurate information is vital to research. Research is a cooperative and cumulative enterprise. Researchers build on the work of others, which means the information they have about other work and the way research is conveyed must be accurate; however, a number of studies suggest that research results are not always conveyed accurately.

- Information presented in abstracts does not

always accurately reflect the information given in the article itself. One study reported major discrepancies in abstracts (inconsistencies or information that was not contained in the body of the article) in 55 of 203 randomly selected articles (32).

- Studies have reported that significant numbers (above 10%) of published articles misuse statistics or contain statistical errors (33).
- Random checks on citations and quotations in published articles have reported error rates well above 10%. Errors were counted as "citation errors" when the names, pages, or other information needed for locating an article was inaccurate (minor) or when the referenced article could not be located based on the information given (major). Errors were counted as "quotation errors" when the reference oversimplified or exaggerated information given in the referenced article (minor) or when the information given in the original article did not support or contradicted claims made in the reference (major) (34, 35).

Inaccuracies in abstracts, the use of statistics, and references do not necessarily invalidate research results. Conclusions or pieces of evidence presented only in an abstract but not in the body of an article could be true. Research results bolstered by inflated or deceptive statistics or inaccurate references to other studies might still be true. At issue, however, is not whether the results are ultimately true or accurate but whether the word (or words in this case) of researchers can always be trusted. The clear answer to this question, unfortunately, is that it (they) cannot.

Peer Review

Inaccuracy and other problems in publication are purportedly reduced, if not eliminated, through peer review. In general, the peer review system enjoys considerable support within the research community and is seen by most as the foundation on which professional self-regulation rests. This does not mean, however, that peer review is above criticism or not in need of further improvement.

- That peer reviewers miss problems in publications has been documented by the fact that different reviewers detect different problems in manuscripts, even when they are in substantial agreement about whether to publish (36) and by studies of how fraudulent publications have made it to press (37). How much effort should be made to improve peer review requires more information about how well it is working and the price of its shortcomings.
- Peer review has been shown to have institutional (38), national (39, 40), methodological (39, 41), gender (42) and outcome biases (43-45). Bias, obviously, runs counter to the value-neutral goal of research.
- Considerable uncertainty exists about the best ways to improve peer review. Traditional approaches, such as blinding, issuing clear instructions, or relying on experienced researchers, have had different measures of success (46-53).
- Studies of peer review have raised questions about whether it helps or hinders innovation (54, 55).

One review of the rich literature on peer review concludes: "Because of the central place of peer review in the scientific community and the resources it requires, more studies are needed to define what it does and does not accomplish" (56). This work will fortunately be fostered by the future Congresses on Biomedical Peer Review and similar efforts.

Self-Correction

Researchers constantly read and check each other's work. The routine process of using the work of others in the day-to-day practice of research provides an additional mechanism for detecting and correcting errors and other problems in research, such as research misconduct. Research is, in other words, self-correcting, which further ensures its integrity. However, research on the effectiveness of self-correction in research has shown that this mechanism is not as vigilant as one might expect.

- Studies of some of the first publicly documented cases of misconduct found that publication of a retraction reduced the citation of fraudulent articles but did not eliminate it (57-59).
- One recent study of articles retracted for a broad range of reasons, from outright fraud to acknowledged experimental errors or later failure to replicate, concluded that retracted

articles continue to be cited and used as a significant rate. Of 299 post-retraction citations listed in the Abridged Index Medicus, only 19 (6%) mentioned the retraction; 17 (6%) explicitly and 263 (88%) implicitly reported the retracted work as "valid" (60).

- Research on the process by which articles are retracted and erroneous information withdrawn has show that it is slow (60, 61) and in some key ways ineffective (60-63).

Findings such as these have important policy implications. In his study of retraction notices, Budd agrees that research is self-correcting, but then he adds: "...there may be a great deal of time, effort, and money spent in discovering that some research is not useful. If erroneous or fraudulent work lives on in the literature, the amount of time, effort, and money to correct work may be even greater" (60, p. 297). At issue, in other words, is not whether research errors are corrected, but when. Failure to correct the literature in a timely and responsible manner is as much a matter of integrity, viewed from the public's investment in research, as a failure to correct at all.

Authorship

In principle, research results are more important than researchers. Who publishes an article should not matter. What matters most are the results. In practice, however, authorship is vitally important to, and significantly influences, the research process. Most research funding today is dependent on productivity. Review panels want to know not only what a researcher is planning to do but what she or he has done. Advancement in academic research is not possible without publication. Getting one's name on research papers is important–so important that as many as one in five aspiring researchers misrepresents publications on résumés in an attempt to improve his or her standings as a researcher (see Table 4).

As with the other research practices discussed in this section, there is considerable evidence to suggest that the ideal standard for determining authorship is not followed in practice and that expected authorship practices in general are sometimes not clearly defined or conveyed.

- Two studies that used the ICMJE criteria (64) for judging authorship found that 19% (65) and 36.4% (66) of papers did not meet these criteria.
- Evidence suggests that the rules for authorship are poorly understood, interpreted differently by different researchers, and not well communicated from senior to junior researchers (22, 67, 68).
- Patterns of authorship and the increase in disputes over authorship suggest that decisions about authorship are significantly influenced by the research environment (69, 70).

The importance of the truthful reporting of research contributions through authorship is widely recognized. The NIH Guidelines for the Conduct of Research note in particular that:

> For each individual the privilege of authorship should be based on significant contribution to the conceptualization, design, execution, and/or interpretations of the research study, as well as a willingness to assume responsibility for the study. Individuals who do not meet these criteria but who have assisted the research by their encouragement and advice or by providing space, financial support, reagents, occasional analyses or patient material should be acknowledged in the text but not be authors. (71, p. 10)

Authors who ask or agree to be listed on papers to which they have not made substantial contribution compromise the integrity of the research environment. The same is true of the 41% of graduate students who report a willingness to list undeserving authors on their papers (see Table 3, above).

Duplicate Publication

In its advice to intramural researchers, NIH research Guidelines caution researchers about duplicate publication:

> Timely publication of new and significant results is important for the progress of science, but fragmentary publication of the results of a scientific investigation or multiple publications of the same or similar data are inappropriate. (71, p. 8)

Despite widespread agreement that duplicate publication is inappropriate, the rate of duplicate publication (publishing the same article twice without reference) seems to hover at about 10% (Table 4) (72-76). Based on his study of publication trends in the *British Medical Journal*, Waldron suggested that duplicate publication was

Study	*Journal*	*Articles*	*Duplicate %*
Waldron (1992)	BMJ	354 published	6-12%
Bernard (1993)	NTvG	172 published	11%
Koen (1994)	NTvG	108 rejected	4%
Blancett (1995)	INJS	642 published	9%
Bloemenkamp (1999)	NTvG	148 published	7%

Table 4. Percent duplicate publication

increasing (72). Bleomenkamp more recently reported that the duplicate publication rate for articles in *Nederlands Tijdschrift voor Geneeskunde* has remained constant over the last 10 years and the number of authors referencing the second publication has increased significantly, from 22% to 73%.(76).

Duplicate publication adversely effects research in a number of ways. It can waste time (editors and reviewers) and resources (library funds and reprint costs). It also makes it difficult to evaluate the productivity of researchers. But perhaps most importantly, in clinical research it has the potential to inappropriately distort or bias findings if the duplicate publications are more prevalent in one treatment regimen.

- In a meta-analysis of post-operative effects of ondansetron, Tramer and Reynolds reported that "17% of published studies and 28% of the patient data were duplicated. Moreover, duplication was more common in studies that reported greater treatment effect. This bias, according to Tramer and Reynolds, "led to a 23% overestimation of ondansetron's antiemetic efficacy" (77).
- Jefferson reports that in a Cochrane review of the effects of Plasma Derived Vaccines, he and his colleagues suspected that 25% (15 of 60) of the trials identified during the first phase of review were duplicate publications. This percentage increased to 43% (3 of 7) when they progressed to the second phase of review. Being aware of the problem of duplicate publication, his group excluded the duplicate studies, but doing so is not common practice (78).

In the final analysis, Jefferson considers only "publishing redundant material with the intention of misleading the public, editors and readers, in order to make them believe the study is different from the original" as a "breach of current ethical tenets" (p. 138). From the public's perspective, however, it makes no difference whether the duplication is intended or not. If researchers do not take steps to ensure that a second or third publication of a body of data is recognized as such, the public could be harmed and the integrity of the research process undermined.

Bias and Conflict of Interest

There has been considerable debate about the role of values and personal interest in research ever since Merton proposed "disinterestedness" as one of four key values on which science rests (79, p. 116). It is now widely recognized that values influence research (80), but there is also a common understanding that the influence of values should be minimized and made public, particularly when financial interests are involved.

Considerable evidence exists to support the contention that personal interest does influence research behavior. Positive-outcomes bias (favoring publications that report positive results over those that report negative results or that do not find results) has been demonstrated in a number of studies (44, 81, 82). The reverse effect has also been reported, that is, slower publication rates for studies that fail to find a particular result (45). Studies are just beginning to assess how these interests affect research and whether they are being properly managed (83-85).

In calling controversial publication, reporting, and other research practices "questionable," the NAS report, *Responsible Science*, highlights an important problem. (30) "Integrity" is not an all-or-nothing proposition. There is a difference between a failure to check the spelling of every author's name or to catch every typo and using improper statistics or delaying the publication of a manuscript to please a sponsor. It is not easy to pinpoint where or when high standards for integrity in research give way to careless research practices, to

irresponsible research practices or to misconduct. The extremes (high standards for integrity and misconduct) can be defined, but behaviors that fall between, to one extent or another, are all subject to interpretation. This, in turn, makes it imperative that these behaviors are well understood and their consequences evaluated, both as part of the process of reassuring the public that its research funds are being spent responsibility and as needed background information for developing responsible conduct of research training programs.

Education

It is commonplace for reports on research misconduct/integrity to emphasize the importance of education. Professions have an obligation to society to educate future generations of professionals, which includes making future professionals aware of the standards for responsible practice. Moreover, if professional ethics education prevents misconduct, it is in a profession's best interest to encourage this education, which most in fact do.

Through the 1980s, research ethics training was commonly relegated to the laboratory and to mentoring. This changed in 1989 when NIH and ADAMHA instituted required "instruction in the responsible conduct of research" (RCR) for all training grants (86). The requirement stipulated that training programs had to have instruction in RCR, which in turn had to be described in the training grant application. Although the requirement technically had no "regulatory teeth," coming as it did in the highly competitive environment of grant-getting, researchers and research institutions quickly complied and instituted a wide variety of research ethics or RCR training programs (87).

The increase in formal RCR training raises an obvious and researchable question: has it or will it make any difference? At the present time, there is no convincing evidence that it does, but this does not necessarily lead to the conclusion that RCR training is ineffective, unnecessary, or unwise. The newness of most programs means that their impact may not yet be apparent. RCR training is delivered in different ways and different settings, making it difficult to isolate the influence this one factor has on the complex process of becoming a responsible researcher. And perhaps most importantly, there is no agreement on the goals of RCR education, making it difficult to judge whether it is succeeding.

RCR training

Straightforward efforts to evaluate the impact RCR training has on attitudes or anticipated behaviors have not reported any clear positive results. Studies by Kalichman et al. and Eastwood et al. compared receiving or not receiving RCR training with anticipated research behaviors. A study by Brown compared receiving or not receiving RCR training with self-reported perceptions of different ethical standards. None of the studies found any significant correlations between attitudes or anticipated behaviors and RCR training (21, 22, 88). Brown's study did report that RCR training increased awareness of options in ambiguous situations (p. 490). However, Eastwood's study reported that fellows who received RCR training were more willing to grant honorary authorship than fellows who did not (p. 95). Overall, direct measures of attitudes and anticipated behavior have pointed to some possible benefits, perhaps one puzzling negative, and a great deal of similarity between those receiving and not receiving RCR training.

Efforts to refine the study of the impact of RCR training have led to a difference of views on appropriate outcome measures. Based on a three-year effort to develop and assess an RCR course at Dartmouth College, Elliot and Stern argue that "if 'ethical behavior' is removed as a basis for the evaluation of teaching ethics," effective assessment tools can be developed. In the place of ethical behavior, they propose using two familiar measures of success in academic courses in general: "the skills and content taught in the course and the learning environment in which the teaching takes place" (89, p. 348). The project allowed them to develop and test various tools for evaluating these ends, which they argue can be accomplished, "but only if [teaching of academic research ethics] is treated as an academic discipline by both faculty and students" (p. 355).

Others believe that striving for some type of behavioral or moral reasoning change is appropriate for professional ethics instruction, including RCR training, and that such change can be measured. In a series of studies of medical, veterinary, and dental education, Self, Baldwin, Bebeau and colleagues have reported that: a) traditional professional education programs may erode and b) the addition of ethics instruction to

traditional programs improves the ability of students to engage in moral reasoning (90-97). Whether changes in the ability to engage in moral reasoning measured in professional education settings generally can be applied to RCR training in particular and whether changes in moral reason have any lasting professional consequences remains to be determined.

The research needed to plan effective RCR programs will clearly need to take into account more than what goes on in the RCR classroom. Studies have shown that environment is closely linked to what students feel they must do as opposed to what they should do (17, 18, 20, 22). Although the 1995 survey of the attitudes and experiences of 2,000 graduate students with misconduct (Table 2, above) indicates "that fraud, plagiarism, and related forms of misconduct are the results of individual predilections or failures of judgement..." (98, p. 225), Anderson et al. in commenting on these results still point to important influences exerted by environment and mentoring relations (p. 226). Without attention to the full context within which integrity is learned and decisions made about right and wrong actions, the goal of ensuring the responsible conduct of research through RCR training could well be negated by influences in the research environment.

Other efforts to educate

In discussions of ways to improve the integrity of research, surprisingly little attention has been given to the role of clear rules and routine monitoring or data audits. If the ultimate goal of research ethics/integrity policy is simply to ensure high standards for publicly supported research, the simplest way to achieve this goal may be to make the rules as explicit and clear as possible and then to check to make sure they are being followed. For each of these approaches to "educating" researchers, there is interesting research that suggests what may or may not work.

Over the last decade, new rules have been formulated for reporting research. Particular attention has been paid to two key areas–journal publication in general and clinical trial reporting. Studies of the effect of new rules suggested that they have had mixed results.

- Two studies that looked at the adoption of specific standards for reporting clinical trials by several medical journals concluded that there was room for improvement (99, 100). Junker suggest that more journals should require authors to follow the Consolidated Standards of Reporting Trials (CONSORT) (101). Clarke and Chalmers conclud that "there is little evidence that journals have adequately implemented the CONSORT recommendation that results of an RCT [randomized controlled trial] be discussed in light of the totality of the available evidence" (p. 280).
- In studies of measures to improve the quality of abstracts, Pitkin found that instructions to the authors had little impact (32, 102, 103).
- In a study of the impact of guidelines published in the *British Medical Journal* for manuscripts on the economics of health care, no difference was found in the quality of manuscripts, although the guidelines were judged to be useful for editorial purposes (104).
- In a comparison of systematic reviews and meta-analyses published following the procedures of the Cochrane Collaboration versus the more open-ended general reviews published in journals, Jadad reported more methodological rigor in the Cochrane reviews (41).
- In a study of the impact of professional codes in physics, Tarnow reported that postdoctoral students were generally not aware of publication rules and spent little time with advisors discussing publication practices (68).

As a group, this research seems to support the perhaps not unexpected conclusion that rules alone will not change behavior and must be accompanied by efforts to both make them known and take them seriously. Simply making information about rules for responsible behavior available is not an effective way to foster responsible behavior.

In contrast, data audits seem to have a significant effect on research behavior. Two studies of major government data audit programs both report that serious misconduct declined over the course of the studies.

- Shapiro and Charrow's study of FDA audits conducted between 1977 and 1988 reported that the rates of specific deficiencies remained about the same throughout but "the

overall level of seriousness of the problems ... declined" (7, p. 130).

- Weiss et al. in their detailed look at the results of audits conducted by the Cancer and Leukemia Group B (CALGB) conclude that: "The CALGB data audit process has been successful in uncovering the very rare instances of scientific misconduct and pressuring group members to improve adherence to administrative requirements, protocol compliance, and data submission. It has also served to weed out poorly performing institutions" (8, p. 464).

If results matter, then one of the most effective ways to educate researchers about their responsibilities may be to check more carefully the work they produce.

Data audits have been resisted because they are allegedly expensive, time-consuming, and perhaps even counter-productive; e.g. too much concern about the bookkeeping required to pass audits might slow the progress of science. There currently are no data to support these concerns. There is evidence, reviewed by Armstrong, that peer review can slow innovation in research (54, pp. 70-71), but no evidence that data audits have a similar effects. Moreover, Glick's rough estimates of the cost of data audits, based on conservative estimates of the amount of careless work and misconduct that may be affecting research results, suggests that over the long term, they will save public dollars. "Data auditing would increase research productivity by 2.5-6% (...), so that each dollar spent on such audits might eventually benefit the public, 20 years later, by an amount equivalent to $25-60" (3, p. 81). These results and estimations will no doubt be challenged, but for now the evidence seems to suggest that research audits might be an effective and efficient way to detect misconduct and reduce the rate of other questionable practices.

Research Literature Overview

As noted in the Introduction, over the last 20 years or longer, several thousand publications have in one way or another addressed the issue of integrity and/or misconduct in research. Most of these publications are based on some research. Reporters do research for news stories. Journal editors investigate problems before writing editorials. Taken to mean simply investigation or study, most if not all that has been written about research integrity is based on some research.

For the purposes of this Report, "research" has been defined as studies that have some element of controlled investigation, which means primarily but not exclusively surveys and quantitative assessments. Limiting the definition of research in this way obviously eliminates many thoughtful articles and books from the literature review, such as editorials, analytical writings, historical and cases studies, and philosophical analyses. The fact that works such as these are not included in this Report should not be taken as suggesting they are not important. They clearly are crucial and in other contexts certainly need to be considered. However, for the purposes of the ORI RRI program, the immediate goal is to gather hard evidence relating to actual research practices, so that policy-making can be based on the way research is conducted as opposed to the way we may think it is conducted.

Controlled quantitative research plays an important role in scholarly investigation. Most significantly, it helps establish reference points for organizing and evaluating other information. For example, historians, journalists, and others have amply documented that misconduct takes place in research. However, without some quantitative assessments, it is difficult to know what to make of individual cases of misconduct or even of the entire body of confirmed cases. Are they typical or atypical? Is misconduct common or rare? Without some controlled counting or surveys, it is difficult to place individual events and behaviors into context.

Locating research on research integrity is not a simple task. Keyword searching for the most part does not separate scholarly analyses from empirical studies. References located through searches for "scientific misconduct," "research ethics" and other keywords need to be evaluated for both relevance and method. The articles summarized in this Report have been located through standard keyword searches in several different databases, checking references listed in bibliographies, and in some cases by searching for publications by scholars with known RRI interests. Major emphasis has been placed on work relating to the biomedical sciences in particular and the hard sciences more generally. Less attention has been paid to research on integrity in the social sciences. The final RRI bibliography contains 136 entries, most of which, but not all, have some empirical or controlled

research component.

That RRI has not yet developed into an organized research field is more than evident from the fact that the 136 articles summarized in this Report appeared in 45 different journals (Table 5) and two books (105, 106). Most journals published only one or two articles. There are, however, three important exceptions.

- Fifty-one of the 136 (37.5%) articles appeared in JAMA. Most of these articles are on integrity in publication and are the product of the three peer review conferences organized by Drummond Rennie.
- Fourteen of the 136 articles (10%) appeared in Academic Medicine. These articles are mostly concerned with student conduct, not research integrity specifically, but have been included because they provide important background on the values researchers may have had as students.
- Eleven of the 136 articles (8%) appeared in *Science and Engineering Ethics*. This group of publications is split nearly evenly between research ethics training and publication practices. SEE is unfortunately not indexed in MedLine®, which limits the knowledge of this important group of publications.

Together, these three journals account for 76 of the 136 articles. Three journals had three research articles; five journals had two, and the remainder published a single research article on research integrity.

The fact that research on research integrity is distributed so broadly through the scholarly literature almost certainly slows research progress. At the present time, the standard search tools simply do not cut across the different disciplines that contribute to RRI. What is "discovered" in one field is thus not easily known in other fields. More importantly, however, is the fact that the absence of a well defined literature and corresponding research community makes interdisciplinary research on research integrity more difficult. This second shortcoming is particularly important for the development of research on research integrity, which of necessity must be interdisciplinary and

Journal of the American Medical Association (51)	Cancer Investigation (1)
Academic Medicine (14)	Cognitive Therapy and Research (1)
Science and Engineering Ethics (11)	Controlled Clinical Trials (1)
British Medical Journal (3)	Image: The Journal of Nursing Scholarship (1)
Journal of Professional Nursing (3)	Journal of Allied Health (1)
Nederlands Tijdschrift voor Geneeskunde (3)	Journal of Bone and Joint Surgery (1)
Accountability in Research (2)	Journal of Clinical Epidemiology (1)
Bulletin of the Medical Libraries Association (2)	Journal of General Internal Medicine (1)
Journal of Dental Education (2)	Journal of Higher Education (1)
Lancet (2)	Journal of Information Ethics (1)
Medical Education (2)	Journal of Investigative Medicine (1)
Medical Reference Services Quarterly (2)	Journal of Medical Education (1)
New Scientist (2)	Journal of Medical Ethics (1)
Tidsskrift for den Norske lægeforening (2)	Journal of the Am. Veterinary Medical Association (1)
AIDS Education and Prevention (1)	Journal of the Royal College of Physicians, London (1)
American Journal of Medicine (1)1	Minerva (1)
American Journal of Public Health (1)	Nature (1)
American Journal of Roentgenology (1)	New England Journal of Medicine (1)
American Scientist (1)	Nordisk Medicin (1)
Annals of Emergency Medicine (1)	Nurse Educator (1)
Annals of Internal Medicine (1)	Research in Higher Education (1)
Cambridge Quarterly of Healthcare Ethics (1)	The Psychological Report (1)
Canadian Medical Association Journal (1)	

Table 5. Journals with RRI articles, listed by number of articles.

broadly inclusive.

The need for interdisciplinary research raises one last observation about the RRI literature and by implication the RRI community. Most of the literature cited in this Report appears in biomedical journals. The only major exception are the eleven articles in Science and Engineering Ethics, which, it should be noted, are not indexed in MedLine® but are in BioEthicsLine, without abstracts. That research on the integrity of biomedical research (the primary focus of this report) appears in biomedical journals is certainly understandable, but the existence of this publication pattern raises serious questions for interdisciplinary research.

To be taken seriously in most academic settings today, researchers must first succeed in their primary research field. This means that sociologists must publish in sociology journals, psychologists in psychology journals, and so on. In addition, they must pursue research that is important to their primary fields of research. Institutional factors such as this unquestionably make the development of interdisciplinary research on research integrity more difficult. When added to the fact that there are few incentives for researchers who are the subject of RRI investigations to study their own integrity, rather than pursuing research in their primary fields of interest, establishing an interdisciplinary RRI initiative and RRI community poses a significant challenge.

Bibliography

1. Woolf P. Fraud in science: How much, how serious?. Hastings Cent Rep 1981;11:9-14.
2. Broad WJ, Wade N. Betrayers of the truth: Fraud and deceit in the halls of science. New York: Simon and Schuster; 1982.
3. Glick JL. On the potential cost effectiveness of scientific audits. Accountability in Research 1989;1(1):77-83.
4. Hals A, Jacobsen G. [Dishonesty in medical research. A questionnaire study among project administrators in Health Region 4]. Tidsskrift for den Norske Lægeforening 1993;113(25):3149-52.
5. Shapiro MF, Charrow RP. Scientific misconduct in investigational drug trials. N Engl J Med 1985;312(11):731-6.
6. Shapiro MF, Charrow RP. The role of data audits in detecting scientific misconduct. JAMA 1989;261(May 5):2505-11.
7. Shapiro MF. Data audits in investigational drug trials and their implications for detection of misconduct in science. In: Lock S, Wells F, editors. Fraud and miscondcut in medical research. London: BMJ Publishing Group; 1993. p. 128-41.
8. Weiss RB, Vogelzang NJ, Peterson BA, Panasci LC, Carpenter JT, Gavigan M, et al. A successful system of scientific data audits for clinical trials. A report from the Cancer and Leukemia Group B [see comments]. JAMA 1993;270(4):459-64.
9. Stern EB, Havlicek L. Academic misconduct: results of faculty and undergraduate student surveys. J Allied Health 1986;15(2):129-42.
10. Perry AR, Kane KM, Bernesser KJ, Spicker PT. Type A behavior, competitive achievement-striving, and cheating among college students. The Psychol Rep 1990;66(2):459-65.
11. Schab F. Schooling without learning: thirty years of cheating in high school. Adolescence 1991;26(104):839-47.
12. McCabe DL. Classroom cheating among natural science and engineering majors. Sci Eng Ethics 1997;3.
13. Stimmel B, Yens D. Cheating by medical students on examinations. Am J Med 1982;73(2):160-4.
14. Bailey PA. Cheating among nursing students. Nurse Educator 1990;15(3):32-5.
15. Rozance CP. Cheating in medical schools: implications for students and patients. JAMA 1991;266(17):2453, 6.
16. Anderson RE, Obenshain SS. Cheating by students: findings, reflections, and remedies. Acad Med 1994;69(5):323-32.
17. Daniel LG, Adams BN, Smith NM. Academic misconduct among nursing students: a multivariate investigation. J Prof Nurs 1994;10(5):278-88.
18. Baldwin DC, Jr., Daugherty SR, Rowley BD, Schwarz MR. Cheating in medical school: a survey of second-year students at 31 schools. Acad Med 1996;71(3):267-73.
19. Dans PE. Self-reported cheating by students at one medical school. Acad Med 1996;71(1 Suppl):S70-2.
20. Satterwhite WM, 3rd, Satterwhite RC, Enarson CE. Medical students' perceptions of unethical conduct at one medical school. Acad Med 1998;73(5):529-31.
21. Kalichman MW, Friedman PJ. A pilot study of biomedical trainees' perceptions concerning research ethics. Acad Med 1992;67(11):769-75.
22. Eastwood S, Derish P, Leash E, Ordway S. Ethical issues in biomedial research: Perceptions and practices of postdoctoral research fellows responding to a survey. Sci Eng Ethics 1996;2(1):89-114.
23. Sekas G, Hutson WR. Misrepresentation of academic accomplishments by applicants for gastroenterology fellowships [see comments]. Ann Intern Med 1995;123(1):38-41.
24. Gurudevan SV, Mower WR. Misrepresentation of research publications among emergency medicine residency applicants [see comments]. Ann Emerg Med 1996;27(3):327-30.
25. Bilge A, Shugerman RP, Robertson WO. Misrepresentation of authorship by applicants to pediatrics training programs. Acad Med 1998;73(5):532-3.
26. Panicek DM, Schwartz LH, Dershaw DD, Ercolani MC, Castellino RA. Misrepresentation of publications by applicants for radiology fellowships: is it a problem?

Am J Roentgenol 1998;170(3):577-81.
27. Dale JA, Schmitt CM, Crosby LA. Misrepresentation of research criteria by orthopaedic residency applicants. J Bone Joint Surg 1999;81(12):1679-81.
28. Tangney JP. Fraud will out—or will not? New Scientist 1987;115(August 6):62-3.
29. Korenman SG, Berk R, Wenger NS, Lew V. Evaluation of the research norms of scientists and administrators responsible for academic research integrity. JAMA 1998;279(1):41-7.
30. Committee on Science Engineering and Public Policy (U.S.). Panel on Scientific Responsibility and the Conduct of Research. Responsible science: Ensuring the integrity of the research process. Washington, D.C.: Nationl Academy Press; 1992.
31. Rennie D, Flanagin A. Congress on Biomedical Peer Review: history, ethics, and plans for the future [editorial]. JAMA 1998;280(3):213.
32. Pitkin RM, Branagan MA. Can the accuracy of abstracts be improved by providing specific instructions? A randomized controlled trial [see comments]. JAMA 1998;280(3):267-9.
33. Gardner MJ. An exploratory study of statistical assessment of papers published in the British Medical Journal. JAMA 1990;263(March 10):1355-57.
34. Eichorn P, Yankauer A. Do authors check their references? A survey of accuracy of references in three public health journals. Am J Public Health 1987;77(8):1011-2.
35. Evans JT, Nadjari HI, Burchell SA. Quotational and reference accuracy in surgical journals. JAMA 1990;263(10):1353-4.
36. Garfunkel JM, Ulshen MH, Hamrick HJ, Lawson EE. Problems identified by secondary review of accepted manuscripts. JAMA 1990;263(10):1369-71.
37. Stewart WW, Feder N. The integrity of the scientific literature. Nature 1987;325:207-14.
38. Garfunkel JM, Ulshen MH, Hamrick HJ, Lawson EE. Effect of institutional prestige on reviewers' recommendations and editorial decisions. JAMA 1994;272(2):137-8.
39. Joyce J, Rabe-Hesketh S, Wessely S. Reviewing the reviews: the example of chronic fatigue syndrome. JAMA 1998;280(3):264-6.
40. Link AM. US and non-US submissions: an analysis of reviewer bias. JAMA 1998;280(3):246-7.
41. Jadad AR, Cook DJ, Jones A, Klassen TP, Tugwell P, Moher M, et al. Methodology and reports of systematic reviews and meta-analyses: a comparison of Cochrane reviews with articles published in paper-based journals. JAMA 1998;280(3):278-80.
42. Dickersin K, Fredman L, Flegal KM, Scott JD, Crawley B. Is there a sex bias in choosing editors? Epidemiology journals as an example. JAMA 1998;280(3):260-4.
43. Dickersin K. The existence of publication bias and risk factors for its occurrence. JAMA 1990;263(10):1385-9.
44. Callaham ML, Wears RL, Weber EJ, Barton C, Young G. Positive-outcome bias and other limitations in the outcome of research abstracts submitted to a scientific meeting. JAMA 1998;280(3):254-7.
45. Misakian AL, Bero LA. Publication bias and research on passive smoking: comparison of published and unpublished studies. JAMA 1998;280(3):250-3.
46. McNutt RA, Evans AT, Fletcher RH, Fletcher SW. The effects of blinding on the quality of peer review. JAMA 1990;263(10, March 9):1371-76.
47. Black N, van Rooyen S, Godlee F, Smith R, Evans S. What makes a good reviewer and a good review for a general medical journal? JAMA 1998;280(3):231-3.
48. Callaham ML, Baxt WG, Waeckerle JF, Wears RL. Reliability of editors' subjective quality ratings of peer reviews of manuscripts. JAMA 1998;280(3):229-31.
49. Cho MK, Justice AC, Winker MA, Berlin JA, Waeckerle JF, Callaham ML, et al. Masking author identity in peer review: what factors influence masking success? JAMA 1998;280(3):243-5.
50. Godlee F, Gale CR, Martyn CN. Effect on the quality of peer review of blinding reviewers and asking them to sign their reports: a randomized controlled trial. JAMA 1998;280(3):237-40.
51. Justice AC, Cho MK, Winker MA, Berlin JA, Rennie D. Does masking author identity improve peer review quality? A randomized controlled trial. JAMA 1998;280(3):240-2.
52. van Rooyen S, Godlee F, Evans S, Smith R, Black N. Effect of blinding and unmasking on the quality of peer review: a randomized trial. JAMA 1998;280(3):234-7.
53. van Rooyen S, Black N, Godlee F. Development of the review quality instrument (RQI) for assessing peer reviews of manuscripts. J Clin Epidemio 1999;52(7):625-9.
54. Armstrong JS. Peer review for journals: Evidence on quality control, fairness, and innovation. Sci Eng Ethics 1997;3(1):63-84.
55. Weber EJ, Callaham ML, Wears RL, Barton C, Young G. Unpublished research from a medical specialty meeting: why investigators fail to publish. JAMA 1998;280(3):257-9.
56. Fletcher RH, Fletcher SW. Evidence for the effectiveness of peer review. Sci Eng Ethics 1997;3(1):35-50.
57. Friedman PJ. Correcting the literature following fraudulent publication. JAMA 1990;263(10, March 9):1416-19.
58. Garfield E, Welljams-Dorof A. The impact of fraudulent research on the scientific literature. JAMA 1990;263(10):1424-27.
59. Pfeifer MP, Snodgrass GL. The continued use of retracted, invalid scientific literature. JAMA 1990;263(10):1420-3.
60. Budd JM, Sievert M, Schultz TR. Phenomena of retraction: reasons for retraction and citations to the publications. JAMA 1998;280(3):296-7.
61. Parrish DM. Scientific misconduct and correcting the scientific literature. Acad Med 1999;74(3):221-30.
62. Snodgrass GL, Pfeifer MP. The characteristics of medical retraction notices. Bull Med Libr Assoc 1992;80(4):328-34.
63. Duggar DC, Christopher KA, Tucker BE, Jones DA,

Watson M, Puckett M, et al. Promoting an awareness of retractions: the Louisiana State University Medical Center in Shreveport experience. Medical Reference Services Quarterly 1995;14(1):17-32.
64. Editors ICoMJ. Uniform requirements for manuscripts submitted to biomedical journals. JAMA 1993;269:2282-6.
65. Flanagin A, Carey LA, Fontanarosa PB, Phillips SG, Pace BP, Lundberg GD, et al. Prevalence of articles with honorary authors and ghost authors in peer-reviewed medical journals. JAMA 1998;280(3):222-4.
66. Hoen WP, Walvoort HC, Overbeke AJ. What are the factors determining authorship and the order of the authors' names? A study among authors of the Nederlands Tijdschrift voor Geneeskunde (Dutch Journal of Medicine). JAMA 1998;280(3):217-8.
67. Shapiro DW, Wenger NS, Shapiro MF. The contributions of authors to multiauthored biomedical research papers. JAMA 1994;271(6):438-42.
68. Tarnow E. The authorship list in science: Junior physicists' perceptions of who appears and why. Sci Eng Ethics 1999;5(1):73-88.
69. Drenth JP. Multiple authorship: the contribution of senior authors. JAMA 1998;280(3):219-21.
70. Wilcox LJ. Authorship: the coin of the realm, the source of complaints. JAMA 1998;280(3):216-7.
71. Health NIo. Guidelines for the Conduct of Research in the Intramural Research Program at NIH. [Bethesda, MD]: [National Institutes of Health]; 1997.
72. Waldron T. Is duplicate publishing on the increase? Br Med J 1992;304:1029.
73. Barnard H, Overbeke AJ. [Duplicate publication of original manuscripts in and from the Nederlands Tijdschrift voor Geneeskunde]. Ned Tijdschr Geneeskd 1993;137(12):593-7.
74. Koene HR, Overbeke AJ. [The ultimate fate of articles rejected for publication in the Nederlands Tijdschrift voor Geneeskunde]. Ned Tijdschr Geneeskd 1994;138(49):2443-6.
75. Blancett SS, Flanagin A, Young RK. Duplicate publication in the nursing literature. Image: The Journal of Nursing Scholarship 1995;27(1):51-6.
76. Bloemenkamp DGM, Walvoort HC, Hart W, Overbeke AJPM. [Duplicate publication of articles in the Dutch Journal of Medicine in 1996]. Ned Tijdschr Geneeskd 1999;143(43):2150-3.
77. Tramer MR, Reynolds DJ, Moore RA, McQuay HJ. Impact of covert duplicate publication on meta-analysis: a case study. Br Med J 1997;315(7109):635-40.
78. Jefferson T. Redundant publication in biomedical sciences: Scientific misconduct or necessity? Sci Eng Ethics 1998;4(2):135-40.
79. Merton RK. A note on science and democrary. Journal of Legal and Political Sociology 1942;1(1-2):115-26.
80. Jasanoff SJ. Is science socially constructed: And can it still inform public policy? Sci Eng Ethics 1996;2(3):263-76.
81. Mahoney MJ. Publication prejudices: An experimental study of confirmatory bias in the peer review system. Cognitive Therapy and Research 1977;1:161-75.
82. Dickersin K, Min Y-I, Meinert CL. Factors influencing publication of research results. Follow-up of applications submitted to two institutional review boards. JAMA 1992;267(3):374-8.
83. Campbell TID. Understanding the potential for misconduct in university-industry relationships: An empirical study. In: Braxton JM, editor. Perspectives on scholarly misconduct in the sciences. Columbus, OH: Ohio State University Press; 1999. p. 259-82.
84. Boyd EA, Bero LA. Assessing Faculty Financial Relationships With Industry : A Case Study. JAMA 2000;284:2209-14.
85. Cho MK, Shohara R, Schissel A, Rennie D. Policies on faculty conflicts of interest at US universities. JAMA 2000;284:2203-8.
86. Health NIo. Reminder and update: Requirement for instruction in the responsible conduct of research in national Research Service Award institutional training grants. NIH Guide for Grants and Contracts 1992;21:2-3.
87. Mastroianni A, Kahn JP. Encouraging accountability in research: A pilot assessment of training efforts. Accountability in Research 1999;7:85-100.
88. Brown S, Kalichman MW. Effects of training in the responsible conduct of research: A survey of graduate students in experimental sciences. Sci Eng Ethics 1998;4(4):487-98.
89. Elliott D, Stern JE. Evaluating teaching and students' learning of academic research ethics. Sci Eng Ethics 1996;2(3):345-66.
90. Self DJ, Wolinsky FD, Baldwin DC, Jr. The effect of teaching medical ethics on medical students' moral reasoning. Acad Med 1989;64:755-9.
91. Baldwin DC, Jr., Daugherty SR, Self DJ. Changes in moral reasoning during medical school. Acad Med 1991;69(9 Suppl):S1-3.
92. Self DJ, Schrader DE, Baldwin DC, Jr., Wolinsky FD. A pilot study of the relationship of medical education and moral development. Acad Med 1991;66(10):629.
93. Self DJ, Schrader DE, Baldwin DC, Jr., Root SK, Wolinsky FD, Shadduck JA. Study of the influence of veterinary medical education on the moral development of veterinary students. J Am Vet Med Assoc1991;198(5):782-7.
94. Self DJ, Baldwin DC, Jr., Wolinsky FD. Evaluation of teaching medical ethics by an assessment of moral reasoning. Med Educ 1992;26(3):178-84.
95. Self DJ, Schrader DE, Baldwin DC, Jr., Wolinsky FD. The moral development of medical students: a pilot study of the possible influence of medical education. Med Educ 1993;27(1):26-34.
96. Bebeau MJ, Tohma SJ. The impact of a dental ethics curriculum on moral reasoning. J Dent Educ 1994;58(9):684-92.
97. Self DJ, Baldwin DC, Jr., Wolinsky FD. Further exploration of the relationship between medical education and moral development. Camb Q Healthc

Ethics 1996;5(3):444-9.
98. Anderson MS, Louis KS, Earle J. Disciplinary and departmental effects on observations of faculty and graduate student misconduct. In: Braxton JM, editor. Perspectives on scholarly misconduct in the sciences. Columbus, OH: Ohio State University Press; 1999. p. 213-35.
99. Clarke M, Chalmers I. Discussion sections in reports of controlled trials published in general medical journals: islands in search of continents? JAMA 1998;280(3):280-2.
100. Junker CA. Adherence to published standards of reporting: a comparison of placebo-controlled trials published in English or German. JAMA 1998;280(3):247-9.
101. Begg C, Cho MK, Eastwood S, Horton R, Moher D, Oking I, et al. Improving the quality of reporting of randomized controlled tirals: the CONSORT statement. JAMA 1996;276:637-9.
102. Pitkin RM, Branagan MA, Burmeister LF. Accuracy of data in abstracts of published research articles. JAMA 1999;281:1129-30.
103. Pitkin RM, Branagan MA, Burmeister LF. Effectiveness of a journal intervention to improve abstract quality. JAMA 2000;283(4):481.
104. Jefferson T, Smith R, Yee Y, Drummond M, Pratt M, Gale R. Evaluating the BMJ guidelines for economic submissions: prospective audit of economic submissions to BMJ and The Lancet. JAMA 1998;280(3):275-7.
105. Lock S, Wells F. Fraud and misconduct in medical research. London: BMA Publishing Group; 1993.
106. Braxton JM, ed. Perspectives on Scholarly Misconduct in the Sciences. Columbus, OH: Ohio State University Press; 1999.

I. Norms and Environmental Issues

1. Students and Mentors

What Would Get You in Trouble: Doctoral Students' Conceptions of Science and Its Norms

Melissa S. Anderson, Educational Policy and Administration, University of Minnesota, USA

Key Words: Graduate education, Graduate students, Norms

Undergirding the academic enterprise is a web of assumptions about how the members of the academic community should conduct their professional lives. These assumptions are expressed in ways ranging from the most explicit directives (legal, institutional, contractual) to the implicit, taken-for-granted understandings that facilitate everyday interactions among members of the profession. They constitute the normative underpinnings of the academic profession.

Braxton and Bayer define norms as "shared beliefs within a particular social or professional group about behavior expected or desired in a given situation or circumstance" (1). In the academic context, the four norms that Robert Merton (2) identified in his 1942 analysis–universalism, communality [to use Barber's (3) term], disinterestedness, and organized skepticism–have framed much of the subsequent research. They figured prominently in Zuckerman's seminal analyses of the social system of science (4, 5). They are also reflected in Mitroff's (6) "counternorms", and they together capture most of the considerable literature that Braxton (7) compiled on the subject of norms.

Others, however, have argued for a more complex understanding of norms. Mulkay, for example, has claimed that norms are best understood as ideologies or "evaluative repertoires" (8). That is, norms constitute a kind of standardized narrative that academics use to describe and evaluate behavior and to prescribe responses to certain behaviors (8). Ajzen and Fishbein have described the significance of "subjective norms" that reflect what others, who are important to an individual, think he or she should do (9). From this perspective, neither an abstract normative system or an individual's own internalized norms are as important as the individual's understanding of others' expectations. Finally, Braxton and Bayer have demonstrated how a combination of inductive and survey-based strategies could uncover a complex set of norms in collegiate teaching (1).

The present study takes a different approach to the norms of the academic profession, with corresponding implications for the design of the study. First, it emphasizes the implicit over the explicit, on the assumption that implicit norms can be particularly powerful in shaping behavior. This study therefore relies on narrative descriptions of norms, instead of on a particular formulation of the normative structure of academia. It is rooted in the proposition that more attention needs to be paid to understanding science and its ethical aspects from the "inside out," that is through the experiences of scientists themselves (10-12). It therefore responds to Braxton's call for study of norms "expressed in the words of the respondents rather than in a priori definitions of possible norms" (7).

Second, it assumes that norms of a group are particularly salient to newcomers during a socialization period (13). The data for this study accordingly come from first-year doctoral students, who are encountering professional norms in intensive ways. Their experiences are likely to produce

Corresponding author: Melissa S. Anderson, Educational Policy and Administration, 330 Wulling Hall, University of Minnesota, Minneapolis, MN, 55455, 612-624-5717 (voice), 612-624-3377 (fax), mand@umn.edu.

"contrast" in the gestalt sense through the process of "sense-making", which highlights the normative insights they acquire (14).

Third, the study assumes no necessary match among students' understanding of the broad norms of the academic profession, the norms that they have internalized and view as most salient, and the behavior of professional colleagues. This study therefore explores levels of consonance and dissonance that students perceive among these three phenomena.

Fourth, this study relies on Durkheim's useful proposition that norms are recognized when they are violated (15). The questions used in this study to elicit students' views of norms, therefore, ask students to contrast their views of general academic norms, as well as the norms to which they subscribe, against the behavior of their colleagues.

Methods

These parameters gave shape to the current study, which is part of a broader project on doctoral education, the Academic Life Project, funded by the National Science Foundation (Grant number 9408S08622). Participants for the current analysis were 30 first-year doctoral students in seven science and social science disciplines at a major research university. (The project will eventually involve over 100 interviewees and will be longitudinal.) Semi-structured interviews of approximately a half-hour yielded narrative data on norms and related topics.

A series of questions in the interviews asked students to consider and comment on relationships between academic norms and behavior (Do you see any conflicts between what people think or say you should do and the way work is actually done?), between their own perspectives and behavior (Do you see people around here acting contrary to your advice [to doctoral students on how to avoid serious mistakes]?), and between their own normative perspectives and academic norms (Are there any ideas or rules about how you should do your work that you don't agree with?). These questions highlighted students' understandings of academic research as a social enterprise whose membership they are entering. Those who articulated a more complex normative perspective showed greater awareness of the social aspects of the scientific enterprise and a more constructivist approach to knowledge development in the sciences. They were also less troubled by dissonance between behaviors and norms, recognizing the inevitable roles played by mistakes, errors of fact and of judgment, and mid-course corrections.

Results

Students' conceptions of norms that underlie their work are presented here in terms of the three contrasts identified above. First, students' conceptions of general academic norms are described in light of the behavior of their colleagues. Then the norms to which they subscribe are seen in contrast, again, to colleagues' behavior. Finally, what they understand to be academic norms are contrasted to their own normative orientations.

Correspondence between academic norms and behavior. The first comparison investigated is between students' conceptions of the norms of their fields and the behaviors of those around them. The interview question was, "Do you see any conflicts between what people think or say you should do and the way work is actually done?"

Approximately two-thirds of those interviewed saw no conflict between prescribed and actual behavior among their colleagues. Most saw no disjuncture; a few were more definite: "No, I mean, an emphatic no with the faculty," and, "They're pretty straightforward, and they'll pretty much hold true to their word." Two students noted that, while they were not aware of conflict between norms and action, they did not really know enough about what people were doing in the department to comment generally about people's behavior; as one put it, "I'm not privy to a lot of the goings on of the department."

Five students noted particular areas of disjuncture between norms and behavior. One mentioned safety rules:

> We all have to go to this safety training before we are allowed to go in the lab. It's just kind of a refresher course every year. And then ... they always say practically nothing is supposed to go down the drain. And sometimes stuff does. But we're not even supposed to put things like ... a simple rinsing agent down the drain ... but it happens all the time.

This student went on to affirm the importance of the safety rules for two reasons: first, that safety supports proper procedures ("if you don't do it right, it doesn't work"), and second, not following these rules is dangerous ("if you don't follow the rules in terms of safety, in terms of

correct procedure, usually that means that the chemist should not work in the lab").

A second point of conflict observed by a student is in admissions procedures for the graduate students in the department. From the vantage point of a place on the departmental graduate admissions committee, the student saw that, though the department touts a highly selective admissions policy, the process is influenced in political ways by influential individuals on the faculty. The result is that the department admits less-qualified people than its policy would suggest.

The third area of dissonance between prescribed and enacted behaviors is in research. One psychology student focused on experiments:

> We talk a lot about being a very experimental field and it's all about experiments, but it's so difficult to run experiments now with getting through the IRB [Institutional Review Board] and getting subjects.... [I]t's so much easier to pass out some sort of survey or some sort of questionnaire. And so we talk about the experiment and how wonderful it is, and then we don't do it.

Two other students also mentioned research, but in a different way. They clearly understood the faculty's focus on research, but they did not see faculty providing enough support to students to get them started on their own research. As one put it, "I think [it's] the absence of direction which is noticeable, which stands out. And I think some students have felt ... you know, they're sort of cast adrift, in some sense, and left to figure everything out for themselves." The other student described her frustration with the research imperative in light of the same kind of lack of direction:

> There almost seems like there's kind of pressure or an expected norm within the department itself that we get involved with research. Yet, in our specific discipline, in our area, there hasn't been very much guidance or, you know, pressure to do that.... I have met with my advisor twice on my own volition — and going to see her and saying, "Okay. Maybe it's time for me to get involved in research," and each time she has not had a specific project that really had any place for me to start.... And I just kind of walked away from it feeling like, just thinking that she had just so much going on already — and really, you know, like almost I kind of felt like I would be a burden to get involved at that point.

Correspondence between subscribed norms and behavior. The second comparison addressed in the interviews is between the norms to which students themselves subscribe and the behavior of their colleagues. Here the question is whether or not students see people around them acting contrary to the way the students think they should act. Employing Durkheim's view that norms are best recognized when violated, the interview protocol invited students to consider what they would advise incoming doctoral students to do to stay out of trouble in their work (15). Responses demonstrate students' personally held beliefs about how first year students should act, identified here as subscribed norms. Students were then asked, as follow-up questions, "Do you see people around here acting contrary to your own advice? What are they doing?"

Responses to these questions fall into three general categories: tasks, relationships, and ethics. Most of the responses addressed the work of graduate students. Several talked about the need for students to take responsibility for their own work and progress. As one put it, "I mean, in our department, it's a problem both with the students not taking the initiative to getting all of their requirements and prelims done and also, with our department, no one says anything if it takes you longer." Others disapproved of student colleagues' not getting their work done, taking too much time to get through their work, or abandoning the work altogether. All of these students clearly demonstrated a strong commitment to hard work and a sense that some others around them acted contrary to this subscribed norm.

Not only do students believe in getting the work done, but several mentioned the need to do independent work. One science student complained,

> I think one of the biggest mistakes that they could make is to do something that is not independent. I see a lot of people that are working with their advisors and really, ... I don't know the best way to describe this without sounding mean, but they just have no interest of their own. They are just a, like a little off-shoot of their advisor, like a little worker.... They're not independent at all.... You know, what they do is what their advisor says, and I think that's a really big mistake, because one day you can look back and be, like, "Oh. This isn't what I wanted to do at all, and if I had the choice I would have done it completely differently."

Taking the initiative for an independent stream of

inquiry is a step beyond responsibility for getting work done, a step that some, but not all, first-year graduate students take.

One student's story about a graduate-student peer illustrates her struggle with maintaining her independence in inquiry. The peer in question is someone she respects.

> But the problem is, he comes from a different undergraduate background, not an American system. He's from a different country, where being the best in the class was very much recognized and very much rewarded, and so he was the best in his class. And so he came here.... Everyone has been asking him for help, and so he would do all of his work way in advance — which was commendable — but then he would — instead of working and taking other people's suggestions and trying to integrate everything when we were working on problem sets — he would be, like, "This is right. I have the answer." And usually he did. Usually he was right. But it was annoying to work with him.... There were times where even though I knew I would probably get a better grade if I worked with him, because he would have the answers, I wouldn't want to do it. And also, you don't want the answers given to you.

Comments about relationships comprise the next category of responses about the contrast between subscribed norms and behavior. Students demonstrate clear ideas about how people should behave toward each other in the graduate school setting. Some mentioned the importance of having an advisor with whom the student can work. They described examples of advisors who were not supportive of their students. This behavior that ran contrary to their beliefs about how advisors are to act met with very strong negative reactions.

Other respondents showed a keen sense of the importance of making a good impression and expressed dismay that some of their peers did not appear to understand this point. A science student said,

> I know there's some people who, whenever there was an exam, they just didn't go into the lab all the time, and I don't think it left a good impression on some people who were working in the lab, working around them.... So if you don't seem very serious about your lab work, then they — someday when you have to go to them for advice or something — they're not necessarily drawn to give you as much time and make as much of a serious effort.

Another student described impression-management in blunt terms as a quid pro quo:

> I guess, just like, you have to do things for people so they'll do things for you later. I guess that doesn't even sound that bad. But more like — I can't think of a particular example — but just basically doing things that you don't want, because you know later it'll get you something you do want.

Not only are students aware of the work imperative, but they are also aware of the need for others to know that they subscribe to it. As the quotes illustrate, the norm bears both sanction and reward. This norm illustrates students' movement toward full acceptance into the academic social world.

The third contrast between behavior and students' own normative orientations was in the area of ethics. Those who mentioned ethics said that they had seen no instances of people acting contrary to what they themselves understood to be appropriate behavior. One said, "I've never seen anyone falsifying data, which is very, very good. And I believe that we don't have the second problem, fishing for data. At least in my group, we don't have that." Another noted, "I haven't seen, I haven't heard of anybody lying about stuff or trying to falsify results." This science student went on to describe how important it is for students to acknowledge mistakes, so that they are not interpreted as more serious offenses: "Everybody makes mistakes.... Everyone's pretty understanding of when your experiments don't work or when you did a stupid mistake or whatever."

The normative understandings that the doctoral students reveal through their comments on the contrast between what peers should do and what they are actually doing thus center largely on their work and their relationships with colleagues. That is, they appear attuned to both functional and social norms of academic life. The next step is to contrast their own normative orientations to what they perceive to be the general norms of their fields.

Contrast between academic norms and subscribed norms. Students' perceptions of prevalent academic norms may not match their own ideas about how they should conduct themselves in the academic world. As both academic norms and subscribed norms can be brought into focus by contrasting them against behavior, so they can be clarified by comparing them to each other. The relevant question on the interview protocol was, "Are there any ideas or rules about how you should do your work that you don't agree with?"

The task-related points of disjuncture fell generally in the category of competition and its attendant work pressures. A student in a social science department commented,

> Everyone's competing for jobs in academic environments primarily.... And I guess what that means for many students is they have to adapt to a competitive type of atmosphere and in some cases be more competitive than they would like to be in order to further their goals further on. And I think that might be disheartening for some students.... And I think all of the students ... try to be good-natured about the entire thing, but I think the pressure of continuing to get as many publications as you can is the reality that dawns on a lot of students — something they didn't anticipate, necessarily, early on.

Another student talked about competitive pressures to publish in terms of "the whole production thing" and the "assembly line production attitude."

Several students complained about the work loads they bear in terms of the mismatch between their professors' views on how much work they should do and their own. A science student talked about peers who never take time off and "work themselves to death" to live up to what they perceive as the standards of work in the field; the student said he would never do that. Another commented on prevalent norms for the quality of a dissertation. In this students' relatively new field in science, it was generally expected, 10 or 20 years ago, that each dissertation would open up a completely new field of inquiry; now, the expansion of the discipline and the far greater competition due to a more crowded field make it much harder to have such an impact through doctoral work. The student noted, though, that normative understandings in the field had not changed in response.

Another point of contrast related to competition is the matter of independent work. Several students mentioned that at least some of their professors require independent, as opposed to collaborative, work on assignments in graduate courses. Many of the students were previously socialized to collaborative norms, and they found the professors' insistence on individual work counterproductive. Here students' normative orientations run counter to the academy's norms of rewarding people on the basis of individual achievement and independent contributions.

Beyond students' attention to task-related disjunctures between academic and espoused norms, the most striking pattern in students' responses is their uncertainty about academic norms in general. Most of them are keenly aware that norms vary by discipline or even from one research group to another. For example, one noted, "Everyone has such different views about how to do things." Another put it this way: "Each professor sort of has their own research policy. And that's academia. They have the freedom to make up the rules of their group, within certain bigger boundaries that the school sets." Yet another respondent said, "I don't think there are very many rules about how we should conduct our research, other than the whole basic 'Be ethical and stuff.' I don't observe very many rules about how we should conduct the research." This student went on to mention that she might change her mind as she got further into her research, when she would have to remember all the rules about where to put commas — thereby illustrating just how far she had to stretch to think of general norms of the field.

Perhaps some of the uncertainty that students expressed about academic norms is related to the ways in which such norms are communicated. The student quoted above who mentioned each professor having his or her own research policy went on to say, "Ideally, it should be talked about as a research group as a whole, but it seems to me that a lot of stuff is just sort of telephone, where one person tells another person, and that person tells the next person." Another talked about his reluctance to ask people how things should be done in the lab:

> The approach towards how you learn your way around the lab is you just go in there and you do it. As far as being taught or having anyone specifically to show you around, you really don't, because everyone in there is really, really busy, because they are doing research. And they don't want to take time out of their research to show you how to work [a machine], because it's such a simple thing to them, and they get really frustrated and impatient with someone who is just learning how to use it. And so, generally you just have to go in there and learn on your own.... I almost felt afraid to go to other people in the group with my stuff, because I don't want to waste their time and I don't want to feel stupid either.

Of course, some students were unable to identify any dissonance between the norms to which they subscribe and the more general academic norms

as they see them. One person wryly commented on the thoroughness of his own socialization to the general normative structure of the field: "Maybe I've been just so well trained that I don't know anything anymore."

The results in this section show, as did the earlier results, that students' normative conceptions are dominated by functional or task-related norms. They also show a general awareness among students of social norms, though their conceptions of norms for interpersonal relations are not as fully developed as their views on functional norms.

Discussion

The findings presented here contribute to our understanding of doctoral students' initial normative orientations. Students' conceptions of normative imperatives are relevant to policy initiatives that are currently receiving a great deal of attention. The federal Office of Research Integrity recently announced a major new initiative that will focus on the promotion of the responsible conduct of research. The American Educational Research Association is currently preparing to publish a book that will direct attention to the AERA Code of Ethics and its use. Dozens of other academic associations are writing or revising their codes of ethics, and virtually every major research university has addressed its institutional policies on ethics and misconduct in the past five years. The federal government is seeking to expand its requirements for formal training in ethics beyond those for trainees covered by National Institutes of Health funding. Most of the attention to expanded training in ethics and related issues focuses on graduate students and other newcomers to the academic profession.

Continued self-regulation by the scientific community depends on the ongoing renewal of normative conceptualizations that, through their generational evolution, continue to reflect the expectations of society for science. Most of the emerging initiatives are driven, however, by a sense of urgency or by federal regulations and directives, without attention to doctoral students' understanding of science, academic life, and the norms of their disciplines. Neither do they reflect ways in which newcomers interact with and shape the normative bases of their fields (16).

This study serves as a window onto the normative assumptions of science, but it furthermore suggests ways in which those norms can be communicated within and beyond the scientific community (17, 18). The doctoral students interviewed reveal the norms of science as they understand them, during a period when they are intensely and reflectively face-to-face with the way science works. They are the future membership of the scientific community, but they are also current participants in the enterprise, struggling with their own ideas of how they should behave as scientists.

The results of the interviews demonstrate intriguing patterns of dissonance among the three phenomena examined. The interview responses show that students' normative conceptualizations are dominated by functional (task-related) norms, as we might expect from earlier work on anticipatory socialization that emphasizes survival in the graduate or professional-school setting (16). Augmenting the functionalist perspective, however, are emergent conceptualizations of social and ethical norms.

The inchoate nature of first-year students' personal normative orientations suggests that approaches to socialization of doctoral students to academic life, particularly in the areas of ethics and related issues, may overestimate the extent of students' understanding of the academic system, the nature of research, and the place of individual academics in the broader context of research. Students interviewed here showed very little awareness of their disciplines, beyond their own work, or of the higher education system, beyond their own departments. The imperatives they identified have to do generally with the work at hand and the people with whom they interact.

Socialization to the field and to the normative bases of research in a discipline should be grounded in the academic world with which these students are familiar, while at the same time introduce them to the broader academic environment. The theme of individual, independent work that runs through these interviews suggests that students might not be subject to as much osmotic group socialization as many faculty assume. It is also clear that the channels by which socialization to the normative aspects of academic life are communicated are primarily informal. Calls for more formal, more deliberate approaches to normative socialization find support in the vagueness with which students conceptualize the norms that underlie academic research.

Aknowledgements

The author gratefully acknowledges the research assistance of Janet M. Holdsworth and Jennifer Milleville.

Bibliography

1. Braxton JM, Bayer AE. Faculty misconduct in collegiate teaching. Baltimore: The Johns Hopkins University Press; 1999.
2. Merton RK. A note on science and democracy. J of Legal and Political Sociology 1942; 1(1-2):115-126.
3. Barber B. Science and the social order. New York: Free Press; 1952.
4. Zuckerman HE. Deviant behavior and social control in science. In: Sagarin E, editor. Deviance and social change. Beverly Hills (CA): Sage; 1977. p. 87-138.
5. Zuckerman HE. The sociology of science. In: Smelser NJ, editor. Handbook of sociology. Newbury Park (CA): Sage; 1988. p. 511-574.
6. Mitroff I. Norms and counter-norms in a select group of the Apollo moon scientists: A case study of the ambivalence of scientists. Amer Sociological Rev 1974; 39, 579-95.
7. Braxton JM. The normative structure of science: Social control in the academic profession. In: Smart JC, editor, Higher education: Handbook of theory and research, v. 2. New York: Agathon Press; 1986.
8. Mulkay, M. Norms and ideology in science. Soc Sci Info 1976; 15(4-5): 637-56.
9. Ajzen I, Fishbein M. Understanding attitudes and predicting social behavior. Englewood Cliffs (NJ): Prentice-Hall; 1980.
10. Anderson MS. (1999). Uncovering the covert: Research on academic misconduct. In: Braxton JM, editor. Perspectives on scholarly misconduct in the sciences. Columbus (OH): Ohio State University Press; 1999.
11. Knorr-Cetina KD, Mulkay M, editors. Science observed: Perspectives on the social study of science. Beverly Hills (CA): Sage; 1983.
12. Latour B, Woolgar S. Laboratory life: The construction of scientific facts. Princeton, NJ: Princeton University Press; 1986.
13. Wanous JP. Organizational entry: Recruitment, selection, orientation, and socialization of newcomers, 2nd ed. Reading (MA): Addison-Wesley; 1992.
14. Louis MR. Surprise and sense making: What newcomers experience in entering unfamiliar organizational settings. Admin Sci Q 1980; 25(2):226-251.
15. Durkheim, E. The elementary forms of religious life. Fields KE, translator. New York: Free Press; 1995[1912].
16. Tierney WG., Rhoads RA. Enhancing promotion, tenure, and beyond: Faculty socialization as a cultural process. ASHE-ERIC Higher Education Report, no. 93-6. Washington, D.C.: The George Washington University; 1993.
17. Lapidus JB., Mishkin B. Values and ethics in the graduate education of scientists. In: Nerad M, June R, Miller DS, editors. Graduate education in the United States. New York: Garland; 1997.
18. Bird SJ. Including ethics in graduate education in scientific research. In: Braxton JM, editor. Perspectives on scholarly misconduct in the sciences, Columbus (OH): Ohio State University Press; 1999.

Data Manipulation in the Undergraduate Laboratory: What are we teaching?

Elizabeth W. Davidson, Department of Biology, Arizona State University, USA

Heather E. Cate, College of Liberal Arts and Sciences, Arizona State University, USA

Cecil M. Lewis, Jr., Department of Anthropology, University of New Mexico, USA

Melanie Hunter, Department of Biology, Arizona State University, USA

Keywords: Biology, Cheating, Chemistry, Motivation, Undergraduate research

Arizona State University (ASU) offers a senior-level course entitled "Professional Values in Science" that addresses a number of topics concerning ethical conduct of research as well as ethical concerns at the intersection of science and society. The course demands active participation by the students. Several years ago, on his own initiative a student in the class developed a questionnaire that explored data manipulation. As most of the students were undergraduates, the questionnaire focused upon manipulation of data in undergraduate science laboratories. We were startled to discover that over 60% of the students openly admitted to manipulation of data in undergraduate laboratories. These results led to the development of a more elaborate survey that has been administered to 7 undergraduate Biology and Chemistry courses, enrolling a total of over 700 students. The courses include both major and nonmajor subjects, at both introductory and upper division level. Arizona State University has approximately 50,000 students, including (in academic year 2000) ca. 1000 majors in Biology and 250 majors in Chemistry. In the fall semester, 2000, 3137 undergraduates are enrolled in Biology courses, while 3355 undergraduates are enrolled in Chemistry courses. Laboratories are therefore limited in available time, are generally supervised and graded by graduate teaching assistants, and many, but not all, of these courses rely upon traditional laboratory exercises.

Methods:

The survey and instructions to students are presented in at the end of the paper. Students were advised by the person administering the survey (who was not their course professor or teaching assistant) that the results would be held anonymous and would not affect their grade. The courses included Chemistry 115: Introductory, non-majors; Chemistry 335: Organic, non-majors; Biology 201: Anatomy and Physiology, non-majors; Biology 100: Introductory, non-majors; Biology 182: Introductory, majors; Biology 193: Introductory, majors, critical thinking focus; Biology 385: Invertebrate Zoology, majors. Seven hundred and two students participated. Institutional Human Subjects committee approval was obtained. Data were analysed by Spearman correlation.

Corresponding author: Elizabeth W. Davidson, Department of Biology, Arizona State University, Box 871501, Tempe, AZ 85287-1501, 480-965-7560 (voice), 480-965-2519 (fax), e.davidson@asu.edu.

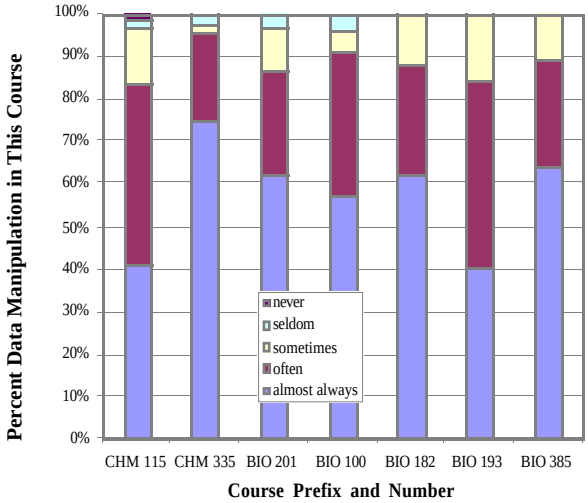


Figure 1. Results of survey, Question 5, "Have you ever manipulated data or made up data in this course?" CHM 115: Introductory, non-majors, N=86; CHM 335: Organic, non-majors, N=44; BIO 201: Anatomy and Physiology, non-majors, N=29; BIO 100: Introductory, non-majors, N=200; BIO 182: Introductory, majors, N=40; BIO 193: Introductory, majors, critical thinking focus, N=57; BIO 385: Invertebrate Zoology, majors, N=64. N= total number of responses to the specific question.

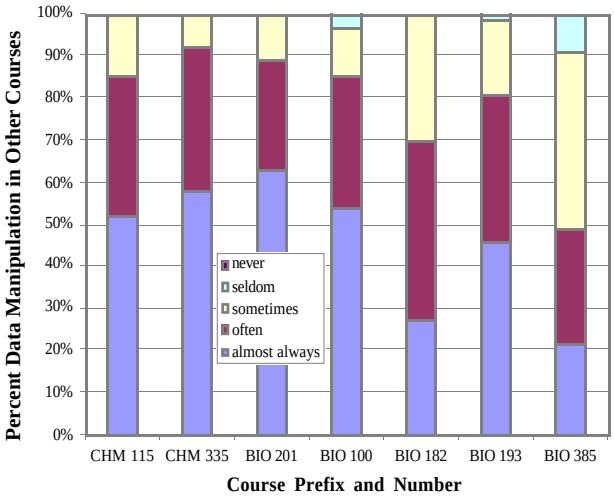


Figure 2. Results of survey, Question 10, "Have you ever manipulated or made up data in any other *science course?" CHM 115, N=87; CHM 335, N=52; BIO 201, N=27; BIO 100, N=81; BIO 182, N=40; BIO 193, N=57; BIO 385, N=66. N= total number of responses to the specific question.*

Results

The key question in this survey was Question 5, "Have you ever manipulated data in this course?" As shown in Figure 1, between 40.4 and 75% of students in the surveyed course admitted to manipulating data "almost always," and another 20-43.9% admitted to such manipulation "often." Students reporting data manipulation "seldom" represented less than 5% of those surveyed, and only one student out of over 500 who responded to this question replied "never." Using correlation analysis, we learned that admission of manipulation in the course surveyed was strongly correlated to admission of manipulation in other courses (Spearman Correlation Sig. (2 tailed) 0.355, significant at 0.01 level) (Figure 2).

We asked whether data manipulation was related to the level (i.e. introductory vs. advanced) of the course, and whether the course was designed for majors or non-majors. No significant difference was found between data manipulation in Introductory Biology BIO 100 (non-majors) and BIO 182 (majors) or between these lower division courses and an upper division course, BIO 385 (Invertebrate Zoology, majors). We compared responses from BIO 182, a traditional introductory course, to BIO 193, an introductory majors course with emphasis on critical thinking. The smallest percentage of students reporting data manipulation "almost always" was in BIO 193, however a large proportion of the remainder reported manipulation "often" (Figure 1). Within the two non-majors chemistry courses surveyed, less data manipulation was found in CHM 115 (Introductory) than in CHM 335 (Organic), and indeed the highest overall reported manipulation (90.5% "almost always" or "often") was reported in Organic Chemistry. Conversations with students in the Professional Values in Science class and elsewhere confirmed that many have

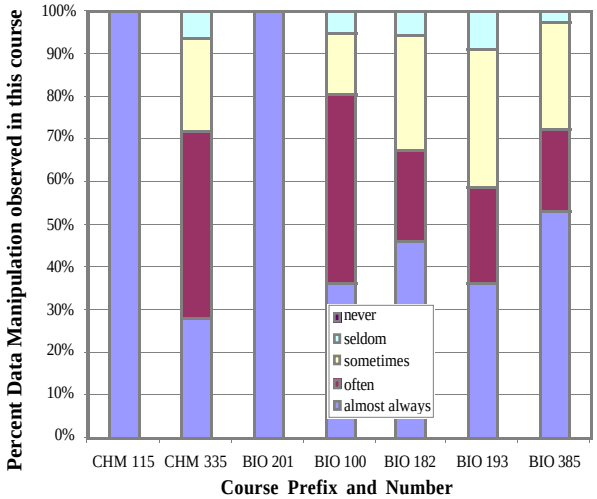


Figure 3. Results of survey, Question 7, "Have you ever observed anyone else *manipulate or make up data in* this course*?" CHM 115, N=91; CHM 335, N=67; BIO 201, N=28; BIO 100, N=237; BIO 182, N=40; BIO 193, N=56; BIO 385, N=66. N= total number of responses to the specific question.*

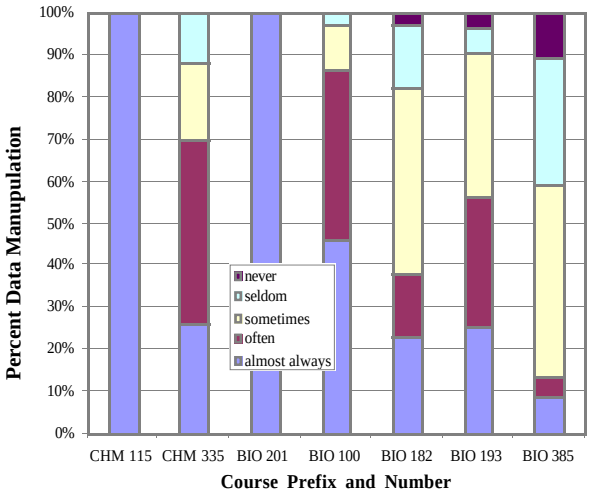


Figure 4. Results of survey, Question 14, "Have you ever observed anyone manipulate or make up data in any *science course?" CHM 115, N=94; CHM 335, N=70; BIO 201, N=30; BIO 100, N=96; BIO 182, N=39; BIO 193, N=55; BIO 385, N=66. N= total number of responses to the specific question.*

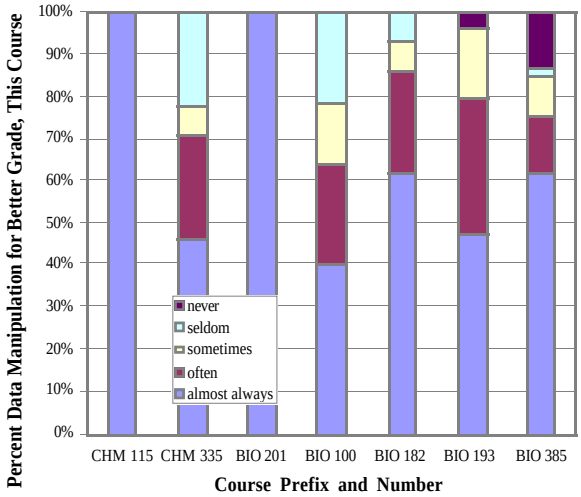


Figure 5. Results of survey, Question 6, "If you have ever manipulated data or made up data, was it motivated by the thought of a better grade?" CHM 115, N=69; CHM 335, N=41; BIO 201, N=17; BIO 100, N=246; BIO 182, N=31; BIO 193, N=55; BIO 385, N=53. N= total number of responses to the specific question.

manipulated data for Chemistry laboratory reports, particularly in Organic. Little difference in data manipulation (Question 5) was found when analyzed by academic year or by gender.

Two other key questions were 7 and 14, which asked whether the student had observed others manipulating data. The results from these questions were less consistent than responses about the students own data manipulation. Two courses (CHM 115 and BIO 201) received an "almost always" response rate of 100%, whereas in other courses a much smaller proportion of students responded "almost always" (Figures 3, 4).

We investigated motivation for data manipulation with questions 6 and 11, which asked whether the students manipulated data in order to get a better grade. Up to 100% of students in some courses replied that manipulation was almost always performed to obtain a better grade (Spearman Correlation Sig. (2-tailed) 0.265, significant at 0.01 level) (Figure 5; data from Question 11 not shown). When asked whether this was because they felt that their grade depended heavily upon the experimental results (Questions 8 and 15), less than half of students felt that their grade in the current course depended on experimental results "almost always", and from 3.0 to 13.6% of the students replied to Question 8 that their grade "seldom" depended on results (Figure 6, data from Question 15 not shown; Spearman (Figure 7, data from Question 16 not shown; Spearman correlation 0.368, significant at 0.01 level). Finally we surveyed student preferences for type of laboratory experiments (Question 17). In all seven courses combined, only 1.7% of students preferred lab experiments which place more emphasis on results, whereas 53.5% preferred more emphasis to be placed upon processes, and 44.7% preferred a balanced combination of both techniques (N=503).

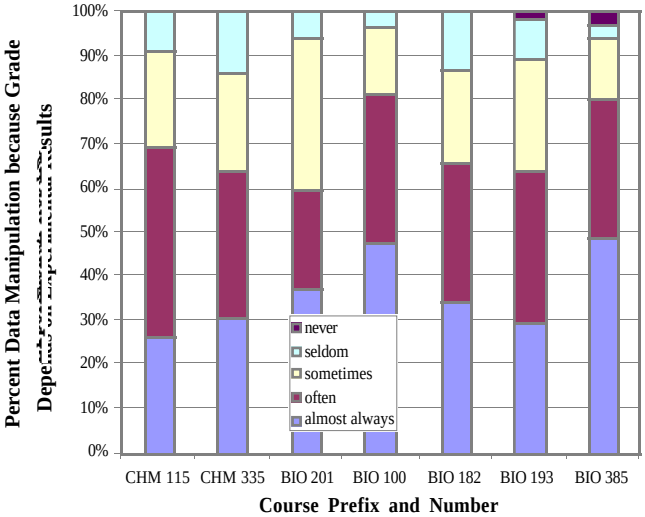


Figure 6. Results of survey, Question 8, "How often have you felt as though your grade in this course depended heavily on your experimental results?" CHM 115, N=102; CHM 335, N=66; BIO 201, N=35; BIO 100, N=218; BIO 182, N=40; BIO 193, N=58; BIO 385, N=66. N= total number of responses to the specific question.

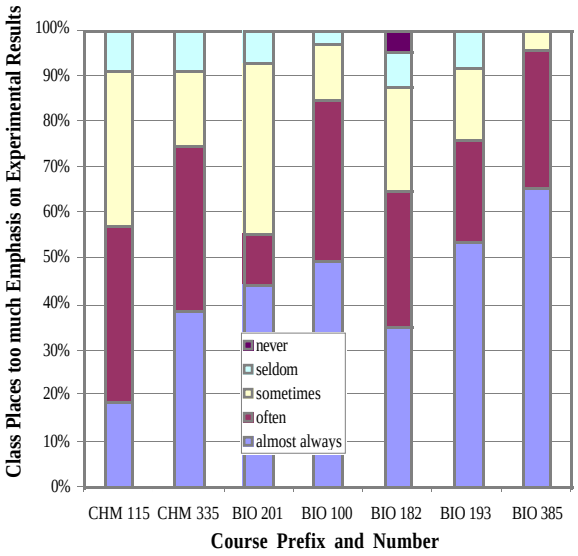


Figure 7. Results of survey, Question 9, "Do you believe this course places too much emphasis on experimental results rather than on the processes used to get the results?" CHM 115, N=98; CHM 335, N=67; BIO 201, N=27; BIO 100, N=194; BIO 182, N=40; BIO 193, N=58; BIO 385, N=66. N= total number of responses to the specific question.

Discussion:

Some precautions should be taken in interpreting these findings. First, the survey was limited to only 7 courses at a single University, which in each case were surveyed only once. We intend to survey additional courses at ASU, and hope to include at least one other large university and a small liberal arts college in our analysis. Second, the survey relies on self reporting. Some of the students did not choose to answer all questions in the survey. The total number responding to each question in each course is presented in the figure caption. Approximately 25% of the students chose not to answer Question 5, for example. Third, the construction of the questions did not permit us to investigate motivations other than that the student felt his/her grade depended upon the experimental results (Questions 8, 9, 15 - 17). Finally, even though students were given a clear definition of "data manipulation" at the beginning of the survey, it is possible that some may not have clearly understood the definition of "data manipulation."

With the above caveats in mind, our results show a very strong tendency among undergraduate science students to manipulate or make up data when writing laboratory reports. As high as these percentages are, they are similar to results observed in surveys of cheating on tests, which Cizek has described as "remarkably and uniformly high" (1). In surveys taken from 1970 to present, from 42% to over 90% of students reported cheating on tests by self or others (reviewed by Cizek, (1)). Out of 6000 students in 31 universities surveyed by Meade, 67% of science majors reported cheating on tests (2). Most surveys of college test cheating ask only whether the student has ever cheated. Our survey expands this question to evaluate how frequently data manipulation in laboratory reports occurs, allowing us to differentiate between occasional events and habitual cheating. Although there are many studies of cheating on college tests, to our knowledge our study is unique in examining manipulation of laboratory data by undergraduates.

Data manipulation apparently does not diminish as the students progress to upper division courses or from non-major to major courses. Commitment to a major subject, presumably because the student intends to continue in this area of science for all or part of his/her professional career, apparently does not diminish this practice.

These results raise some important questions, which include: How can this data manipulation be reduced or eliminated? What are the implications of data manipulation in the undergraduate laboratory to the future careers of these students? In other words, when do the students stop manipulating data? In graduate, professional or medical school? When they begin doing "real" research? When the research is published?

In response to the first of these questions, the faculty and the system itself must take significant responsibility. Faculty must recognize that this data manipulation occurs, and not turn a blind eye to this practice. We must examine the reason why we require laboratory reports in the first place, and whether there is another method of assessing whether the student has learned the necessary laboratory skills. Numerous laboratory manuals are structured to provide "cook book" procedures in which students are expected to verify known biological, chemical, or physical laws (3). However, these verification exercises give students a false notion of the deductive investigative process. They begin their training with the preconceived notion that a "right" answer exists and should be found. They are therefore willing to adjust their laboratory results for fear that the "wrong" answer would affect their grade (4).

We must change the common perception among undergraduate students that their grade often depends upon producing the "right" answer (Figure 6). This change will involve not only the laboratory experimental design, but also the training of graduate teaching assistants and elimination of grading based on achieving a preconceived result. Although students must still be evaluated on whether they are using laboratory instruments accurately, we must consider whether a given laboratory can be designed for training in the hypothetical-deductive process in addition to the specific laboratory technique (4, 5).

Unfortunately, the number of students enrolled in science laboratory courses at large universities in many ways promotes cook-book laboratory exercises. The limited time allowed for experiments, inability to repeat an experiment, and disinterest of many teaching assistants in spending adequate time to grade reports all contribute to the perception on the part of students that making up data is more profitable than writing up what really happened.

Faculty must rethink the reasons for requiring laboratory reports. If the reasons include determining whether the student was present, completed the tasks, understood the experiment, and learned the techniques, then the results presented here suggest that these goals are not being accomplished by the current mechanism of laboratory reports graded based upon achieving the "right" answer. Other mechanisms for discovering whether students have learned the important aspects of the exercise may include laboratory-based questions on exams, and building later experiments upon earlier laboratory exercises. Instructors must be willing to address this problem bluntly with the students and teaching assistants.

At ASU some laboratories have been redesigned to emphasize the inquiry approach to laboratories in Biology and Chemistry. Students generate alternative hypotheses and design experiments themselves, and concepts are introduced after, not before, results are obtained. Teaching assistants are trained to lead students into open-ended and thought-provoking questions (4, 5). In spite of these efforts, our data suggest that data manipulation occurs in these laboratories as well. As students commonly state, "everybody does it." The students themselves overwhelmingly prefer laboratory exercises which emphasize processes or a balance between process and results.

The second concern, whether undergraduates continue data manipulation as professional scientists, has even greater implications. In the frequently-cited study by Swazey et al., 72% of graduate students and 59% of faculty reported to have observed or had direct evidence of some form of scientific misconduct (6). Data falsification, however, was reported by a much smaller proportion of respondents, ranging from 2% to 20%. Apparently, then, data manipulation does decrease when the student becomes a "professional" and becomes dedicated to the science.

Over the last 5 years approximately 400 undergraduates at ASU have been engaged in research programs funded by the Howard Hughes Medical Foundation, the National Institutes of Health and the National Science Foundation, in addition to individual faculty grants. Conversations with these students reveal that once the research becomes "their own" and important to them personally, they have far less motivation to manipulate data, particularly if they have a respectful relationship with the faculty mentor. Hands-on undergraduate research experience may therefore be important in molding the ethical practices of students who will go on to become professional scientists.

When we emphasize the importance of getting the "right" answer, we are teaching undergraduates that their hypothesis must be supported. In truth, the function of an experiment should be to *allow for a fair test* of the hypothesis. We recognize that there exists temptation for graduate students and professional scientists to manipulate data in order to finish research before a deadline, to obtain the next grant, or to have an outstanding publication record. We must take serious responsibility that we do not teach data manipulation techniques at the undergraduate level that will continue to be used in later professional careers.

Acknowledgements

We are grateful to Dr. Anton Lawson for helpful suggestions.

Bibliography

1. Cizek, GJ. Cheating on tests: How to do it, detect it, and prevent it. L Erlbaum Assoc, Mahwah, NJ. 1999.
2. Meade J. Cheating: Is academic dishonesty par for the course? Prism 1992;1: 30-32.
3. Germann PJ, Haskins S, Auls S. Analysis of nine high

school biology laboratory manuals: promoting scientific inquiry. J Res Science Teaching 1996; 33: 475-499.
4. Lawson AE, Lewis CM Jr., Birk JP. Why do students "cook" data? J College Science Teaching 2000; 29(3): 191-198.
5. Lawson AE, Rissing SW, Faeth SH. An inquiry approach to nonmajors biology. J. College Science Teaching 1990;19(6): 340-346.
6. Swazey JP, Anderson MS, Lewis KS. Ethical problems in academic research. Amer Scientist 1993; 81: 342-553.

Data Manipulation Survey

Instructions to students:
Space shuttles blow up, bridges fall, and planes crash and not all are due to natural disasters. An undergraduate student at ASU has been conducting a research project for the last year and a half. During his undergraduate career, he found that in some laboratory settings, there appears to be a great deal of pressure to get the "right" result rather than an emphasis on the scientific and experimental process. In one of his labs he found that 80% of the students manipulated data in some way during the semester. He became concerned: where do students learn scientific ethics? Should we have faith that human morality will overcome pressures to manipulate data in the hopes of a better grade in our college career, or a publication in our professional career?

The purpose of this survey is to collect data on the extent to which undergraduates feel pressured to manipulate, change, or make up data acquired in the laboratory. For example, if you only have a 30% yield of a particular reaction, have you ever felt pressured to say you had more to get a better grade? Moreover, how did you respond to that pressure? Alternatively, has the lab concentrated on experimental process rather than actual results?

Data Manipulation**: To change or omit acquired data or to make up data without confession to those evaluating your performance.**

1. What is your TA's name?

2. What is you major and what year are you (freshman, sophomore, etc.)?

3. Are you:
   A. Female B. Male

4. How many science labs have you taken?
   A. 1 B. 2-5 C. 6 or more

5. Have you ever manipulated data or made up data in this course?
   A. Almost Always B. Often C. Sometimes D. Seldom E. Never
6. If you have ever manipulated data or made up data, was it motivated by the thought of a better grade?
   A. Almost Always B. Often C. Sometimes D. Seldom E. Never

7. Have you ever observed anyone else manipulate or make up data in this course?
   A. Almost Always B. Often C. Sometimes D. Seldom E. Never

8. How often have you felt as though your grade in this course depended heavily on your experimental results?
   A. Almost Always B. Often C. Sometimes D. Seldom E. Never

9. Do you believe this course places too much emphasis on experimental results rather than on the processes used to get the results?
   A. Almost Always B. Often C. Sometimes D. Seldom E. Never

10. Have you ever manipulated or made up data in any other science course?
   A. Almost Always B. Often C. Sometimes D. Seldom E. Never

11. If you have manipulated or made up data, was it motivated by the thought of a better grade?
A. Almost Always B. Often C. Sometimes D. Seldom E. Never

12. If you have manipulated or made up data, was (were) the course(s):
A. Lower Division (100-200 level) B. Upper Division (300 or 400 level) C. Both A & B

13. If you have manipulated or made up data, what department was (were) the course(s) in? (Please circle all that apply.)
A. Biology B. Physics C. Chemistry D. Zoology E. Botany F. Microbiology

14. Have you ever observed anyone manipulate or make up data in any science course?
A. Almost Always B. Often C. Sometimes D. Seldom E. Never

15. How often have you felt that your grade in a science course depended heavily on you experimental results?
A. Almost Always B. Often C. Sometimes D. Seldom E. Never

16. Do you believe that science courses place too much emphasis on experimental results rather than on the processes used to get those results?
A. Almost Always B. Often C. Sometimes D. Seldom E. Never

17. Would you like to see lab experiments:
A. Place more emphasis on results. B. Place more emphasis on processes.
C. Have a balanced combination of both.

Preliminary Observations on Faculty and Graduate Student Perceptions of Questionable Research Conduct

Ravisha Mathur, Department of Psychological Sciences, Purdue University, USA

Stuart I. Offenbach, Department of Psychological Sciences, Purdue University, USA

Keywords: Ethical dilemmas, Mentoring, Perceived misconduct, Research ethics training

When thinking about how graduate students learn the values and standards of science, most universities and departments utilize an apprentice model. In this model, students learn values and ethics by observing their mentor and through working with the mentor–learning via a kind of "osmosis" process. However, the mentoring relationship between faculty advisor and graduate student is one of the most difficult and complex relationships in academia. This sometimes professional, sometimes personal relationship is generally beneficial to both individuals. Advisors usually help students develop their careers and develop professionally, as well as help students network and give them guidance with advice, support, and knowledge. Graduate students help their advisors by assisting with projects, increasing research productivity, increasing professional visibility through the student's research, and can provide their mentors with personal satisfaction and a sense of competence (1, 2). Despite this mutually beneficial relationship, vital for a graduate student's career in graduate school and beyond, faculty members receive very little, if any, training about mentoring. In fact, given this lack of formal preparation, some suggest the mentoring relationship can cause as much potential harm as it does benefits (1).

As a mechanism to transmit ethical codes and standards, the mentoring-apprentice model is, according to some investigators, not very effective (e.g., 3, 4). In order to provide faculty and graduate students with more effective methods of training and educating students about the responsible conduct of research, it would be useful to determine which aspects of the practice of research are most vulnerable to be misperceived, skewed, or violated. In this study, our definition of the responsible conduct of research includes (but is not limited to) honesty, reporting all collected data, using appropriate statistical analyses, and fairly recruiting research participants. Although there is some research describing the types and frequency of scientific misconduct by faculty members and by graduate students, there is little research examining both faculty and graduate student perceptions of violations of the responsible conduct of research. Nor do we know how concordant or discordant these "pairs" are. One purpose of this study was to assess these faculty and student perceptions. A second purpose of this study was to examine the training that students receive from their faculty advisors and departments. We hope to pinpoint how training can be improved and enhanced by examining faculty members' and students' perceptions of training and regulations (at both the department and university level).

Corresponding author: Ravisha Mathur, Department of Psychological Sciences, 1364 Psychological Sciences Building, Purdue University, West Lafayette, IN 47907-1364, 765-494-6928 (voice), 765-496-1264 (fax), ethicssurvey@psych.purdue.edu.

In order to investigate these issues, we sent a survey to faculty members and to graduate students in each of 30 Purdue University departments from the schools of Agriculture, Consumer and Family Sciences, Engineering, Liberal Arts, Science, and Veterinary Medicine. Faculty members were certified to chair students' doctoral committees and graduate students were certified by the graduate school as doctoral students. 733 faculty and 242 graduate students received copies of the survey, and we received a total of 241 surveys from faculty (of which 225 contained usable data) and 47 surveys from students (all of which were usable data).[1] Although the participation rate in this survey was comparable to previous research on similar issues with mail-in surveys (e.g., 5), we were disappointed that we did not receive more responses from students (which limited the analyses and results reported below). The distribution of returns by Gender and by Discipline are in Tables 1 and 2, respectively.

	Female	Male
Faculty	47	162
Grad. Student	16	29

Table 1: Number of responses by gender

The percentage of responses from both male and female faculty members and graduate students matched the gender distribution for the entire faculty (faculty: 22% female and 78% male; graduate student: 35.5% female and 64.5% male). Equivalent comparisons of responses from the different disciplines were more difficult to make since different numbers of departments from each discipline were asked to participate. As Table 2 indicates, more responses were received from the Schools of Agriculture, Engineering, and Science. Only a few graduate students from Consumer and Family Sciences and from Liberal Arts participated. Most of the student responses were from Agriculture and from Engineering.

There were three parts of the survey. Part 1 addressed how information about the responsible conduct of research is exchanged (Item 2 of Part 1 is shown in Figure 1). The questions in Part 1 focused on how and where students learned about the responsible conduct of research and if students and faculty knew of others who had been involved in ethical conflicts. The main section of the survey, Part 2, consisted of 38 hypothetical dilemmas (each included a proposed action to resolve the dilemma). The dilemmas were written to cover the following types of problems (which were supported the confirmatory factor analysis described below):

1) Information Sharing in the Lab;
2) Truth/ Completeness in Writing up Research Results;
3) Misleading the Audience (Plagiarism);
4) Seeking Credit for doing the Research; and
5) Consent Issues.

(Examples of the dilemmas for each factor are shown in Figure 2.) Participants responded by rating each dilemma on a five point Likert scale (Strongly Disagree to Strongly Agree). The third and final section of the survey examined participant's perceptions of university and departmental policies on the responsible conduct of research and whether the faculty member or graduate students would feel comfortable reporting incidents of suspected misconduct.

Which of the following are ways that graduate students learn about professional values and ethical standards? (Circle all that apply).

1. Brown bag/colloquium
2. Special courses devoted to this topic
3. Interaction with faculty in research work
4. Codes of ethics and professional standards provided by professional organizations
5. Informal discussion of ethical problems when they occur
6. Department policies for teaching and research
7. Discussion of ethics and values in regular course work

Figure 1: Item 2 from Part 1 of the Survey

	Agriculture	CFS	Engineering	Liberal Arts	Pharmacy & Medical Sci.	Science
Faculty	52	23	32	27	20	38
Grad. Stud.	13	4	10	2	7	7

Table 2: Number of Responses by School

a. **Sharing Information**: Grant is in his office one day and sees his officemate's lab notebook open. While paging through the notebook, he discovers that Anli has found a way to metabolize ABC enzyme. Grant has been working for two months to discover a way to metabolize this enzyme for his dissertation. After thinking about it for a few days, Grant decides to use the same process to keep his dissertation on track. He does not bother to tell Anli because she is in his lab group and probably would not mind anyway. Do you agree with his decision?

b. **Writing**: Mei has been collecting data for a long-term study for the past two years. Although she still is in the middle of the data collection phase, the trends she sees in her data are very exciting. She decides to write up her results and present them as a complete study and continue to collect data for the full term of the study. She plans to publish those data in at least two "follow-up" reports. Do you agree with her decision?

c. **Misconduct**: Angelo has written an article in which he included a full paragraph from a paper written by a student for a class Angelo was teaching. Do you agree with Angelo's decision to include the paragraph?

d. **Seeking Credit:** John has written an article in which he included a full paragraph from a pre-publication version of an article reviewing the research literature in his area of interest. The author of the article was planning to submit it to a prominent journal that publishes such reviews. Do you agree with John's decision to include the paragraph?

e. Consent Issues: Professor Gleeson is conducting a research project concerned with social customs in a village in rural South Africa. The village consists of members of a single tribe, and is led by a tribal chief and council of elders who make all decisions for the village. The tribal chief insists that he will decide if his villagers can participate in Professor Gleeson's research project, and that he (the Chief) will distribute the payment to the villagers. Professor Gleeson may not ask the villagers whether they want to participate because that would be an insult and challenge to the Chief and Elders of the village. Do you agree that Professor Gleeson can go ahead with the research project if the Chief and Elders approve?

Figure 2: Sample Hypothetical Dilemmas from Part 2 of the Survey

(Two of these items are shown in Figure 3.) Items from both Part 1 and Part 3 were adapted from Judith Swayze and coworkers' survey of faculty and students (6). Items for Part 2 were written by the authors and were based on real events and scenarios gleaned from reading and teaching about the responsible conduct of research for the past five years.

Participants were given a response sheet to use as their answer sheet and were asked to return the response sheet in a self addressed envelope we provided them. Once we received the survey, a third party removed any identifying information. The responses on each survey form were entered into a computer file separately by the two authors. All coding errors then were reconciled by the authors.

Results

Part One. The first questions focused on settings in which respondents learned some or all of their professional values. Seventy-two percent of faculty members and 60% of graduate students believed supportive faculty members provided such information. Sixty-seven percent of faculty members believed professional organizations provided such information compared to only 15% of graduate students (t = 28.377; Only t-values significant at .05 or less are reported). This difference probably reflected a lack of contact with such organizations by graduate students. Graduate students also relied more on other students as a source of information (51%), a source not considered by faculty members (15%, t = 16.97).

Interactions with faculty in research work and informal discussions of ethical problems were considered effective settings to learn professional values by 90% or more of students and faculty. Brown bag discussions, colloquia, and courses, on the other hand, were not seen as effective settings by most respondents (percentages all less than 30%).

We also asked whether respondents ever discussed with peers value issues related to external sources of research funding or the application of research findings. Eighty percent

2. How active a role does your department actually take in preparing graduate students to recognize and deal with ethical issues in your field as part of their training?

Very active	Active	Somewhat active	Not very active	Not at all active

3. Could you report cases of suspected misconduct in your department without expecting retaliation?

Misconduct by a faculty member: Yes definitely, Yes, Probably, Probably not, Definitely not
Misconduct by a graduate student: Yes definitely, Yes, Probably, Probably not, Definitely not

Figure 3: Items #2 and #3 from Part 3 of the Survey

of faculty members and 47% of the graduate students (t = 18.263) did so. In addition 38% of faculty members and 11% of graduate students actually knew someone who had refused to participate in a research project because of personal reservations about funding sources. These faculty-student difference probably reflects differences in age and experience in the field.

What is clear from these analyses is that faculty members and students do have different views of the best place or way to learn about professional standards and to learn to recognize ethical research issues.

Part 2: Hypothetical Dilemmas. A confirmatory factor analysis of the hypothetical dilemmas produced five factors: 1) Information Sharing in the Lab; 2) Truth/ Completeness in Writing up Research Results; 3) Misleading the Audience (Plagiarism); 4) Seeking Credit for doing the Research; and 5) Consent Issues. The alphas for these variables were moderate, ranging from .47 - .61. We recognize that not all of the dilemmas applied equally to all of the disciplines sampled in this survey, but we were pleased that some general factors appeared. The nature of the five factors can be explained in several ways. First (and probably foremost) is the construction of the scenarios by the principle investigators. Construction of these scenarios was not a random process, and the factors extracted from the analysis may simply confirm biases and predispositions that entered into our construction of the items. On the other hand, the areas represented by the five factors have been identified by many investigators as areas of concern vis-a-vis research ethics. The fact that these items hang together at all may be a confirmation of the concerns many investigators and ethicists have about the process of research.

Although we could not adequately examine the faculty-student differences on the responses to the Hypothetical Dilemmas because of the disparity in the number of responses from each group, we were able to draw some tentative conclusions. Faculty members clearly took "more extreme" views than did students. That is, faculty members were more likely to indicate strong disagreement or agreement with the action taken in a dilemma than were graduate students. For example, on the 20 dilemmas that contributed to the five factors, more faculty members responded "strongly agree" (or "strongly disagree") on every dilemma. Graduate students had more moderate responses. Actually, there were no faculty-student differences in the number of combined "strongly agree" and "agree" (or "strongly disagree" and "disagree"). Thus for the second item in Figure 2, of the 98% faculty members who disagreed with the action, 80% checked "strongly disagree." All of the graduate students disagreed with the action, but only 43% expressed strong disagreement. Perhaps faculty members' greater experience with ethical issues has led them to be more certain of their views (or the students' lack of experience led them to be more tentative).

Finally, while the responses to the hypothetical dilemmas made intuitive sense, the construction of the dilemmas is more complex than we thought. Respondents often commented that they saw some items as dealing with multiple ethical issues or that there was not enough information presented to make a judgement. This may be one reason alpha levels were low for the five factors. More thought must go into the development of items that have a more specific focus (and are less complex) for a survey of this type.

Two sets of analyses were not computed.

Analyses to compare factor scores for students with those of faculty were not conducted because the factor scores have not yet been corrected for directionality differences. That is, some factors include items with which most respondents agree and items with which most respondents disagree. The point values for these items needs to be on the same scale or have the same valence in order to examine factor scores. The other analyses not yet conducted would have compared student responses with those of their mentors. These analyses depended on both the student and his or her mentor actually submitting a survey, and having the student identify his or her mentor. Unfortunately, we were able to identify only five faculty-student pairs, precluding any analysis of whether the two are concordant or discordant.

Questions about department and university policies

The questions in Part 3 focused on respondents perceptions of the role that departments should take and actually do take in preparing students to recognize and deal with ethical issues (see Tables 3 and 4). Significantly more students than faculty (70% vs. 45%) reported almost no effort by their departments to train them to recognize and deal with ethical issues in science (it also is interesting that 16% of faculty members thought their departments were active, but only 6% of the students shared that perspective). Thus both faculty and students believe academic departments should take a more significant role in training graduate students to recognize and deal with ethical issues (we only asked about academic departments, faculty members and students may actually ascribe greater responsibility to larger academic units — e.g., schools, graduate school, etc.). There is a mismatch here – faculty and students wanting departments to take a role and departments not doing that. And there is no formal structure at the university level for training in the responsible conduct of research. Thus, the student is left to his or her own devices. The most frequent choice made by students seems to be to ask another student or to ask the advisor.

The next two questions asked whether one could report misconduct by a faculty member or by a graduate student without expecting retaliation. The results in Table 6 show that 89% of faculty members believed they could report misconduct by a graduate student "safely." They would expect no retaliation. The graduate students also seemed less concerned about retaliation if they reported misconduct by another student. Seventy-three percent thought it was safe to report misconduct by another graduate student. Reporting misconduct by faculty members was another matter. Fewer faculty members were comfortable about reporting misconduct by a colleague (73%). Only 55% of students thought they could report misconduct by a faculty member "safely." In contrast, 28% of the faculty members who responded said they would not feel safe reporting misconduct by a faculty colleague. Almost half of the graduate students, 44%, were concerned about retaliation for reporting a faculty member's misconduct. These results seem consistent with anecdotal data. A cursory review of comments from the electronic list-serve Sci-Fraud reveals a concern by many participants that to make a good faith allegation that a faculty member has engaged in misconduct is to place one's career in jeopardy.

Finally, we asked about knowledge of university and departmental policies on misconduct. Half of graduate student respondents did not know that the University has a research misconduct policy and 72% do not know if their department has such a policy. The faculty were more knowledgeable – 63% knew there was a university policy. However, only half of them were familiar with the policy's contents.

	Very active	Active	Some-what active	Not very active	Not at all active
Faculty	37	45	14	03	01
Grad. Stud.	22	52	22	04	00

Table 3: Role a department should take (percent agreeing)

	Very active	Active	Some-what active	Not very active	Not at all active
Faculty	02	14	38	34	11
Grad. Stud.	02	04	26	51	17

Table 4: Role a department does take (percent agreeing)

Respondents	Misconduct by	Definitely Yes	Probably Yes	Probably Not	Definitely Not
Faculty Members	Faculty	32	41	23	05
	Students	48	41	09	01
Graduate Students	Faculty	04	51	40	04
	Students	11	62	23	04

Table 5: Reporting Misconduct Responses (percentage agreeing)

Conclusions:

The hypothesis that graduate students learn to identify and deal with ethical situations in research from their mentors without specific instruction or discussion could not be tested using the data collected in the first "pass" of our study. We received too few mentor-student data pairs to make any analysis. Our failure to obtain data relating mentor's values directly to that of their specific students was disappointing – only five student - mentor pairs were identified (we hope to rectify this situation by increasing the size of the student data pool). However, we believe the modeling or osmosis hypothesis probably will not be supported because of the different perceptions graduate students and faculty members have of how scientific values are transmitted. Faculty members and students do rely on other faculty members, but only the students rely on their student peers. At the same time, both faculty and students believed that interactions in the work or lab settings would be useful in learning to recognize and deal with ethical situations. Unfortunately, this expectation means that people seem to want to learn from "personal experience," but no one wants to have that kind of personal experience.

One thing is certain, things will not continue in the same way. Actions by the National Institutes of Health to require specific education on the responsible conduct of research generally specifically will require universities to do a better job. That better job might be facilitated with a more complete understanding of how students are learning now and by determining not only what they are learning , but also by determining what they are NOT learning.

Notes

1. These numbers differ from the totals in Tables 1 and 2 as some participants did not answer the gender or discipline questions.

Bibliography

1. Johnson, W. B., & Nelson, N. (1999). Mentor-protégée relationships in graduate training: Some ethical concerns. Ethics Behav, 9, 189-210.
2. Wright, C. A., & Wright, S. D. (1987). The role of mentors in the career development in young professionals. Fam Relations, 36, 204-208.
3. Brown, S., & Kalichman, M. W. (1998). Effects of training in the responsible conduct of research: A survey of graduate students in experimental sciences. Sci Eng Ethics, 4, 487-498.
4. Folse, K.A. (1991). Ethics and the profession: Graduate school training. Teach Sociol, 19, 344- 350.
5. Mearns, J., & Allen, G. J. (1991). Graduate students' experiences in dealing with impaired peers, compared with faculty predictions: An exploratory study. Ethics Behav, 1, 191-202.
6. Swazey, J. P., Anderson, M. S., & Louis, K. S. (1993). Ethical problems in academic research: A survey of doctoral candidates and faculty raises important questions about the ethical environment of graduate education and research. Am Scientist, 81, 542-553.

Constructing a Personal Model of Research: Academic Culture and the Development of Professional Identity in the Professorate

L. Earle Reybold, Adult and Higher Education, The University of Texas at San Antonio, USA

Keywords: ***Academic culture, Faculty development, Graduate socialization, Professional identity, Research ethics, Research training, RCR***

Doctoral students pursuing academic careers are educated in awkward and mostly tacit apprenticeships. As students, they are expected to learn professional knowledge and the technical skills associated with their program of study. Yet, they must simultaneously absorb the culture of academe and learn their future roles as faculty members. Because learning and thinking are situated in a social milieu, socialization is a process initiated and established in contexts that construct knowledge through activity (1). In other words, academic culture and educational knowledge "act together to determine the way practitioners see the world" (p. 33).

Generally, socialization studies have investigated academic culture as context for student learning and development. Many of these studies focus on the social aspects of academic culture, particularly relationships between students and their colleagues or professors (2, 3, 4, 5). These socialization studies concentrate on students' experiences *as students* in higher education and are centered on classroom modality.

Likewise, inquiry into new faculty socialization segregates faculty roles and responsibilities into particular genres of experiences such as teaching success (6) and tenure and promotion processes (7). Unfortunately, faculty socialization studies fail to address how graduate school experiences, particularly as they are situated in an academic culture, affect the development of professional identity and ultimately professional decision-making and activity.

When the concept of professional identity and competency *is* addressed in the faculty socialization literature, the discussion surveys the development of the faculty teaching roles but ignores the complex faculty identity as teacher, researcher, and service provider. This lack of attention to an integrated identity that begins to emerge during graduate studies portrays faculty socialization in perfunctory terms. For example, Boice discusses new faculty success in terms of teaching style and mastery (6). The author notes the characteristics of "quick starters," but these are teaching characteristics of new faculty, with no attention to the development of these characteristics. Pollard, Pollard, & Rojewski also investigate the college teaching experience of new faculty (8). They argue that doctoral students are academically prepared for their careers in higher education, but their study concentrates only on the impact of higher education culture on new faculty.

Purpose of Study and Research Focus

The purpose of this study is to describe the role of academic culture in determining a personal model

Corresponding author: L. Earle Reybold, Adult and Higher Education, College of Education and Human Development, The University of Texas at San Antonio, 6900 North Loop 1604 West, San Antonio, TX 78249-0654, 210-458-5429 (voice), 210-458-5848 (fax), ereybold@utsa.edu.

of ethical research practice in the professorate. While little is known about the construction of faculty identity and role expectations during graduate studies, even less is understood about the impact of student experiences on professorial activities and decision-making, particularly research competence and reasoning. Two questions demand consideration. First, how are doctoral students socialized into the practice of academic research? Further, how do these students construct a model of research standards and ethics that will inform their future practice as faculty members?

Two general assumptions guide this inquiry:

- Socialization into the professorate is a developmental rite of passage rather than two discrete phases of socialization marked by graduation and/or faculty appointment..
- Preparation for the professorate is situated in an academic culture that shapes one's personal understanding of the professorate and professional identity and perceived roles.

This study initiates a two-phase longitudinal qualitative investigation. Using case study methods (9), this study focuses on doctoral students' perceptions of research ethics in education. Interview questions concentrated on emergent definitions of research ethics, training in research ethics, and experiences of ethical dilemmas.

Case study research is uniquely geared toward description and understanding of institutional culture and its impact on perspective. Merriam describes case study research as an ideal design for exploring participants' understanding and perspective (9). Further, she says case study is appropriate when inquiry is interested in "process rather than outcomes, in context rather than a specific variable, in discovery rather than confirmation" (p. 19).

Sampling for this phase of the study is network sampling, which locates participants through recommendations of initial participants and key informants based on selected criteria (10). Participants were located at three universities in Georgia and Texas, including institutions identified as Research I, Research II, and Doctoral II. Participants were doctoral students in education preparing for a faculty career in academe.

Data were collected through in-depth interviews with doctoral students and faculty members at three universities in two states, and through archival data such as program materials and reflection journals supplement the interview data. Interviews were conducted using a semi-structured format to allow comparison of data across participants (11). In general, interview questions addressed student and professional identity, academic culture, training in teaching and research, and ethical decision-making as a professional. Journaling allowed students to explore and document their process of decision making as relevant issues arose, the entries were guided by the following statement: *Describe your decisions that are most important to your preparation for the professorate*.

Standards for Quality Research

Emphasizing qualitative inquiry as a creative process, Patton (10) reminds researchers of the "technical side to analysis that is analytically rigorous, mentally replicable, and explicitly systematic" (p. 462). Merriam (9) adds that qualitative research findings "are trustworthy to the extent that there has been some accounting" (p. 198) for quality. In general, the criteria for trustworthy qualitative research include rigorous and systematic data collection and analysis techniques, credibility of the researcher, and belief in naturalistic inquiry (10). The quality of this study is enhanced by several factors. First, I have experience as a qualitative researcher and have taught qualitative methods at the graduate level. Further, triangulation of methods and peer review of data and analysis will enhance the trustworthiness of the data. Finally, the multi-site design encourages usability of the findings beyond the university settings included in the study.

Situating Faculty Identity Development in Academic Culture

This study is framed by the concepts of research ethics and integrity, faculty socialization and enculturation, and professional identity development.

Research Ethics and Integrity.

Research is often messy and complicated. Best-case scenarios of theoretical contributions and improvement of practice are weighed against questionable issues of right and wrong research behavior. In these cases, research decisions may evolve as uneasy guesses with no obvious consequence. Confronted with uncertain choices,

how do researchers define and respond to ethical dilemmas?

Ultimately, ethical decision-making reaches beyond the local boundaries of specific research projects. Because research is fundamental to higher education, it could be argued that research decisions symbolize the moral character of higher education. Under the guise of exploration and discovery, research is a noble enterprise. But research agendas are realized within the "publish-or-perish" mentality of higher education in which ethical dilemmas may become stumbling blocks to promotion and tenure. This is the context where doctoral students are socialized toward the professorate; this is the culture that trains future faculty members as future researchers.

Faculty Socialization and Enculturation. Tierney & Rhoads (12) remind us that "organizations exist as social constructions" (p. 1) that revolve around shared understandings. This organizational culture shapes behavior and expectations, bounding faculty socialization. Tierney & Rhoads define faculty socialization as "the process through which individuals acquire the values, attitudes, norms, knowledge, and skills needed to exist in a given society" (p. 6). Their definition of faculty socialization as transmission of culture complements this study of professional identity development.

Tierney & Rhoads (12) describe academic culture as the nexus of five forces: national, professional, disciplinary, individual, and institutional. Although these are conceptualized as distinct subcultures, these forces are synergistic and do not operate independently of one another. Professional identity is an aggregate sense of self that develops across these subcultures. This process of socialization occurs in two overlapping stages: anticipatory socialization and organizational socialization. The anticipatory stage "pertains to how non-members take on the attitudes, actions, and values of the group to which they aspire" (p.23). The organizational stage, on the other hand, involves initial entry and role continuance. Noting the importance of the transition process, Tierney & Rhoads comment that when anticipatory socialization and organizational socialization are consistent, the socialization process is affirming. When socialization experiences are not consistent, the organization will attempt to modify or transform the individual's values and beliefs to fit the "cultural ethos of the institution" (p. 25). Tierney and Bensimon continue this emphasis on socialization in academe, focusing on the tenure process as the locus of organizational socialization (7). Although they offer strategies for anticipatory and organizational socialization, the authors do not focus their attention on the transition process.

Bergquist examines academe within the framework of organizational culture, concluding that there are four distinct cultures: collegial, managerial, developmental, and negotiating (13). Culture, he says, "provides meaning and context for a specific group of people," adding "the culture holds the people together and instills in them an individual and collective sense of purpose and continuity" (p. 2). Further, Bergquist says culture defines the nature of reality for members of a given culture, providing the "lenses through which its members interpret and assign value to the various events and products of the world" (p. 2). Although there are four distinct cultures within academe, one will usually be dominant. Bergquist notes that the interaction among the four unequal cultures helps "to produce the often confusing and paradoxical conditions in which contemporary faculty find themselves" (p. 7).

Both Bergquist (13) and Tierney & Rhoads (12) note the influence of academic culture on faculty perspectives, decisions, and behavior; also, they agree that cultural differences create a backdrop of conflict for members within a given culture. This study extends their conclusions to graduate education, adding that students also are influenced by academic culture. Further, the transition process from doctoral studies to the professorate adds another layer of possible conflict between academic cultures.

Developing a Professional Identity. Marcia defines identity development as a self-constructed organization of drives, abilities, beliefs and individual history (14). Bruss & Kopala (15), building on Marcia's definition, define "professional identity "the formation of an attitude of personal responsibility regarding one's role in the profession, a commitment to behave ethically and morally, and the development of feelings of pride for the profession" (p. 686). This definition directly connects professional identity to professional behavior.

While the identity development literature is

concerned predominantly with the psychological aspects of self, identity may be viewed as both personal *and* social. Social identities result in identity relationships within a given culture, and these identity relationships determine identity status and role expectations (16). For the purpose of this study, status and role expectations will be examined as cultural aspects of professional identity development, particularly as they relate to *anticipatory socialization* during the graduate school experience (7).

Graduate training is expected to nurture the development of professional identity. In their discussion of psychology students, Bruss and Kopala (15) described graduate school training as *professional infancy* and "the training institution . . . as an environment wherein it is the task of the faculty and training staff to nurture and promote growth" (p. 686). However, academic culture is not always nurturing; structural problems in graduate education are potentially harmful to students' self-esteem (17). Attitudes—good and bad—about professional responsibility, ethical behavior, and professional pride are constructed within the cultural context of graduate training. These attitudes produce social identities and role expectations that persist through a graduate student's transition into the professorate. In short, academic culture exerts directive force over professional decision-making and activities.

Chickering & Reisser, in their study of college student development, define identity as a sense of self (18). The process of identity development results in "a solid sense of self [that] emerges, and it becomes more apparent that there is an *I* who coordinates the facets of personality, who 'owns' the house of self and is comfortable in all of its rooms" (p. 49).

Findings

To describe the role of academic culture in determining ethical research practice, data were analyzed within four concentrations: the perceived role of research in higher education, the perceived standards for ethical research, the actual ethical dilemmas experienced by graduate student researchers, and the factors that hinder or support ethical research.

What is the perceived role of research in higher education? Participants in this study experience research and subsequent publication as an institutional priority and a personal badge of prestige. While one participant views the professorate as a delicate balance of professorial roles, most participants emphasized the preeminence of becoming a researcher, and only one participant noted a teaching role being more important than a research role. For example, Betsy says, "research is painful and boring, but the doctorate is about what the university considers important—getting published!" Echoing this sentiment, Claire says the "doctoral degree is mainly trying to get us into the research part of being a professor and much less teaching; it is indoctrination into the research aspect of being a professor."

While some participants came in with considerable research experience, most are concerned that they don't "know what to do with the research" after the dissertation process. Post-dissertation concerns include translation of theory into educational practice, establishing a research agenda, and getting published.

What are the perceived standards for ethical research and who defines ethics in academic settings? Coursework in research ethics is almost nonexistent. As students, participants expect professors to guide them through the process of learning and implementing ethical research, but they are relying instead on their own sense of right and wrong. Julia says she relies on her "internal gyroscope" to guide her decisions; and Claire relies on her "personal ethics and personal morals." Grace adds that "ethics is about power differences." Her professors talked about collaboration and high quality, but their practice expressed a disregard for the Institutional Review Board (IRB), quality research, and research integrity.

More than a lack of definition of ethical research, participants are concerned and confused about "grey" areas of research ethics and believe they must define ethical research according to their own experiences and standards. Interestingly, the two participants with training in medical ethics find research ethics easier to define. The other participants have scattered definitions of research ethics, with most positioning ethical research as a data collection and/or analysis issue. However, a couple of participants have a complex, comprehensive definition of research ethics, including researcher attitude and choices throughout the research process. One participant noted that article readers have an ethical responsibility to read the results thoroughly. Another participant, Grace, is quite concerned with the power issues that

impact ethical decision-making: "power issues come into play, whether we like to admit it or not…these are times we just have to make a mental note, 'this is not right'…. But I'm at a point where I have no power to address this."

One participant has a collaborative relationship with her major professor. Kelly says her discussions with her major professor about research methods and ethics have been invaluable, even to the point where she feels comfortable mentoring other students with research problems. Although Betsy claims to have a collaborative and mentoring relationship with her major professor, she often finds herself involved in ethical dilemmas with others in the same department. For the participants in this study, the most beneficial contribution to ethics and methods training is involvement in actual research projects, particularly pilot studies of own research and collaborative efforts as research partners with professors, but only when that contribution is valued and rewarded as equal.

What types of actual ethical dilemmas do graduate student researchers experience? While most participants define ethical dilemmas in terms of research methods, their experiences of ethical dilemmas focus more on relationships and issues of power and coercion. One participant reports her professor "uses" students to review his own material prior to publication. Student assignments in non-related courses revolve around this professor's research agenda, and students are expected align their work to match that agenda. Several participants report being forced to manipulate data to yield desired outcomes; if a student refuses, he or she is no longer funded as a research assistant. Kelly, a research assistant on a grant-funded study, voiced disapproval of research decisions being made by professors on the grant:

> I've been vocal, but I wasn't a threat or anything. I was unhappy with the way the professors were doing things . . . . I was just going along, and it hit me. Did I feel free to leave? No! To a certain extent, this is part of being a graduate student. I mostly feel free to voice my concerns, but in this case, it was an ultimatum—or I was off the grant! I never want to do this in my own research.

Another participant, Grace, reports working on presentations and articles with more than one professor and negotiating authorship—but the articles were published without her name or with a different authorship order than negotiated. This is particularly troublesome at conference presentations because funding for student travel to conferences depends on authorship. Grace did try to confront the professor, but to no avail. The professor was on the editorial board of the journal that published the article, and she believed the issue would not be taken seriously. Participants report that even when the research situation is uncomfortable, they "don't want to sacrifice the relationship" by removing themselves from the project.

Another type of dilemma involves committee make-up. One participant had approval for a mixed design dissertation, but her committee politicized her design and held up her research. She decided "to write it one way for the dissertation" and then publish it using her mixed design approach. Other participants experienced negative "shaping" of their research based on professors' interests. As one participant reports, "professors stay in their comfort zones" and won't head committees outside their personal interests. This is particularly problematic in small departments with few faculty members.

What factors hinder or support ethical research? Several factors hinder ethical research: institutional/structural, relational/positional, and technical. First, the culture of academe encourages ambivalence toward the issue of ethical research. Institutions reward research productivity, even at the expense of other professorial roles, perpetuating the adage, *publish or perish*. While some professors "nudge" their students to commit ethical violations, others ignore the need for training and guidance in ethical research practice. Dan, looking toward a future career in academe, acknowledges that "tenure is political, so go way beyond their expectations!"

A second factor hindering ethical research is the role of hierarchy in academic relationships. Graduate students are afraid to report ethical violations; they fear losing their assistantships and professorial support. As a student, one participant notes that "it's important to know where your allegiances lie; the only way you'll get lobbied for is if you are clearly in someone's camp." Only one student, Kelly, says her professors treat her as a peer. Her major professor, she says, "got me involved with his projects, but told me to 'find your own thing—academia isn't just doing other people's work.'" Several participants alluded to expecting a similar role as junior faculty; coercion will continue to force them to make ethical decisions

that might not be supported by academic expectations.

A third factor that hinders ethical research is the lack of training and exposure to guidelines. Only those participants with medical backgrounds had any courses in ethics, and those courses dealt with medical ethics rather than research ethics. Only one participant reports research ethics discussed in her doctoral research classes. None of the participants in this study knew of any guidelines for education research practice other than IRB guidelines.

Only one participant, Kelly, reports preparation for ethical research. Her major professor and a network of research professors provide guidance through formal and informal mentoring and involvement in various research projects. This particular participant has published with several professors, and her contributions are valued as equal to those of the professors. In fact, this professor reminds the participant that she is his "primary responsibility" and that she is to develop her own line of research separate from his. Another participant feels secure in her relationship with her major professor, but says her other experiences with faculty members in the same department make her bitter and wary. She notes there are two levels of culture in the department, and a "lot of politics gets played below the surface" even though "we represent ourselves as a united front."

Summary and Conclusion

Almost all participants in this study raised themes of power and censorship. The impact of coercion and fear on the research process must be explored. Graduate students believe their research is censored on several levels: personal, institutional, and systemic. First, graduate students expressed fear of retaliation if they resisted their faculty advisor's management of their research. Further, these students believe they are bound by the dissertation committee structure and the institutional support of highly productive faculty members. Finally, censorship is systemic, according to these students' experiences, because certain topics are "favorites" of funding agencies. Likewise, these students believe journal editors and blind reviews control the emergence of new knowledge.

The goal of higher education is the preparation and personal development of competent, well-trained professionals. While much of higher education focuses on the development of technical competence and critical thinking skills, the transformation from student to faculty member is too often left to chance.

Future inquiry will explore the development of professional identity throughout preparation for the professorate, and how this emerging identity impacts professional decision-making as a scholar.

Bibliography

1. Brown JS, Collins A, Duguid P. Situated cognition and the culture of learning. Educational Researcher 1989; 18(1):32-42.
2. Heinrich, KT. Doctoral advisement relationships between women. Journal of Higher Education 1995; 66(4):447-469.
3. Martinez Aleman AM. Understanding and investigating female friendship's educative value. Journal of Higher Education 1997; 68(2):119-158.
4. Schroeder DS, Mynatt CR. Female graduate students' perceptions of their interactions with male and female major professors. Journal of Higher Education 1993; 64(5):555-573.
5. Stalker J. Athene in academe: Women mentoring women in the academy. International Journal of Lifelong Education 1994; 13(5):361-372.
6. Boice R. Quick starters: New faculty who succeed. New Directions for Teaching and Learning 1991; 48:111-121.
7. Tierney WG, Bensimon EM. Promotion and tenure: Community and socialization in academe. New York: State University of New York Press; 1996.
8. Pollard C, Pollard RR, Rojewski JW. The college teaching experience: A qualitative study. College Student Journal 1991; 25(2):149-154.
9. Merriam SB. Qualitative Research and Case Study Applications in Education. San Francisco: Jossey Bass; 1998.
10. Patton MQ. Qualitative evaluation and research methods. 2nd ed. Newbury Park, CA: Sage; 1980/1990.
11. Bogdan RC, Biklen SK. Qualitative research for education: An introduction to theory and methods. 2nd ed. Boston: Allyn and Bacon; 1992.
12. Tierney WG, Rhoads RA. Faculty socialization as a cultural process: A mirror of institutional commitment. ASHE-ERIC Higher Education Report No. 93-6. Washington, D.C.: The George Washington University, School of Education and Human Development; 1993.
13. Berquist WH. The four cultures of the academy. San Francisco: Jossey-Bass; 1992.
14. Marcia J. Identity in adolescence. In: Adelson J, editor. Handbook of adolescent psychology. New York: Wiley; 1980. p. 159-187.
15. Bruss KV, Kopala M. Graduate school training in psychology: Its impact upon the development of professional identity. Psychotherapy 1993; 30 (4); 685-691.

16. Goodenough WH. (1969). Rethinking 'status' and 'role': Toward a general model of the cultural organization of social relationships. In: Tyler SA, editor. Cognitive anthropology. New York: Holt, Rinehart and Winston; 1969. p. 311-30.
17. Egan JM. (1989). Graduate school and the self: A theoretical view of some negative effects of professional socialization. Teaching Sociology 1989; 17, 200-208.
18. Chickering AW, Reisser L. Education and identity. 2nd ed. San Francisco: Jossey-Bass; 1993.

Undergraduate Academic Cheating as a Risk Factor for Future Professional Misconduct

Julio F. Turrens, Department of Biomedical Sciences, University of South Alabama, Mobile, AL, USA

Irene M. Staik, Department of Behavioral and Social Sciences, University of Montevallo, Montevallo, AL, USA

D. Kristen Gilbert, Department of Behavioral and Social Sciences, University of Montevallo, Montevallo, AL, USA

W. Curtis Small, Department of Biomedical Sciences, University of South Alabama, Mobile, AL, USA

John W. Burling, Department of Behavioral and Social Sciences, University of Montevallo, Montevallo, AL, USA

***Keywords**: Academic misconduct, Cheating, Undergraduate education*

"T'is education forms the common mind:
Just as the twig is bent the tree's inclined."
Alexander Pope (1688-1744)

Scientific misconduct may be more acceptable in the minds of those professionals who grew accustomed to lower academic standards during their formative undergraduate years. The hypothesis proposed in this paper is that the recent increase in cheating at the undergraduate level is likely to result in an increase in the number of future professionals involved in scientific misconduct.

Twenty years ago, academic misconduct at the undergraduate level was considered by the great majority of both students and faculty as unacceptable and dishonest behavior. Currently, not only are most undergraduate students aware that misconduct is very common but most of them by their Junior year have participated or witnessed more than one event. Even those students who do not engage in academic misconduct have become more skeptical of the need to be personally responsible for their own academic work and accept this lowering of standards as a fact of life.

Because of these changes in the environment of higher education, the incidence and prevalence of cheating by college students has been an area of intense concern for educators and researchers since the 1970s. A vast number of articles in the literature indicate that cheating or academic dishonesty is at epidemic proportions within academia (1-7). A representative sampling of articles documenting this

Corresponding author: Julio F. Turrens, Department of Biomedical Sciences, University of South Alabama, Mobile, AL 36688, 251-380-2714 (voice), 251-380-2711 (fax), jturrens@usouthal.edu.

Eric # or Journal	Year	Sample size	Institutions	Reported cheating
ED427355	1998	203	four years	78 %
			two years	57%
EJ351071	1986	380		> 50%
ED334921	1990	232	Rutgers	88%
ED347931	1992	87		81%
EJ449186	1992	6000	31 top-ranked	business: 87% engineering: 74% science: 67% humanities: 63%
EJ489082	1994	480	2 colleges	89%
EJ518822	1995	300		83%
Res. High Ed.	1984	380	mid size	54.1%
37:487-502, 1996	1994	474	liberal arts	61.2%

Table 1. Studies showing increased cheating by undergraduate students.

recent increase in cheating by students is shown in Table 1. Estimates in the literature reveal that 75% to 98% of college students cheat at least once during their college career (8, 9). Students, also reported that they are deterred from cheating only by a fear of getting caught and public embarrassment (2, 10). High achievers and students who have too little time to study for tests are particularly vulnerable to cheating (11, 12).

Students also report that their perception of faculty reactions to cheating is one of apathy. Faculty members often do not report a case of student cheating to the institutional justice system, either for fear of legal repercussions or to prevent hurting the reputation of the student. Instead, many faculty members prefer to handle each case on an individual basis, sending a signal to students that the repercussions for cheating are minimal (6, 13). This signal is tantamount to acceptance of academic dishonesty as a fact in higher education by both faculty and students.

An added problem is that faculty and students often do not agree on what actions constitute cheating in and out of the classroom (14-17). The literature recommends that college teachers should be very specific in their definition of academic dishonesty, giving concrete examples, and then following through on consistent discipline when cheating occurs (18, 19). In an effort to determine the level of potential disagreement and/or confusion as to what constitutes cheating behaviors in and out of the classroom, the students and faculty of the University of Montevallo were presented with a variety of examples of academic misconduct, and then asked to rank their perceived severity on a scale from 1 to 5 (1 = Not Severe to 5 = Very Severe) (14). The results of this study are shown in Table 2. In several cases (see questions 22-24) there was almost a full point difference between the student and faculty perception indicating a lack of communication between faculty and students. Some of the most problematic areas of disagreement (see questions 3, 5, 12, 14, and 15) indicate a educational moral laxity on the part of the students.

One may interpret these results in two different ways. One possibility is that the results reflect stricter standards developed by faculty members as they moved in their careers. In other words, their perception reflects a more mature evaluation of the scenario being considered. If this interpretation is correct, one also would expect students to improve their moral standards as they mature. In other words, the students' perception of what constitutes misconduct, should not have any influence in their future professional conduct. This hypothesis, however, does not take into consideration that the faculty members polled in this study already had a different perception of what constituted cheating

	Question	Faculty	Student	P
1	Looking directly at another persons paper to copy an answer during a test	4.88 ± 0.67	4.38 ±1.29	0.0017
2	Using "crib notes" or a "cheat sheet" during a test or class assignment	4.83 ± 0.70	4.32 ± 1.31	0.0016
3	Getting a copy of the test prior to taking it	4.80 ± 0.76	3.94 ± 1.23	0.0001
4	Having/paying someone to do homework or at-home projects for you	4.76 ± 0.81	4.06 ± 1.20	0.0001
5	Copying someone's homework	4.63 ± 0.87	3.77 ± 1.19	0.0001
6	Using answer book or keys to get homework answers	3.95 ± 1.24	3.10 ± 1.34	0.0001
7	Leaving the test to go to the restroom/or another place to get answers	4.77 ± 0.84	4.24 ± 1.33	0.0022
8	Answering "here" or signing someone's name when he/she is absent	4.71 ± 0.79	3.55 ± 1.32	0.0001
9	Copying someone's paper to work and putting your name on it	4.82 ± 0.73	4.17 ± 1.30	0.0001
10	Trying to influence a teacher to give you a better grade	3.46 ± 1.31	2.83 ± 1.30	0.0011
11	Using sorority/fraternity test files	3.56 ± 1.46	3.05 ± 1.47	0.0178
12	Finding someone's idea and using it as your own	4.36 ± 1.00	3.77 ± 1.32	0.0009
13	Asking for answers with gestures or sign language during an in-class assignment	4.54 ± 1.01	3.93 ± 1.39	0.0010
14	Plagiarism of resource materials or documented work	4.76 ± 0.75	4.06 ± 1.39	0.0010
15	Using another's research for your own benefit	4.31 ± 1.13	3.67 ± 1.40	0.0008
16	Watching someone cheat without reporting it	3.51 ± 1.23	2.88 ± 1.26	0.0007
17	Not carrying your weight in a group project for which everyone gets the same grade	3.93 ± 1.17	3.62 ± 1.36	0.0991
18	Using sources on homework which the professor told you not to use	4.15 ± 1.16	3.59 ± 1.26	0.0526
19	Getting a teacher's copy of a test to sell	4.62 ± 1.03	4.22 ± 1.31	0.0072
20	Conducting group sessions to swap or check the accuracy of answers	2.71 ± 1.35	2.15 ± 1.34	0.0166
21	Giving answers with gestures or sign language during an in-class assignment	4.50 ± 1.18	3.83 ± 1.30	0.0017
22	Lying to a teacher about why you are not prepared in class	4.22 ± 1.98	3.27 ± 1.31	0.0000
23	Taking money for doing someone's work	4.58 ± 1.01	3.62 ± 1.33	0.0001
24	Glancing at another paper and seeing something to jar your memory	4.40 ± 1.15	3.49 ± 1.24	0.0000
25	Working with someone else on a take-home exam	3.92 ± 1.37	3.06 ± 1.37	0.0004

Table 2. Perception by Faculty and Students of Cheating Behavior in College. 140 students and 108 faculty members were asked to assign a value to the perceived severity of the behavior on a scale of 1 to 5, with 5 being most severe. The results are presented as average ± SD. The study was carried out at the University of Montevallo during the Fall of 1997.

when they were in college. They grew up with a different set of standards, in an environment in which cheating was not as prevalent. Thus, accepting this hypothesis would imply that regardless of the predominant moral values among college students at any given point in history, they will always develop the correct moral values as they become professionals.

An alternative hypothesis is that, although the moral standards of most individuals increase through life, some of these individuals do not see any need to change their values. For them the

concept of "misconduct" disappears. The concern of those interested in maintaining high post-secondary educational standards is that the habits established by some college students will continue to be their habits in graduate school, employment and research in the future. Therefore, an increase in the proportion of an undergraduate students involved in academic misconduct is likely lead into an increased incidence of professional misconduct in the future.

The current situation is likely to deteriorate even more. The development of the Internet at the end of the 20th century has also increased the number of cheating episodes by providing tools that were not available even 10 years ago. Students may now download an enormous amount of information in seconds, which may be incorporated into a paper with a couple of keystrokes. Moreover, several virtual companies have proliferated offering term papers in all disciplines on a per page cost (see for example, www.schoolsucks.com, www.ezwrite.com, www.academictermpapers.com, etc.). In the last two years there has been a increase in number of cases of plagiarism by students who simply download text from the internet, not just at the University of South Alabama and the University of Montevallo but also at many other institutions. When confronted by the faculty, these students are dismayed at getting caught, but many will repeat similar behaviors in the future. The only tools available to faculty to identify these cases is to search the web for a specific (unique) paragraph in the paper or to contract the services of commercial search engines (for example, www.plagiarism.org) that can look for the papers sold to students by Internet companies. The first procedure is time-consuming and limited. Hiring the services of a company to track these papers down still requires someone to enter the text in the Internet and also the becomes too expensive.

Since the formative years of college are important in setting many of our standards, as the students' academic standards decrease future professionals may find it easier to engage in scientific misconduct as they will perceive it to be less immoral and more expedient. For example, a study done with 2,459 sophomore medical students showed that 4.7% admitted to cheating while 66.5% admitted to having heard of cheating among their peers (20). About 70% of the students that admitted having cheated in medical school also admitted to cheating in high school and college. Thus we see a moral laxity beginning at the high school level (or before) and progressing, probably with more cheating occurring rather than less, as the level of the academic workload increases.

One of the established patterns of human development is the relative stability of personality traits and behavioral habits over the life span. Thus, traits of dishonesty in the face of hard or demanding intellectual work in college, will, in all likelihood, remain stable characteristics as these college students grow older. One cognitive/moral development theorist, Kohlberg, proposed a universal set of discrete stages of moral development based on Piaget's theory of cognitive development (21, 22). As a child develops more complex and different modes of thinking and reasoning, the child should also be able to make more complex and adaptive moral judgments. Kohlberg proposed a six-level moral developmental sequence. At Levels 1 and 2, there is a basic desire to escape punishment and to win some level of approval from significant others. At Levels 3, 4, 5, and 6, the individual may progress from living up to others' expectations, to following rules to maintain the social order and avoid chaos, to adhering to a social contract only when it appears to be valid to the individual, and, finally, to upholding moral judgments and principles despite potential harm or threat to oneself because of their intrinsic worthiness.

Kohlberg proposes that rarely do most individuals progress in moral development past Level 3 or perhaps 4 (21, 22). We do the "right" thing in any given situation to garner favor and approval from others who expect a substantial effort from us. And, if we perceive the rules that are in place for us to follow to be unfair or nonsensical, we may make a judgment to avoid complying with those rules on what we call moral grounds.

With Kohlberg's postulations in mind, it is then easy to hypothesize that an individual who learned to cheat in academic situations without active reprisal from faculty or a school administration, would tend to repeat those cheating behaviors in future learning/academic/research situations as a way to gain approval for completion of the assignment or project. In addition, if the adult who participated in academic dishonesty all the way through graduate school may view the demands of a thesis or dissertation committee as non-valid, that

individual may engage in academic dishonesty with an almost-clear conscience. The requirements of "publish or perish," then, in the post-academic world may become "non-valid" in the professional's mind, and the individual may continue to participate in dishonesty in research.

In summary, the correlation between cheating in high school, college and in medical school supports our hypothesis that future professional misconduct will also show a positive correlation with previous history. Thus, we propose that part of the efforts to promote integrity among future professionals should be devoted to curbing cheating at the undergraduate level since an increase in one is likely to increase the other.

Bibliography

1. Aaron RM, Georgia RT. Administrator perceptions of student academic dishonesty in collegiate institutions. NASPA Journal 1994; 31: 83-91.
2. Diekhoff GM, LaBeff EE, Clark RE, Williams LE, Francis B, HainesV J. College cheating: Ten years later. Research in Higher Education 1996; 37: 487-502.
3. Haines VJ., Diekhoff GM, LaBeff EE, Clark RE. College cheating: Immaturity, lack of commitment, and the neutralizing attitude. Research in Higher Education 1986; 25: 342-354.
4. Holm TT. Ignorance of the crime is no excuse: Do public speaking students even know what cheating is? Paper presented at the 84th Annual Meeting of the National Communication Association (New York, NY, November 21-24, 1998). ERIC NO: ED427355.
5. Lord T, Chiodo D. A look at student cheating in college science classes. Journal of Science Education and Technology 1995; 4: 317-324.
6. Maramark S, Maline MB. Academic dishonesty among college students. Issues In Education. 1993; U.S. Department of Education, Office of Research, p. 5, ERIC NO: ED360903.
7. Whitley BE Jr. Factors associated with cheating among college students: A review. Research in Higher Education 1998; 39: 235-274.
8. Graham MA, Monday J, O'Brien K, Steffen S. Cheating at small colleges: An examination of student and faculty attitudes and behaviors. Journal of College Student Development 1994; 35: 255-260.
9. Genereux RL, McLeod BA Circumstances surrounding cheating: A questionnaire study of college students. Research in Higher Education 1995; 36: 687-704.
10. Moffatt M. Undergraduate cheating. 1990; ERIC NO: ED334921.
11. Greene AS, Saxe L. Everybody (else) does it: Academic cheating. 1992; Paper presented at the Annual Meeting of the Eastern Psychological Association (Boston, MA, April 3-5), ERIC NO: ED347931.
12. Meade J. Cheating: Is academic dishonesty par for the course? Prism 1992; 1: 30-32.
13. Jendrek MP. Faculty reactions to academic dishonesty. Journal of College Student Development 1989; 30: 401-406.
14. Fleming KD, Keegan DM, Staik IM,, Burling JW. Faculty perception of cheating in college: Toward a campus consensus. 1998; Paper presented at the 1998 NCUR (National Conferences on Undergraduate Research) meeting, Salisbury State University, MD.
15. Livosky M, Tauber RT. Views of cheating among college students and faculty. Psychology In The Schools 1994; 31: 72-82.
16. Roig M, Ballew C. Attitudes toward cheating by college students and professors. 1992; ERIC NO: ED349895.
17. Falleur D. An investigation of academic dishonesty in Allied Health. Journal of Allied Health 1990; 19:313-324.
18. Davis SF, Ludvigson HW. Additional data on academic dishonesty and a proposal for remediation. Teaching of Psychology 1995; 22: 119-121.
19. Beemsterboer PL. Academic integrity: What kind of students are we getting and how do we handle them once we get them? Journal of Dental Education 1997; 61: 686-688.
20. Baldwin DC Jr., Daugherty SR; Rowley BD, Schwarz MR. Cheating in medical school: a survey of second-year students at 31 schools. Acad. Med. 1996; 71: 267-273.
21. Kohlberg L Essays on moral development: Vol. II. The psychology of moral development. San Francisco: Harper & Row; 1984.
22. Kohlberg L. The development of moral judgment and moral action. In L. Kohlberg (ed.), Child psychology and childhood education: A cognitive-developmental view. New York: Longman; 1995, pp. 259-328.

2. Institutions and Professions

Comprehensive Guidelines for the Responsible Conduct of Researchers

Gregory Brock, Department of Family Studies, University of Kentucky, USA

Sandra Sutter, Department of Family Studies, University of Kentucky, USA

Ada Sue Selwitz, Office on Research Integrity, University of Kentucky, USA

Keywords: Conduct guidelines, Research ethics, Researcher ethics

In 1989, the Department of Health and Human Services (DHHS) through the Public Health Service defined research misconduct and established regulations for reporting scientific misconduct among awardee and applicant institutions (1). The focus of this regulation was on fabrication, falsification, and plagiarism. More recently DHHS has shifted emphasis toward preventing misconduct and to the promotion of Responsible Conduct in Research (RCR).

Success in implementing regulatory initiatives on research integrity has been stymied by several factors. There is disagreement about the extent of research misconduct. Steneck (2) reported that fewer than 200 cases of misconduct have been documented by federal government research investigation offices over the past 20 years. Indirect evidence also cited by Steneck, however, suggests that misconduct may occur far more frequently.

Additionally, there is a lack of clarity about what amounts to research misconduct. In 1989, the term focused on, "…fabrication, falsification, plagiarism, or other practices that seriously deviate from those that are commonly accepted within the scientific community for proposing, conducting, or reporting research."(1). Defining deviant practice as well as what is common practice is particularly challenging in view of the rapid development now occurring within many scientific disciplines—what was deviant can become common practice. Plus, collaboration among academic disciplines, between universities and industry, between universities and government, and between international research teams creates new syntheses that further complicate our understanding of what constitutes common practice. In an effort to address these issues, regulators have turned to requiring training of researchers as one means of communicating that the incidence of misconduct is troubling. Training objectives also clarify what amounts to misconduct.

On December 1, 2000, the DHHS Office of Research Integrity adopted and published the final PHS Policy on Instruction in the Responsible Conduct of Research that delineates RCR training requirements to all research investigators applying for or using PHS funds and their institutions (3). Although nine core areas of instruction are specified, the policy does not establish the exact content in the form of standards and principles within each area. In complying with this mandate, each institution will be responsible for its own content.

Much attention in the RCR literature has been directed to standards within specific areas, such as authorship, peer review, and collaborative practices. Presentations at national conferences and

Corresponding author: Gregory Brock, Ph.D., Department of Family Studies, University of Kentucky, 316 Funkhouser Building, Lexington, KY 40506-0054, 859-257-7742 (voice), 859-257-321 (fax), gwbrock@pop.uky.edu.

institutional committees have addressed RCR practice standards. As well, many professional associations have established standards of conduct within their ethical codes. Institutional policies such as *Guidelines for the Conduct of Research at the National Institute of Health* have also incorporated a selection of RCR topics (4). However, no single set of principles encompassing all aspects of responsible conduct of research exists in unified form.

Grinnell (5) pointed out that "…promoting responsible conduct of science requires a clear description of what doing science entails." In addressing why standards are important, Frankel (6) discussed the need of the general public for accountability in science, and how a set of standards not only meets this need but also increases trust in the scientific community. Frankel noted specific benefits to establishing ethical standards: Standards provide an enabling document, professional socialization, public accountability, gain public trust/support, improve public relations, self-preservation, deterrence, professional support, and are a source of public policy. Standards also provide guidance when an ethical course of action is unclear. Mastroianni and Kahn (7) point out that training students in the basics of RCR is crucial to the continued maintenance of public trust in the scientific community by cultivating the integrity of research practices. However, results on the effectiveness of RCR training thus far are inconclusive (8, 9). Brown and Kalichman (9) offer the interpretation that a lack of consensus on what constitutes misconduct may contribute to the lack of clarity on the effectiveness of training.

Frankel (10) advocates the development of research standards as the single most important step in promoting scientific integrity and handling misconduct. Faced with the new training requirements established by the PHS, this step is particularly important for promoting and supporting a climate of integrity at the organizational level that can function in a reciprocal fashion to influence and be influenced by individual actions.

Initially, the purpose of the document presented here was to provide a comprehensive set of guiding principles to serve as a basis for RCR training at the University of Kentucky. Content analysis was applied to an exhaustive list of behavioral guidelines identified in a thorough review of the research integrity literature including ethics codes of professional associations. Guidelines were then sorted into discrete thematic categories. These categories were called principles because they identified core values of research practice. Three groups of principles emerged from the analysis: General, Professional, and Focused. Subprinciples also were defined that served to elucidate contemporary issues rather than merely exemplifying situations in which the principles might apply. A series of revisions were made after obtaining feedback from research colleagues and university administrators.

What emerged was a comprehensive set of guidelines for the conduct of researchers more akin to a code of conduct for a profession (see attached guidelines). These guidelines provide a broad-based foundation for the safe and effective practice of research across disciplines, settings, methods, and questions. Our intent in presenting them here is to increase the awareness and sensitivity of institutional policy makers to the many issues that researchers must attend to in the conduct of their professional responsibilities. By presenting the results of our analysis, we wish to further the discussion about the content of RCR training.

Acknowledgements:

Support was provided by the Vice President for Research and Graduate Studies, University of Kentucky.

References

1. Department of Health and Human Services, Public Health Service. Responsibilities of Awardee and Applicant Institutions for Dealing With and Reporting Possible Misconduct in Science. Federal Register Vol. 54 No. 151 Tuesday, August 8, 1989.
2. Steneck, N. Assessing the Integrity of Publicly Funded Research: A Background Report. Paper prepared for the Department of Health and Human Services, Office on Research Integrity Research Conference on Research Integrity on November 2000, Bethesda, MD.
3. PHS Policy on Instruction in the Responsible Conduct of Research, Department of Health and Human Services, Public Health Service, Office of Research Integrity. Available at URL: http://ori.dhhs.gov/html/programs/rcrcontents.asp
4. American Association of Medical Colleges. Developing a code of ethics in research: A guide for scientific societies, executive summary. Conference Materials of the Public Responsibility in Medicine and Research; 1997 May 13-14; Bethesda, Maryland.
5. Grinnell, F. Ambiguity, trust, and the responsible conduct of research. Science and Engineering Ethics, 1999; 5 (2): 205-214.

6. Frankel, M. S. Developing ethical standards for responsible research: Why? Form? Functions? Process? Outcomes?, Journal of Dental Research 1996; 75 (2): 832-835.
7. Mastroianni, A. C., & Kahn, J. P. The importance of expanding current training in responsible conduct of research. Academic Medicine 1998; 13 (12): 1249-1254.
8. Eastwood, S., Derish, P., Leash, E., & Ordway, S. Ethical issues in biomedical research: perceptions and practices of postdoctoral research fellows responding to a survey. Science and Engineering Ethics 1996; 2: 89-114.
9. Brown & Kalichman. Effects of training in RCR: A survey of graduate students in experimental sciences. Science and Engineering Ethics 1998; 4: 487-498.
10. Frankel, M. S. Scientific community must set the standards. Forum for Applied Research and Public Policy, Spring 1998

Guidelines for the Responsible Conduct of Researchers

Preamble: Advancing the scientific record is the noble task of those who conduct research. In large part the quality of that record is the product of inquiry. Ranging well beyond the conduct of research however is the realm of activities constituting the work of researchers that influences the public trust, that affects global well-being, and that indirectly affects the scientific record. The guidelines presented here define expectations so that researchers uphold the highest ethical standards by practicing within the bounds of both effectiveness and safety.

Important, sustaining values that support humankind and global well-being serve as the basis for three groups of principles and sub-principles. (1) General principles apply to all research contexts. (2) Professional principles define relations among researchers and practices that constitute the scientific method. (3) Focused-principles address discrete aspects of research practice for particular investigations, research contexts, or scientific disciplines. Sub-principles elucidate contemporary issues rather than identifying the component issues of any principle.

Where governmental laws contradict these guidelines, researchers are cautioned to seek consultation from appropriate authorities and colleagues. Resolution is not always possible, consequently, researchers act so as to benefit the greater good even if that path demands personal sacrifices.

In an effort to create a research climate worthy of the public trust, it is incumbent upon researchers to report any breech of these guidelines to an appropriate authority. Where there is no relevant authority, researchers are obliged to focus public media attention on wrong doing.

These guidelines apply to professional and amateur researchers, students, research technicians, research administrators, as well as private, public, and governmental research agency personnel.

General Principles

General Principle 1: Commitment to Society and to Global Well-being

Researchers protect the interests of society within a broader commitment to global well-being. They recognize that the public has entrusted them to uphold the integrity of the scientific record.

1.1 Researchers do not obligate themselves to withhold research findings that may jeopardize the health or well-being of others.

1.2 Researchers take active steps to prevent the misuse of their findings that may jeopardize the well-being of others.

1.3 Researchers take active steps to correct errors or oversights in proposing, conducting, or reporting research.

1.4 Researchers present themselves to the public in a competent, sincere, and trustworthy manner.

General Principle 2: Commitment to Competency

Researchers are aware they are responsible for maintaining professional competency and remaining knowledgeable within their areas of expertise.

2.1 Researchers conduct their work within the scope of their own training and knowledge base.

2.2 Researchers recognize they are vulnerable to stress and impairment. When stress or impairment interferes with their ability to conduct professional responsibilities, researchers seek assistance.

2.3 Researchers ensure that all persons who assist in the conduct of their research are adequately trained and perform their responsibilities competently.

2.4 Researchers inform their work with views, values, and co-workers from diverse sources.

2.5 Researchers foster a scientific community in which discrimination based on gender, race age, sexual orientation, religious affiliation, ethnic or national origin does not occur.

General Principle 3: Understanding Laws, Regulations, and Policies

Researchers are aware of and stay informed of professional, institutional, and governmental regulations and policies in proposing, conducting, and reporting research.

3.1 Researchers take active steps to resolve discrepancies when policies or regulations are unclear or contradict one another.

General Principle 4: Conflicts of Interests

Researchers are cognizant that conflicts of interest occur in the context of professional activities and they recognize and avoid them.

4.1 When researchers cannot avoid real or perceived conflicts of interest, they seek consultation and take active steps to minimize bias, flawed judgment, harm, or exploitation.

Professional Principles

Professional Principle 5: Peer Review

Researchers respect others' rights to have work reviewed in a confidential, timely, and objective manner.

5.1 Researchers assess and disclose multiple roles or allegiances which may undermine the confidential and fair review of others' work.

5.2 Researchers take active steps to protect the integrity of review materials and guard the intellectual property of others.

Professional Principle 6: Research Management and Data Access

Researchers clearly and authentically record data and methods. They protect the integrity of their research materials. They make data, methods, and materials available to others for analysis or replication.

6.1 Researchers select materials appropriate for data acquisition, recording, and storage.

6.2 Researchers stay informed of and implement policies for appropriate storage and disposal of research materials.

6.3 Researchers take active steps to select methods and materials that protect research participants' right to privacy.

6.4 Researchers take active steps to safeguard data when using electronic or Internet-based methods.

6.5 Researchers are cognizant of the ownership of their research data, methods, and findings.

Professional Principle 7: Commitment to Credibility

Researchers engage in practices that are currently accepted within the scientific community to propose, conduct, and report research.

7.1 Researchers practice honest stewardship of their research resources and use recognized accounting methods.
7.2 Researchers do not conduct their professional responsibilities in a manner that is intentionally deceitful or with reckless disregard for the truth.
7.3 Researchers who witness or suspect fraud or misconduct follow established procedures to preserve the integrity of the scientific record.
7.4 Researchers accused of fraud or misconduct do not harass those believed or known to have made accusations against them.
7.5 Researchers do not misrepresent their work by omitting data that changes the meaning or significance of their findings.
7.6 Researchers do not fabricate or falsify data.
7.7 Researchers do not present or publish component findings of a larger body of work if misunderstanding may result or to conceal findings.

Professional Principle 8: Mentoring, Training, and Supervisory Relationships

Researchers nurture the intellectual, technical, ethical, and career development of their trainees, supervisees, and students.

8.1 Researchers recognize that trainees, supervisees, and students have needs unique to their individual strengths and limitations. Researchers provide guidance, constructive feedback, and assistance that matches the changing needs of each trainee, supervisee, or student.
8.2 Researchers establish clear and appropriate rules and boundaries in their relationships with trainees, supervisees, and students.
8.3 Researchers do not engage in sexual harassment, disrespect the character of, or impede the progress of their trainees, supervisees, and students.
8.4 Researchers recognize that exploitation is a risk in relationships where differences in power exist. They avoid conflicts of interest and dual relationships. Sexual interaction with subordinates is avoided.
8.5 Researchers take active steps to inform trainees, supervisees and students of supervisors' responsibilities to avoid dual relationships.

Professional Principle 9: Authorship and Publication Practices

Researchers respect the intellectual property rights of others.

9.1 Researchers attribute credit for others' words and/or ideas in proposing, conducting, or reporting their own work.
9.2 Researchers facilitate discussion and set ground rules early in collaborative relationships regarding authorship assignment.
9.3 Researchers assume responsibility for the accuracy of research reports for which they claim full or co-authorship.
9.4 Researchers preserve the integrity of the scientific record by taking active steps to correct errors in the publication of their findings.
9.5 Researchers do not submit or publish previously published materials without appropriate citation.
9.6 Researchers respect the privacy of others' unpublished work.

Professional Principle 10: Responsibilities to Colleagues and Peers

Researchers recognize they are members of the scientific community and respect the contributions of others to the scientific record.

10.1 Researchers clarify early in a collaborative project the expectations and responsibilities among those involved.

10.2 Researchers do not impede the progress of others' work.
10.3 Researchers protect the integrity of intellectual property and research materials when reviewing others' work.
10.4 Researchers take active steps to maintain positive relations among team members and to seek consultation if necessary to resolve interpersonal conflicts.

Focused Principles

Focused Principle 11: Protection of Human Participants

Researchers respect the dignity of human participants and take active steps to protect their well-being. They follow institutional, professional association, and governmental ethical and regulatory guidelines.

11.1 Researchers ensure that each participant gives voluntary and informed consent regardless of age, race, gender, ethnic or national origin, sexual orientation, mental or physical health status, or incarceration.
11.2 Researchers take active steps to evaluate and to minimize potential risks to participants.
11.3 Researchers respect each participant's right to privacy, and they take active steps to protect confidentiality of data or other disclosures.
11.4 Researchers take active steps to achieve an equitable balance of benefits and risks to each participant.
11.5 Researchers honor fairness and equity in the selection of research participants.

Focused Principle 12: Care and Use of Animals for Research

Researchers are stewards of animals used for research. They follow institutional, professional association, and governmental ethical and regulatory guidelines.

12.1 Researchers substitute inanimate materials and processes for animals where appropriate. When this is not possible, researchers make active efforts to use species that may be less susceptible to pain and distress.
12.2 Researchers take active steps to use procedures which reduce the incidence and/or severity of pain and distress experienced by animals.
12.3 Researchers take active steps to reduce the use of animals to the minimum number necessary to yield valid answers to their research questions.

Focused Principle 13: Commitment to Native Populations and Other Identifiable Groups

Researchers respect the rights and protect the interests of Native populations and other identifiable groups.

13.1 Researchers who work with Native populations and other identifiable groups recognize that to minimize risks and to maximize benefits to individuals and to populations themselves there is value in obtaining the advice, participation, and viewpoints of those individuals and populations in formulating research questions, designing research methods, collecting and analyzing data, and in reporting results.
13.2 Researchers recognize that consent from or consultation with group authorities or representatives is sometimes necessary before obtaining consent from individuals within Native populations or other identifiable groups.
13.3 Researchers take active steps to distinguish individual property both tangible and intangible from collective property owned by Native populations or other identifiable groups.
13.4 Researchers take active steps to reduce the risk to Native populations or other identifiable groups that result from misuse of their research findings.

Focused Principle 14: Genetic Research and Technology

Researchers strive to preserve and protect global well-being from the unintended consequences of genetic research.

14.1 Researchers involved in genetic research take active steps to identify potential risks and benefits to research participants. They inform participants of the possibility that risks may not yet be identified.

14.2 Researchers take active steps to protect the confidentiality of genetic materials collected from human participants and do not allow the use of these materials for purposes which may discriminate against or harm an individual or group of individuals.

14.3 Researchers are sensitive to social, physical, psychological and environmental factors that may influence individuals' consent to participate in genetic research.

14.4 Researchers inform individuals, their families, and Native and other identifiable populations of the disruptive influence that genetic research may have on their lives. They take active steps to minimize disruptions.

14.5 Researchers are cognizant of the increasing complexity of the ethical concerns about genetic research. They stay informed of the developing research guidelines as well as the public discourse about genetic research.

14.6 Researchers actively participate in the development and refinement of ethical standards in this area.

Research Integrity in Social Work: Status, Issues, and Future Directions

Margaret Gibelman, Wurzweiler School of Social Work, Yeshiva University, USA

Keywords: Scientific integrity, social work, code of ethics, scientific misconduct, institutional review boards

This paper explores the issue of scientific integrity in social work and its implications for the training of social work researchers. Data concerning a growing body of cases in which allegations have been made and/or violation of legal and ethical research standards have been substantiated illustrate that the integrity of research in social work and related fields is a growing concern. However, mechanisms to review and monitor social work research are under-developed compared to other disciplines. A research agenda is offered to assess the status of institutional systems to review and monitor research in social work and, concurrently, determine social workers' familiarity with the profession's ethical code as it relates to research integrity. Implications for faculty and practitioner education and training and the development and enforcement of systems to review the integrity of research protocols are explored.

Scientific misconduct or, more positively, appropriate conduct in the realm of research inquiry, is a topic that has received very little attention in the social work literature. Unfortunately, this is because social workers have not, historically, been strong contenders in the successful competition for federal research grants, particularly large-scale research protocols (1, 2, 3, 4). Social work research is still in its infancy compared to research in other disciplines. However, there is a professional commitment to increase the capacity and productivity of social work research, as evidenced by the burgeoning number of social work research centers and a growing empirical social work literature base. This expansion of social work research is not without risks. Although the majority of publicized cases of scientific misconduct have centered largely on bio-medical research and the applied sciences, the circumstances associated with these cases have strong implications for the preparation of students and the standards to which social work researchers will be held. The growing number of cases in fields related to social work, as discussed below, highlight areas of potential vulnerability.

The Status of Social Work Research

Unlike most of the social and behavioral sciences, social work is a practice-based profession rather than an academic discipline or field. Social work has been defined as the "applied science of helping people achieve an effective level of psychosocial functioning and effecting societal changes to enhance the well-being of all people" (5). Historically, its knowledge base has been predicated upon a liberal arts perspective and has drawn from psychology, psychiatry, sociology, political science, economics, and other disciplines to formulate applied practice principles. However, within the past two decades, social work has striven to define its own unique body of knowledge, an effort incorporated into the purposes of social work itself, one of which is "the development and testing of professional knowledge and skills..." (6).

Corresponding author: Margaret Gibelman, Yeshiva University, Wurzweiler School of Social Work, 2495 Amsterdam Avenue, New York, NY 10033, 212-960-0840 (voice), 212-960-0822 (fax), gibelman@ymail.yu.edu.

Although research has always had a place within the purposes of the profession, the larger socio-political environment has, in recent years, profoundly affected the priority afforded to research. There is a growing mandate for all social workers to incorporate research into their practice, a phenomenon underscored by the demands of funding bodies, oversight agencies, and consumer choice movements for hard data documenting that programs of service lead to tangible results. A leading force has been that of managed care, which has brought with it heightened demands for accountability, with particular emphasis on documenting the successful outcomes of service (7).

At the same time that external demands to provide empirical evidence of the impact and outcomes of services grow, social workers, to better protect the interests and well-being of the people they serve, are seeking to empirically examine the consequences of the managed care movement, itself. This has translated to a concern about documenting the effects of managed care (e.g., short-term hospitalization; short-term treatment; limited provider choice). These developments have led to the need for a new or enhanced repetoire of research skills on the part of not only academics and researchers, but among the totality of social workers directly providing, supervising, or managing the delivery of human services.

The long and ongoing admonishment that the profession must develop an internal research capacity has borne fruit. In fact, a notable number of studies have been conducted on the status of research productivity and the scholarly contributions of social workers (8, 9, 10, 11, 12, 13). Perhaps the most significant influence, however, on the growing social work research enterprise has been the shift in criteria for tenure and promotion within academia, which remains the richest source of social work research (14, 15). Longevity of academic careers now rests firmly on scholarly productivity and standards related to both quality and quantity continue to rise as social work is increasingly held to the same standards as other academic and professional units within the university (4). A related factor in the emphasis on research productivity is the growing sophistication of faculty in identifying funding sources and competing successfully for publicly supported research dollars.

The emergence of schools of social work as major research centers and the increased productivity of social work researchers has been long in coming. The mandate to create a coordinated research infrastructure had been echoed for two decades (16, 17, 18, 19). National Institute of Mental Health funding has been a major impetus to establish social work research centers at academic institutions. In this process, however, the profession faces a host of issues and challenges, foremost among them the preparation of future researchers, including socialization to the ethos of scientific integrity.

Ethical Guidelines

The latest revision of the Code of Ethics of the National Association of Social Workers (NASW)(20) emphasizes the central role of research: "social workers should contribute to the knowledge base of social work and share with colleagues their knowledge related to practice, research, and ethics. Social workers should seek to contribute to the profession's literature and to share their knowledge at professional meetings and conferences" (Section 5.01(d), p. 24). Section 5.02 (b) of the Code (1996) encourages social workers to "promote and facilitate evaluation and research to contribute to the development of knowledge" (p. 25).

The Code of Ethics not only seeks to establish an obligation on the part of social workers to engage in knowledge building through empirical research, but also provides the basic guidelines for how such research is to be conducted. Specific provisions pertain to risk-benefit analysis, voluntary and written informed consent, protection from harm, confidentiality, and accurate reporting of findings. Further, the Code sets forth the obligation of social workers to educate themselves and for programs of social work education to provide relevant education concerning responsible research practices.

An important caveat about ethical guidelines exists that is idiosyncratic to the profession — the limited application of the Code to social workers. The Bureau of Labor Statistics (21) estimates that there are approximately 542,000 professional educated social workers in the United States (at the bachelor's, master's and doctoral levels). At the same time, current membership of the National Association of Social Workers is approximately 155,000. The Code of Ethics is a product of the National Association of Social Workers and, upon joining,

members must pledge to abide by the Code. But what about the more than 387,000 social workers who are not members of NASW and not committed to abiding by the provisions of the Code? These social workers may belong to other professional associations which have their own ethical guidelines, but data to support this contention are lacking (22). Social work researchers based in institutions of higher education may have their own review and oversight procedures, separate from university-wide IRBs, but again there is an absence of substantiating empirical data. An unknown, but impressionistically high proportion of social work research is outside the purview of federal funding, which may mean that IRB review procedures are not applied. (It should be noted, however, that such research is now selectively being reviewed by IRBs to conform with their own internal procedures, partially reflecting the prevalence and influence of the growing number of studies sponsored by private sources, including pharmaceutical companies, in areas such as genetic testing (23).)

Finally, in some instances, social work research may be absent of any oversight by any source. This latter scenario is most likely to prevail among those working in service organizations which have not yet established review and oversight procedures and may, indeed, not even recognize the need to do so. Of particular concern is the mandate for practice agencies to engage in research without assurances of appropriate procedures and absent collaborations with educational institutions from which such protocols may be borrowed.

Learning from the Mistakes of Others

To date, public disclosure of cases of scientific misconduct within the social work research community have been absent. Over a 10 year period of vigilant reporting of scientific misconduct, the *Chronicle of Higher Education* referenced only one situation involving a social worker. This case concerned a researcher who submitted bogus articles to professional journals as part of an experiment to test peer-review practices (24). Because the research did not involve the use of Federal funds, review of allegations of ethical misconduct remained within the purview of the adjudication process of the NASW. Ultimately, NASW dismissed the complaint, arguing that the issue involved a disagreement over research methods rather than ethics and that there had not been an explicit violation of the Code of Ethics (24). However, as social workers increasingly compete successfully for federal research funds, they become subject to the same level of scrutiny as researchers in other disciplines. Similarly, as IRBs extend their purview to include privately supported research, more diligent reviews of social work research protocols can be expected.

As the social work profession seeks to enhance its research capability in a credible and responsible manner, there is much to be learned from the experience of related disciplines and professions. In recent years there has been a growing number of cases of scientific misconduct among allied health-related industries (e.g., nursing, psychology, and psychiatry), the predominant theme of which concerns plagiarism and/or falsification or fabrication of data (25, 26, 27, 28, 29). Eight cases from the helping professions over the last decade were identified from media reports, out of an unknown universe of substantiated cases of misconduct. Unlike many cases of misconduct substantiated in the bio-medical fields, these cases were absent allegations of human subjects violations. However, findings of misconduct highlight the diligent reviews to which research reports are subject and the serious penalties that are levied when ideas are appropriated or results falsified. Sanctions include forced resignations, criminal prosecution, ineligibility from receiving publicly supported grants or serving on review panels, and remedial courses in ethics. These sanctions have widespread and serious implications for how research is conducted and highlight the potential consequences that may ensue when procedural and ethical breaches are uncovered.

Emerging Issues

The mistakes of researchers of allied disciplines suggest the scope and magnitude of potential areas of scientific misconduct that may similarly affect social work. Further, the record on misconduct shows that attention to the initial review of protocols is only a beginning step in an ongoing process necessary to ensure scientific integrity. Although a systematic process for reviewing research proposals, including attention to scientific validity of the study design, can alleviate many potential problems, it is in the reporting of research findings, at least to date, that the allegations of scientific misconduct are

most likely to occur. Reports of research are, in fact, reviewed; how research is carried out and findings reported are subject to scrutiny, and, sometimes, reprisals. This fact presents a formidable problem in balancing the traditional academic freedom associated with the pursuit of research and institutional responsibility to ensure accountability of the outcomes of such research. The extent to which a school of social work can monitor the work of its faculty and students is inherently limited.

While only about 30% of the cases of scientific misconduct are eventually determined to be founded, the impact of the allegations is profound (30). The investigation of allegations consumes significant institutional resources and can ruin careers, even if the allegations are unfounded. If allegations are confirmed, it is lethal to a researcher's career (see, for example, 31), causes reputational damage to the university, and may affect public perceptions of the integrity of all research. Worse, human lives and well-being may be compromised (4).

Internal systems to prevent and, when necessary, address scientific misconduct are not without their critics. There are enormous workload implications, particularly for senior faculty who may not have the time or desire to spend their time monitoring junior faculty. There are also those who argue that when schools/universities serve as the "scientific validity police" of their own colleagues, they will either join ranks in defense, or, to the other extreme, find against their colleagues for fear of accusations of institutional bias (32, 33).

Current Review Mechanisms

Since allegations and, in some cases, findings of scientific misconduct are, by definition, after-the-fact of the activity, the most significant lesson from these cases is the importance of ensuring that research review and monitoring procedures are uniformly followed. The integrity of scientific research is monitored by two main and distinct sources: professional associations and their applicable ethical codes and institutional review boards (IRBs). In social work, these mechanisms for ensuring research integrity are less firmly entrenched. As discussed earlier, there is no one body with the authority or jurisdiction to oversee the entirety of the social work research enterprise. The guidelines detailed in the profession's Code of Ethics about ethical research conduct are, however, limited by their lack of applicability to a large proportion of social workers. Social work educators, who are the major producers of research, are ill-represented among the membership of NASW and are thus outside of its monitoring and adjudication provisions. Thus, the question of what mechanisms govern academic social work research remains unanswered.

The majority of schools of social work are housed in research universities which have their own IRBs and the logical source of research review and oversight lies with IRBs. However, the focus of many, if not most, IRBs on bio-medical research, with the composition of IRBs reflecting this emphasis, has limited the informed review of social work protocols. Social and behavioral science research protocols, including those of social work, are often "expedited" and/or are reviewed by researchers who are unfamiliar with the nature of such scientific inquiries. (An analogy holds when social and behavioral scientists are asked to participate on IRBs in the review of bio-medical research.) Without the procedures in place and a cadre of trained researchers available and able to review social work research protocols, social work may well be vulnerable to some of the questionable research practices that have been unearthed in related fields.

The expanding boundaries of what constitutes scientific integrity are of particular relevance to social work researchers. The research conducted by social workers, both students and faculty and agency-based practitioners, involves interaction with populations that are often classified as vulnerable and confidentiality of data is often an issue. Direct observations, the administration of questionnaires, review of existing case records, or the introduction of therapeutic interventions and the use of control groups that do not receive interventions may be innocuous or, alternatively, may pose risks to the emotional, social, or economic well being of participants (4). Deception, invasion of privacy, lack of informed consent, mandatory reporting requirements (such as cases in which potential child abuse is identified), or the loss of economic benefits (as may apply, for example, to the disabled or welfare recipients) are all examples of harm that may result from faulty research designs or misconduct in the implementation of research protocols (4). Although substantiated cases to date fall outside of these human protection areas,

the nature of the research conducted within the helping professions suggests the potential of such misconduct.

A Call for Research on Social Work Research

Given the relatively undeveloped, but now rapidly expanding research enterprise in social work, there is a clear need for information about how research is monitored and reviewed. The number of publicized cases of wrongdoing in fields closely allied with social work suggest that programs of social work education need to formulate or revise their procedures for research review and oversight. Institutional mechanisms are needed to ensure that: (1) researchers are cognizant of the ethical issues involved; (2) the protocols meet university and Federal standards; and (3) findings are based on systematic and valid research. The question then becomes whose responsibility it is to monitor such protocols and review the research conducted and how mechanisms can be established which significantly reduce the potentiality of scientific misconduct.

Some schools have assembled their own committees to review and pass judgment about compliance with university and/or federal research requirements. However, such reviews usually focus on issues of methodology and/or informed consent. This is not sufficient given the broadened definition of scientific misconduct, which has been extended beyond the initial focus on informed consent, risk levels, and coercion (34). The definition of misconduct now includes fabrication, falsification, or plagiarism in proposing, performing, or reviewing research, or in reporting research results" (35, p. 4). The extent to which social work education programs maintain their own review and oversight procedures is also unknown. Anecdotal evidence suggests that such internal program mechanisms are the exception. Given the limited applicable of the professional Code of Ethics, the unknown degree of inclusion of social work within the purview of IRBs, and the similarly unknown degree of school-specific procedures, the need for "research on the status of social work research" is suggested. Possible areas of inquiry include:

- An analysis of the social work education curriculum to ascertain the degree to which ethical conduct is a component of research courses.
- An assessment of social workers' familiarity with ethical provisions regarding appropriate scientific conduct, through such means as: (1) an "exit" test of graduating BSW, MSW and doctoral students; (2) a sample survey of agency practitioners; and (3) a sample survey of agency administrators charged with responsibility to collect, analyze, and report on client-sensitive data.
- An analysis, perhaps through the use of focus groups, of issues and obstacles to the conduct of ethical research which result from the demands of external accountability bodies.
- An investigation of the procedures used by schools of social work to review and monitor faculty and student research, including the scope of such reviews and the extent to which the validity of the science itself is considered.
- A survey of social work faculty concerning their level of participation in university-wide institutional review boards.
- A survey of deans and directors of social work education programs to identify the frequency, nature, and types of issues and problems that have arisen in regard to studies, once approved and implemented.
- A content analysis of material covered in federally prescribed training of researchers and an assessment of the applicability of such training to the social and behavioral sciences.

The data emanating from such studies would provide a basis for an informed assessment of the extent to which mechanisms for research review and monitoring are in place and how well they operate. Such information could form the basis for developing or revising review procedures through university IRBs, through separate IRBs potentially established for the social and behavior sciences, or through social work education-specific structures. Further, such information could be used to develop targeted educational programs about research integrity to the social work community.

Conclusion

Research about social work research has tended to be descriptive, often focused on admonishments about the under-developed state of the art or analyses of what content areas have been researched and what gaps exist. Ethical research conduct has, by and large, been ignored, in part because of the early stage of development of the research enterprise. However, the issue of research integrity takes on increasing importance as social work gains a legitimate role in the conduct of scientific inquiry. The profession is likely to experience a stronger imperative to engage in research as demands for accountability and documentation of the outcomes of human services continue to grow.

Strategies to ensure research integrity depend, first, on a clearly formulated agenda based on an assessment of the current status of review and monitoring systems. Based on hard data, the professional schools of social work and their universities can assume the task of modifying and strengthening procedures in a manner that is reflective of the burgeoning social work research enterprise. Means of prevention as well as amelioration need to be developed, codified, and enforced. In this process, there is a need to define the parameters of both appropriate scientific conduct and what constitutes misconduct as it relates to social work research and to elaborate on its meaning with some degree of precision. Clear university and school standards, widely publicized, and ongoing education regarding appropriate scientific conduct would help alleviate actual or potential problems as social work secures a more extensive and important role in the production of research.

Bibliography

1. Austin, D. Research and social work: educational paradoxes and possibilities. J Soc Sci Res 1979; 2: 159-76.
2. Fanshel, D. National Conference on the future of social work research: Final report. Washington, DC: National Association of Social Workers, 1979. Contract # 278-77-0079[SM]. Sponsored by the National Institute for Mental Health.
3. Rubin, A., Gibelman, M. Social work research in mental health: the state of the art. Rockville, MD: Division of Human Resources/Division of Prevention and Special Mental Health Programs, National Institute of Mental Health; 1984. Report no.: 83M055981901D & 83M052517801D.
4. Gibelman, M., Gelman, S.R. Learning from the mistakes of others: a look at scientific misconduct within academia. J Soc Work Educ 2001; 37(2): 241-254..
5. Barker, R.L. Social work dictionary. 4th ed. Washington, DC: NASW Press; 1999.
6. Council on Social Work Education. Handbook of accreditation standards and procedures. Alexandria (VA): CSWE; 1994.
7. Gibelman, M., Whiting, L. Negotiating and contracting in a managed care environment: considerations for practitioners. Health Soc Work 1999; 24: 180-90.
8. Baker, D.R., Wilson, M.V.K. An evaluation of the scholarly productivity of doctoral graduates. J Soc Work Educ 1992; 28: 204-12.
9. Bloom, M., Klein, W.C. Publications and citations: a study of faculty at leading schools of social work. J Soc Work Edu 1995; 31: 377-87.
10. Gelman, S.R., Gibelman, M. A quest for citations?: An analysis and commentary on the trend toward multiple authorship. J Soc Work Educ 1999; 35: 203-13.
11. Green, R.G., Baskind, F.R., Best, A.M., Boyd, A. Getting beyond the productivity gap: assessing variation in social work scholarship. J Soc Work Educ 1997; 33: 541-53.
12. Ligon, J., Thyer, B.A., Dixon, D. Academic affiliations of those published in social work journals: a productivity analysis, 1989-1993. J Soc Work Educ 1995; 31: 369-76.
13. Smith, S., Baker, D., Campbell, M., Cunningham, M. An exploration of the factors shaping the scholarly productivity of social work academicians. J Soc Serv Res, 1985; 8: 81-99.
14. Marsh, J.C. Should scholarly productivity be the primary criterion for tenure decisions? Yes! J Soc Work Educ 1992a; 28: 132-4.
15. Marsh, J.C Response to Dean Weick. J Soc Work Educ, 1992b; 28: 138-9.
16. Austin, D. Findings of the NIMH task force on social work research. Res on Soc Work Pract 1992; 2: 311-322.
17. Council on Social Work Education. Research utilization in social work education. New York: CSWE; 1981.
18. Council on Social Work Education. A planning proposal to establish a National Institute for Social Work Research. New York: CSWE; 1983.
19. Estes, R.J. Social work research centers. Soc Work Res Abstr 1979; 15: 3-16.
20. National Association of Social Workers. Code of ethics. Washington, DC: NASW; 1996.
21. U.S. Department of Labor, Bureau of Labor Statistics. Tabulations from the Current Population Survey, 1998 annual averages. Washington, DC: BLS; 1998.
22. Gibelman, M., Schervish, P. Who we are: a second look. Washington, DC: NASW Press; 1997.
23. IRBs and human stem cell research protocols. Hum Res Rep 2000 Jan 15; 3-4.
24. Social workers' group plans no disciplinary action against researcher who submitted bogus articles. Chron High Educ 1989 Apr 15; Sect. A: 5.
25. Burd, S. Nursing professor found guilty of scientific misconduct. Chron High Educ 1995 July 7; Sect. A:23.
26. Burd, S. Research powerhouse under the microscope. Chron High Educ 1994 Jun 15; Sect. A: 24.
27. Campbell, P.W. Researcher found guilty of misconduct. Chron High Educ 1999 Jan 15; Sect. A:34.
28. Walker, P.V. Public health service finds 3 scientists guilty of misconduct. Chron High Educ 1996 Jan 12; Sect. A: 26
29. Wheeler, D.L. Suggestions offered for investigating reports of biomedical-research fraud. Chron High Educ 1991 Nov 27; Sect. A: 6.
30. Goodman, B. Scientists exonerated by ORI report lingering wounds. Scientist 1997, June 9; 11: 124.
31. Schmidt, P. U. of Arizona fires prominent researcher over charges of scientific misconduct. Chron High Educ 1998 Aug 14; Sect. A: 12.

32. Block, L.E. Science's credibility depends on effective peer review. Chron High Educ 1991 July 31; Sect. B:2.
33. Silbergeld, E.K. Annotation: Protection of the public interest, allegations of scientific misconduct, and the Needleman case. Am J Public Health 1995; 85: 165-6.
34. Broad, W.J. Data tying cancer to electric power found to be false. New York Times 1999 July 24; Sect. A:1, A10.
35. New definition proposed for research misconduct. Hum Res Rep 1999 Dec 14; 4.

Organizational Influences on Scientific Integrity

Michael D. Mumford, Department of Psychology, University of Oklahoma, USA

Whitney B. Helton, University of Oklahoma, University of Oklahoma, USA

Keywords: Ethical behavior, Integrity, Organizational influences, Scientific integrity, Situational influences

Our image of the working scientist remains inherently romantic (1). We envision an individual, working alone, pursuing knowledge in an area solely for its intrinsic interest. As attractive as the image may be, it has little to do with the realities of current work in the sciences (2, 3, 4). Scientists work in a distinctly social setting, conducting their work in both collaboration and competition with others (5, 6). This work, moreover, occurs in organizational settings, including business, government and academia. Thus, the pressures that face people working in any organization – pressures of time, conformity, resources, and production – also confront scientists.

Although one might argue that scientists, by virtue of their work, are granted more autonomy and are carefully buffered from the more "ugly" demands of organizational life, the conditions currently confronting most scientific endeavors are such that we can expect organizational pressures to become a progressively more important influence on scientific work. The emerging forces of the new economy, where innovation is the true competitive edge, move scientists from the periphery of the business world to the heart of the industrial enterprise (7). Academia, moreover, under the financial pressures imposed by funding cutbacks, has placed a new emphasis on responding to the needs of the business community (8). Finally, academia has begun a slow process, for good or ill, of learning how to manage itself differently, and manage itself like a business.

Given these pressures, there is a need to understand how organizational variables influence scientific integrity. Unfortunately, systematic studies of scientific integrity are virtually nonexistent. However, a number of scholars have sought to understand the variables that influence integrity in organizational settings as a general phenomenon. Accordingly, our intent in the present study is to examine prior studies of integrity with respect to their implications for understanding organizational influences on scientific integrity. We will begin by considering the findings obtained in one line of research concerned with the individual and situational factors that influence integrity in organizational settings. Subsequently, we will examine the kind of organizationally-based situational variables that might influence scientific integrity using a multi-level perspective that considers situational variables operating at the individual, group, and organizational levels of analysis (9).

Studies of Integrity

Psychological studies of integrity have typically employed one of two broad approaches (10). The first approach holds that integrity, or the lack thereof, is primarily a function of certain characteristics of the situation in which people find themselves. Thus, studies along these lines examine the

Corresponding author: Michael D. Mumford, Department of Psychology, University of Oklahoma, 455 W. Lindsey, #705 Dale Hall Tower, Norman, OK 73019, 405-325-5583 (voice), 405-325-4737 (fax), mmumford@ou.edu.

opportunities provided for dishonest behavior (11), the reinforcements and punishments associated with unethical acts (12), perceptions of procedural justice (13), and stress and authority norms (14). The second approach holds that a lack of integrity is primarily a function of certain characteristics of the individual. Scholars applying this second approach have sought to develop global measures of integrity (15, 16), and identify certain unique characteristics of people that are associated with a lack of integrity (17, 18).

Individual Variables

In one series of studies along these lines, Mumford and his colleagues (19-21) sought to develop a general model of the individual characteristics likely to promote destructive or unethical acts. To identify the characteristics of individuals related to the propensity for unethical acts, Mumford and his colleagues reviewed relevant studies in the clinical (22-24), management ethics (12, 18, 25), social-personality (26-28), and criminology (29-31) disciplines. This review resulted in the identification of seven individual characteristics that might plausibly be related to socially destructive unethical behavior: 1) narcissism, 2) fear, 3) outcome uncertainty, 4) power motives, 5) object beliefs, 6) negative life themes, and 7) lack of self-regulation.

These differential characteristics were held to operate as a dynamic syndrome in shaping unethical acts. It was held that narcissism, or extreme self-absorption and overevaluation of the self leads to a motivated defense of a weak self-system (22, 32). This perception of threat, in turn, induces outcome uncertainty and activates power motives as a defensive strategy. Fear, or anxiety, is also held to lead to perceptions of threat, thereby leading to outcome uncertainty (33). When people are uncertain about their capacity to attain desired outcomes, self-protective tendencies will activate power motives, although the activation of power motives may be somewhat inhibited by the tendency of fearful individuals to withdraw.

Once activated, power motives induce a tendency to harm or exploit others which, with the resulting desensitization, may lead to the emergence of object beliefs, or the view that others can be used as tools for personal gain (14, 22). In harming others, unless such effects are inhibited by self-regulation, people are likely to acquire negative images of others and their relationships with others. Thus, object beliefs, along with fear, may lead to the emergence of negative life themes. Negative life themes, along with object beliefs, power motives, self-regulation and outcome uncertainty reflect beliefs and motives held to exert direct effects on people's willingness to engage in destructive unethical acts. Figure 1 provides a summary of the key structural relationships specified in this model.

In an initial test of the plausibility of this model, O'Connor, Mumford, Clifton, Gessner, and Connelly obtained biographies for 82 notable historic leaders (21). They content-coded the "rise to power" chapters included in each biography for leaders' expression of behaviors indicative of the seven characteristics included in this model (e.g., object beliefs, narcissism, etc.), and obtained indices of the harm done to society by leaders' policies. In a subsequent causal modeling effort, not only was support obtained for the ability of these variables to predict harm done by leaders' policies, it was found that the a priori structural model presented in Figure 1 provided adequate fit to the observed data. The resulting model is shown in Figure 2.

In the second set of investigations, Mumford, Connelly, Helton, Mowry, and Osburn sought to determine whether the variables included in this model could account for scores on standard measures of integrity (34). Here 292 subjects were asked to complete two overt measures of integrity, the Reid Report (35) and the London House PSI or Personnel Selection Inventory (36).

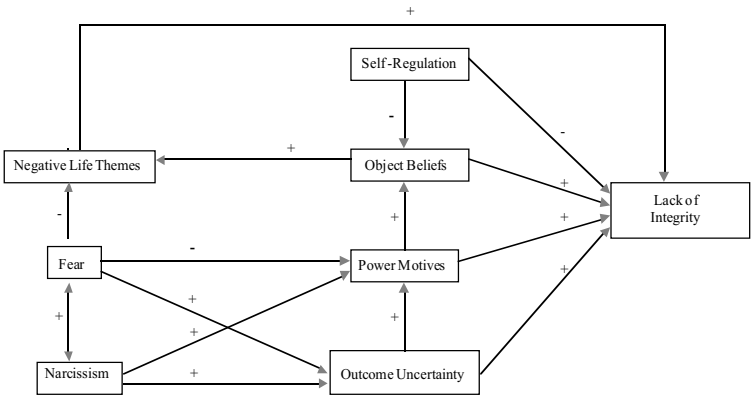
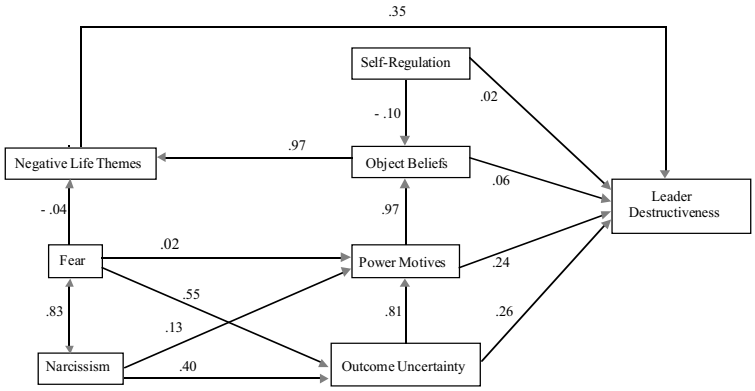


Figure 1. General structural model for individual influences on integrity.



Figures 2. Test of structural model for individual influences with respect to leader destructiveness.

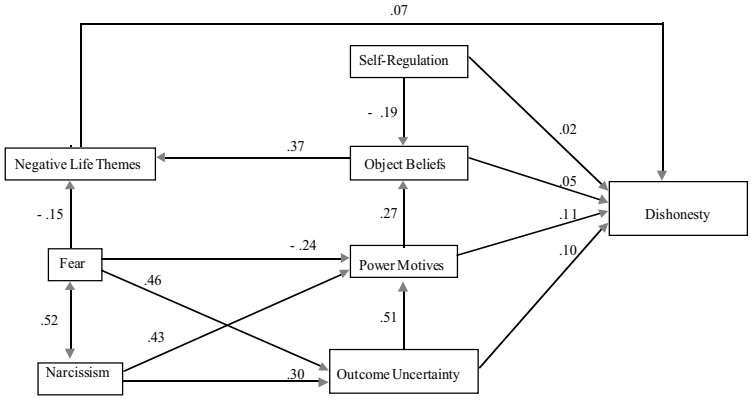
Both these measures examine theft, dishonesty, and punitive attitudes as direct markers of integrity. In addition, 400 subjects were asked to complete two commonly used personality based measures of integrity (37) – the Socialization and Delinquency scales of the California Psychological Inventory (CPI). Here background data scales were developed to measure each of the characteristics included in this model using the procedures suggested by Mumford, Costanza, Connelly, and Johnson (38). Again, it was found that the structure of the a priori model was confirmed. However, here it was found that although scores of these differential variables yielded effective prediction of integrity test scores ($\underline{r}$ = .32), the obtained prediction was not of overwhelming power. Figure 3 illustrates the nature of the results obtained in this study, while Table 1 describes the items used to measure these variables.

A potential explanation for the limited, albeit significant, impact of these variables on integrity test scores may be found in a study conducted by Mumford, Gessner, Connelly, O'Connor, and Clifton (20). In this study, 152 Masters of Business Administration (MBA) students were asked to work on an in-basket exercise which presented 32 decisions that might be made by regional sales managers. On half of the items included in this in-basket exercise, the MBA students were presented with ethical decisions where the actions selected might result in harm to others or harm to the organization.

Prior to starting work on this task, the MBA students were asked to complete the background data scales measuring the beliefs and motives relevant to integrity (e.g., object beliefs, power motives, etc.). Additionally, manipulations were made in the conditions of task performance, specifically authority norms, psychological distance, and feelings of self-efficacy. It was found that MBA students who expressed individual characteristics held to influence the occurrence of unethical acts would take unethical actions when feelings of self-efficacy were low. However, they would not necessarily make unethical decisions unless they had reason to believe that the actions taken would be supported by people in authority. Thus, it appears that situational variables might influence ethical decisions potentially interacting with individual predispositions in conditioning the occurrence of unethical behavior or, alternatively, by creating unique effects on unethical behavior.

Situational Variables

In fact, beginning with the work of Hartshorne and May (11), many scholars have argued that situational variables might exert strong effects on unethical behavior. In an initial investigation intended to identify the kind of situational variables that might influence the occurrence of unethical acts, Gessner, O'Connor, Mumford, Clifton, and Smith developed a set of life history items intended to capture exposure to situations likely to influence development, or expression of, the various individual characteristics held to



Figures 3. Test of structural model for individual influences with respect to integrity.

Individual Scales	Example Items
Object Beliefs	Surprised by how much people invest in friendships; did not do favors for people who could not return them; told white lies to get own way; viewed dealing with people as a game; has not gotten emotionally involved when dealing with people.
Power Motives	Frustrated when could not convince friends to adopt one's view; was important to be on the winning side; was willing to make a scene to get compliance from others; enjoyed making others do things; liked to have the last word.
Negative Life Themes	Enjoyed parties where people were really out of control; was not upset by media violence; spending time with family was not important; has not reflected upon one's purpose in life as much as others.
Outcome Uncertainty	Often planned for things that never happened; wished things would slow down or remain the same; worried about the future; annoyed by people who claimed something was a sure thing; wished there were more guarantees in life.
Fear	Friends thought they worried too much; often agonized over decisions; often woke up at night for no apparent reason; was bothered by things that could go wrong when things were going well; had difficulty making decisions about the future.
Narcissism	Tried to make self look good; was important to receive praise from others; spend a lot of time worrying about appearance; did not talk about things not of interest to them; did not spend time with others whose opinions were different.
Lack Of Self-Regulation	Not hard on one's self; rarely said the right thing at the right time; not important to identify own limitations; took long to fit in with an unfamiliar crowd; did not express opinions according to the situation at hand.

Table 1: Examples of Items Included in the Individual Scales

influence unethical behavior (e.g., object beliefs, outcome uncertainty, etc.) (39). A subsequent factoring of these items after they had been administered to 285 undergraduates, lead to the identification of seven situational factors: 1) alienation, 2) non-supportive family, 3) negative role models, 4) life stressors, 5) competitive pressure, 6) exposure to negative peer groups, and 7) financial need. Table 2 illustrates the nature of the items used to measure these variables.

To examine the impact of these variables on integrity, Mumford, Connelly, Helton, Mowry, and Osburn, administered the life history items measuring exposure to these situational factors to the 292 subjects asked to complete the two overt integrity tests, the Reid Report and the PSI, and the 400 subjects asked to complete the two personality-based tests, the CPI socialization and delinquency scales (34). In this study, scores on the overt and personality based measures of integrity were both correlated with, and regressed on, the seven situational scales.

The first major finding to emerge from these analyses was that the situational scales were correlated with scores on the measures of individual characteristics held to influence unethical behavior (e.g., negative life themes, object beliefs, etc.) yielding bivariate correlations in the .40s. The second major finding indicated, however, that the situational variables were strongly related to integrity test scores producing relationships in the mid-.20s to low-.50s. Of these variables, exposure to negative peer groups, alienation, and financial need appeared to produce the strongest relationships across the four measures of integrity. The third major finding to emerge in these analyses indicated that the situational variables yielded better prediction of scores on the four integrity tests than the individual variables while yielding significant gains in prediction when added to the individual variables. The results obtained in this third analysis are summarized in Figure 4 which indicates that the situational variables accounted for far more variance in integrity test scores than the individual variables.

Although these findings underscore the fundamental importance of understanding situational influences in attempts to understand and control unethical acts. These findings leave two crucial questions unanswered. First, they do

Situational Scales	Example Items
Alienation	Had worked in a setting where they saw discrimination; had superiors who were condescending; worked with people and withheld information; had belonged to organizations in legal trouble; lost something because others took advantage of status or position; often worked in situations where they could not keep up with demand.
Non-Supportive Family	Parents were not consistent in praise of punishment; parents did not explain why they punished; parents and teachers did not praise work; did not have input into important family decisions; parents and siblings did not help with schoolwork.
Negative Role Models	Parents broke promises; parents openly criticized others; often witnessed violent arguments among adults in household; parents gave harsh punishments; parents lost temper for no apparent reason; family had different standards than other families.
Life Stressors	Unable to go to school due to health; had to cope with large unexpected expenses; teachers made unrealistic work demands; had serious illness; schoolwork effected by problems of family members; was in situations where they could not keep up with work.
Competitive Pressure	Often experienced competition among coworkers; concerned about finding a good job after graduation; frequently sought recognition for work; had to be competitive to get ahead at work or school; selected people for membership in clubs; was involved in team projects.
Negative Peer Group	Friends had a cynical attitude towards society; high school and college friends had trouble with law; friends and family were heavy users of drug and alcohol; observed people breaking rules while growing up; saw people taken advantage of; witnessed verbal/physical violence.
Financial Need	Many families in neighborhood they grew up in received some type of public assistance; lost mother or father; regular schedule was not emphasized in family; members of family had been in trouble with law; people could take things away from them because of family position.

Table 2: Examples of Items Included in the Situational Scales

not tell us exactly how unethical acts are influenced by situational variables. For example, situational variables might constrain unethical behavior, interact with individual variables or, alternatively, compel unethical behavior in their own right. Second, these findings do not tell us about the specific kinds of situational variables that act to influence unethical behavior in the kind of organizational settings in which scientists are likely to work. Accordingly, in the following sections, we will examine the specific kinds of situational variables operating at the individual, group, and organizational levels that might influence scientific integrity.

Individual Level

Of the situational variables found to be related to integrity, stress seems to be the variable most likely to be linked to integrity in research work. Scientific work is known to be demanding and stressful resulting from multiple commitments, deadlines, the need to acquire resources, and uncertainty about project outcomes (40). When these occupational demands are combined with the intense focus characteristic of those engaged in scientific work (41), it seems plausible to argue that stress represents an endemic feature of life in the sciences. Although, up to a point, stress may contribute to productivity, high levels of stress may not only prove debilitating, but, more centrally, may contribute to incidents of unethical conduct through two distinct mechanisms (42). First, high levels of stress may lead people to take more risky actions than they might under other conditions due to the negative effects of stress on self-regulation (27). Second, stress reduces the cognitive resources available for reasoning and analytical problem solving (43). This loss in cognitive capacity is noteworthy because effective moral reasoning inhibits the occurrence of unethical acts (18, 44, 45). These observations, in turn, lead to our first

	Personality Based Tests			*Overt Tests*		
	CPI	**CPI**	**PSI**	**Reid**	**PSI**	**Reid**
	Socialization	*Delinquency*	*Honesty*	*Honesty*	*Theft*	*Theft*
INDIVIDUAL SCALES						
Multiple Correlations	.42	.38	.36	.27	.25	.30
Cross Validated Multiple Correlation	.36	.31	.29	.20	.07	.17
SITUATIONAL SCALES						
Multiple Correlation	.62	.58	.57	.43	.35	.28
Cross-Validated Multiple Correlation	.49	.51	.40	.38	.26	.12
SITUATIONAL SCALES ADDED TO INDIVIDUAL SCALES						
Multiple Correlation	.67	.61	.61	.47	.41	.40
Cross-Validated Multiple Correlation	.62	.50	.58	.38	.27	.17
Change in R Square	.26**	.23**	.24**	.17**	.11**	.07*

Figure 4: Comparison of Individual and Situational Variables with Respect to the Prediction of Integrity Test Scores.
**P < .05 ** P < .01*

two propositions.

- Proposition One: Incidents of unethical behavior will be more frequent when individuals experience stress and overload.
- Proposition Two: Attempts by organizations to reduce stress by minimizing time pressure, managing overload, clarifying goals, and providing requisite resources will reduce incidents of unethical behavior.

Actions taken to reduce work demands, of course, are not the only steps that might be taken to reduce stress and unethical behavior in organizational settings. Both stress and uncertainty about outcomes are influenced by people's feelings of competence and their ability to exert positive, effective control over their work environment. In keeping with this observation, Weeks, Moore, McKenney, and Longnecker administered vignettes calling for ethical decisions to managers with greater and lesser experience (46). They found that experienced managers were more likely than their less experienced counterparts to make ethical decisions. Other studies by Arlow and Uhlrich (47), Chonko and Hunt (48), Kidwell, Stevens, and Bethke (49), and Teal and Carroll (50) also indicate that more experienced successful workers, workers with greater expertise, are less likely to engage in unethical activities or make unethical decisions. As noted above, one potential explanation for these findings is the ability of experienced, competent workers to handle stress and uncertainty. Experienced, competent workers, however, may also feel less need to take shortcuts. Regardless of the explanation used to account for these effects, however, it is clear that organizations may take a number of steps to build competence and expertise through educational and mentoring programs, careful selection of employees, and providing people with time to pursue continuing education projects (2).

Competence and expertise, of course, also allow people to induce effective control over their work environment. Given the impact of stress, outcome uncertainty, and fear on unethical acts, one would expect that control beliefs would be related to unethical behavior in organizational settings. In fact, studies by Hegarty and Sims (12), Trevino and Youngblood (18), and Reiss and Mitra (51) all indicate that people who have a strong internal locus of control are less likely to engage in unethical acts than people who believe their actions are controlled by external forces. What is important to recognize here, however, is that organizations can build feelings of control by assigning people to tasks commensurate with their capabilities, allowing input to critical decisions, and buffering people from uncontrollable events. Taken as a whole, these observations imply the following three propositions.

- Proposition Three: Less skilled or less experienced scientists will be more likely to engage in unethical acts and will be more sensitive to organizational pressures that promote unethical acts.

- Proposition Four: Organizational actions intended to develop expertise and maximize feelings of competence will inhibit unethical acts.
- Proposition Five: Organizational actions intended to maximize people's control of their environment will inhibit unethical acts.

As important as competence and control may be to the management of stress and the minimization of unethical behavior, some consideration should be given to family and social relationships. Family and social relationships, specifically supportive relationships, help people cope with stress while the implied commitment to others embedded in these relationships promotes a prosocial outlook. Accordingly, Mumford, Connelly, Helton, Mowry, and Osburn (34) found that exposure to a non-supportive family environment was related to a lack of integrity. Unfortunately scientists, in part due to their introversion (52) and, in part due to their work commitments (53), appear to have some difficulty in establishing viable family and social relationships. By the same token, however, scientists do appear to establish viable, long-term collaborative relationships and create social connections through their network of enterprise (5, 54). These observations, in turn, suggest that incidents of unethical behavior will occur less frequently among scientists who have a rich extensive network of supportive professional colleagues. Moreover, by co-locating scholars with similar interests, encouraging collaborative work, recognizing the value of multiple-authored publications, and providing time for collegial interactions, organizations can reduce incidents of scientific misconduct. Thus:

- Proposition Six: Individuals lacking collaborative networks will be more likely to be involved in incidents of scientific misconduct.
- Proposition Seven: Organizational actions intended to facilitate and recognize the value of collaborative activities will minimize incidents of scientific misconduct.

Our foregoing observations with regard to collaboration point to another factor likely to be involved in incidents of scientific misconduct – alienation. Alienation among scientists is not a strictly social phenomenon. Alienation from the work, and the work's potential contributions to society, appear particularly significant with regard to scientific misconduct because scientific work is often motivated by intrinsic interest in the work for its own sake and an abiding belief in the potential contribution of its' worth to society as a whole (55, 56). As Bowie points out, this intrinsic motivation buffers individuals from situational pressures likely to promote unethical acts (57) . He notes, furthermore, that a variety of organizational policies might influence alienation and intrinsic motivation including explicit recognition of social contributions as well as contributions to the "bottom line", allowing individuals to pursue personally interesting work, and maximizing autonomy in decision-making. These observations suggest the following proposition.

- Proposition Eight: Attempts by the organization to recognize and reward social contributions and allow individuals to pursue their unique interests will reduce incidents of scientific misconduct.

Eisenberger and Cammeron, however, remind us that creative work, including scientific work, is not simply a matter of intrinsic motivation (58). People's work as scientists is also motivated by extrinsic factors such as pay, recognition, and status. At first glance, it might seem plausible to argue that extrinsic rewards lead to unethical behavior. However, the relationship between the pursuit of extrinsic rewards and unethical behavior appears somewhat more complex with the pursuit of extrinsic rewards contributing to unethical acts only when people expect that the unethical behavior will be rewarded, the unethical act will not be detected, and the act, if detected, will not be sanctioned by the organization (12, 18, 59). One implication of this expectancy model is that high performers will sometimes engage in unethical acts because they believe they are less likely to be sanctioned by the organization (60, 61)–potentially resulting in a culture that seems to condone such acts. Another implication of this expectancy model is that ethical behavior will decrease when extrinsic rewards such as pay and promotions are based on immediate short-term production demands rather than long-term contributions to others (62).

In considering the impact of production demands, however, it is necessary to bear in mind a unique characteristic of scientific work. Scientists' rewards are often explicitly tied to production such as journal publications, patents, and fielding new software (63, 64). By expressly tying extrinsic rewards to production counts, however, one can expect that misconduct will increase whenever ambitious, extrinsically motivated individuals, individuals motivated by

financial needs, status concerns, and recognition, encounter significant reverses in the production process. Thus, organizations might minimize misconduct by rewarding progress towards goals as well as production output, recognizing alternative indices of performance such as impact and innovation, and providing a minimal degree of security and visibility for all group members based on their unique strengths.(65) Taken as a whole, our preceding observations about extrinsic motivation suggest the following four propositions.

- Proposition Nine: Organizational reward systems that stress long-term innovation and impact will tend to minimize incidents of unethical behavior.
- Proposition Ten: Organizational reward systems that recognize progress as well as output will tend to minimize incidents of unethical behavior.
- Proposition Eleven: Scientific misconduct will occur more frequently when extrinsic rewards are based on production and people are treated harshly for production setbacks.
- Proposition Twelve: Scientific misconduct will occur less frequently in organizations where all incidents of misconduct are treated similarly, regardless of the past performance of the people involved.

Groups

The Mumford, Connelly, Helton, Mowry, and Osburn study not only points to the influence of individual level situational influences on integrity, such as stress, relational support, alienation, and financial need, it also underscores the importance of certain group level influences (34). In this study, three variables operating at the group level, role models, exposure to negative peer groups, and competitive pressure, were found to influence integrity. Again, all three of these situation variables appear to represent important influences on integrity in organizational settings.

In organizations, role modeling is commonly subsumed under this broader area of leadership (66), and there is, in fact, reason to believe that the behavior of people assigned to formal organizational leadership roles will influence the manifest integrity of their "followers". In one study along these lines, Schminke and Wells had 81 business students participate in a four-month long strategic planning simulation (67). During the course of this simulation, measures of ethical decision-making were obtained along with measures of group process variables and leadership styles, specifically consideration and initiating structure. They found that the leaders' emphasis on initiating structure contributed to ethical decision-making, presumably because the initiation of structure led group members to focus on task accomplishment rather than personal concerns. In another study along these lines, Zabid and Alasgoff found that the behavior of people's immediate superior exerted stronger effects on the occurrence of unethical acts than other putative organizational influences such as climate and codes of conduct (68).

Leaders appear to influence ethical behavior through a variety of different mechanisms, some of which may inhibit unethical acts and some of which may promote such acts. Sims, in a study of leadership in financial services firms, identified four ways leadership behavior contributes to or promotes integrity (69). He argues that leaders promote ethical behavior by a) focusing the attention of people on ethical issues, b) responding to crises based on ethical, productive concerns rather than self-protection, c) allocating rewards based on long-term contributions rather aggressive self-promotion, and d) applying sanctions for incidents of unethical behavior. Along similar lines, Minkes, Small, and Chatterjee have argued that leaders' articulation and communication of personal, ethical, and moral values will promote integrity on the part of group members (70). Contrawise, it appears that leaders who articulate poor values or exhibit self-serving, narcissistic behavior implicitly encourage unethical behavior on the part of subordinates (71, 72). Vredenburgh and Brender point out, moreover, that leaders who consistently abuse power through arbitrary actions, a focus on personal control, and inequitable decisions, induce stress, fear, and outcome uncertainty while activating the power motive linked to unethical acts (73).

Although it seems clear that leaders have an impact on ethical behavior in general, the question remains as to whether leaders have a similar impact on the ethical behavior of scientists. One might argue that, due to their greater autonomy and specialized professional expertise, scientists are less susceptible to leader influence (66, 74). Although this argument seems plausible, the available evidence indicates that leaders exert notable effects on people's behavior in research settings (75). A case in point may be found in Hounshell's analysis of research on synthetic fabrics in Dupont's Pioneer

research laboratories where the vision defined by founders in the 1920s continued to shape the laboratories' research programs well into the 1990s (76). Nonetheless, the autonomy and expertise of scientists suggest that leader influences on ethical issues will be less evident in day-to-day direction and more evident in the leaders'

a) definition of a coherent constructive research vision, b) focus on production as opposed to status relationships, and c) articulation of ethical values in interactions with staff. When these observations are considered with respect to the findings sketched out above, they suggest the following three propositions:

- Proposition Thirteen: Scientific misconduct will be less common in groups where leaders have the expertise needed to define a coherent vision for the work.
- Proposition Fourteen: Scientific misconduct will be less common in groups where the leader actively articulates ethical values, potential social contributions of the work, and enhancement of the work rather than career status.
- Proposition Fifteen: Scientific misconduct will be less common in groups where the leader focuses on effective direction of production activities rather than personal professional recognition, maintenance of control, or social acceptance.

Leadership, of course, is not the only group level variable that might influence integrity in organizational settings. For example, Mumford, Connelly, Helton, Mowry, and Osburn found that competitive pressure was related to a lack of integrity (34). The effects of competition on ethical behavior, however, appear to be quite complex in organizational settings. One way competition appears to influence ethical behavior may be found in the tendency of people to discount the relevance of moral considerations to decision-making in competitive situations (77). Another way competition influences ethical behavior is that negative perceptions of competitors' intentions provide a justification of unethical acts (78). Still another way competition influences ethical behavior is by inducing feelings of stress and uncertainty (39).

These varied mechanisms by which competition influences ethical behavior are all clearly applicable to scientists. In the case of scientists, however, it is quite possible that these negative aspects of competition represent particularly important influences on unethical acts. Scientists have been found to be highly competitive evidencing not just competitive intensity but some degree of hostility and arrogance (79)–all dispositional factors likely to make scientists particularly susceptible to the negative effects of competitive pressure. Competitive pressure, however, may not always be destructive provided it is managed effectively by the organization (80). More specifically, when competition is accompanied by respect for competitors, people feel that they have sufficient technical competence to compete effectively, and competition is viewed as a depersonalized, professional challenge, then competition may contribute to performance and ethical behavior (81, 82). These observations, in turn, suggest the following three propositions.

- Proposition Sixteen: Unethical acts are more likely to be observed when ambitious, highly competitive people are placed in competitive settings where they lack requisite skills.
- Proposition Seventeen: Organizations that take actions to reduce personalized competitive pressure by evaluating performance on an absolute rather than relative basis or by encouraging collaborative work among potential competitors are less likely to experience incidents of unethical behavior.
- Proposition Eighteen: Unethical behavior is less likely to occur when leaders, or organizational practices, encourage people to analyze and identify the merits in competitors' work.

Personalized competition within-groups, of course, may result in conflict and a lack of cohesiveness. In this regard, the Schminke and Wells study cited earlier is noteworthy. In addition to examining leadership styles and their influence on ethical decision-making, they also examined the effects of group cohesiveness (67). Here it was found that cohesiveness influenced ethical decision-making both directly with more cohesive groups making more ethical decisions and indirectly with cohesive groups evidencing higher performance which, in turn, led to more ethical decision-making. These findings suggest that actions taken to induce cohesiveness through development and articulation of a shared, common vision, use of group as well as individual rewards, integration of members work activities, and encouragement of within-group collaborative efforts will all contribute to ethical behavior. Thus, the following three propositions seem indicated.

- Proposition Nineteen: Unethical acts are more likely to occur in non-cohesive conflict-laden groups.
- Proposition Twenty: Cohesiveness within a group will reduce scientific misconduct both by enhancing performance and minimizing the negative effects of within-group competition.
- Proposition Twenty-One: Organizational actions that lead to higher cohesiveness, such as development of a shared vision on the allocation of group, as well as individual, rewards, will reduce incidents of scientific misconduct.

Although it appears that cohesiveness may contribute to integrity, a cautionary note seems in order. Many prior studies of groups, including destructive behavior on the part of groups, indicate that conformity pressures can induce destructive, unethical behavior when the primary concern is maintenance of harmonious group relations and the goals being pursued by the group are likely to result in destructive, unethical behavior (24, 83). Hence:

- Proposition Twenty-Two: When high levels of cohesiveness prohibit questioning of group actions, cohesiveness may be related to unethical acts.

As implied by our foregoing proposition, exposure to the behaviors of, and expectations imposed by, other group members may influence ethical behavior in organizational settings (34). Exposure to peer groups is commonly held to influence integrity through the models for appropriate behavior provided by other group members and the normative expectations imposed on people by other members of the group (39, 84) . Accordingly, Murphy has argued that anomie, or normlessness, will engender unethical behavior because group members lack models for appropriate behavior and sanctions are not imposed for unethical acts (10). In keeping with argument, Leede, Nijhof, & Fisscher, note that when groups are experiencing conditions of rapid change the resulting breakdown in extant normative structures may lead to an increase in the frequency of unethical acts (85). Thus,

- Proposition Twenty-Three: When groups are experiencing rapid changes in personnel, technology, or productions processes, incidents of unethical behavior will increase.

The notion that normlessness will contribute to the occurrence of unethical acts also implies that the presence of normative expectations for ethical behavior among group members will contribute to integrity. As might be expected, the bulk of the available evidence does indicate that ethical norms within a group lead to ethical behavior. For example, studies by Barnett (86), Kawathatzopoulos (87), Verbke, Ouwerkerk, and Peelen (88), and Weaver and Farrell (89) indicate that when groups communicate expectations for ethical behavior, and sanction violations by group members, ethical decision-making improves and unethical acts become less frequent. In this regard, however, it is important to bear in mind a point made by Fritz, Arnett, and Conkel (90), Grimalda (91), and Schokkaert and Sweeney (92). More specifically, the effects of group norms on ethical behavior will vary with people's commitment to the group. Accordingly, the following three propositions seem indicated.

- Proposition Twenty-Four: Ethical behavior will be more common in groups that have, and actively apply, positive normative standards in group decision-making and the application of sanctions.
- Proposition Twenty-Five: The effects of ethical norms on integrity depend on building feelings of commitment to the group, the organization, or the profession.
- Proposition Twenty-Six: the creation and articulation of normative ethical standards by leaders on professional organizations will prove less effective when groups are experiencing rapid change and commitment is low.

Organizations

The Mumford, Connelly, Helton, Mowry, and Osburn study focused primarily on situational factors operating at the individual or group level (34). As a result, this study does not directly address the various organizational level variables that might be related to integrity. Nonetheless, the nature of the individual and group based situational influences on integrity do suggest that certain organizational level variables will also influence integrity. One set of organizational level influences suggested by our foregoing observations is the organization's operating environment – specifically three features of the organization's operating environment turbulence, munificence, and interdependence.

Environmental turbulence refers to rapid changes in technology, business processes, product markets, and competitors (93). Of course, turbulence will lead to normlessness as well as uncertainty about the requirements for

effective performance, both conditions that can be expected to promote unethical acts. Accordingly, Morris, Marks, Allen, and Perry found that ethical values were less evident among people working for organizations operating in a turbulent environment (94). Along similar lines, Rossouw has argued that the turbulence induced by social disruption can lead to unethical acts on the part of organizations (95). Among scientists, however, it seems likely that turbulence will exert larger effects when its impact is evident in their immediate technical environment or in employment practices. These observations, in turn, lead to the following two propositions.

- Proposition Twenty-Seven: As turbulence increases in the organization's operating environment the frequency of unethical acts will increase.
- Proposition Twenty-Eight: Scientific misconduct will increase in periods of rapid change in technological paradigms and employment practices.

In contrast to turbulence, munificence refers to the availability of resources and the low degree of competitive pressure evident in the organizations' operating environment. In fact, the available evidence indicates that munificence is related to ethical conduct in organizational settings. For example, Verschoor (96), in a study of Fortune 500 companies, found that ethical conduct with regard to organizational shareholders increased with financial performance while Judge (97), in a study of hospitals, found that scarcity of financial resources was negatively related to social contributions. In still another study along these lines, Zarkada-Fraser found that collusion in government project bids was related to project desirability and competition (98). Among scientists, where resources are critical to conducting requisite research work, non-munificent environments may encourage unethical acts as a way of insuring resource availability. Thus,

- Proposition Twenty-Nine: As the munificence of the organizations operating environment decreases, unethical behavior and incidents of scientific misconduct will increase.

A third, and final, environmental variable commonly linked to ethical behavior in organizational settings is interdependence, or the extent to which organizational success depends on maintaining viable relationships with other organizations including suppliers, alliance partners, or government agencies. As might be expected, high interdependence appears to promote ethical behavior (99, 100, 101). Although it is unclear exactly what mechanisms shape the influence of interdependence on ethical behavior the following proposition does seem indicated:

- Proposition Thirty: Unethical behavior occurs less frequently in organizations where performance depends on the support, or goodwill, of other entities.

The organization's operating environment is, of course, one influence on the structure of the organization. Structure, or the manifest division of labor in an organization, has not commonly been studied as an influence on integrity. However, the available evidence indicates that unethical acts are less likely to occur in small organizations (102, 103) and in organizations where roles and responsibilities are clearly defined (85, 104). One explanation for this pattern of findings may be found in diffusion of responsibility and its derivative effects or alienation. In keeping with this alienation and diffusion of responsibility notion, Dooley and Fryxell found that diversification was related to corporate pollution levels (105). These observations imply the following proposition:

- Proposition Thirty-One: As organizational structures become more complex, and roles and role accountability are less clearly defined for individuals, unethical acts will become more frequent.

While structure refers to the organization of the work, climate refers to people's perceptions of social interactional expectations with their work environment (106). Relative to structure, climate has received substantially more attention as a potential influence on ethical behavior in organizational settings. In one study along these lines, Sims and Keon administered five business scenarios calling for an ethical decision to 245 business students who were also asked to complete a survey describing the company for which they were currently working (107). It was found that perceptions of their work environment were related to ethical decision-making. Similar findings have been obtained by Baumhart (59).

Although there is reason to believe that organizational climate influences ethical behavior, more debate surrounds the nature of the specific climate dimensions involved. Agarwal and Malloy identify five climate dimensions related to ethical behavior:

1) individual caring 2) social caring, 3) independence, 4) Machiavellianism, and 5) law and code (108). Vidaver-Cohen proposes a different model of ethical climate which stresses the importance of 1) social responsibility, 2) social support, 3) avoiding harm of others, 4) task support, and 5) equity of reward procedures (109). Still another model, one proposed by Key, views climate as a function of: 1) day-to-day reinforcement of ethical conduct, 2) punishment of unethical conduct, and 3) management role modeling (110). Finally, Argadona and Hartman, Yrle, and Galle argue that trust and perceptions of distributive and procedural justice represent key organizational climate dimensions influencing ethical behavior on organizations (111,112).

While a variety of models of ethical climate are available, it seems likely that some of these dimensions will prove more important than others in shaping the ethical behavior of scientists. Given the hostility and competitiveness characteristic of scientists (79), it seem plausible to argue that climates stressing trust and social support while maintaining perceptions of procedural and distributive justice will prove particularly important in minimizing misconduct (7). The demands of creative work, moreover, suggest that climates reinforcing autonomy, openness, and minimization of premature criticism will also prove useful in enhancing ethical behavior (75, 113). Thus, the following two propositions seem indicated.

- Proposition Thirty-Two: Organizational climates that promote perceptions of trust and fairness will minimize incidents of scientific misconduct.
- Proposition Thirty-Three: Organizational climates that are open and not overly critical of new ideas will minimize incidents of scientific misconduct.

The climate literature, however, also underscores the importance of day-to-day reinforcement on ethical conduct. In the case of scientists, the importance of ethical standards implies that professional codes, as well as their acceptance and embodiment by the organization, will also influence incidents of scientific misconduct. In fact, studies by Weaver and Farrell (89) of American Marketing Association members, and Gotterbarn (114) of software engineers, indicate that professional codes are viewed as important influences on ethical behavior in the sciences and may lead to improvements in ethical decision-making.

On the other hand, however, there is no assurance that professional ethical codes will be adopted by organizations in their day-to-day practices. This point is nicely illustrated in a study by Etheredge who examined attitudes toward ethical behavior in business managers and identified two dimensions: a) the importance of ethics and social responsibility, and

b) subordination of ethics and social responsibility to organizational effectiveness (115). Thus, organizations in their quest for efficiency and control, may reject professional ethical standards that conflict with organizational needs. When organizations reject these professional standards, however, it can be expected that the resulting organizational-professional conflict will induce some stress as people are forced to choose between these competing expectations. Although a number of considerations will influence how this conflict is resolved, it appears that investment in the organization, as opposed to the profession, is of critical importance (116). Accordingly, the following three propositions seem indicated.

- Proposition Thirty-Four: Incidents of scientific misconduct will be less common among individuals who are more invested in the profession rather than the organization they are working.
- Proposition Thirty-Five: Incidents of scientific misconduct will be less common in organizations that rely on their professional technical reputation for market advantage and view organizational needs as consistent with professional ethical codes.
- Proposition Thirty-Six: Professional ethical codes will prove most effective in reducing scientific misconduct when codes are actively supported by the organization.

Conclusions and Directions

Figure 5 summarizes the various propositions we have proposed with respect to the situational variables influencing ethical behavior at the individual, group, and organizational levels. In reviewing these propositions, however, an important caveat seems in order. More specifically, although all of the propositions were formulated based on a review of the organizational literature as it relates to the situational variables influencing integrity. Few, if any, studies have directly examined the influence of organizational, situational variables on research integrity. Thus, these propositions should not be viewed as well established

Individual Level

1) Incidents of unethical behavior will be more frequent when individuals experience stress and overload
2) Attempts by organizations to reduce stress by minimizing time pressure, managing overload, clarifying goals, and providing requisite resources will reduce incidents of unethical behavior
3) Less skilled or less experienced scientists will be more likely to engage in unethical acts and will be more sensitive to organizational pressures that promote unethical acts
4) Organizational actions intended to develop expertise and maximize feelings of competence will inhibit unethical acts
5) Organizational actions intended to maximize people's control of their environment will inhibit unethical acts
6) Individuals lacking collaborative networks will be more likely to be involved in incidents of scientific misconduct
7) Organizational actions intended to facilitate and recognize the value of collaborative activities will minimize incidents of scientific misconduct
8) Attempts by organizations to recognize and reward social contributions and allow individuals to pursue their unique interests will reduce incidents of scientific misconduct
9) Organizational reward systems that stress long-term innovation and impact will tend to minimize incidents of unethical behavior
10) Organizational rewards that recognize progress as well as output will tend to minimize incidents of unethical behavior
11) Scientific misconduct will occur more frequently when extrinsic rewards are based on production and people are treated harshly for setbacks
12) Scientific misconduct will occur less frequently in organizations where all incidents of misconduct are treated similarly regardless of past performance

Group Level

13) Scientific misconduct will be less common in groups where leaders have the expertise needed to define a coherent vision for the work
14) Scientific misconduct will be less common in groups where the leader actively articulates ethical values, potential social contributions of the work and enhancement of the work rather than career status
15) Scientific misconduct will be less common in groups where the leader focuses on effective direction of production activities rather than personal professional recognition, maintenance of control, or social acceptance
16) Unethical acts are more likely to be observed when ambitious, highly competitive people are placed in competitive settings where they lack requisite skills
17) Organizations that take actions to reduce personalized competitive pressure by evaluating performance on an absolute rather than relative basis or by encouraging collaborative work among potential competitors are less likely to experience incidents of unethical behavior
18) Unethical behavior is less likely to occur when leaders, or organizational practices, encourage people to analyze and identify the merits in competitors' work
19) Unethical acts are more likely to occur in non-cohesive, conflict-laden groups
20) Cohesiveness within a group will reduce scientific misconduct both by enhancing performance and minimizing the negative effects of within group competition
21) Organizational actions that lead to higher cohesiveness such as development of a shared vision or the allocation of group as well as individual rewards will reduce incidents of scientific misconduct
22) When high levels of cohesiveness prohibit questioning of group actions, cohesiveness may be related to unethical acts
23) When groups are experiencing rapid changes in personnel, technology, or production progress, incidents of unethical behavior will increase
24) Ethical behavior will be more common in groups that have, and actively apply, positive normative standards in group decision-making and the application of standards
25) The effects of ethical norms on integrity may depend on building feelings of commitment to the group, organization or profession
26) The creation and articulation of normative ethical standards by leaders in professional organizations will prove less effective when groups are experiencing rapid change and commitment is low

Organizational Level

27) As turbulence increases in the organization's operating environment, the frequency of unethical acts will increase
28) Scientific misconduct will increase in periods of rapid change in technological paradigms and employment practices
29) As the munificence of the organization's operating environment decreases, unethical behavior and incidents of scientific misconduct will increase
30) Unethical behavior will occur less frequently in organizations where performance depends on the support, or goodwill, of other entities
31) As organizational structures become more complex, and roles and role accountability are less clearly defined for individuals' unethical acts will become more frequent
32) Organizational climates that promote perceptions of trust and fairness will minimize incidents of scientific misconduct
33) Organizational climates that are open and not overly critical of new ideas will minimize incidents of scientific misconduct
34) Incidents of scientific misconduct will be less common among individuals who are more invested in the profession rather than the organization for which they are working
35) Incidents of scientific misconduct will be less common in organizations that rely on their professional or technical reputation for market advantage and view organizational needs as consistent with professional ethical codes
36) Professional ethical codes will prove most effective in reducing scientific misconduct when codes are actively supported by the organization

Figure 5. Summary of Propositions at Individual, Group, and Organizational Levels

conclusions but, instead, as a set of hypotheses that might be used to guide further research.

The need for further research along these lines becomes even more salient when one takes two other considerations into account. First, although the propositions presented in the present effort all seem plausible, evidence is not available examining the relative importance of these various situational variables on scientific misconduct and research integrity. For example, given the known dispositional characteristics of scientists (79), it seems attractive to argue that

competition, conflict, and a lack of cohesiveness will have a greater impact on misconduct than the direction provided by a leader. Unfortunately, however, evidence allowing us to evaluate the relative importance of various situational influences within and across three levels of analysis is, at this juncture, simply not available.

Second, in formulating these propositions we have examined organizations as a general phenomenon drawing heavily from past research in the "for profit" business arena (18, 107). What must be recognized here, however, is that scientists' work occurs in a variety of settings aside from the business arena including universities, government agencies, and non-profit research institutes. As a result, the unique characteristics of these non-business settings may influence the relative importance of the various situational variables identified in the present effort. A case in point can be found in our observations about organizational conflicts with professional codes of ethics since such conflicts maybe less pronounced outside the business setting. Thus, there is a need to assess the generality of these propositions across work settings.

Even bearing these caveats in mind, however, we believe that the present study does lead to some noteworthy conclusions about research integrity. To begin, we tend to attribute incidents of misconduct to characteristics of the individual. Although the importance of the scientist's character is not to be underestimated, the results obtained in the present effort suggest that situational variables have a large, perhaps a larger, impact on integrity than individual variables. Although this argument is by no means unique (11), it does suggest that future studies of research integrity should give as much attention to situational and individual influences.

The present effort, moreover, has served to identify an initial set of situational variables that should be examined in studies of research integrity. The Mumford, Connelly, Helton, Mowry, and Osburn study underscores the importance of stress, alienation, support, need, role models, peer groups, and competitive pressure (34). In this paper we have provided some evidence that these same situational pressures might also be operating in organizational settings. For example, stress appears to be a potentially significant influence on incidents of misconduct at the individual level while competitive pressure appears to influence integrity at the group level. These individual and group level situational influences, moreover, appear to be associated with a coherent set of organizational level influences such as turbulence and munificence.

In identifying the situational variables operating at the individual, group, and organizational levels, moreover, it becomes possible to draw inferences about the conditions under which incidents of misconduct are most likely to be observed and the actions that might be taken by organizations to reduce incidents of misconduct. For example, support appears to be related to misconduct with individuals lacking collaborative networks and broader social support being more vulnerable to misconduct. Organizations, however, by encouraging people to collaborate and build a strong network of professional connections, may do much to minimize misconduct. Similarly, while competitive pressure apparently plays a notable role in scientific misconduct, such simple strategies as avoiding person-to-person comparisons and insuring adequate resources are available may do much to minimize the occurrence of misconduct. Hopefully, the present effort will serve not only as a framework for further research examining the impact of situational variables on scientific misconduct but will provide a basis for formulating new policies that will help insure the integrity of the research process. In fact, given the changes occurring in many scientific fields, there may well in the future be an even more pressing need for practical guidelines along these lines as the rarefied world of science comes into ever closer contact with the manifold demands and pressures of the modern organization.

Acknowledgements

We would like to thank Shane Connelly, Ted Gessner, Jennifer O'Connor, and Howard Timm for their contributions to the present effort. Parts of this effort were supported by a series of grants from the United States Office of Naval Research, Michael D. Mumford, Principal Investigator.
Correspondence concerning this article should be addressed to Dr. Michael D. Mumford, Department of Psychology, The University of Oklahoma, Norman, OK 73019, or *mmumford@ou.edu* .

Bibliography

1. Mumford MD, Baughman WA, Sager CE. Picking the right material: Cognitive processing skills and their role in creative thought. In: Runco MA, ed. Critical Creative Thought. Cresskill, NJ: Hampton; 2000.
2. Mumford MD, Gustafson SB. Creativity syndrome: Integration, application, and innovation. Psychol Bull. 1985;103:27-43.
3. Policastro E, Gardner H. From case studies to robust generalizations: An approach to the study of creativity. In: Sternberg RJ, ed. Handbook of Creativity. Cambridge, England: Cambridge University Press; 1999.
4. Weber RJ, Perkins DN. Inventive Minds: Creativity in Technology. Cambridge, England: Cambridge University Press; 1992.
5. Abra J. Collaboration in creative work: An initiative for investigation. Creativity Res J. 1994;7:1-20.
6. Dunbar K. How scientists really reason: Scientific reasoning in real world laboratories. In: Sternberg RJ, Davidson JE, eds. The Nature of Insight. Cambridge, MA: MIT Press; 1995:365-96.
7. Ruppel CP, Harrington SJ. The relationship of communication, ethical work climate, and trust to commitment and innovation. J Business Ethics. 2000;17:1563-71.
8. Press E, Washburn J. The Kept University. Atlantic. 2000;285:39-55.
9. Dansereau F, Yammarino FJ. Introduction and overview. In: Dansereau F, Yammarino FJ, eds. Leadership: the Multiple Level Approaches. Stanford, CT: JAI Press; 1998:xxv-li.
10. Murphy KA. Honesty in the Work Place. Pacific Grove, CA: Brooks/Cole; 1993.
11. Hartshore H, May MA. Studies in Deceit. New York: MacMillian; 1928.
12. Hegarty WH, Sims H. Some determinants of unethical decision behavior: An experiment. J Appl Psychol. 1978;58:111-21.
13. Victor BC, Trevino K, Shapiro DL. The influence of justice evaluations and social context factors. J Business Ethics. 1993;12:253-63.
14. Darley JM. Social organization for the production of evil. Psychol Inquiry. 1992;3:199-218.
15. Ones DS, Viswesvaran C, Schmidt FL. Comprehensive meta-analysis of integrity test validities: Findings and implications for personnel selection and theories of job performance. J Appl Psychol Monograph. 1993;78:674-703.
16. Sackett PR, Waneck JE. New developments in the use of measures of honesty, integrity, conscientious, dependability, trustworthiness, and reliability for personnel selection. Personnel Psychol. 1996;49:787-829.
17. Hogan J, Brinkmeyer K. Bridging the gap between overt and personality-based integrity tests. Personnel Psychol. 1997;50:587-600.
18. Trevino L, Youngblood SA. Bad apples in bad barrels: A causal analysis of ethical decision making behavior. J Appl Psychol. 1990;75:378-85.
19. Holt RW, Clifton TC, O'Connor JA, Smith TC, Gessner TC, Mumford MD. Influences of destructive personality information on juror decision making. J Appl Social Psychol. 1997;27:781-99.
20. Mumford MD, Gessner TE, Connelly MS, O'Connor JA, Clifton TC. Leadership and destructive acts: Individual and situational influences. Leadership Q. 1993;4:115-48.
21. O'Connor JA, Mumford MD, Clifton TC, Gessner TE, Connelly MS. Charismatic leaders and destructiveness: A historiometric study. Leadership Q. 1995;6:529-55.
22. Fromm E. The Anatomy of Human Destructiveness. New York, NY: Holt, Rhinehart, & Winston; 1973.
23. Gough HG. A sociological theory of psychopathy. Am J Sociology. 1998;53:359-66.
24. Staub E. The Roots of Evil: The Origins of Genocide and Other Group Violence. Cambridge, MA: Cambridge University Press; 1989.
25. Robin DP, Reidenbach RE, Forrest PJ. The perceived importance of an ethical issue on the ethical decision making of ad managers. J Business Ethics. 1996;35:17-28.
26. Blass T. Understanding behavior in the Milgram Obedience experiment. The role of personality, situations, and interactions. J Pers Soc Psychol. 1991;60:348-413.
27. Leith KP, Baumeister RF. Why do bad moods increase self-defeating behavior? Emotion, risk-taking, and self-regulation. J Pers Soc Psychol. 1996;71:1250-67.
28. Oyserman D, Markus HR. Possible selves and delinquency. J Pers Soc Psychol. 1990;59:112-25.
29. DiLalla LF, Gottesman II. Heterogeneity of causes for delinquency and criminality: Lifespan perspectives. Dev Psychopathology. 1990;1:339-49.
30. Reid JB, Patterson GR. The development of antisocial behavior patterns in childhood and adolescence. Eur J Pers. 1989;3:107-19.
31. Walters GD. The Criminal Lifestyle: Patterns of Serious Criminal Conduct. Newberry Park, CA:Sage; 1990.
32. Emmons RA. Relationship between narcissism and sensation seeking. Psychol Rep. 1981;48:247-50.
33. Riskind JH, Kelly K, Harmen W, Moore R, Gaines HS. The loomingness of danger: Does it discriminate all phobia and general anxiety from depressing? Cognitive Ther Res. 1992;16:602-22.
34. Mumford MD, Connelly MS, Helton WB, Mowry JR, Osburn HK. On the construct validity of integrity tests: Individual and situational factors as predictors of test performance. Int J Selection Assessment. [In press].
35. Ash P. Predicting dishonesty with the Reid Report. J Am Polygraph Assoc. 1985;5:139-45.
36. Harris MM, Sackett PR. A factor analysis of an employee honesty test. J Business Psychol. 1987;2:122-35.
37. Gough HG. The California Psychological Inventory. In: Newmark CS, ed. Major Psychological Assessment Instruments: Volume II. Boston, MA: Allyn and Bacon; 1989:67-98.

38. Mumford MD, Costanza DP, Connelly MS, Johnson JF. Item generation procedures and background data scales: Implications for construct and criterion-related validity. Personnel Psychol. 1996;49:361-98.
39. Gessner TE, O'Connor JA, Mumford MD, Clifton TC, Smith JA. Situational variables influencing the propensity for destructive acts: Taxonomy development and validation. Curr Psychol. 1995;13:303-25.
40. Wild JT. The origin of soft-tissue ultrasonic echoing and early instrumental application to clinical medicine. In: Weber RJ, Perkins DN, eds. Inventive Minds: Creativity in Technology. Cambridge, England: Cambridge University Press; 1992.
41. Gruber HE, Wallace DB. The case study method and evolving systems approach for understanding unique creative people at work. In: Sternberg RJ, ed. Handbook of Creativity. Cambridge, England: Cambridge University Press; 1999:93-115.
42. Vosberg SK. The effects of positive and negative affect on divergent thinking performances. Creativity Res J. 1998;11:165-72.
43. Fiedler FE, Garcia JE. New Approaches to Effective Leadership: Cognitive Resources and Organizational Performance. New York: Wiley; 1987.
44. Chang MK. Predicting unethical behavior: A comparison of the theory of reasoned action and the theory of planned behavior. J Business Ethics. 1998;17:1825-34.
45. Goolsby JR, Hunt SD. Cognitive moral development and marketing. J Marketing. 1992;56:55-68.
46. Weeks WA, Moore CW, McKinney JA, Longenecker JG. The effects of gender and career stage on ethical judgement. J Business Ethics. 1999;20:301-13.
47. Arlow P, Ulrich TA. Business ethics and social responsibility of business students: An empirical comparison of Clark's study. Akron Business Economic Rev. 1980;11:17-23.
48. Chonko LB, Hunt SB. Ethics and marketing management: An empirical examination. J Business Res. 1985;13:339-59.
49. Kidwell JM, Stevens RC, Bethke RL. Differences in the ethical perceptions between male and female managers: Myth or reality. J Business Ethics. 1987;6:489-93.
50. Teal EJ, Carroll AB. Moral reasoning skills: Are entrepeneurs different. J Business Ethics. 1999;19:229-40.
51. Reiss MC, Mitra K. The effects of individual difference factors on the acceptability of ethical and unethical workplace behaviors. J Business Ethics. 1998;17:1581-93.
52. Feist GJ, Gorman ME. The psychology of science: Review and integration of a nascent discipline. Rev Gen Psychol. 1998;2:3-47.
53. Feldman DH. The development of creativity. In: Sternberg RJ, ed. Handbook of Creativity. Cambridge, England: Cambridge University Press; 1999:169-88.
54. Csikszentmihalyi M. Implications of a system's perspective for the study of creativity. In: Sternberg RJ, ed. Handbook of Creativity. Cambridge, England: Cambridge University Press; 1996:313-38.
55. Amabile TM. Motivational synergy: Toward new conceptualizations of intrinsic and extrinsic motivation in the work place. Hum Resource Management Rev. 1993:3;185-201.
56. Collins MA, Amabile TM. Motivation and creativity. In: Sternberg RJ, ed. Handbook of Creativity. Cambridge, England: Cambridge University Press; 1999:297-312.
57. Bowie NE. A Kantian theory of meaningful work. J Business Ethics. 1998;17:1083-92.
58. Eisenberger R, Cammeron J. Detrimental effects of reward: Reality or myth? Am Psychol. 1996;51:1153-6.
59. Baumhart R. How ethical are businessman. Harvard Business Rev. 1961;39:6-19.
60. Bellizi JA, Hite RE. Supervising unethical sales force behavior. J Marketing. 1989;53:36-47.
61. Deconinck JB. How sales managers control unethical sales force behavior. J Business Ethics. 1992;11:789-98.
62. Harrison JS, Fiet JO. New CEOs pursue their own self-interests by sacrificing stakeholder value. J Business Ethics. 1999;19:301-8.
63. Pelz DC, Andrews FM. Scientists in Organizations: Productive Climates for Research and Development. New York: Wiley; 1966.
64. Simonton DK. Historiometric studies of creative genius. In: Runco MA, ed. The Creativity Research Handbook: Volume One. Cresskill, NJ: Hampton Press; 1997:3-28.
65. Mullins RF, Sherman R. Creativity and performance appraisal: Shall never the twain meet. Creativity Res J. 1993;6:425-34.
66. Yukl G. Leadership in Organizations. Englewood Cliffs, NJ: Prentice-Hall; 1998.
67. Schminke M, Wells D. Group processes and performance and their effects on individual's ethical frameworks. J Business Ethics. 1999;18:367-81.
68. Zabid ARM, Alasgoff, SK. Perceived ethical values of Malaysian managers. J Business Ethics. 1993;12:331-7.
69. Sims RR. Changing an organization's culture under new leadership. J Business Ethics. 2000;25:65-78.
70. Minkes AC, Small MW, Chatterjee SR. Leadership and business: Does it matter? Implications for management. J Business Ethics. 1999;20:327-35.
71. House RJ, Howell JM. Personality and charismatic leadership. Leadership Q. 1992;3:81-108.
72. Meyer AG, Brown T, Browne MN, Kubasek N. Do we really want more leaders in business? J Business Ethics. 1998;17:1727.
73. Vredenburgh D, Brender Y. The hierarchical abuse of power in work organizations. J Business Ethics. 1998;17:1337-1347.
74. Baron F, Harrington D. Creativity, intelligence, and personality. Annu Rev Psychol. 1981;32:439-76.
75. Mumford MD, Whetzel DC, Reiter-Palmon R. Thinking creatively at work: Organizational influences on creative problem solving. J Creative Behav. 1997;31:7-17.

76. Houndshell DA. Invention in the industrial research laboratory: Individual act on collective process? In: Weber RN, Perkins DN, eds. Inventive Minds: Creativity in Technology. Cambridge, England: Cambridge University Press; 1992:273-390.
77. Reall M, Bailey JJ, Stoll SK. Moral reasoning on hold during a competitive game. J Business Ethics. 1998;17:1205-10.
78. Schwepker CH. Understanding salespeople's intention to behave unethically: The effects of perceived competitive intensity, cognitive moral developments, and model judgement. J Business Ethics. 1999;21:303-16.
79. Feist GJ. The influence of personality on artistic and scientific creativity. In: Sternberg RJ, ed. Handbook of Creativity. Cambridge, England: Cambridge University Press; 1999:273-96.
80. Zuckerman H, Cole JR. Research strategies in science: A preliminary inquiry. Creativity Res J. 1994;7:391-406.
81. Pelton LE, Chowdhury J, Vitell SJ. A framework for the examination of relational ethics: An interactionist perspective. J Business Ethics. 1999;19:241-53.
82. Rallapalli KC, Vitell SJ,, Szeinbach S. Marketers' norms and personal values: An empirical study of marketing professionals. J Business Ethics. 2000;24:65-75.
83. Cote J, Goodstein J. A breed apart? Aecurity analysis and herding behavior. J Business Ethics. 1999;18:305-14.
84. Gessner TE, O'Connor JA, Clifton TC, Connelly MS, Mumford MD. The development of moral beliefs: A retrospective study. Curr Psychol. 1993;11:236-59.
85. Leede JD, Nijhof AHJ, Fisscher OAM. The myth of self-managing teams: A reflection on the allocation of responsibilities between individuals, teams, and the organization. J Business Ethics. 1999;21:203-15.
86. Barnett T. A preliminary investigation of the relationship between selected organizational characteristics and external whistle blowing by employees. J Business Ethics. 1992;11:949-59.
87. Kawathatzopoulos I. Development of cognitive skill in solving business ethics problems: The effect of instruction. J Business Ethics. 1993;12:379-86.
88. Verbke WC, Ouwerkerk C, Peelen E. Exploring contextual and individual factors on ethical decision-making of sales people. J Business Ethics. 1996;15:1175-87.
89. Weaver KM, Farrell OC. The impact of corporate policy in reported ethical beliefs and behavior of marketing practicioners. AMA Proc. 1977;27:477-81.
90. Fritz JMH, Arnett RC, Conkel M. Organizational ethical standards and ethical commitment. J Business Ethics. 1999;20:289-92.
91. Grimalda G. Participation versus social exclusion. J Business Ethics. 1999;21:269-79.
92. Schokkaert E, Sweeney J. Social exclusion and ethical responsibility: Solidarity with the least skilled. J Business Ethics. 1994;21:251-67.
93. Castrogiovani GT. Pre-start-up planning and the survival of new small businesses: Theoretical linkages. J Management. 1996;22:801-22.
94. Morris MH, Marks AS, Allen JA, Perry NS. Modeling ethical latitudes and behaviors under conditions of environmental turbulence: The case of South Africa. J Business Ethics. 1996;15:1119-30.
95. Roussouw GT. Establishing moral business culture in newly formed democracies. J Business Ethics. 1998;17:1563-71.
96. Verschoor CC. A study of the link between a corporation's financial performance and its commitment to ethics. J Business Ethics. 1998;17:1509-16.
97. Judge WO. Correlates of organizational effectiveness: A multilevel analysis of a multi-dimensional outcome. J Business Ethics. 1994;13:1-10.
98. Zarkada-Fraser A. A classification of factors influencing participating in collusive tendering agreements. J Business Ethics. 2000;23:269-82.
99. Enderle G, Tavis LA. A balanced concept of the firm and the measurement of its long-term planning and performance. J Business Ethics. 1998;17:1124-44.
100. Koehn D. Can and should businesses be friends with one another and their stakeholders. J Business Ethics. 1998;17:1755-63.
101. Valenzuela JLD, Villaconta FS. The relationship between companies and their suppliers. J Business Ethics. 1999;22:273-80.
102. Jones MT. The institutional determinants of social responsibility. J Business Ethics. 1999;20:163-79.
103. Vitell S, Festervand TA. Business ethics: Conflicts, practices, and beliefs of industrial executives. J Business Ethics. 1987;6:111-22.
104. Dubinsky AJ, Ingram TN. Correlates of salespeople's ethical conflict: An exploratory investigation. J Business Ethics. 1984;3:343-53.
105. Dooley RS, Fryxell GE. Are conglomerates less environmentally responsible? An empirical examination of diversification strategy and subsidiary pollution in the U.S. chemical industry. J Business Ethics. 1999;21:1-14.
106. Reichers AE, Schneider B. Climate and culture: An evolution of constructs. In: Schneider B, ed. Organizational Climate and Culture. San Francisco, CA: Josey-Bass; 1990:5-39.
107. Sims RL, Keon TL. Determinants of ethical decision making: The relationship of the perceived organizational environment. J Business Ethics. 1999;20:393-404.
108. Agarwal J, Malloy DC. Ethical work climate decisions in a not-for-profit organization: An empirical study. J Business Ethics. 1999;20:1-14.
109. Vidader-Cohen D. Moral climate in business firms: A conceptual framework for the analysis of change. J Business Ethics. 1998;17:1211-26.
110. Key S. Organizational ethical culture: Real or imagined? J Business Ethics. 1999;20:217-25.
111. Argandona A. Sharing out in alliances: Trust and ethics. J Business Ethics. 1999;21:217-28.

112. Hartman SJ, Yrle AC, Galle WP. Procedural and distributive justice: Examining equity in a university setting. J Business Ethics. 1999;20:337-51.
113. Amabile TM, Griskiewicz N. The creative environment scale: The work environment inventory. Creativity Res J. 1989;2:231-54.
114. Gotterbarn D. Not all codes are created equal: The software engineering code of ethics, a success story. J Business Ethics. 1999;22:81-9.
115. Etheredge JM. The perceived role of ethics and social responsibility: An alternative scale structure. J Business Ethics. 1999;18:51-64.
116. Brierley JA, Cowton CJ. Putting meta-analysis to work: Accountants' organizational-professional conflict. J Business Ethics. 2000;24:343-53.

3. Medical Practice and Clinical Research

Waiving Informed Consent: Long-Term Consequences for the U.S. Military

Mary L. Cummings, Engineering Fundamentals, Virginia Tech, USA

Keywords: Anthrax, Gulf War, Informed consent, Investigational drugs, Military, Waiver

In December 1990, the Department of Defense (DoD), anticipating the invasion of Kuwait for Operation Desert Storm, petitioned the Federal Drug Administration (FDA) to waive the federally mandated informed-consent requirements in the case of two investigational drugs: pyridostigmine bromide (PB) and botulinum toxoid (BT). PB, administered orally, was thought to be an effective pre-treatment against the nerve agent soman. The BT vaccine was potentially effective against the bacterium causing botulism (1). Fearful of the possibility that Saddam Hussein would conduct chemical and biological warfare against American troops, the Joint Chiefs of Staff felt that these two investigational drugs could protect U.S. soldiers. The concerns of military leadership were well-founded. Saddam Hussein had used chemical nerve agents and mustard gas against his own people in the Iran-Iraq War (2). However, while military intelligence confirmed that Iraq had the capability to make biological and chemical (nerve agent) weapons, no evidence indicated Iraq had ever made a weapon with soman (3).

FDA did not approve PB and BT. They were considered experimental and fell under the category of investigational new drug (IND). Federal regulations stipulate that if any Federal agency, including the military, desires to use an unapproved drug, that agency must first fully brief the individuals receiving the IND. This briefing must include mention of associated drug use hazards, and the potential recipients' written consent must be obtained. Prior to the Gulf War, informed consent for INDs could only be waived in extreme emergencies, even for the military. However, the U.S. military determined that it was not feasible to seek the informed consent of 700,000 personnel deployed to the Middle East. In 1990, in the months preceding the Gulf War, the military petitioned the FDA to waive the informed consent regulations. The FDA, not wishing to intervene in national security policy and with the approval of an Institutional Review Board (IRB), issued the waiver in an interim ruling in December 1990 (4). However, as part of the approval for the waiver, the military was required to provide information sheets about PB and BT to the recipients detailing the possible side effects. In addition, the military was expected to carefully document the use of the INDs as well as any adverse reactions.

Approximately 300,000 military personnel received the PB pills and 8000 individuals received the BT vaccine during the Gulf War (5). Despite the specific requirement by the FDA that the military track data on both drugs, no procedure was ever established to document which personnel received the drugs and if any adverse side effects were noted (1). Many military personnel experienced systemic medical problems both during and after the Gulf War that were not combat related. Such problems have been termed as the Gulf War Syndrome (GWS). Most notably, over 100,000 Gulf War veterans complained of maladies ranging from chronic fatigue to paralysis in the

Corresponding author: Mary L. Cummings, 332 Randolph Hall,Virginia Polytechnic Institute and State University, Blacksburg, VA 24061, 540-231-6555 (voice), 540-231-6903 (fax), cmngs@vt.edu.

years immediately following the war (3), and of these, 20,000 reported debilitating symptoms (6). In preliminary studies, PB has now been implicated as the primary catalyst of the GWS, however the research is still in its early stages (3).

Waiving Informed Consent

The Federal regulations that govern informed consent for human subjects fall under the purview of the Department of Health and Human Services (DHHS). The regulations state that informed consent may be waived when using INDs, but a number of conditions must be met. No more than minimal risk can exist for the patient, and after the treatment is concluded, the participants must be notified of both the procedure and the possible risks (7). FDA, bound by the DHHS regulations, established their own framework of rules regarding INDs. Prior to the Gulf War waiver, FDA maintained that the informed consent process could be waived only in a life-threatening emergency with the patient unable to communicate and without time to obtain consent from patient's legal representative (7).

The Joint Chiefs of Staff decided it was not feasible to obtain the informed consent of 700,000 military personnel deployed to the Gulf War region and that the pending conflict was essentially an emergency situation by FDA standards. However, prior to granting the military informed consent waivers for the use of PB and BT, FDA required the military to convene an IRB (1). To meet this Federal requirement for the BT vaccine, the military actually convened two IRBs. The first IRB, the U.S. Army Medical Research Institute of Infectious Diseases (USAMRIID) Human Use Committee, was the panel typically used by Army research personnel to consider protocols involving human subjects. The USAMRIID concluded that it was unethical to waive the informed consent of military personnel who would receive BT (8). They further recommended that oral, not written, consent be obtained because oral consent was feasible, and it also respected the rights of the soldiers. Six days later, for reasons not stated in any DoD documents or in any IRB minutes, the DoD then convened a second, entirely different IRB, the Surgeon General's Human Subjects Research Review Board (HSRRB). The HSRRB approved the BT protocol as submitted and recommended that informed consent be waived (9).

Even though FDA waived the requirement for obtaining informed consent for the use of PB and BT in the Gulf War, the approval was contingent upon the military providing those service members who received the INDs with information sheets describing the PB and BT treatments in detail. The sheets were to explain the reasons for using the INDs, the symptoms of botulism and a nerve agent attack, and most importantly any potential side effects or reactions. In addition, the soldiers were also asked to report any of these side effects or reactions. Apparently, the information sheets never made it to the Gulf War theater, so the personnel who received the treatments did not receive any written information about the INDs. However, even a cursory glance at the information sheets that were approved by the Army for dissemination shows that they were at best superficial.

Ethical Issues

In 1978, the National Commission for the Protection of Human Subjects of Biomedical and Behavioral Research issued the Belmont Report that identified three principles that are fundamental in determining whether a research protocol is ethical. They are: respect for persons, beneficence, and justice. These are the primary ethical considerations of an IRB when evaluating a research protocol (10). The crux of the respect-for-persons principle is the preservation of a person's autonomy when making decisions about his/her own medical care. It is this aspect of the Belmont Report that is at issue in waiving informed consent. By swearing an oath to the military and the nation, service members willingly sacrifice some autonomy concerning decisions about their own lives. Enlisting in the military is a supreme sacrifice and highly commendable, but should soldiers lose all rights to autonomy, especially when it comes to their health? The DoD defends its actions in waiving informed consent for INDs by stating, "Allowing a soldier to refuse treatment would endanger him/her as well as those who would try to save their lives and ruin mission success"(5). This paternalistic approach by the DoD overlooks one critical aspect: What exactly constitutes "treatment?"

There has been much debate as to whether the military's use of PB and BT constitutes research or treatment. In the clinical trials held

months before the Gulf War, only a select group of male human subjects were tested with PB and BT. There was no testing for interactions with other chemicals or drugs likely to be used with the INDs, and no long-term studies were conducted (5). Additionally, persons with health problems typical of military populations were never studied in conjunction with the drug testing, and women never participated in any trials (2). Is it ethical and reasonable to maintain that military members receiving drugs tested on a very small, isolated population were receiving "treatment?" Despite the fine line between treatment and research with investigational drugs, FDA's own regulations clearly state that informed consent is required even when the unapproved drug is to be used in a therapeutic manner because the drug has not yet passed full FDA efficacy and safety trials (11).

The respect-for-persons principle was again violated when the information sheets for the INDs were "lost" (5, 12). These sheets should have been paramount in the minds of military medical professionals overseeing the PB & BT programs. The IRB approval and FDA authorization for PB and BT were contingent on the investigators adhering to the approved protocols, which included the distribution of the information sheets. The INDs found their way successfully to the Gulf War theater, and if DoD leadership had considered the sheets a similar priority, they would have been delivered also. Did the military view the information sheets as "not feasible" just as they did for informed consent? When FDA later evaluated the military's use of INDs during the Gulf War, it identified "significant deviations from Federal regulations published in Title 21, Code of Federal Regulations (CFR), parts 50 and 312." (1). FDA cited several areas in which the military was not in compliance. Most notably FDA admonished the military for not disseminating the information sheets prior to the use of INDs in the Gulf War. FDA also issued DoD a stern reprimand for not keeping detailed records on who received the drugs and, most importantly, any adverse reactions suffered by military personnel.

Lastly, the most glaring ethical issue was DoD's use of two different IRBs. When the Army's first IRB found that it was unethical to administer BT to military personnel without their informed consent, the DoD convened a second IRB that produced the desired result of recommending the waiver of informed consent with no impediments. The military was clearly circumventing the system and in doing so trivialized the IRB process and violated Federal regulations. It appears the military was only seeking IRB approval as a formality in an administrative procedure and lost sight of the purpose of the review. FDA, very concerned about the military's use of multiple IRBs when seeking informed consent waivers, censured the military in October of 1999 for this violation and changed the federal regulations regarding military IRBs (1). As a result, IRBs convened by the military to evaluate IND protocols are now required to include at least three members who are not employees or officers of the federal government and are not affiliated with the protocol in any way.

Long-Term Consequences

In December 1997, DoD announced plans to vaccinate all 2.4 million U.S. troops against the biological threat of anthrax. If not treated in its initial stages, anthrax is deadly (13). The current anthrax vaccine is approved by the FDA and was originally designed for agricultural workers and veterinarians. It is a six-shot protocol that is administered over a period of 18 months. Because of this extended treatment period, DoD decided that it must vaccinate all 2.4 million personnel in the unlikely event that all U.S. forces faced a biological threat.

Almost immediately after DoD made its announcement, military members began to protest, based in part on the revelation that service members were given experimental drugs without their knowledge in the Gulf War. Military, medical, and legal critics of the anthrax-vaccine decision were not satisfied that the vaccine was approved by the FDA (13 -15). The sole manufacturer of the anthrax vaccine, Michigan Biologic Products Institute (now Bio-Port) has failed numerous FDA inspections. Most recently, Bio-Port was cited for 23 violations, some of which included sterility and potency deviations, and some microbial contamination (14, 15). In fact, to date the Michigan plant still has not passed an FDA inspection (15, 16).

There have never been any published studies of human efficacy or long-term effects for the anthrax vaccine (15). Moreover, according to an April 1999 General Accounting Office (GAO) report, long-term effects of the anthrax vaccine have never been studied. To further add to the

debate over the efficacy of the anthrax vaccine, the Institute on Medicine has stated that the licensed anthrax vaccine is only effective against cutaneous anthrax and furthermore has never been tested for pulmonary anthrax, which would be the method of delivery in a combat arena (13). A chief Army biological researcher wrote in a 1994 textbook on vaccines that "the current vaccine against anthrax is unsatisfactory" (14). Despite the military's assertions that it is only interested in protecting the welfare of its soldiers, GAO charges that DoD is extremely negligent in tracking adverse reactions to the anthrax vaccine, which was a significant problem with the INDs used in the Gulf War. In fact, many military personnel have reported adverse reactions to the anthrax vaccine. However, in the absence of any established tracking and monitoring system, there is no way to accurately identify any percentages.

With the data supporting the questionable status of the anthrax vaccine and considering DoD's past history, it is not unreasonable to expect military personnel to have doubts about both the efficacy of the anthrax vaccine and the military's plans for implementation. To combat potential insubordination, DoD court-martialed those personnel who refused the vaccine, stating that allowing soldiers to refuse the vaccine would undermine discipline and be prejudicial to good order. Many military members, outraged at DoD's response and facing involuntary inoculation, chose to resign from the service rather than risk their health. The military is already facing serious retention and recruiting problems, and DoD's refusal to make the anthrax vaccine voluntary is only adding to an already critical personnel shortage.

Prior to the mandated anthrax vaccination of all U.S. troops, the military's policies against the threat of chemical and biological warfare were deterrence, containment of the enemy, and use of other defensive measures such as protective suits and warning devices (13). It was not until the Gulf War that troops were inoculated against the threat of possible biological warfare, and it was not until 1997 that troops were forcibly inoculated in peacetime. There has been much criticism directed toward DoD for implementing the anthrax vaccine in peacetime. DoD responded that even though there is no threat of war, the 18-month treatment period for the anthrax vaccine requires that it must prepare its forces for any future contingencies. However, GAO asserts that based on military intelligence data, the biological warfare threat for U.S. troops has not changed since 1990 (14).

A Final Note on Accountability

Accountability is an imperative moral trait required of all military personnel and is considered the cornerstone for military command and leadership. By court-martialing military personnel who refuse the anthrax vaccine, DoD is holding these people accountable for their actions. For those court-martialed, this accountability will not cost them just their jobs within the military. In addition, they are dishonorably discharged and lose all their veterans' benefits as well as their retirement benefits. The nation recognizes the right to make autonomous health-related decisions for all citizens, but it appears, not for military personnel who pay a high price for both autonomy and accountability.

This exacting level of military discipline and accountability is unfortunately glaringly absent from DoD's use of INDs in the Gulf War. Especially troubling are the following:

- DoD convened a second IRB for an IND protocol when the first did not produce the desired recommendation to waive informed consent.
- No one was held accountable for the lost information sheets in the Gulf War. If military officers lost strategic documents protecting troops' safety, they would most definitely face severe punishment.
- No one was held accountable for the incredible lack of record keeping including tracking adverse reactions during and after the Gulf War. Not only did military personnel suffer from a lack of treatment information, but also the entire medical field suffered from the loss of critical data.

This clear double standard in accountability will only continue to haunt the military. Public reports on the military's use of experimental drugs on troops without their knowledge and the anthrax debacle will only continue to exacerbate personnel issues. FDA has recently issued more stringent rulings to prevent some of these ethical transgressions from occurring in the future and to compel the military to abide by the laws they are supposedly defending. However, not until DoD embraces the Federal policies designed to respect basic human rights and autonomy will the

military regain some of its medical credibility and confidence in leadership.

Acknowledgments

I would like to extend a special thanks to Dr. Doris T. Zallen for her guidance and technical support.

Bibliography

1. 64 Federal Register at 54,180; October 5, 1999.
2. Milner CA. Gulf War guinea pigs: Is informed consent optional during war? The Catholic University of America Journal of Contemporary Health Law & Policy 1996 Fall; 13: 199-232.
3. Myers SL. Drug May Be Cause of Veterans' Illnesses. New York Times 1999 Oct 18; Sect. A:18, (col. 4).
4. 55 Federal Register at 52,187 (1990) codified at 21 CFR 50.23(d).
5. Ryan RP. Should combat troops be given the option of refusing investigational drug treatment? The Food and Drug Law Institute Food and Drug Law Journal 1997; 52(4): 377-400.
6. Tuthill KA. Human Experimentation: Protecting Patient Autonomy Through Informed Consent. The Journal of Legal Medicine 1997 Jun; 18: 221-250.
7. Code of Federal Regulations: 45 CFR 46.116(c), (1999).
8. U.S. Army Medical Research Institute of Infectious Diseases memorandum SGRD-UIZ-H. Fort Detrick, Maryland: Department of the Army; October 5, 1990.
9. Office of the Surgeon General memorandum SGRD-HR (15-1a). Falls Church, VA: Department of the Army; October 11, 1990.
10. The Belmont Report: Ethical Principles and Guidelines for the Protection of Human Subjects of Research. Washington DC: DHEW Publication (OS) 78-0011; 1978.
11. Code of Federal Regulations: 21 CFR 312.35 (2000).
12. The Pill. "60 Minutes." CBS television; September 29, 1996.
13. Havrilak GT. The Pentagon's Anthrax Vaccination Immunization Program. Military Readiness Project; 1999 Dec.
14. Rempfer TL. Why Am I Resisting the Vaccine? The Military Trained Me To. The Washington Post 2000 Jan 3; Sect. B:01.
15. Meryl Nass MD. Written testimony before the Subcommittee on National Security, Veterans' Affairs, and International Relations, U.S. House of Representatives. April 19, 1999.
16. The Court-Martial of Major Sonnie Bates: "60 Minutes." CBS television; February 6, 2000.

Promoting Scientific Integrity: The Long Road Ahead–Some Considerations from Espírito Santo, Brazil

Jaime Roy Doxsey, Department of Social Sciences, Federal University of Espírito Santo, Brazil

Keywords: Brazil, Research integrity, Social and human sciences, Teaching research ethics

> "We live in an historical moment of transformation of the scientific paradigm which questions the criteria that scientific rigor in and of itself is ethical." BB Sawaia, 1999.

While the promotion of research integrity has tended to receive widespread governmental and institutional support in the United States and Canada, the responsible conduct of research, including preventing and handling of misconduct, are not always prominent issues in many developing countries such as Brazil. This paper examines the need to stimulate institutional awareness and debate on major issues such as production and communication of scientific knowledge as well as the ethical challenges for developing responsible research practices in the human and social sciences.

A lack of Federal or state legislation, institutional policies or public concern regarding the quality and the ethics of scientific research do not exempt researchers or universities from establishing programs to insure research integrity. The institutional context of a medium-sized Federal government university, the Federal University of Espírito Santo, is examined in an attempt to describe work conditions, the institutional culture and other obstacles for establishing a program to promote research integrity.

In Brazil, recent Federal resolutions in the areas of health, medicine and medical research have established guidelines for human protocol, research integrity, and the protection of human subjects and have determined a local project review procedure along the lines of North American legislation. These guidelines extend themselves to all scientific or academic research activities that involve human subjects. The Brazilian university system and the National Council for Research (CNPQ), however, have neither acknowledged the relevance of these resolutions for research practices nor incorporated them into grant procedures.

At the local level, universities, research institutes, academic centers, departments and graduate programs establish their own policies for research projects and scientific production. Institutional procedures seldom exist for handling allegations of scientific misconduct or establishing protocols for human subjects.

The recent expansion of the number of graduate programs also has increased the need for programs to promote the teaching of research integrity, the ethics of mentoring, and academic career pressures. Further, data management, recording, retention, etc., require pro-active policies to anticipate conflicts and incidents of misconduct.

Corresponding author: Jaime R. Doxsey, Department of Social Sciences, Federal University of Espírito Santo, Rua Pernambuco, 81/1102, Praia da Costa, Vila Velha, Espírito Santo 29.101-335, Brasil, 55-27-3349-8992 (voice), 55-27-3335-7614 (fax), jaime@npd.ufes.br.

What are the implications of these conditions for research with human subjects in Brazil? Is the Brazilian population unduly exposed to doubtful research practices and scientific misconduct, particularly the lower population strata (over 50% of the total population) and more specifically, vulnerable sectors of this population?

At first glance, the answer would be an uncategorical "no". Even considering the lack of a more systematic analysis of actual research practices, there is no direct or indirect evidence that medical, health, human, or social sciences research in Brazil is unethical. What could be considered unethical is the lack of priority for such research at all levels of government in light of the rising indices of preventable social diseases, human violence, drug abuse, and the subsequent decline of living conditions/quality of public services for the lower strata of the population.

With financial support and investment in social policies at an astonishingly low level, social research tends to be descriptive, exploratory, or action-oriented. Academic research seldom receives external or internal financing, and most funding is limited to scholarships for undergraduate trainees or the support of field work.

The lack of a regulatory system of project approval and norms for the protection of human subjects should not be misinterpreted as a lack of research ethics. In a country like Brazil, the few individuals actively engaged in research with human subjects do so with great dedication and considerable respect for their human subjects. Ethical values are not necessarily culturally ascribed or limited by adverse institutional and social conditions.

Nevertheless, what are the actual circumstances in which the social and human sciences are being practiced in Brazil? In what institutional context might it be necessary to initiate the promotion of research integrity and at least provide guidelines for misconduct regulation? How may this promotion of research integrity be best approached?

Design

This paper is a descriptive essay based on personal observations and a review of scientific journals, research methodology textbooks published in Portuguese, Internet homepages, records of research projects available in the Pró-Rectory for Graduate Study and Research, Federal University of Espírito Santo and the annual reports of the Office of Research Integrity, Department of Health and Human Services, U.S. Office of Public Health and Science. The journal editions of the Cadernos de Ética em Pesquisa [Notebooks of Research Ethics], published by the Brazilian National Commission of Research Ethics were specially useful in providing background information for this text.

Results–The Brazilian Context

In Brazil, Federal resolutions first established the National Commission of Research Ethics (CONEP) in 1996 and determined guidelines for human protocol, research integrity, and the protection of human subjects in 1997. The 1997 resolution determined a project review procedure in the areas of health, medicine, and medical research by local Committees of Ethics and Research. At the present time, there are approximately 266 Committees of Ethics and Research (CEPs), the majority of which are located in institutions related to medical instruction or university-associated hospitals.

Although the guidelines extended themselves to all scientific or academic research activities that involve human subjects, the Federal Brazilian university system and the CNPQ have neither acknowledged the relevance of these resolutions for research practices nor incorporated them into institutional procedures.

Data from CONEP reveal the registration of 559 projects in 1999. In a classification by Specialty Topics, most of these projects were grouped under the topic of "international cooperation" (78.3%), and a majority within this category (80%) involved new medications. Distribution in other topical areas included human genetics (7.8%), reproduction (5%), indigenous populations (1.6%), new medical procedures, and equipment (5.3%) (1).

In observance of the data cited above, it is not surprising to conclude that medical and health research formally lead the way in establishing human protocols for research with human subjects. Also, it is not accidental that the majority of the projects reviewed involve international funding and/or cooperative agreements. A recent review of the literature available within Brazil points exclusively toward bioethics and medical and health ethics as dominant topics in the field of ethical considerations (2).

In the humans sciences, there is little to report. However, in 1997, the Federal Council of Psychology determined that new methods or procedures in the field could be utilized if presented as research following research norms for human subjects. The Committee of Ethics in Research at the Catholic University of São Paulo (Catholic University – SP) was implemented through the work of a sociologist who lead discussions to delimitate general principles regarding research ethics, which "took into consideration the specificity, plurality and scientific creativity of the production of knowledge in the human sciences" (3).

Unlike the CEPs created in the medical area, at the Catholic University-SP, the Committee has developed educational functions to represent the ethical principles of the institution, serving as a review board for special recourses. Research projects that are considered to have special ethical questions are sent to the Committee by academic orientators, or by dissertation, thesis, or research commissions for educational evaluations. This university understood that ethical evaluations were already occurring at other institutional levels and that the centralization of the approval process in one committee would be not only impossible but would fail to capture the different optics of research ethics.

Another indicator of the extent of concern for research integrity was presented in a study entitled: "Analysis of ethical aspects of research in human beings contained in the authors' instructions of 139 Brazilian scientific journals". (4) Although the study was limited to a review of scientific journals in the areas of medicine, nursing, odontology, and the general sciences, the authors discovered that 79% of the journals made no reference to ethical considerations in their notes to potential contributors. Only 12% of the journals made reference to the necessity of approval or analysis of the research project by a Committee or Commission of Ethics in Research.

This author has no knowledge of instructions to authors in the area of the social and human sciences. With the growing number of scientific publications in Brazilian universities, there is some concern for establishing selection processes for articles and the evaluation process of the journals. During May, the Faculty of Education at the University of São Paulo organized a conference to discuss the publication policies of scientific journals in education. Discussion was focused on the role of the journals in improving research quality, technical aspects of the journals, and proceedings for evaluation/selection of articles. The last session included an item on scientific and ethical aspects of journal editing.

Increased public concern with electoral opinion polling has attracted attention in the last national elections for president and congress, and most recently in municipal elections. The concern voiced by media and politicians is directed, however, to the possible undue influence of the poll results on the voter and the political system. No ethical concern for poll subjects has been registered. Issues regarding informed consent, the use of the poll results, or the subjects' knowledge of the funding sources have not been publicly evaluated.

Although the lack of governmental support for scientific and technological research and development is a constant criticism throughout the Brazilian society, there is no strong public support for financing academic research. Resources from private and international foundations are centered on corporate interests with little direct university participation. In short, there is little grant money, private or public, which might warrant an institutional policy being created in order to qualify for grant applications.

While international funding or "cooperation" might be instrumental in aligning research interests in the biomedical sciences to installing parallel regulatory proceedings for research ethics, there are no similar external stimuli for the human and social sciences in Brazil. With no public pressure or support for human research, little or no funding, and a lack of issues that might stimulate institutional response tend to neutralize the need for more relevant, modernized research policies in the Brazilian University system.

A Short Case Study–the UFES

Current research policies at the Federal University of Espírito Santo deal principally with the administrative approval of faculty involvement in research as well as release time from academic classroom schedules. Authorization to conduct research is granted by the department council, after a written evaluation often by a research commission of peers. A simplified regulatory system presently requires project approval by the council of department heads at the level of the academic center and

eventual registration of the project in the Pró-Rectory for Graduate Studies and Research.

Details of the project must be outlined on a basic form that specifies the usual information regarding the nature of the study, authors, methods, objectives, and bibliography. No human protocol is required. References to study samples, human subjects, and data collection procedures, when indicated, usually are located in a section on "methodology."

Research projects involving human subjects must have the approval of the Committee on Ethics in Research only for professors from the Biomedical Center. This Committee was registered in March of 1997. No communication from this committee to other academic centers has been documented by the institution. The potential institutional role of this committee could be to distribute and discuss the present regulations, which affect other areas of knowledge.

The lack of information on the necessity for compliance with existing regulatory standards for human protocol or the absence of academic/ administrative requirements for recognizing the ethical consideration of data collection with human subjects are seen as substantial obstacles for promoting research integrity. However, the implications for dealing with possible misconduct are the most serious.

The first dilemma is the extreme negligence with which most universities treat their internal problems of human communication and academic relationships among faculty and students, with no viable procedures or mechanisms to identify, solve, or prevent such problems. In the case of the public Federal universities, professors and university functionaries are classified, by law, as federal public servants, subject to Federal legislation. The legislation is basically a disciplinary regime where duties and obligations are specified. Denouncements of irregularity/misconduct are treated administratively in an process that can consume a year or more.

These laws as well as the university statues and internal regulations date from the years of the military dictatorship in Brazil, seldom having been reformed to establish a less authoritarian academic administrative structure. These instruments refer to problems with faculty or student behavior in terms of order and discipline, keywords common to public policy of the military government. Academic problems involving misconduct in research, plagiarism, misrepresentation of academic production or other problems of research integrity can only be handled administratively under the existing legislation and institutional procedures (5).

In synthesis, academic or research integrity as a terminology or concept plays little part in the actual institutional culture, or at least is not configured as a formal organizational principle in the university culture. This is not to say that academic integrity is not present in many of the pedagogical and academic actions of students and faculty, nor in the daily practices of this institutional culture. Nevertheless, the fact that academic/scientific ethics or research integrity are not explicitly registered in formal university institutional norms considerably complicates the institutional capacity to develop scientific integrity and deal with ethical problems of any nature.

Conclusions

These results confirm the necessity for urgent institutional action to establish normative standards that promote a responsible research environment and a critical consciousness of the need for training/research in scientific integrity in all areas of knowledge. However, the advancement of academic/scientific ethics depends upon a critical analysis of present research practices and the recognition of the protection of human subjects as one component of research integrity inherently connected to the ethical production of knowledge.

Institutional research is needed to identify academic areas with accessibility for a new approach to teaching research integrity as well a current researchers' concerns with research ethics. Institutional support for such curriculum reform is vital, but must occur with a greater strategy to set university goals for excellence in research with human subjects and to reform regulations that are obsolete and ineffective in dealing with problems of academic/scientific integrity.

Caution is necessary to avoid "overdeveloped" procedures that do more to serve the rule makers than to protect the victims of unethical research practices. Perhaps, instead of taking the long road and merely reproducing regulations and administrative procedures for projects review, or awaiting federal legislation, local universities such as the UFES should consider the middle road, one which is not a

short cut or dodges vital issues, but one which stimulates a process that provides access to information, provides debate about research integrity, and acknowledges institutional needs for guidelines to avoid scientific misconduct and to safeguard human subjects, particularly those subjects in situations of cultural or social risk.

Bibliography

1. Bontempo de Freitas C, Oliviera Lôbo M. CONEP faz balanço positivo [National Commission of Ethics in Research – CONEP makes a positive evaluation]. Cadernos de Ética em Pesquisa. 2000;4;4-7.
2. A ética e a bioética em livros [Ethics and bioethics in books] [Special article]. Cadernos de Ética em Pesquisa. 2000;5;17-19.
3. A ética nas ciências humanas [Ethics in the human sciences] [interview]. Cadernos de Ética em Pesquisa. 2000;4;14-7.
4. Sardenberg T, Müller SS, Pereira HR, Oliveira RA, Hossne WS. Estudo mapeia ética em revistas [Study maps ethics in scientific journals] Cadernos de Ética em Pesquisa. 2000;5;11-14.
5. Doxsey JR. Ética academica e cientifica numa perspectiva psicossociológica: a escuta ativa do ouvidor [Academic and scientific ethics in a psycho-sociological perspective: the active listening of an ombudsman]. In: Pinto Lyra R, organizer. A ouvidoria na esfera pública brasileira [The ombudsman in the Brazilian public domain]. João Pessoa (PB): Ed. Universitária, 2000. p. 143-57.

Ethical Research Practice with Human Participants: Problems, Procedures, and Beliefs of Funded Researchers

Elana Newman, Department of Psychology, University of Tulsa, USA

Victoria McCoy, Department of Psychology, University of Tulsa, USA

Anthony Rhodes, Department of Psychology, University of Tulsa, USA

Keywords: Cognition, Confidentiality, Ethics, Health, Informed consent, Mental health.

Although Federal and local guidelines provide general advice as to inform researchers regarding ethical practice (1 - 3), little information is available regarding how researchers carry out such ethical procedures. Despite the use of Institutional Review Boards (IRBs) to monitor ethical practice, there is great variability in how these boards operate and what types of policies are deemed acceptable (4). Similarly, it appears that psychopathology researchers greatly differ in their practices on how to assess and handle participant distress or injury (5 - 7). In some specialty areas, such as depression, there is preliminary evidence that most researchers routinely give referrals (8). Nevertheless, the range of practice is not known.

The need to document how different biomedical researchers implement ethical research policies is important in order to generate and develop viable and informed research policy. For example, it is helpful to understand how researchers recruit participants, train staff, obtain informed consent, and debrief participants (9). Furthermore, specific policies about response and compensation with regard to responding to participants' distress, worsening of conditions, confidentiality issues, informed consent, and other ethical dilemmas across different groups of human research participants is also needed. Sharing such information among researchers from different disciplines, who use different methodologies and research samples, can help to identify the range of options and the need for training initiatives. Finally as technology makes research more global, local community standards of practice may no longer be adequate to understand good research practice (10). To compound this issue, distinctions between research and clinical work and research and organizational consulting are blurring with the trends in program evaluation. Finally, advances in science have made human experimentation itself more complex. Hence there is a need to share information and understand the range of ethical practice in the field so we are better able to respond to these challenges and equipped to create policy in the future.

Currently it is unknown how often research-related injuries and problems occur in the course of routine research protocols. Although flagrant violations are reported or receive media attention, there has been no attempt to quantify the prevalence of such problems in routine practice (11). In order to understand participants' responses it is also important to ascertain the actual prevalence rates of research-related costs and injury across a wide range of samples to determine what groups need

Corresponding author: Elana Newman, Ph.D., Department of Psychology, University of Tulsa, 600 South College Avenue , Tulsa, OK 74104, 918-631-2248 (voice), 918-631-2833 (fax), elana-newman@utulsa.edu

additional safeguards. These risks must be quantified to include both minor costs (abrasions, emotional distress) and major costs (death, disability, and needed hospitalization). Identification of the subgroups at greatest risk for research related harm could help inform policy (12).

Finally the expertise of researchers and opinions need to be shared. As documented, opinions and assumptions about possible risks and benefits of research participation shape ethical appraisals of research (13 - 17). Documenting experienced scientists' opinions and attitudes toward IRBs and research risk, can help establish a clearer understanding of the values that may shape research and research policy.

The goal of the current study is to delineate the rates and types of potential research-related injuries as well as the range of ethical practices and beliefs. This is important since several studies document the range of ethical research practice, but none of them actually assess the prevalence and types of risks (8).

First, it was hypothesized that there is considerable variability of research policies and procedures both within and across types of research and sample characteristics with those researchers working with psychiatric illness being more protective than researchers in other areas. Policies and procedures were defined as (a) level informed consent policy, (b) emergency policies, (c) determination of research- related risk, (d) debriefing procedures, (e) use of referrals, and (f) follow-up procedures.

Second, it was hypothesized that the research risks experienced by psychiatric health groups will be significantly greater than those experienced by the medical physical health group. In addition, it was hypothesized that researchers who studied psychiatric and medical samples were expected to report significantly greater rate of research risks than the non-psychiatric or medical samples. Research risk was defined as (a) Incidence of confidentiality violations for suicide, homicide, and abuse status; (b) Incidence of participants' condition worsening; and (c) Incidence of complaints and or suits filed against researcher or institution.

Method

We generated a list of 3,684 investigators who received federal funding for research projects pertaining to four at-risk groups. Specifically, researchers who studied humans with schizophrenia (n = 264), cardiovascular disease (n = 1472), major affective disorder (n = 899), and traumatic stress (n = 564) were identified from relevant NIH institutes using the Community of Science National Institute of Health database of funded grants (http://cos.gdb.org/best/fedfund/nih-select/inst.list.html) and the Veterans Administration Medical Center grant database (http://www.va.gov/research/research.html). These groups were chosen to represent medically and psychiatric samples that are hypothesized to be at greater risk for research-related injuries. In addition, we identified a pool of 485 federally funded investigators who study cognition in non-patient samples to represent a group hypothesized to be a relatively lower risk for research-related research.

Relevant grant proposals were identified by conducting a search of all proposals that had titles which contained a relevant key word. For example for studies on depression, depression needed to be in the title. For traumatic stress studies, PTSD, trauma or stress needed to be in the title. A detailed listing of key words and the systematic manner in which certain protocols were eliminated is available from the first author. Studies that crossed topic domains, used minors, used animals, or were post-mortum human studies were eliminated from the pool of studies. All treatment studies were eliminated, since they have unique risks and benefits that were not assessed in this study. All projects that were funded as multi-site collaborative studies were also eliminated since it was assumed the ethical considerations might vary across site. Ultimately, 69 funded researchers who study cognition, 79 who study schizophrenia, 61 who study lung-cardiovascular disease, 56 who study affective disorders, and 49 who study violence/PTSD were contacted.

A cover letter, 7 page survey form[1], and return envelope were sent to 314 researchers. A reminder card was sent one month later to all responders and non-responders. The survey began with general information about the respondent's demographics, and research and clinical experience. The researcher was asked to complete the questionnaire in regard to the most recent funded grant. Questions pertained to the setting, sample, type of research, number of sessions, participant characteristics, staff/training and supervision. Then questions about informed

consent, confidentiality issues encountered, participants' reactions, emergency policies, and injuries were attached.

Results

A total of 101 surveys were returned yielding a 32% response rate. Eleven surveys were dropped from the analysis because they were post-mortem studies (n = 4), used minors exclusively (n = 1), focused on substance abuse, HIV, or personality disorders (n = 4), animal studies (n = 1) or couldn't be classified into the groups based on the responses (n = 1). Of the 9 researchers who participated, 52.2% studied mental health (PTSD n = 12, schizophrenia n = 16, major affective disorders = 19), 24.4% studied cardiac or health problems and 23.3% studied "normal" cognition.

Participants

The 90 principal investigators were comprised of primarily Ph.D. trained researchers (73%) and M.D.s (19%). There were more males (63%) than females (37%) represented, and the majority of respondents were Caucasian (94%). The respondents' experience with research ranged from 2 to 49 years and had received a mean of 2.8 (SD = 1.8) federally funded grants in the 5 years prior to the study. The group of researchers reported a mean of 70 peer-reviewed publications, a median of 44 and a mode of 150. Only 20% reported completing a course in research ethics during advanced training. Despite this lack of formal training, 73% felt that they kept current with ethical issues and 50% felt they kept current with legal issues in research. Only 6% and 22% felt they were not current regarding ethical and legal research issues, respectively.

Research Procedures

Informed Consent Policy. With respect to informed consent, the majority of the sample (97%) provided written informed consent and 48% endorsed using methods to assess participants' comprehension of the consent form. Of the 39 respondents who provided open ended descriptions of these methods, 25 asked participants if they had questions, 3 had the interviewer certify person heard and understood, 3 used independent monitors, 2 relied on other indicators (fluency, literacy, neurological status), 1 used family consent, 1 used structured consent, 2 asked the respondent to repeat questions, and 2 relied on signature to indicate comprehension. Although 85% reported no need to investigate if the identified participant could legally provide consent, the remaining 15% reported a need ranging from once (7%) to eighty-five times (1%).

With respect to informed consent, 53% of these researchers indicated that there were instances in which the confidentiality of the research participant might be broken. As predicted, this policy differed by type of sample group [x^2 (2, n = 85) = 10.75 p =<.05], with 66% of those who worked with mental health groups, 55% of those who worked with physical health groups, and 21% of those who studied cognition stating instances in which the research team would consider breaking the confidentiality of the research record. Among the group who informed participants about confidentiality issues, 55% reported communicating this in specific rather than general terms.

Emergency Policy. Seventy-eight percent (n = 61) of the researchers endorsed having a protocol in place a priori to respond to emergencies. The groups significantly differed in this policy [x^2(2, n =78) =32.15, p <.05] such that 95% of mental health researchers, 90% of physical health researchers, and 28% of cognitive researchers reported such emergency policies in place. Among the 47 who provided open ended descriptions of these policies, 15 described use of emergency on-call personnel, 8 cited they had "written policies," 6 used standard local protocols, 6 cited immediately contacting the project director or principal investigator, 5 trained staff in Cardio Pulmonary Resuscita tion (CPR), and 3 discussed continuous monitoring during research. The remaining four described emergency medication, medical response plan in lab and for evacuation, methods for handling high blood pressure, and one general training how to respond to a variety of situations.

Determination of Research-Related Risk. Seventy-eight percent (n = 62) of the researchers sampled reported keeping records regarding the "frequency to which individuals experienced negative and noticeable reactions." Mental health researchers reported significant greater documentation than health or cognitive researchers [x^2 (2, n = 81) = 19.79, p < .05] such that 88% of mental health researchers, 79% of physical health researchers, and 52% of cognitive researchers kept such records.

Debriefing Procedures. Sixty-four percent (n = 57) of the researchers conducted debriefings

	Ranking			
Factors	Least important	Important	Fairly Important	Most Important
Manipulation check	24 (63%)	5 (13%)	8 (21%)	1 (3%)
Educate participants	1 (2%)	18 (33%)	7 (13%)	28 (52%)
Check on participant	7 (14%)	12 (24%)	10 (20%)	21 (42%)
Express gratitude	6 (11%)	9 (16%)	26 (46%)	15 (27%)

Table 1. Number (and percentage) of participants ranking relative importance of 4 factors in planning debriefing.

after the research protocol. In fact, 70% of mental health professionals, 42% of health researchers, and 71% of cognitive researchers used such debriefings [x^2 (2, n = 80) = 5.06, p =.08]. The majority (80%) of these debriefings were conducted orally, although 6% were conducted in writing, with 14% conducted in both formats; there was no statistically significant difference among the groups regarding format [x^2 (4, n = 51) = 4.48, p = .34]. The majority of these debriefings were done in individual sessions (88%) rather than group (4%), varied (6%) or family formats (2%); this did not vary significantly among groups format [x^2 (6, n = 51) = 9.05, p = .17]. As can be seen on Table 1, investigators felt debriefings were most important for educating participants and checking on participants. It is interesting to note that manipulation checks were deemed least important.

Use of Referrals. Forty-one researchers (46% of the sample) responded to the item about referral policy. Among those who responded, 20% reported providing referrals to all participants, 12% to those participants who indicated interest, 17% to only those in distress, 42% to those either interested or distressed, and 10% in "other" circumstances. Three researchers described such other circumstances as "offered to all deemed appropriate, but given to those interested;" "two found to have physical disorders," and "all those screened with high blood pressure."

Given this practice, the number of referrals for non-emergencies ranged from 0 to 90 (mean = 4.76, s.d. =13.02; mode =0). The mean number of referrals for the mental health, health and cognitive research teams were 8.56 (S.D. = 17.83), 2.29 (S. D. = 4.10) and .40 (S.D. =1.05) respectively, but these differences did not meet criteria for statistical significance [F (2, 65) = 2.9, p = .062].

With respect to actual practice regarding referral for immediate hospitalization, 6 researchers recommended immediate referral for a condition or concern, (with two researchers recommending it once, and the rest experiencing it twice, three times, four times and 10 times). It is unknown if these referrals were based on research-related injuries, or other conditions uncovered during the protocol.

Follow-up procedures. Fifty-four percent (n = 41) of the researchers reported follow-up efforts to determine if participants experienced a worsening of condition. These efforts significantly differed across groups [x^2 (2, n = 76) = 14.35, p <.01] such that 67% of mental health researchers, 55% of health researchers, and 8% of cognitive researchers used such methods. In terms of actual numbers, 24 researchers reported conducting a follow-up at least once to check on a participant.

	Never	**Infrequently**	**Sometimes**	**Regularly**	**Always**
Suicidality	58 (64%)	20 (24%)	4 (4%)	2 (2%)	1 (1%)
Homicide	76 (91%)	5 (6%)	2 (2%)		1 (1%)
Child abuse	72 (85%)	9 (11%)	2 (2%)		2 (2%)
Elder abuse	78 (94%)	4 (5%)			1 (1%)
Abuse of the disabled	78 (94%)	4 (5%)			1 (1%)
HIV status	64 (77%)	9 (11%)	8 (10%)	2 (2%)	
Substance abuse	49 (59%)	10 (12%)	14 (17%)	9 (11%)	1 (1%)
Criminality	68 (83%)	9 (11%)	1(1%)	3 (4%)	1 (1%)
Violence toward partner	67 (80%)	11 (13%)	2 (2%)	3 (4%)	1 (1%)
Other	50 (94%)	3 (6%)			

Table 2. Number and (Percentage) of researchers who faced confidentiality issues.

Research Risks

Incidence of confidentiality violations. The research staff occasionally faced confidentiality dilemmas as shown in Table 2, with substance abuse being the most frequently encountered issue. However, only 8 researchers actually broke confidentiality. Of these 8, 6 studied mental health (n = 3 mood disorders, n = 2 schizophrenia, n =1 PTSD), 1 studied normal cognition, and 1 studied health conditions. Among those researchers who described the specific circumstances, two reported needing to hospitalize at least one participant against his/her will, three reported having to file at least one report to the authorities, and two reported needing to warn at least one person in danger.

Incidence of participants condition worsening. During the protocol a range of emotional and physical experiences were encountered (See Table 3); clearly crying appeared most often. Although it was rare that a participant became medically compromised, it did occur. Twelve researchers (13%) reported at least one research-related injury. Two researchers reported that at least one participant had a research-related infection. Five researchers reported at least one case of temporary disability, and none reported research-related death. It should be noted that only 53% of researchers reported knowing how many participants experienced an immediate worsening of condition (research related injuries) after completing the research protocol; Knowledge of research-related injuries was not related to type of research conducted [x^2 (2, n = 73) = .42, p = .81]

Incidence of complaints filed against a researcher or institution. In this sample, 18% reported infrequent complaints about research staff's conduct. Two percent (n =2) reported complaints filed against the institution however none resulted in legal proceedings. On the other hand, 77% of researchers reported that participants thanked them, with 33% reporting this occurring sometimes, and 12% reporting this as a regular occurrence.

Discussion

In this preliminary study, 90 federally funded researchers who work with human participants responded to a survey about ethical research practice. There seems to be a great variation in ethical practice among distinguished researchers, although all these research participants were sensitive to research-related ethical dilemmas.

Policies

There is a great deal of variation in research policy implementation. Although nearly all use written informed consent, researchers varied in the detail that they provide participants about the limits of confidentiality. Although the majority of researchers developed emergency policies and debriefing procedures, the nature of these procedures also varied. Although often required, 32% did not keep records of participants' negative and noticeable reactions. Approximately half the researchers reported

	Never	Infrequenty	Sometimes	Regularly	Always
Cried	35 (42%)	24 (29%)	16 (19%)	7 (8%)	1 (1%)
Became hostile or angry	33 (43%)	35 (42%)	13 (16%)	3 (2%)	0
Experienced Panic Attacks	59 (71%)	17 (21%)	6 (7%)	1 (1%)	0
Expressed extreme fear	55 (66%)	16 (20%)	8 (9%)	4 (5%)	0
Reported feeling spacey	51 (62%)	18 (22%)	12 (15%)	1 (1%)	0
Became medically compromised	66 (81%)	14 (17%)	2 (2%)	0	0
Threatened the research staff	71 (87%)	10 (12%)	1 (1%)	0	0
Other	33 (86%)	2 (5%)	1 (3%)	1 (3%)	1 (3%)

Table 3. Number and percentage of researchers who encountered participants' emotional or physical response to research.

using follow-up methods to check on participants' condition. However, less than half the sample responded to the item regarding the use of referrals and those that did respond indicated a range of practices with respect to referring to other agencies. As anticipated, researchers working with psychiatric illness being more protective and explicit about policies for emergencies, risk documentation, and follow-up procedures but not for debriefing.

Risks

With respect to research risk, a minority of researchers reported having to deal with confidentiality issues, worsening of conditions, and complaints from participants. However, emotional and physical symptoms were encountered. In particular, 58% (n = 48) experienced crying, and 12 researchers (13%) reported temporary research-related injuries. Given that several of these studies were about health conditions, it is difficult to evaluate if these reactions were elicited by research participation, or were symptoms that individuals experienced irrespective of research participation. These reactions need to be examined in future studies in the context of baseline functioning of individuals to further understand if they meet the requirements of minimal risk. Nonetheless, the data are consistent with claims that the physical hazards of being a research participant are minimal even among medical procedures (18). Although, these risks appear minimal, they might be an underestimate given that about half the researchers did not document or know the number of participants whose condition worsened.

Finally, very few researchers received formal training in research ethics although the majority were confident that they were up to date in ethics, and half felt prepared for legal challenges. Given that researchers thought highly of their respective IRBs, continuing education may be best implemented through local IRBs.

There are several limitations to this study. First sample bias and demand characteristics may have affected the generalizability of these results. Although the extensive comments written on those returned surveys suggest that researchers were interested in sharing their experiences, sample bias may have affected the results. Second, while this study reveals a diversity of ethical practices, the quality of ethical implementation is not examined. Hence it is not known if this diversity suggests unsuccessful or successful flexibility of methods in responding to the needs of human participants.

Although the participation rate precludes generalizing to all researchers, these preliminary results provide information that can be useful in designing training and compliance policy. In particular, the diversity of responses suggests the need for cross-training across subspecialties to share perspectives. Individuals with risk factors may not only present for studies of health and mental health problems, so it can be helpful to share approaches across specialties. For example, although the majority of research-injuries were identified among those mental heath studies, they were not exclusively there. Furthermore it is unclear, given the lack of documentation and investigation, if this reflects better preparedness of mental heath researchers or greater risk in these studies. Future studies may be able to better examine this by ongoing quality control (19).

Acknowledgements

The research for this paper was supported by a University of Tulsa Summer Faculty Development Grant. A portion of this paper was presented at Research Conference on Research Integrity, Office of Research Integrity, Bethesda, MD, November 2000.

Notes

1. A copy of the survey is available from the first author.

Bibliography

1. American Psychological Association. Ethical principles of psychologists. Am Psychol 1981; 36: 633-8.
2. Protection of Human Subjects. 45 C. F. R. Subtitle A, pt. 46. (2001).
3. National Commission for the Protection of Human Subjects of Biomedical and Behavioral Research. The Belmont Report: Ethical principles and guidelines for the protection of human subjects of research (DHEW Publication No OS 78-0012). Washington DC: Department of Heath, Education, and Welfare; 1978.
4. Goldman J, Katz MD. Inconsistency and institutional review boards. JAMA 1982; 248(2): 197-202.
5. Barton JM, Macmillan MS, Sawyer L. The compensation of patients injured in clinical trials. J Med Ethics 1995; 21(3): 166-9.
6. Burbach DJ, Farha JG, Thorpe, JS. Assessing depression in community samples of children using self-report inventories: Ethical considerations. J Abnorm Ch Psychol 1986; 14(4): 579-89.

7. McCrady BS, Bux DA. Ethical issues in informed consent with substance abusers. J Consult Clin Psychol 1999; 67: 186-93.
8. Stanton AL, New MJ. Ethical responsibilities to depressed research participants. Prof Psychol 1988; 19: 279-85.
9. Silverman HJ. Ethical considerations of ensuring an informed and autonomous consent in research involving critically ill patients. Am J Respir Crit Care Med 1996; 154(3): 582-6.
10. Edgar H, Rothman DJ. The institutional review board and beyond: Future challenges to the ethics of human experimentation. Milbank Q 1995; 73: 489-96
11. Hilts PJ. Safety Concerns Halt Oklahoma Research. The New York Times 2000 Jul 11; Sect. A:12.
12. Newman E, Kaloupek DG, Keane TM, Folstein S. Ethical decision making among trauma researchers. In Kantor GK, Jasinski J, editors. Out of the darkness: contemporary perspectives on family violence. Newbury Park (CA): Sage Press; 1997; 271-81.
13. Ceci, SJ, Peters D, Plotkin J. Human subjects review: personal values, and the regulation of social science research. Am Psychol 1985; 40(9): 994-1002.
14. Oldham JM, Haimowitz S, Delano SJ. Protection of persons with mental disorders from research risk. Arch Gen Psychiatry 1999; 56: 688-93.
15. Scarr S. Race and gender as psychological variables: Social and ethical issues. Am Psychol 1988; 43: 56-9.
16. Sieber JE. Issues presented by mandatory reporting requirements of researchers of child abuse and neglect. Ethics Behav 1994; 4: 1-22.
17. Stanley B, Sieber JE, Melton GB. Empirical studies of ethical issues in research. A research agenda. Am Psychol 1987; 42(7): 735-41.
18. Levine M. Ethics and regulation of clinical research. Baltimore: Urban & Schwarsenberg; 1981.
19. Newman E, Willard T, Sinclair R, Kaloupek D (forthcoming). The costs and benefits of research from the participants'view: the path to empirically informed research practice. Accountability Resear 2001; 8: 27-47.

*Balancing Risks and Benefits of Deception in Assessing Genetic Screening**

Dennis L. Thombs, Department of Adult Counseling, Health and Vocational Education, Kent State University, USA

Colleen A. Mahoney, Mahoney Consulting Group, Kent, Ohio USA

R. Scott Olds, Health Education and Promotion, Kent State University, USA

Keywords: Alcoholism, Deception, Genetic screening

The Human Genome Project is a massive international research program designed to map the human genome sequence(1). The fundamental purpose of the program is to spur a transition to DNA sequence-based biology and biomedical science(2). In addition to revolutionizing medical diagnostics and therapy, the Human Genome Project will create new challenges in a variety of fields including law, medical ethics, public health, and health services administration(3). The anticipation of these changes does not represent a distant concern. A "working draft" of the entire human sequence is expected by the end of 2001(2).

Against the backdrop of the Human Genome Project, this article critically examines the use of intentional deception to assess (and anticipate) the utilization of genetic screening for alcoholism susceptibility. For some time, the manipulation of study participants by deception has been controversial in experimental social psychology(4). This controversy has emerged in health behavior research as a consequence of the remarkable progress made by the Human Genome Project. Little is known about the public's interest and utilization of clinical genetic testing(5). In the specific area of *predictive* genetic screening, a deception paradigm (described below) has been found useful for assessing utilization. This paradigm helps estimate utilization when such tools are on the horizon, but not yet available to the consumer. Intentional deception appears to be necessary because "hypothetical testing,"(6, 7) honestly described to research subjects as available "sometime in the future," generates inflated interest compared to testing described as "currently available"(8, 9).

In an editorial that appeared in the *Journal of American College Health*,"Hard Questions About Research Procedures: The Search for Authenticity"(10), Dr. Richard Keeling objected to the use of deception in a quasi-experimental study conducted by the authors. The report of this investigation appears in the same issue of that publication "Application of a Bogus Testing Procedure to Determine College Students' Utilization of Genetic Screening for Alcoholism"(11). Interested readers may turn to that article for a full description of the study methods, including the fabricated story concocted to test student interest in genetic screening for alcoholism susceptibility.

Dr. Keeling's editorial is an example of a conservative, but perhaps increasingly common position

* Reprinted with permission from the *Am J Health Behav* 2001; 25(2): 100-105. R. Scott Olds was invited to present this paper at the RRI conference under the title: Responsible Research Conduct that Balances Risks and Benefits of Deception in Assessing Genetic Screening Utilization.
Corresponding author: R. Scott Olds, H.S.D., Health Education and Promotion, 316 White Hall, Kent State University, Kent, OH 44242-0001, 330-672-0679 (voice), 330-672-3063 (fax) , rolds@kent.edu.

on human subjects protection that exaggerates risk to study participants and discourages potentially valuable inquiry. The conservative position is based on the following beliefs: 1) deception is inherently harmful; and 2) deception research is not carried out under realistic conditions and therefore is not of value. The authors believe their views are based on an ethic of measured and reflective discourse, instead of a "knee-jerk" response fashioned to serve a particular ideology.

According to Aronson and colleagues (4), when considering the use of deception in research, investigators must weigh the psychological discomfort participants may experience against the value of the study. There is no single set of rules that can be applied to resolve this dilemma, and reasonable professionals will arrive at different judgments in this difficult analysis. To determine college student interest in genetic screening for alcoholism susceptibility, it was reasonable to expose them to what was believed to be modest psychological and social risks. The Institutional Review Board at Kent State University concurred, and with certain stipulations gave approval to conduct the study.

The subjects in this study were deceived about the availability of a genetic screening test. For up to seven days, 181 students thought they could schedule a predictive screening test for alcoholism that does not yet exist. The authors did not believe that this lie harmed the students in any substantial way. In broad-brush comments, Dr. Keeling (10; see page 101 of his editorial) claims that today's college students are often exploited by society and that any challenge to their "search for authenticity" poses an unacceptable risk to their mental health and/or future social functioning. It seems that this view is not unusual in academia today. Such a position represents "politically correct" discourse that exaggerates the risks of deception in this study and casts a broad net of condemnation over all uses of deception in research. Clearly, humans have been mistreated in research that employed deception (e.g., the Tuskegee Syphilis Study), but distinctions can and should be made in its application.

In this era of heightened concern about compliance with Federal regulations on research involving human subjects, "minimal risks" in behavioral science research have sometimes been subtlety redefined as "unacceptable risks." The authors have no data to support or dispute such speculation, but wonder whether the balancing of risks and benefits has tilted toward the former in recent years. If so, does this shift represent increased concern for human subjects? An iconoclastic interpretation is that the conservative analysis of risk has been motivated by fears of lawsuits and a desire to protect the university from legal action. In addition, doubts about the quality and usefulness of behavioral science research in general, may be in operation in some quarters which only further discourages full consideration of the potential benefits of such work.

No data were collected in this study to support the claim that the students were not harmed by the deception. However, it should be noted that the empirical literature does not support the view that research using deception is any more harmful than non-deception research (4). One review of the literature concluded that it was rare for participants to feel that they had been harmed by intentional deception (12). Though empirical studies on the effects of deception are few, those that have been conducted generally have found that participants report greater enjoyment from having participated in a deception experiment than in a nondeception experiment (13). This is probably due to deception studies being less boring (4). To address these concerns, in the future, investigators should follow up with participants to determine their reactions to research deceptions.

It is noted that the source of discomfort in deception research is not only learning later that one has been deceived, but equally, if not more important is that the person often learns something painful about themselves or others (14). Again, data were not collected to support this hypothesis, but it is strongly suspected that among those students who were uncomfortable in this study, the primary source of their discomfort was their current drinking behavior. As noted, the sample was over-represented by heavy drinking students. Participation in the study required them to reflect on their own alcohol use as well as that of their family members. Indeed, it was sensed by the authors that some students were uncomfortable while responding to the questionnaire and watching the presentation. In other words, the discomfort that some experienced appeared to occur *before* the debriefing, rather than after it (when they learned they had been deceived). Some students actually appeared amused during the debriefings.

The level of discomfort experienced by

students was probably comparable to being asked to participate in an anonymous self-report survey of alcohol use, and probably no greater than sitting in routine lectures and discussions in health education courses that deal with any number of sensitive issues. The discomfort that some may have experienced was not considered to be detrimental or bad. Good health education "shakes up" students by confronting biased perceptions of risk and challenging existing social norms. It also is consistent with the traditional view of higher education, which is to challenge conventional thinking and behavior and to engage students in debate about controversial issues.

Dr. Keeling (10) also was critical of the contention that the study conditions were "realistic." The authors agree with his observation that if (or when) genetic testing for alcoholism susceptibility becomes available, protocols very likely will require extensive patient counseling before and after the procedure. So by this benchmark, the study's procedure was not realistic. The authors should have been more precise by stating that "our method was more realistic than using a procedure that described screening as a future possibility." However, at the same time, introducing extensive patient counseling into the study procedure would have required us to employ a far greater level of deception. Such a research design would be considered unethical by virtually all professionals and would justify Dr. Keeling's response. This study protocol, however, does not.

As the study was carried out, participants were deceived for no more than seven days. They were debriefed and offered the opportunity to withdraw their data without penalty. In his editorial, Dr. Keeling (10) stated,

> . . . Having watched a computer-generated presentation (for 7 minutes) and heard a brief explanation of the study itself, they were then required to state their intentions about being tested immediately. There was little time for them to ponder the issues and submit a formal request to be tested. . .(p. 100).

This description of the study's methods is not accurate. Careful reading of the methods clearly stated that students were told they did *not* have to make a decision immediately after the presentation. A questionnaire item allowed them to respond *I am uncertain about whether or not to be tested* (see p.106 of our article)(11). Further, their participation was always voluntary and invitational. They were able to cease participation at any time without penalty. Dr. Keeling was accurate in describing that over the next seven days, students were not given counsel or additional information about the test. In this respect, the procedure was not as realistic as future testing probably will be, but neither was it as unrealistic as described by Dr. Keeling in his editorial. It is acknowledged that in the future, people may contemplate the testing decision for extended periods of time, perhaps even many years. Obviously, this study does not address readiness to seek testing over extended time intervals, but it does provide marketing information about what to expect if promotion of genetic screening for alcoholism susceptibility among high-risk drinkers becomes a public health goal.

The preliminary findings from this study suggest that among college students, there may be little enthusiasm for seeking alcoholism screening if (or when) it becomes available. Certainly this issue deserves further investigation. The authors believe the health promotion profession has an obligation and responsibility to conduct research that anticipates and informs the development of sound public health policy. If future public health policy supports genetic testing for alcoholism susceptibility, ethical questions need to be raised by the professions concerned with public health. This study is part of the foundation needed to address these questions.

These debates are important and healthy, but they are not easy. The issues surrounding genetic testing are complex. Billions of dollars are being spent on genome research for the purpose of developing effective technologies to treat and prevent disease. Yet, relatively little attention is being given to the behavioral, social, and health service implications of this technology. There is a need to better understand the utilization of predictive screening for a variety of disorders, including alcoholism. This study should stimulate discussion among health promotion professionals about these aspects of genetic testing.

Bibliography

1. National Human Genome Research Institute. The Human Genome Project. *www.nhgri.nih.gov,* 2000; April, 14.
2. Collins FS, Patrinos A, Jordan E, Chakravarti A, Gesteland R, et al. New goals for the U.S. Human Genome Project: 1998-2003. *Science* 1998;282:682-689.

3. Khoury MJ and Genetics Work Group. From genes to public health: The applications of genetic technology in disease prevention. *Am J Pub Hlth* 1996;86:1717-1722.
4. Aronson E, Ellsworth PC, Carlsmith JM, Gonzales MH. Methods of Social Psychological R esearch (Second Edition). New York: McGraw-Hill; 1990.
5. Croyle RT, Lerman C. Interest in genetic testing for colon cancer susceptibility: Cognitive and emotional correlates. *Prev Med* 1995;22:284-292.
6. Kessler S, Field T, Worth L, Mosbarger H. Attitudes of persons at risk for Huntington's disease toward predictive testing. *Am J Hum Genet* 1987;26:259-270.
7. Williamson R, Allison MED, Bentley TJ, Lim SMC, Watson E, et al. Community attitudes to cystic fibrosis carrier testing in England: A pilot study. *Prenat Diag* 1989;9:727-734.
8. Craufurd D, Dodge A, Kerzin-Storrar L, Harris R. Uptake of presymptomatic predictive testing for Huntington's disease. *Lancet* 1989;2:603-605.
9. Tambor ES, Bernhardt BA, Chase GA, Fadden RR, Geller G, et al. Offering cystic fibrosis carrier screening to an HMO population: Factors associated with utilization. *Am J Hum Gen* 1994;55:626-637.
10. Keeling RP. Hard questions about research procedures:The search for authenticity. *J Am Coll Health* 1998;47:99-102.
11. Thombs DL, Mahoney CA, Olds, RS. Application of a bogus testing procedure to determine college students' utilization of genetic screening for alcoholism. *J Am Coll Health* 1998;47:103-112.
12. Baumrind D. Research using intentional deception: Ethical issues revisited. *Am Psychol* 1985;40:165-174.
13. Christensen L. Deception in psychological research: When is it justified? *Personality and Social Psychology Bulletin* 1988;14:664-675.
14. Smith SS, Richardson D. Amelioration of deception and harm in psychological research: The importance of debriefing. *Journal of Personality and Social Psychology* 1983;44:1075-1082.

Research Integrity and the Direct Involvement of Persons with Disabilities

Nancy B. Robinson, Communication, Arts and Sciences Department, California State University, Chico, California, USA

Kathleen C. Sadao, Curriculum and Instruction Department, University of the Pacific, Stockton, California, USA

Keywords: Responsible conduct of research, Preservice training, Problem-based learning, Persons with disabilities, Affective learning

Teaching students in the health, human service and education professions to be responsible in their interactions with persons with disabilities, as service providers and researchers, poses unique challenges to educators to move beyond imparting knowledge to impacting attitudes, values and ethics. Recent emphasis on outcomes of professional education programs most frequently focuses on indices of cognitive achievement and performance of specific skills or competencies. Measures of affective learning, or student attitudes and values toward the persons they serve, are less frequent and more difficult to document. Universities need to educate professionals who are capable of the responsible conduct of research. Pre-service education models are shifting from a traditional didactic approach to the use of case studies and problem solving, in an effort to influence affective learning and the application of knowledge and skills in real-life simulations. Studies of effective teaching methods to prepare professionals in the area of responsible conduct of research with human subjects are clearly needed. Person-focused learning approaches developed from interactive teaching models, used increasingly in pre-service education in disability services and programs. The use of case studies tends to promote application of theoretical knowledge and positive changes in affective learning, or students' attitudes and values.

Person-focused learning approaches move beyond case studies and directly include persons with disabilities and family members as partners. Research and teaching-involving people with disabilities assume that validity is strengthened through the direct involvement of people who experience disability daily (1). Kvale and Burns discuss threats to validity and the need to reconceptualize validity in qualitative research (2, 3). Due to the integral involvement of the researcher to conduct qualitative research, Kvale argued that qualitative research requires attentiveness to the concept of validity and its social construction with constant reference to the values, attitudes and experiences of the researcher and participants (2). Further, qualitative research methodology applies to interactive teaching, in which themes are explored and developed based on real-life scenarios (4). Participatory action research, a qualitative research process, directly involves key stakeholders in all phases of investigation (5, 1). In the present study, partnerships with persons with disabilities and family members began and continued throughout the design, implementation, and evaluation of co-teaching activities.

Corresponding author: Nancy B. Robinson, CSU Chico, SPPA Program, AJH 107B, Chico, CA 95929-0350, 530 898-6848 (voice), 530-898-6612 (fax), nrobinson@csuchico.edu

The goal of the present study is to demonstrate and evaluate an interactive teaching method that directly involves people with disabilities and their family members and the impact of this model on students' attitudes and values, or on affective learning. Although the use of case study approaches in college level teaching, particularly with persons with disabilities, produces positive student learning outcomes, the differences in approaches to the uses of case studies are not explored. Specifically, the researchers sought to examine the effectiveness of person-focused learning to promote the responsible conduct of research among graduate, post-graduate and doctoral students.

Three major developments in policy, program development and teaching practices led to the development of person-focused learning. First, shifts in legislation and policy began in the 1950's and 1960's in the US, which continues today with increasing emphasis and advocacy for the rights of people with disabilities to have equal access to all arenas of community life. Second, increasing focus on rights and advocacy for people with disabilities contributed to the self-determination movement that places decision-making and life choices with the people affected, people with disabilities. Third, teaching practices in higher education shifted from traditional didactic models to interactive, problem-solving models that strive to establish critical thinking skills among students in preprofessional training programs. The combined influences of these broadly defined trends in policy, program, and professional practice are particularly relevant in higher education, where the forming of future professionals' values, attitudes, knowledge, and skills are critical for future practice and partnership with people with disabilities.

Teaching methodology in professional training programs is changing from a didactic approach to an interactive model that requires students to take responsibility for their own learning (6). Medical education first developed problem-based learning (PBL) to create a student driven learning model. PBL was since adapted to curricular content in several health, human service, and education disciplines. Beginning with PBL, four approaches to interactive and problem-solving approaches to teaching are briefly described in this paper. The strengths and contributions of each model are addressed and the person-focused learning model is highlighted as the focus of this study and context for participatory action research.

Problem-Based Learning. As stated above, PBL began within medical education to increase the application of medical theory and information with specific patient case studies and has since extended to nursing, occupational therapy, and other fields (7-11). Cockrell, Hughes, Caplow, and Donaldson described problem-based learning as a "collaborative learning approach" (12). Collaborative learning is premised on Vygotskian concepts that define learning as the social construction of knowledge. The cooperation and shared resources that take place in PBL learning reflect tasks in "real world" settings. These authors outlined six basic phases in PBL: (a) encounter with the problem; (b) free inquiry; (c) identification of learning issues; (d) peer teaching; (e) knowledge integration and (f) problem resolution. Based on their investigation of student's perspectives of PBL, Cockrell et al. found three key areas of student perspectives of PBL: ownership, group dynamics, and tutor feedback (12). Students reported a deeper level of understanding and retention in the PBL process compared to more traditional teaching approaches and increased awareness of team building skills. Students stated a preference for tutors who were non-directive and non-obtrusive. Students reported that the benefits of collaborative learning included: a) learning to become part of a learning community, and b) learning to speak the language of the community of professionals within the discipline.

Inquiry-based learning. Inquiry-based learning (IBL) uses a case-study process to encourage student responsibility for learning outcomes. Inquiry-based learning is similar to PBL in teaching methodology and includes presentation of case studies and the application of a problem-solving process that students use to identify relevant issues that require further research. However, rather than resolving the case through a diagnosis, IBL focuses on the inquiry process using issues that are relevant to the case (13, 14). As in PBL, students take ownership from the beginning, as in PBL and work in small, tutorial groups guided by a faculty member. The case is discussed and analyzed based on what information is known, further information needed, and the identification of learning issues that require further research. The cases provide a structure and format that guide students to

explore potential solutions to posed problems. Casebooks are now an accepted technique in preservice teacher training programs (15). As is indicated in PBL, the use of a case encourages group work that inevitably models collaborative communication skills found in the field. The paper case leads learners to apply skills learned to field projects (16). Students then conduct independent research and at a later session, present the results of their research that originated from the initial case study. Faculty members with the focus on critical analysis of relevant policy, program, advocacy, financial, cultural, facilitate summary and wrap-up discussion and community issues related to the case.

Family-focused learning. Family-focused learning (FFL) formed in the context of interdisciplinary education for health professionals to provide a model of direct involvement of family members in the teaching process (17). Family-focused learning follows the inquiry based approach through a series of sessions that begin with identification of issues around a particular family with an individual member with a disability, and close with student presentation of research issues related to the particular family that is participating in the teaching and learning process. The key difference in the FFL, compared to the previous models described, is that actual families and people with disabilities participate in the teaching process with faculty, interact with faculty and students throughout the development of case information to be presented and provide supportive critique to students in their work. Similar to PBL and IBL, the FFL model requires an initial session to present concerns and information that guide student inquiry. In contrast to the other two models, FFL involves actual family members who present the "family story" to students through video and written media. The development of the video is a joint venture for the family and participating faculty members that can require two or more sessions. When the family is satisfied with the video presentation, the tape is shared with students of several health, human services and education disciplines that identify key issues in a problem-solving process similar to the two models already described. Following completion of independent research, students prepare issue papers and present them to the family and/or individual for critique in a closing session. Family members and individuals with disabilities attend the closing session for the purpose of providing feedback to students on the scope of their work, relevance to their particular case, and quality in addressing the particular issue selected. As in the IBL closing session, faculty assist students in summarizing their analyses their individual research and relate students' findings to broad issues affecting families and persons with disabilities.

Person-focused learning. Person-focused learning (PFL) incorporates teaching and learning methods included in the previous models, but builds on elements found in each preceding approach. The elements of problem-solving and critical thinking that are hallmarks of PBL and IBL approaches are also essential to person-focused approaches. As in the FFL model, person-focused learning is designed and implemented with the participation of families and persons with disabilities. A new element is the service-learning aspect of PFL. In the PFL approach, students are required to complete a project that responds to needs and concerns identified by the family or individual (18). The involvement of persons with disabilities, families, faculty, and students in the development and implementation of the teaching experience produces a qualitative shift in teaching methodology and creates an action research model (4, 19-21). In the case-study approach, students respond to the issues presented for the primary purpose of advancing their own learning. In the person-focused model, students are placed in an interactive relationship with family members and individuals from the outset of the experience. The student learning goals, from the faculty perspective, involve: a) application of theoretical knowledge with real families and individuals with disabilities; and b) development of resources that respond to the needs expressed by families and individuals.

In the current study, the authors were concerned with the qualitative impacts of the PFL model on the people involved: students, families, and persons with disabilities. The unique features of the PFL model which incorporate problem solving in a real-life context and service to families and individuals require systematic evaluation. The assumption that direct involvement of actual family members and people with disabilities increases validity and thus applicability of the teaching process required empirical investigation and

IHE	CSU Chico, CA	UOP, Stockton, CA	Disability Studies University of Hawaii
Course	Speech Pathology: AAC	Special Education: Methods	Disability Studies: Team Work
Level	Upper Division & Graduate	Upper Division/ Graduate	Upper Division & Graduate
Dept.	Speech Pathology	Special Education	Interdisciplinary Disability Studies
Students	18 students	40 students	13 students

Table 1. Student participants in Person-Focused Learning at three universities.

consideration of the ethics involved. In this study, the authors sought to systematically evaluate the reciprocal impact of interactive teaching on student learning outcomes and people with disabilities, specifically with people with disabilities in direct interaction with students for the duration of semester-long courses.

The foci of investigation centered on three questions:

1. What are student perceptions of the PFL process, both in the process of interacting with families and individuals and in learning outcomes?
2. What are family member and individual perspectives of the PFL process, regarding their partnership role in teaching students and project outcomes?
3. What are ethical and logistical considerations for the replication of PFL in human service training programs, particularly related to disabilities?

Methods

The study was completed in the context of three interdisciplinary courses at three different university sites, with 71 students and 7 families including persons with disabilities. While course content differed across the three sites, teaching methods were similar. Teaching partnerships used principles of "Family Centered Care," in which family concerns drive professional interventions (22, 14, 23). Key steps in the teaching partnership included: (a) determination of family priorities; (b) adaptations to meet family and individual needs; (c) family input in project development; and (d) evaluation of completed projects by family members and persons with disabilities. Student learning outcomes were evaluated with qualitative surveys completed independently. Family and individual outcomes were identified through semi-structured interviews completed with the investigator.

The courses that provided the context for the study included a core special education course, an elective course in augmentative and alternative communication (AAC), and interdisciplinary teamwork course. Family members and individuals with disabilities participated as teaching partners with faculty members. Courses were located at California State University, Chico; the University of the Pacific in Stockton, California; and the University of Hawaii. Students who participated in the courses included three groups, shown in Table 1.

Characteristics of the seven individuals and families who participated in the study are listed below:

- Three adults, three children
- Communication disorders and physical disabilities in all subjects
- Two individuals with Asian/Pacific Islander ethnicity
- Five individuals were Caucasian

Course content and learning objectives differed across the three sites. However, key variables were held constant in teaching methodology. All courses included persons with disabilities and/or family members who participated in the design and implementation of the curriculum. The major requirement in each course included direct interaction with persons with disabilities and family members in the design and development of adaptive equipment or technology to meet needs identified by the individual and family.

Students engaged in a common process that included identification of needs by persons with disabilities and/or family members adapted from participatory action research (5, 1). Eight steps were completed in the person-focused learning teaching process. First, faculty developed

curriculum information about individuals in partnership with identified families and persons with disabilities. Second, students reviewed available information about the family and/or individual determine an initial developmental or environmental concerns identified by the family and/or individual. Third, student groups conducted brainstorming regarding potential family and individual concerns. Fourth, students prepared interviews based on guidelines provided by faculty. Fifth, students conducted interviews with individuals and/or family members. Sixth, the working group met to identify adaptation or support project based on results of prior information and interviews with individual and family members. Seventh, student groups presented completed projects to individuals and family members. Finally, student evaluations of the process and projects were completed.

The qualitative effectiveness of the person-focused learning process was evaluated by: (a) student perceptions of learning outcomes; and (b) perceptions of family members and persons with disabilities. Methods of evaluation included student's self reports and family/individual interviews.

Self-Report. Students were requested to complete qualitative comments in response to questions designed by the investigators. Questions addressed students' perceptions of the learning process and outcomes related to direct interaction with family members and persons with disabilities.

Family/Individual Feedback. Individuals with disabilities and family members were asked to evaluate their participation in the courses in a teaching/consultant role. Perceptions of these participants were also requested regarding the quality of student projects and interaction with family members and persons with disabilities. As the focus of teaching included adaptations and assistive technology, participants were requested to evaluate benefits and changes related to adaptations or resources developed by students.

Results and Discussion

Results of the study are discussed in relationship to perceptions of student learning outcomes and impacts on family members and persons with disabilities.

Student Problem-Solving. Student responses to qualitative questions were analyzed to determine recurring themes related to investigative and problem-based learning in direct interaction with people with disabilities and family members. Analysis of student surveys identified seven themes: (a) attitudinal change; (b) authentic engagement; (c) critical thinking; (d) sensitivity to families and individuals; (e) collaborative teamwork; (f) preparation for inclusion; and (g) self–efficacy/skills to adapt materials. Examples of student comments are included below related to each theme:

Attitudinal Change.

"There are many things that disabled students are able to do…most important to focus on those strengths." 18c

"I realized how many aspects of a person's life can be affected by a disability." 18c

"It made me realize how difficult it must be to have a child with a disability, or to be a child with a disability; everyday actions are so difficult!" 19c

"I find myself constantly looking at isles in stores, toys, elevators, etc. to see how they could possibly be adapted to better suit the needs of children with disabilities—more awareness." 7c

"I think it helped me look at adapting equipment as a fun responsibility instead of a required duty." 8c

"It has helped me to realize that children with disabilities have a vast amount of needs, and that each child's needs are unique. Adapted equipment may still need further adaptations to meet a specific child's needs." 10c

Authentic Engagement.

"The hands-on work helped me to develop a better understanding of a family's needs and wishes for their children. Though most of all…learning the true-to-life reality of the processes involved in working with a family." 12c

"Actually making the adaptations brings more involvement and thus more interest, which lead to more learning." 12c

"I think with the case study, it is each to maintain the same frame of reference and not to expand on ideas or think about new things. With the adapted equipment, new ideas or problems are presented and we brainstormed." 10c

Critical Thinking.

"This assignment makes you think about aspects of disabilities that normally one wouldn't consider." 2c

"We had discussed the written assignment a lot,

even before we knew what the questions were. We were always thinking, how it would help B." 6c

Sensitivity to Families and Individuals.
"Meeting in an informal setting allows both sides of the team to get to know each other with out the pressure of a meeting…with the family relaxed we can start building relationships." 16c
"Getting to know the family was an important milestone for us." 16c
"It has made me realize that the parents are very important in identifying the a child's needs." 16c
"I thought it was very useful to ask T. [the parent] our questions because we had to know exactly what her situation was so the outcome would be helpful." 5c

Collaborative Teamwork.
"Yes, because we need each other's specialized skills along with knowledge and creativity." 14c
"It was a great idea to work in a group because everyone has different ideas which we can bring together. Then everyone has different talents which were utilized in the production process." 12c

Preparation for Inclusion.
"This is something I will have to do in my classroom so I appreciate the preparation." 2c
"To find different ways to teach someone the ABCs and how slow the song needs to be so that the child can learn." 9c
"It has made me realize that each child with a disability is an individual; helping each child can be done only if that child is looked at as an individual." 15c

Self-Efficacy and Adaptive skills.
"The most important part of this assignment was that it opened a door for me and pretty much told me that I had the potential to help any child with a disability." 3c
"I learned that I take my skills and abilities for granted. From meeting B., I realized that many aspects of daily living would be difficult for her, and in order for them to function at her level, more things would need to be adapted." 10c
"Yes, because it provides hands on time that I will remember more than any case study. It is also more fun than any case study." 9c
"I liked the developmental framework and the way this was all set up. It was very realistic to what we deal with in our real jobs and it was very hands on." 20c
"It makes me become more aware of the types of things; a lot of things that I would have never thought of." 13c

Family and individual interviews revealed four themes: (a) interaction with students; (b) self-validation; (c) support networks; and (d) alternatives to meet individual needs. Families and individuals commented that they would participate again. Table 2, below demonstrates representative feedback provided by family members and person with disabilities.

Ethical issues identified included the need to (a) respect individual choice in participation; (b) confidentiality; (c) honor individual priorities and (d) respect family differences. Comments provided by families and individuals at the completion of each class indicated the possibility

Theme Identified	**Family Comments**
Interaction with students	"Having students come to our home was a highlight of the week for J., he looked forward to it all week."
	"Students gave S. attention and made us appreciate his importance."
	"I am getting braver to ask for what my son needs."
Self-validation	"I always knew that J. knows more and the students helped to document that."
Support networks	"It is wonderful for our whole family to participate with the students…going to the beach together was a first for us."
	"All of the time and support has given S. a chance to get out more."
Alternatives to meet needs	"The help provided by the class gave S. a way to communicate that he did not have before."
	"We want S. to learn with the other kids and he shows the book to every one who comes over."

Table 2. Qualitative themes and family comments regarding Person-Focused Learning Outcomes.

of initial reluctance to participate. One parent commented that she initially was nervous when meeting the students for the first time, particularly due to cultural differences between them. However, this parent later reported that her feelings changed later after realizing how much attention and support the students demonstrated toward her son. This mother's comment highlights the need to honor individual family priorities that may be based on cultural styles, educational background, language differences, and other variables. Related to this is the need to respect and understand family differences and follow the lead of the individual or family to determine the most appropriate time and place to conduct interviews and project activities.

The results revealed positive qualitative student learning outcomes. People with disabilities and family members reported that their participation provided important benefits that included perceptions of increased self-efficacy and competence when interacting with students. Risks were not specifically identified by families or persons with disabilities, but inferred from their feedback. The responsibility to consider risk, which may include risks to privacy of participants, remains with the researcher who embarks on teaching partnerships with families and persons with disabilities. Comments provided by students in all thematic areas reported revealed increased awareness and respect for the life experiences of persons with disabilities and family members, thus establishing a foundation for ethical behavior in future professional roles with persons with disabilities, including teaching, service, and research.

Summary

The results of the present study support the effectiveness of interactive teaching, specifically Person-Focused Learning, to promote student learning outcomes that demonstrate respectful and responsible professional attitudes and behavior with persons with disabilities and family members. The specific student learning outcomes were found in both cognitive and affective domains, as seen in students' evaluations of the learning experience. These findings have implications for preservice training of health, human service, and education professionals to establish a foundation for ethical behavior with human subjects in the career contexts of service and research.

The qualitative evaluation results of student learning outcomes indicate that involvement of persons with disabilities in the teaching process provides authentic learning that cannot be replicated with more traditional didactic methods. Further, involving family members in the teaching and evaluation process at all levels follows a participatory action research process and allows "checkpoints" for subjects to be fully cognizant of the research agenda and purposes. Thirdly, including people with disabilities in the research/teaching process strengthens validity as recommended by Kvale and Burns (2, 3). Further, reciprocity in the learning setting is achieved where students learn the needs of families and the value their knowledge when designing materials and technologies to assist them in the learning environment. The research participants are valued by the researchers and the students involved in the assignment and the student-made products are valued by the families.

The demonstration of a pre-service training approach that teaches reciprocal relationships with subjects is perhaps the key finding with implications for training future professionals in the area of responsible conduct of research. Not only did students demonstrate qualitative evidence of critical thinking in the learning process, the direct interaction with subjects in the action research model employed in Person-Focused Learning showed an effect on the students' sensitivity toward persons with disabilities and family members. The demonstrated effect on students' sensitivity with subjects could effect future professional ethics and conduct. While, further study is needed to determine attitudes and values that are directly related to the responsible conduct of research with human subjects, student attitudes toward subjects are considered a critical variable of ethical behavior. The question of what particular teaching model effectively trains professionals who are prepared to implement responsible conduct of research was only partially addressed by the present study. The attitudes and skills required for responsible conduct of research are clearly a constellation of knowledge and ethics that require further explication.

This qualitative study explored person-focused learning principles in several preservice courses and revealed positive findings for students and the families who shared their

stories. The "realness" of the learning setting allowed researchers to identify multiple learning outcomes and ethical issues when involving people with disabilities in a teaching setting and research endeavor. Bowen identified the need to strengthen internal validity through the integration of qualitative and quantitative research methodology (24). Further research in PFL is needed to a) specify affective and cognitive student learning outcomes; b) quantify changes in student attitudes; b) compare PFL teaching to other problem-solving approaches; c) identify long range impacts on student learning; d) develop guidelines for replication; and e) explore the use of PFL to teach responsible conduct of research. The philosophical attitude and the research model in the present study provide a framework for preservice education and further research to determine specific professional attributes that lead to affective, cognitive, and ethical foundations for the responsible conduct of research, particularly with persons with disabilities.

Bibliography

1. Guerrero, JL. Research paradigm shift: Participatory action research. Working Paper No. June 1995. Rehabilitation Research and Training Center of the Pacific. San Diego State University, Interwork Institute.
2. Kvale, S. The social construction of validity. Qualitative Inquiry 1995; 1 (1): 19-40.
3. Burns, PA. Securing internal validity: How to avoid the threats [serial online] 1997 [cited 2000 Nov] Available from: URL:http://trochim.human.cornell.edu/tutorial/burns/int.html
4. Newman, JM. Action research: A brief overview [serial online] 2000 Jan [cited 2000 Nov] Forum Qualitative Sozialforschung / Forum: Qualitative Social Research [On-line Journal]. Available from: URL: http://qualitative-research.net/fqs.
5. Whyte, WF. Participatory action research. Newbury Park (CA): Sage Publications; 1991.
6. Casey, MB, & Hwoson, P. Educating preservice students based on problem-centered approach to teaching. Journal of Teacher Education 1993; 44: 361-369.
7. Chabon, S. & Wilkerson, DL. Discussions on diversity: A problem-based instructional model to facilitate cultural competence. Poster presented at American Speech, Language and Hearing Association Annual Conference, Washington, D.C., November 16, 2000.
8. Branda, LA. Implementing problem-based learning. Journal of Dental Education 1990; 54: 548-549.
9. Bridges, EM. & Hallinger, P. Problem-based learning in leadership development. 1995; Portland: University of Oregon. ERIC Clearinghouse on Educational Management.
10. Gist, GL. Problem-based learning: A new tool for environmental health education. Journal of Environmental Health 1992; 54: 8-13.
11. Royeen, C.B. A problem-based learning curriculum for occupational therapy education. Journal of Occupational Therapy 1995; 49: 338-346.
12. Cockrell, K.S.; Hughes Caplow, J.A.; Donaldson, J.F. A context for learning: Collaborative groups in the problem-based learning environment. The Review of Higher Education 2000; 23: 347-363.
13. Geln, S. Towards a new model of nursing education. Nurse Education Today 1995; 15: 90-95.
14. Matteoli, RF.; Verhey, MP. & Warda, MR. Collaboration between families and university faculty: A partnership in education. Journal of Psychosocial Nursing 1994; 32: 17-20.
15. Shulman, J., Lotan, RA., & Whitcomb, JA. Groupwork in diverse classrooms: A casebook for educators. New York: Teachers College Press; 1998.
16. Boyle, JR. & Danforth, S. Cases in Special Education. Boston: McGraw-Hill. 2001.
17. Ratliffe, K., Stodden, N., Guinan, M., & Robinson, N.R. Family partnerships in teaching: family-focused learning. Journal of intellectual disability research 2000; 44 (3 & 4): 416.
18. Stevens, L. Academic Service Learning within an AAC Curriculum. 2000 Poster presented at American Speech, Language and Hearing Association Annual Conference, Washington, D.C., November 16, 2000.
19. Henry, SK., Scott, JA., Skobel, B., Jones, A., Cross, S., Butler, C., & Blackstone, T. Linking university and teacher communities: A Think Tank Model of professional development. Teacher Education and Special Education 1999; 22: 251-268.
20. Hutchinson, NL. & Martin, AK. Fostering inclusive beliefs and practices during preservice teacher education through communities of practice. Teacher Education and Special Education 1999; 22: 234-250.
21. Perry, NE., Walton, C. & Calder, K. Teachers developing assessments of early literacy: A community of practice project. Teacher Education and Special Education 1999; 22: 218-233.
22. Harry, B., Rueda, R., Kalyanpur, M. Cultural reciprocity in sociocultural perspective: Adapting the normalization principle for family collaboration. Exceptional Children 1999; 66:123-136.
23. Gartner, A., Lipsky, DK., Turnbull, A. Supporting families with a child with a disability. Baltimore: Paul H. Brookes Publishing; 1991.
24. Bowen, K. The Sin of Omission -Punishable by Death to Internal Validity: An Argument for Integration of Qualitative and Quantitative Research Methods to Strengthen Internal Validity[serial online] 1996 [cited 2000 Nov]. Available from: URL: http://trochim.human.cornell.edu/gallery/bowen/hss691.htm.

4. Conflict of Interest

What is Driving Policies on Faculty Conflict of Interest? Considerations for Policy Development

Mildred K. Cho, Center for Biomedical Ethics, Stanford University, USA

Ryo Shohara, Institute for Health Policy Studies, University of California, San Francisco, USA

Drummond Rennie, Institute for Health Policy Studies, University of California, San Francisco, USA

Keywords: Academic policies, Academic-industry ties, Conflict of interest, Research integrity

There are several factors driving policies on conflict of interest of faculty at academic research institutions in the United States today. The first is that researchers and institutions have a greater number, and a wider variety of financial conflicts of interest, especially in the area of biomedical research. Sometimes, these financial interests appear to lead to very bad outcomes, and when that happens, public scrutiny of the financial interests increases. Sometimes, this leads to new policy.

What is the current state of academic-industry ties in biomedical research? In 2000, the NIH's budget is $17.8 billion (1), while the pharmaceutical industry's R&D budget is $22.4 million (2). Krimsky found that 34% of research articles published in the top 14 biomedical research journals in 1992 had undisclosed financial ties of a lead author. These ties included holding a patent on an invention related to the published research, or being on an advisory board or a major shareholder in a company whose activites were related to the published research (3). In a review of FDA records, USA Today reported that 54% of the time, experts hired by the FDA to advise on safety and effectiveness of drugs have a direct financial interest in the drug or topic they are asked to evaluate (4). Therefore, academic-industry ties are now the norm, rather than the exception.

Academic-industry ties have been the apparent cause of bad outcomes, including censorship of data (5, 6), publication bias (7-10), lower quality of research (11), and harm to research subjects, including death (12). Although it is impossible to determine a causal link between financial interest and adverse outcome in individual situations, systematically gathered evidence suggests that, in the aggregate, academic-industry ties can have adverse effects on the scientific process and outcome in the aggregate (13).

One bad outcome in particular has led recently to public scrutiny and re-examination of policies on conflicts of interest — the death of Jesse Gelsinger, who was a research subject in a Phase I clinical trial of gene transfer at the University of Pennsylvania (12). Much attention focused on the financial ties of investigators and the investigators' institution with a company that was, in part, sponsoring the trial. Although, again, it is impossible to prove that there was a causal link between the financial ties and the death of Mr. Gelsinger, it was a link that was inevitably made, time and

Corresponding author: Mildred K. Cho, PhD, Stanford Center for Biomedical Ethics, 701A Welch Road, Suite 1105, Palo Alto, CA 94304, 650-725-7993 (voice), 650-725-6131 (fax), micho@stanford.edu.

again. A quote from a recent newspaper article sums up the public perception:

> Paul Gelsinger, Jesse's father, said yesterday he had undergone a painful change of heart in the year after his son's death, at first fully trusting the researchers and holding them blameless and then gradually, as disclosures of apparent wrongdoing emerged, concluding that he had been duped by scientists who cared more about profits than safety. (14)

After Mr. Gelsinger's death, the National Institutes of Health (NIH) held a public meeting this year to re-examine some aspects of conflict of interest policy, and several professional organizations, including the National Academy of Sciences, the American Association of Medical Colleges (AAMC), and Association of Academic Health Centers (AHC), the American Association of Universities (AAU), and the American Association of University Professors have all assembled internal groups to do the same.

What are the current policies on faculty conflict of interest?

Current policies on faculty conflict of interest exist at several levels, including federal, state, institutional regulations, editorial policies at research journals, and statements by professional societies. All are limited, however, in different ways. The most widespread federal rules include the "Objectivity in Research" regulations (15). These are applicable only to researchers who apply for research funding from the National Science Foundation and the Public Health Service (PHS), which includes the NIH. These regulations are limited to disclosure of financial ties that could be construed to affect the publicly-funded research, and to financial ties that exceed $10,000 annually or 5% equity interest. Thus, financial ties in the context of industry-funded research, where more serious conflicts of interest might be found, are not covered under these regulations.

In addition to federal regulations, there are state laws that might apply to faculty at public institutions. For example, some states prohibit or require full disclosure of gifts to public employees, which include faculty of state universities. These state laws often do not apply to private universities, and are not uniform from state to state.

Institutional policies are mandated by the federal regulations, which require that institutions whose faculty apply for PHS or NSF funding develop and implement their own written rules for faculty conflicts of interest. These institutional policies must conform to, but need not be limited to, federal regulations. Indeed, the majority of institutional policies go beyond federal regulations in scope and management of conflicts of interest, but most do not state specific limits on financial interests, even when in conjunction with company-sponsored research (16). Most of these policies imply or state that conflicts of interest are dealt with on a case-by-case basis, and seem to rely heavily on disclosure as a primary mechanism for dealing with conflict of interest.

Some research journals have developed policies that require disclosure of authors' financial interests to editors and reviewers . However, such disclosures often do not surface on the pages of the published articles, so their effects are limited (Krimsky, this volume).

The AAMC, AHC, and the AAU created guidelines for faculty conflict of interest long ago (17-19), and although they thoughtfully outline policy considerations, they are not specific and are not enforced. Finally, in the wake of Jesse Gelsinger's death, two professional societies (the American Society of Gene Therapy and the American Society of Human Genetics) have put forward statements that faculty having financial interests in companies sponsoring their gene transfer research is inappropriate and should be avoided (20, 21). These statements only apply to gene transfer research, however, and also have no enforcement power.

What should we do about conflicts of interest?

The answer to the question, "what do we do about conflicts of interest?" depends upon the answers to the questions, "what is conflict of interest?", "what is the primary interest of academic institutions and the government?", and "what are the secondary interests we are concerned about?"

What is conflict of interest? Opinions are diverse. Many make the distinction between "actual" and "potential" conflicts of interest. Others call it scientific misconduct (22). Depending on how one defines conflict of interest, one may be led to base policy on evidence of bad outcomes or on ethical or professional values. We define conflict of

interest as the co-existence of a primary interest or duty (such as research integrity, patient welfare, or education) and a secondary interest (such as financial gain or recognition) (23). The policy concern is that the secondary interest exerts undue influence on the judgements made in the course of executing the primary interest, leading to adverse outcomes (such as research bias or adverse effects on research subjects).

It is important to remember that conflict of interest rules are activated in the absence of a "crime" (24). Stark likens them to speed limit laws. In contrast to laws against murder, which are aimed at activities that, in themselves, are deemed immoral and are not in the public interest, speed limit laws are aimed against conditions that predispose to the activities that are not in the public interest. So, while driving at 70 miles per hour may not in itself be wrong in the way that murder is wrong, high-speed driving may enhance the chances of causing harm to others. Some drivers might be quite capable of avoiding crashes at even 200 miles per hour, but because it would be difficult and impractical to determine who they are and whether they are so capable under all circumstances, the laws are aimed at preventing the situation rather than particular outcomes. However, there may be certain speeds that would be considered "reckless" in almost any circumstances, and thus immoral – and there may be analogous financial interests.

However, there is an important difference between speed limit laws and conflict of interest regulations, in that speed limit laws apply to everyone, whereas conflict of interest laws apply to groups that have a fiduciary relationship to the public, such as public officials or professionals. This distinction is important, because it means that there are reasons to set the rules by criteria other than probability of harm to the public, namely in order to earn or preserve the right to occupy the special position in society (25).

This definition of conflict of interest implies that there can be no distinction made between "actual" and "potential". The conflicting interests simply either exist or they do not. They are, in themselves, not scientific misconduct, although they may lead to misconduct. The current definition of scientific misconduct carries with it the notion of wrongdoing with intent (26), which is based on the proven existence of a bad outcome, and is therefore incompatible with a definition of conflict of interest that is based on the characteristics of a situation rather than the outcome.

What is the primary interest? Lack of clarity about the primary interests of researchers and their institutions will lead to bad policy, because one of the points of having the policies is to protect the primary interests. So, the question is, what are the roles of academic institutions and the government in the conduct of science? The passage of the Bayh-Dole Act gave the government a new role in academic research, namely, "to promote the marketing of inventions developed under federally supported research and development projects by nonprofit organizations and small business firms." (27)

Government specifically encouraged academic institutions to be involved in the marketing of inventions. Universities have taken this encouragement to heart, "… shifting from ivory tower to revving economic engine." (28) The new role of universities as economic engines leads to expectations that they create jobs and even whole industries. In fact, the government has implicitly adopted the values of the business world, where money is an incentive for employees to work in the interests of shareholders. In this model, the secondary (financial) interest is considered to be in alignment with the primary interest, rather than acting as a competing interest. By contrast, the model of professionalism says that the Bayh-Dole Act and related legislation specifically put not only faculty but institutions in a position of conflict of interest. If academic institutions and their faculty are expected to add economic goals to their primary missions, can those institutions be expected to be effective at developing and enforcing conflict of interest rules for their faculty? This seems to be a dangerous thing to ask.

We must be clear about whether academic institutions should take on economic health as a primary interest. We must also be clear about whether we are concerned only with or more concerned about certain kinds of primary interests. For example, is only federally-funded research of concern, or all research? That is, should policies be directed only at interests that conflict with government-funded research, or should they also be directed at interests that conflict with industry-funded activities, too? Finally, we should also ask whether clinical research is of more concern than other research. There are good ethical reasons to distinguish

research that involves human subjects from other research, primarily that human subjects are subjected directly to risks from the research itself.

What is the secondary interest? Lack of clarity about the secondary interests that are of concern will also lead to bad policy. Current regulations focus on financial interests, rather than other, less-tangible interests such as academic recognition and fame, or personal ties. This is appropriate for the time being, not because the intangibles are less damaging, but because the financial interests are avoidable and because avoiding them is consistent with the role of a professional, and enhances public trust. Financial interests have also increased to a high level and are deserving of attention merely because of their frequency. Furthermore, those who point to the unfairness of concern about financial interests seem to imply that financial interests merely replace the non-financial interests, so that there is no need for special consideration of the financial interests. However, the literature suggests that the effect of financial interests on biomedical research can be detected as an independent factor, above the background "noise" of the want for academic recognition and fame (assuming that it exists uniformly among researchers).

There is less clarity about what specific kinds of financial ties are of concern. Current regulations focus on personal financial ties such as consulting fees, honoraria, royalties and equity holdings. They generally do not consider company-sponsored research per se to be a conflict of interest, but a growing body of literature suggests that industry sponsorship in itself biases research and publication (7-9, 13, 29).

How do we manage conflicts of interest?

Standard methods of managing, or mitigating, conflicts of interest include (1) disclosure (e.g., publication of a secondary interest), (2) mediation (e.g., a blind trust, which puts a secondary interest under the control of a third party, or oversight, which puts a primary interest under the review or control of a third party), (3) abstention (e.g., recusal from a primary interest), (4) divesti-ture (e.g., removal of a secondary interest), and (5) prohibition (e.g., permanent withdrawal from a whole category of secondary interests) (23). At first glance, these five methods seem to be organized smoothly along a continuum of stringency. However, closer examination reveals that there is actually a qualitative difference between these strategies, because they are based on different assumptions.

In theory, all of these methods act by modifying the conflict of interest situation through either the primary or secondary interest. However, disclosure is distinct from all the other methods. It is supposed to act not by virtue of supplying information to the disclosee, but because the release of this information is supposed to make the discloser more aware of the potential effects and thus affect the discloser's behavior (24). Clearly this is a weak method because of its indirectness. In practice, the information rarely gets out to a wide audience, and the discloser knows it, limiting effectiveness. More importantly, this method allows the discloser to feel that the act of disclosing has let him or her off the hook, and places the burden of management on the disclosee. Stark points out that disclosure is based on a model where the role of the discloser is as an "agent", or delegate, rather than a trustee. By this model, the disclosee is assumed to have a large degree of control over the activities of the discloser.

In contrast, the other management methods are based on a trustee or fiduciary model. By this model, the disclosee is assumed to have little control over the activities of the discloser and therefore depends on the discloser to act in the best interests of the disclosee. Mediation and abstention carry with them the notion that the fiduciary position is a role that can be filled by interchangeable individuals. That is, the protagonist can be replaced by a third party such as an oversight committee or another researcher. Divestiture and prohibition imply that the protagonist is not replaceable, and therefore the mitigation of the conflict of interests requires removal of the secondary interest.

How we deal with conflicts of interest depends on how we view the players. Are researchers delegates or trustees? People who hold elected public office may better fit the delegate or agency model, since the public has the power to remove them from office if their performance is unsatisfactory. Researchers, however, are more like trustees (especially clinical researchers) because it is understood that the public supports their training and activities to perform tasks that others are not qualified to

perform, and the public is not in a strong position of control over these activities. The professional role of scientists and clinicians is fiduciary in nature, and requires that public interests be placed ahead of self-interest.

How we deal with conflicts of interest also depends on how broadly we define the interests and the conflicts. The goal of academic-industry ties is to maintain the ability to conduct good science and to enhance technology transfer for public good, while preserving research integrity (including the direction of research) and, in the case of clinical research, protecting human subjects from harm. In order to achieve any of these goals, it is essential to maintain the public trust and a sense of professionalism, in the original sense of the word (25, 30, 31), which includes strong self-regulation (32).

Recommendations for policy development

What are the implications of these definitions of interests and conflicts of interest for policy development? First, conflicts of interest should be defined by characteristics of situations, rather than by outcomes. This allows taking into account professional values as well as evidence that certain situations tend to lead to bad outcomes. Second, we should not rely on disclosure as a primary mechanism for mitigating conflicts of interest. Instead, we should acknowledge that researchers have professional responsibilities that are fiduciary in nature. As trustees, they should be trustworthy. Third, institutions should remember that institutional interests play a role in individual conflicts of interest, as well as the administration of policies about individual conflicts of interest. Therefore, institutions should not use policies only as administrative tools, but also as mechanisms for communicating institutional values to the public (24, 31), because the nature of professionalism is to profess a vow to place the interests of the public above self-interest (33). The goal is to provide reassurance to the public that the institutions have also accepted their fiduciary role.

Bibliography

1. National Institutes of Health, About NIH: Overview, . *2000.*
2. Pharmaceutical Research and Manufacturers of America, Pharmaceutical Industry Profile 2000, . 2000.
3. Krimsky, S., et al., Financial interests of authors in scientific journals: a pilot study of 14 publications. *Science and Engineering Ethics,* 1996. 2: p. 395-410.
4. Cauchon, D., FDA advisers tied to industry., *in* USA Today. 2000. p. 1A.
5. Kahn, J., et al., Evaluation of HIV-1 Immunogen, an immunologic modifier, administered to patients infected with HIV having 300-549 x 10^6/L CD4 cell counts: a randomized controlled trial. *JAMA,* 2000. 284: p. 2193-2202.
6. Rennie, D., Thyroid storm. *JAMA,* 1997. 277: p. 1238-1243.
7. Cho, M. and L. Bero, The quality of drug studies published in symposium proceedings. *Annals of Internal Medicine,* 1996. 124: p. 485-489.
8. Davidson, R., Source of funding and outcome of clinical trials. *J Gen Intern Med,* 1986. 1: p. 155-158.
9. Rochon, P., *et al.,* A study of manufacturer-supported trials of nonsteroidal anti-inflammatory drugs in the treatment of arthritis. *Arch Intern Med,* 1994. 154: p. 157-163.
10. Stelfox, H., et al., Conflict of interest in the debate over calcium-channel antagonists. *New Engl J Med,* 1998. 338: p. 101-6.
11. Rochon, P., *et al.,* Evaluating the quality of articles published in journal supplements compared with the quality of those published in the parent journal. *JAMA,* 1994. 272: p. 108-113.
12. Weiss, R. and D. Nelson, Teen dies undergoing experimental gene therapy., *in Washington Post*. 1999: Washington, DC. p. A1.
13. Blumenthal, D., *et al.,* Withholding research results in academic life science. *JAMA,* 1997. 277: p. 1224-1228.
14. Weiss, R. and D. Nelson, Penn settles gene therapy suit., in *Washington Post.* 2000: Washington, DC. p. A4.
15. U.S. Department of Health and Human Services, Objectivity in Research, 45 CFR 94, 1995, U.S. Government: Federal Register. p. 35810.
16. Cho, M., *et al.,* Policies on faculty conflicts of interest at US Universities. *JAMA,* 2000. 284: p. 2203-2208.
17. AAMC Ad Hoc Committee on Misconduct and Conflict of Interest in Research, Guidelines for faculty involvement in commercially supported continuing medical education. *Academic Medicine,* 1992. 67: p. 617-621.
18. AHC Task Force on Science Policy, Conflicts of interest in academic health centers policy paper #1, 1990, Association of Academic Health Centers: Washington, DC.
19. AAU Clearinghouse on University-Industry Relations, University policies on conflict of interest and delay of publication, 1984, American Association of Univerisities: Washington, DC.
20. The American Society of Gene Therapy, Policy on Financial Conflict of Interest in Clinical Research. 2000.
21. The American Society of Human Genetics, Statement on Gene Therapy. 2000.
22. Bodenheimer, T., Conflict of interest in clinical drug trials: A risk factor for scientific misconduct. 2000: Bethesda, MD.
23. Thompson, D., Understanding financial conflicts of

interest. *New Engl J Med, 1993. 329: p. 573-6.*
24. Stark, A., Conflict of interest in American public life. 2000, Cambridge, MA: Harvard University Press. 271.
25. Buchanan, A., Is there a medical profession in the house?, in *Conflicts of interest in clinical practice and research*, R. Spece, D. Shimm, and A. Buchanan, Editors. 1996, Oxford University Press: New York. p. 105-136.
26. Office of Science and Technology Policy, Fact Sheet: Research misconduct - A new definition and new procedures for federal research agencies. October 14, 1999.
27. US Senate, University and Small Business Patent Procedures Act, in Bill S. 414. 1979.
28. Goldberg, C., Across the US, universities are fueling high-tech booms., in *New York Times*. 1999: New York. p. A1.
29. Blumenthal, D., et al., Participation of life-science faculty in research relationships with industry. *New Engl J Med*, 1996. 335: p. 1734-39.
30. DeAngelis, C., Conflict of interest and the public trust. *JAMA*, 2000. 284: p. 2237-2238.
31. Korn, D., Conflicts of interest in biomedical research. *JAMA*, 2000. 284: p. 2234-2237.
32. Rothman, D., Medical Professionalism — Focusing on the Real Issues. *New Engl J Med*, 2000. 342: p. 1281-6.
33. Wynia, M., Medical professionalism in society. *New Engl J Med*, 1999. 342: p. 1612-1616.

The Commercialization of Academic Science: Conflict of Interest Policies and the Faculty Consultant

Lisa M. Jones, Postsecondary Education Policy Studies Center, University of Minnesota, USA

Karen Seashore Louis, Educational Policy and Administration, University of Minnesota, USA

Keywords: Academic scientists, Conflict of interest, Faculty consultants

Social scientists have studied the effects of faculty consulting on academic productivity - teaching, research, and service (1- 6) – and used productivity as a proxy for conflict of interest. Most recently, writers in both the disciplinary and popular literature have addressed conflict of interest and faculty consultants. However, little empirical research that investigates the connection between entrepreneurial behavior, consulting, and conflict of interest, exists. This study identifies four specific behaviors that could compromise scientific objectivity and thus, be classified as conflicts of interest: research agenda bias, prior review, withholding, and secrecy.

These conflict of interest behaviors are grounded in the norms and counternorms of science proposed by Merton and Mitroff (7-8). Four norms dominate the roles of scientific researchers: universalism, dissemination, disinterestedness, and organized skepticism.

Universalism suggests that science is open to all individuals regardless of their personal traits. The scientific method is used to pursue truth. Dissemination allows for research to become open to all challenges, subject to verification, and widely disseminated, the antithesis of prior review. Research advances knowledge and resides in the public domain. Results become communicated so that others may build upon previous work to move knowledge forward. The purpose of communication also allows for research to become open to all challenges, subject to verification, and widely disseminated (9).

The disinterested search for truth enables scientists to explore all information regardless of where it might lead. Science's reliance on verification and reliability reflect institutionalized controls to ensure that knowledge benefits humanity and allows the researchers to proceed objectively. Although knowledge advancement is the institutionalized role of scientists, some desire credit for their discoveries vis-à-vis election to the National Academy of Sciences or a trip to Stockholm (e.g., Nobel Prize). Conflicts then arise over the priority of discovery that further fuels secrecy. Furthermore, academic science is a competitive industry that encourages researchers to withhold results for personal aggrandizement either through enhanced reputation or financial gain. Entrepreneurial behavior is a perceived threat to the researchers' disinterestedness in the pursuit of knowledge for its own sake. Burton Clark views entrepreneurialism as "a characteristic of social systems...taking risks

Corresponding author:Lisa M. Jones, Ph.D., Postdoctoral Associate, Postsecondary Education Policy Studies Center, University of Minnesota, 330 Wulling Hall, 86 Pillsbury Drive S.E., Minneapolis, MN 55455-0208, 612-625-0773 (voice), 612-624-3377 (fax), jones049@tc.umn.edu.

when initiating new business practices where the outcome is in doubt...(10)" The scientist maintains a vested interest in the research outcomes. When individual scientists establish research agendas based on profitability, science is not served. The payoff between basic research discoveries and economic profitability often requires time that neither society nor the marketplace are willing to grant academics. This creates the appearance that basic research projects compete with commercially viable proposals for funds.

Finally, Merton described organized skepticism as the "temporary suspension of judgment and the detached scrutiny of beliefs" that affords scientists with the opportunity to examine results using empirical or logical criteria (11).

The search for truth rests upon the foundations of basic research. When academic scientists pursue lines of inquiry regardless of their commercial viability, the public interest is served. However, shifting political forces place equal or even greater importance on commercially viable academic science that could stimulate economic growth expeditiously (12).

This study examines life sciences faculty who report earning additional income by consulting for non-profit organizations, industry, and government and their engagement in actual conflict of interest behaviors. This study limits the definition to consulting activities for financial remuneration, and examines individuals who select consulting as a major source of supplemental income from nonprofit organizations or government agencies, private enterprise, or both public and private. Furthermore, the study examines behaviors of those who consult exclusively with one company.

Methods

The data source used for this study is part of the Academic-Industry Research Relationships Study in Genetics and Other Life Sciences. The analyses here are based on data from the broader study's 1994-1995 national survey of 3,169 U.S. faculty in the life sciences. Fifty research-intensive institutions were selected based on the levels of National Institutes of Health funding for 1993. All medical-school departments and other academic life-science departments and graduate programs were identified using the 1994 *Peterson's Guide to Graduate Programs in Biological and Agricultural Sciences.* One medical clinical department, one non-medical clinical department, and two non-clinical departments were randomly selected from each institution. Both the Peterson's Guide and University Bulletins identified 4,000 faculty that included non-clinical, clinical, and researchers funded by the Human Genome Project (HGP). A stratified random sample of faculty, half of whom were clinical and half of whom were non-clinical, were selected from a list of faculty across the 200 departments. Special provisions were made to include the HGP researchers because of the broader study's interest in behaviors of genetics researchers. Ineligible faculty (those who were deceased, retired, or not located) were omitted from the sample, leaving a final sample size of 3,169 faculty.

Data Collection

The data collection process occurred from October 1994 through April 1995 by the Center for Survey Research at the University of Massachusetts. Each participant was mailed a survey packet, which included a cover letter, coded postcard, and questionnaire. The questionnaire and postcard were to be returned separately to protect respondent anonymity. Reminder/thank you postcards were mailed shortly after the initial mailing. Follow-up calls conducted from late November to mid-February to non-respondents generated an additional 190 cases for analysis. We received useable responses from 2,052 faculty, for a total response rate of 65 percent.

For this substudy, the sample consists of the 1,032 non-clinical faculty respondents. Selection of the individuals was assured by including only faculty who do not conduct clinical trials on " drugs, devices, or diagnostic or therapeutic technologies." The non-clinical faculty was chosen because previous research conducted using the complete sample shows that these individuals are on the "front end" (entrepreneurial) of the commercialization process. Furthermore, the industry relationships between clinical faculty and corporations are structured around clinical trials rather than new discoveries (12).

Variables

Faculty gender, academic rank, average annual research budget, average level of entrepreneurial behavior, and average income earned above salary were used as independent variables in the

statistical analyses. The entrepreneurial behavior scale constructed consists of the following survey items: "Has the research that you do at your university resulted in....*(Check one for each item)*...patents applied for, a startup company." Individuals could check either yes (coded as "1") or no (coded "0"). The next question used for this scale was: "For the firm with which you currently have the greatest involvement, which of the roles listed below do you have? *(Check all that apply)*...equity holder, company owns or licenses a patent based on your research." If the respondent left the item blank, it was coded as "0" for no. A check mark was coded as "1" for yes. The reliability for the entrepreneurial behavior scale offered a standardized alpha of .69 (n = 1032).

Conflict of Interest measures

Research agenda bias. One conflict of interest measure concerns external influence on research topics: "To what extent has your selection of research topics been affected by...*(Check one for each item)* a) the likelihood of commercial application of the results." Participants were offered the following response options: Not at all (coded as "0"); very little (coded as "1"); to some extent (coded as "2"); or, to a great extent (coded as "3"). The results were collapsed into a dichotomous variable coded "1" for yes and "0" for no.

Prior review. Another conflict of interest measure considers the publication relationship between faculty and the sponsor. The following item measured prior review: "Have you personally conducted any research at your university, the results of which are the property of the sponsor and cannot be published without the sponsor's review or consent?" Yes was coded as "1" and no as "0".

Secrecy. This variable identifies the relationship between commercial science and publication of results. "Has your university research resulted in findings that were *never* published for proprietary reasons?" was the item used to measure secrecy. Yes was coded as "1" and no as "0".

Withholding. The final conflict of interest measure for this study considers the sharing relationships between academic researchers. This item asks individuals to report their denial of others' requests for research tools: "In the last 3 years, have any other university scientists requested any results or materials that you did not provide?" Yes was coded as "1" and no as "0".

Statistical analysis

Unless otherwise noted, statistical significance and the direction of reported relationships between consulting and conflict of interest behaviors were tested by multivariate linear and logistic regressions. The equations were adjusted for academic rank, gender, institutional control (public or private), academic program ranking, institutional location (metropolitan versus non-metropolitan), supplemental income amount, and levels of entrepreneurial behavior.

Results

Sixty percent (n =616) of this sample (n =1032) report that they have consulted with either public (35.2%) or private (24.5%) enterprises at least once. This contrasts with the 26% of the respondents who consult with either group as a major source of supplemental income. Table 1 shows the consultants' characteristics broken down by gender, academic rank, average research budget, average level of entrepreneurial behavior, and average income earned above salaries. Males account for 82% of the sample, thus it is not surprising to see them represented more than females in the consulting categories (x^2 = 24.74 p < .001). Full professors represent 54% of the total sample and are also consult more than assistant and associate professors (x^2 = 16.88 p < .05). However, the assistant professors that consult work more with private enterprise than the public sector. One possible explanation for this finding is that assistant professors may have established relationships with companies during their graduate training. The results further indicate that those who consult exclusively with one company tend to be male, full professors. Furthermore, private enterprise consulting faculty have larger research budgets than non-consultants, which supports a Louis et al. (13) earlier study that suggested that research budget reflects entrepreneurial behavior as it indicates a commitment to large-scale research. Private enterprise consultants also report more entrepreneurial behaviors. The analysis indicates the specific entrepreneurial activities of these individuals: 65% have applied for patents (x^2 = 63.99 p < .01); 20% have started new companies (x^2 = 15.19 p < .01); 23% hold equity in a company (x^2 = 82.87 p < .001); and 15% are involved with companies that own patents from

their university research (x^2 = 31.94 p < .001).

When faculty who consult exclusively with one company were compared with those who do not (including non-consultants), exclusive consultants report higher levels of entrepreneurial behavior, research budget, and amount earned above their institutional salaries. Table 2 shows the mean differences between these groups. Exclusive consulting offers greater financial rewards for the academic scientist, which should increase the potential for them to defy research behavioral norms for self-aggrandizement.

The analysis indicates the specific entrepreneurial activities of those who consult exclusively with one company: 72% have applied for patents (x^2 = 30.41 p < .001); 35% have started new companies (x^2 = 33.65 p < .001); 35% hold equity in a company (x^2 = 83.61 p < .001); and 30% are involved with companies that own patents from their university research (x^2 = 70.09 p < .001).

Conflict of interest variables. When consultants were asked to report on the conflict of interest variables used in this study, we found that of those who answered "yes", the majority were private enterprise consultants. Table 3 shows these results. Private enterprise and nonprofit/government consultants were most represented in research agenda bias (x^2 = 26.58 p < .001); prior review (x^2 = 37.15 p < .001); withholding (x^2 = 11.49 p < .01); and trade secrets that resulted from university research (x^2 = 10.61 p < .05). The results for secrecy were not statistically significant.

Logistic regression analyses. Entrepreneurial behavior level (0 to 4) is associated with private enterprise consulting when gender, academic rank, teaching, publication numbers, service, research budget, and amount of supplemental income are held constant. The most meaningful

	Characteristics							
	Gender		Rank			Research Budget	Entrepre-neurial Behavior	Income over Salary
Consulting:	Male	Female	Assist.	Assoc.	Full			
No Consulting	79%	21%	13%	24%	63%	239,752	.43	4,995[1]
Public Consulting	80%	20%	8%	29%	63%	355,494	.47[2]	3,880[3]
Private Consulting	96%	4%	17%	22%	61%	397,337[4]	1.14[5]	1,5201[6]

Table 1. Consultant characteristics (N=1032) reported in percentages and means. [1] Difference between non- and public consultants (p < .001) [2] Difference between public and private consultants (p < .001) [3] Difference between public and private consultants (p < .05) [4] Difference between public and private consultants (p < .001) [5] Difference between non- and private consultants (p < .001) [6] Difference between public and private consultants (p < .001)

Consulting:	Research Budget	Entrepreneurial Behavior	Income over Salary
Exclusive	365,568[1]	1.76[2]	22,170[3]
All Others in Sample	269,196	.48	5,595

Table 2. Mean differences between exclusive consultants and all others in the sample on research budget, entrepreneurial behavior, and amount earned over income. [1] (p < .05) [2] (p < .001) [3] (p < .001)

	Behaviors				
Consulting:	Research Bias***	Prior Review***	Withholding**	Secrecy	Trade Secrets*
No Consulting	23%	11%	9%	ns	6%
Public Consulting	24%	9%	8%	ns	7%
Private Consulting	43%	29%	18%	ns	12%

*Table 3. Consultant reports (N=1032) of conflict of interest behaviors. ***p < .001 ** p < .01 * p < .05*

variable in the equation is the private enterprise consultant status ($t = 9.32$, $p < .001$), followed by publication numbers ($t = 4.48$, $p < .001$). The strength indicates that private enterprise consultants appear more likely to engage in entrepreneurial activities than either public consultants or non-consultants. The full model, which explains 15% of the variance, suggests that faculty who consult with private industry and who have higher publication numbers are more likely to engage in entrepreneurial behaviors than others.

There is a modest correlation between supplemental income and private enterprise consulting ($r = .32$, $p < .001$), and exclusive consulting ($r = .32$, $p < .001$). Supplemental income amount was not regressed on consulting, however, because of these correlations. The model, which accounts for 15% of the variance, indicates that publication numbers, service levels, and total research budget from all sources is closely aligned with supplemental income amount. The most salient independent variable is service ($t = 5.86$, $p < .001$), followed by publications ($t = 3.73$, $p < .001$) and overall research budget ($t = 3.61$, $p < .001$).

Correlations show weak relationships between private industry consulting and research agenda bias ($r = .16$, $p < .001$), withholding ($r = .09$, $p < .01$), and prior review ($r = .18$, $p < .001$). Additionally, those who consult exclusively with one company are correlated with research agenda bias ($r = .08$, $p < .001$) and prior review ($r = .15$, $p < .001$).

Logistic regressions were conducted to test whether or not consulting with private enterprise affects research agenda bias, prior review, secrecy, and withholding. The models to test private enterprise consulting effects included the following control variables: faculty attributes, institutional characteristics, academic productivity measures, and entrepreneurial behavior levels.

The first regression shows that the level of entrepreneurial behavior ($x^2 = 74.05$, $p < .001$) of the faculty member as well as academic program ranking and metropolitan location affects whether or not they allow commercial potential or funding opportunities to determine their research agenda. This finding suggests that faculty in highly ranked programs in metropolitan areas are less likely to allow external factors such as commercial viability and funding to affect their research topics. However, as levels of entrepreneurial behavior increase, the odds that they define research topics according to non-research-related dynamics increase by a factor of 1.65.

The second regression tests the relationship between consulting and prior review. The results indicate that private enterprise consulting has a negative effect on prior review, while supplemental income amount and level of entrepreneurial behavior has a positive effect ($x^2 = 68.16$, $p < .001$). The probability that private enterprise consultants will publish results only after sponsor's review decreases by a factor of .50. However, the likelihood of prior review increases by a factor of 1.59 for rising entrepreneurial behavior levels and 1.24 for supplemental income amount. Essentially, a private enterprise consultant is less likely to conduct research not published without the sponsor's consent. But, increased entrepreneurial behavior and supplemental income do affect prior review.

Private enterprise consulting does not appear to affect withholding research tools from other scientists who request them in either tested model. Faculty in private institutions are less likely to withhold (by a factor of .59), while supplemental income increases the likelihood of withholding (by a factor of 1.26). When entrepreneurial behavior level is added, the negative effect of institutional control remains constant, while the supplemental income effect is slightly lessened ($x^2 = 34.90$, $p < .001$). Levels of entrepreneurial behavior increase the chance that one will withhold from others by a factor of 1.37. The results indicate that faculty in private institutions are less likely to withhold from other scientists even when controlling for levels of supplemental income and entrepreneurial behavior.

Finally, academic program ranking decreases the likelihood that a scientists' university research results in trade secrets by a factor of .56 while level of entrepreneurial activity increases it by a factor of 2.67 ($x^2 = 58.30$, $p < .001$). This model accounts for 21% of the variability for this variable.

The models generated to explain why some scientists conduct research that is never published for proprietary reasons were not statistically significant. Thus, issues related to secrecy as defined in this study were not examined in this analysis.

Analyses on the effects of exclusive

consulting on the conflict of interest variables showed results that are similar to the private enterprise consultant for research agenda bias (no effect), prior review (negative association), and withholding (no effect). These important findings suggest that even the faculty member who consults exclusively with one company is unlikely to violate the research norms of the academic enterprise.

Discussion

The results do not indicate that conflicts of interest occur with any significant frequency; to the contrary, the results show that academic scientists are able to balance their institutionalized scholarly roles with commercial science. Faculty remain embedded in their own social organizations which in the case of the consultant includes the university, the discipline, and the government, organization, or company with whom one consults. Rowan and Miskel argue that these social organizations generate the norms that direct individual behavior (15). Although conventional wisdom suggests that when the faculty consultant serves multiple masters, academic roles and norms are sacrificed for self-interest, the results imply that the consultant maintains an allegiance to the norms of teaching, research, and service. Given these criteria, the faculty in this study can be perceived as actors within the institution of academic science, rather than simply as a set of actors who operate within a single organizational entity. This argument is founded on the capacity of faculty members to interact in a variety of situations that appear to have competing interests and values while they perfect their craft. If academic science is the institution, the institutionalized roles and norms embedded in the scientific method become the criteria consultant-scholars use to make decisions in their relationships with commercial scientists.

University faculty have a societal contract that affords researchers with academic autonomy in exchange for a commitment to improve social welfare through teaching, research, and service (16). The question that drives university conflict of interest policies is whether or not faculty fulfill these institutionalized roles without serving their own self-interest. If they fail to fulfill their duties or pursue their own self-interest in the course of their academic activities, critics would argue that they are involved in a conflict of interest. However, the conflicts that academic scientists face are complex and do not allow for a simple explanation.

Despite the lack of a positive relationship between private enterprise consulting and the conflict of interest variables tested in this study, the need to protect universities, disciplines, and the public from academic renegades remains. Current methods such as disclosure to both academic journals and universities provide an important mechanism to alleviate conflict of interest. However, these policies should be grounded in conflict of interest behaviors, rather than potentials, and enforced by individuals in the academic community. Emanuel and Stein reported that one out of three authors of journal articles held financial stakes in reported research outcomes and failed to disclose such in their publications (17). If self-regulation of the academic enterprise should continue without external interference, enforced disclosure becomes an important tool to prevent conflicts of interest from bleeding into research activities.

The results of this study offer some important implications for how academic policies should be conceived. First, policy development and implementation should rest upon data. Empirical data provides a foundation for the formulation of effective and enforceable policy. The policies developed in this arena span the boundaries between the disciplines, funding agencies, academic institutions, and private sector companies. Rather than establish guidelines in isolation of one another, policies could become aligned across these boundaries to establish both consistency and clarity. Ultimately, compliance becomes evaluated at both the department and disciplinary levels. Consistency and clarity across boundaries will permit faculty to make informed choices.

Second, policymakers should develop clear guidelines within their institutional and agency sectors. Policies that guide rather than constrain faculty behavior could aid faculty understanding of specific behaviors that constitute conflict of interest. Furthermore, clearly articulated guidelines should identify the consequences of individual action so faculty will understand the ramifications of their behavior.

Finally, academic institutions could identify consulting as a component of the faculty reward structures. Boyer and Lewis suggested that consulting could become a means for faculty to involve themselves in both community and institutional service (1). Consulting activity

could become an element of faculty development programs that stimulate faculty vitality and, ultimately, productivity.

Acknowledgements

This research made possible by a grant funded by the National Center for Human Genome Research, National Institutes of Health (Grant Number HGO0724-01), administered by the University of Minnesota and Harvard University. For inquiries contact: Dr. Lisa M. Jones, Postsecondary Education Policy Studies Center, University of Minnesota, 330 Wulling Hall, 86 Pillsbury Drive S.E., Minneapolis, MN 55455-0208, 612-625-0773 (voice), 612-624-3377 (fax).

Bibliography

1. Boyer CM, Lewis DR. Maintaining Faculty Vitality Through Outside Professional Consulting. In SM Clark, DR Lewis (editors). Faculty Vitality and Institutional Productivity: Critical Perspectives for Higher Education. 1985.
2. Fairweather JS. Faculty Work and Public Trust: Restoring the Value of Teaching and Public Service in American Academic Life. Needham Heights (MA): Allyn and Bacon Publishers; 1996.
3. Patton CV, Marver JD. Paid Consulting by American Academics. Educational Record 1979: 175-184.
4. Rebne D. Faculty Consulting and Scientific Knowledge: A Traditional University - Industry Linkage. Educational Administration Quarterly 1989; 25(4): 338-357.
5. Teague GV. Faculty Consulting: Do Universities Have "Control"? Research in Higher Education 1982; 17(2): 179-186.
6. Marsh HE, Dillon KE. Academic Productivity and Faculty Supplemental Income. Journal of Higher Education 1980; 51(5): 546-555.
7. Merton RK. The Sociology of Science: Theoretical and Empirical Investigations. Chicago: University of Chicago Press; 1973.
8. Mitroff II. Norms and Counternorms in a Select Group of the Apollo Moon Scientists: A Case Study of the Ambivalence of Scientists. American Sociological Review 1974; 39(4): 579-595.
9. Louis KS, Anderson MS. The Changing Context of Science and University-Industry Relations. In H Etzkowitz, A Webster, P Healey (editors). Capitalizing Knowledge: New Intersections of Industry and Academia. Albany (NY): State University of New York Press; 1998.
10. Clark BR. Creating Entrepreneurial Universities: Organizational Pathways of Transformation. United Kingdom: Pergamon; 1998: 4.
11. Merton, RK. op cit. 1973: 277.
12. Etzkowitz H, Webster A. Entrepreneurial Science: The Second Academic Revolution. In H Etzkowitz, A Webster, P Healey (editors). Capitalizing Knowledge: New Intersections of Industry and Academia. Albany (NY): State University of New York Press; 1998.
13. Louis, KS, Jones, LM, & Anderson, MS. Bought and Sold: Academic Ideas in the Marketplace. Paper presented at the meeting of the Association for the Study of Higher Education, Miami, FL. (1998, November 6).
14. Louis, KS, Blumenthal, D, Gluck, ME, & Stoto, MA. Entrepreneurs in Academe: An Exploration of Behaviors Among Life Scientists. Administrative Science Quarterly, 1989; 34, 110-131.
15. Rowan B, Miskel CG. "Institutional Theory and the Study of Educational Organizations". In J Murphy, KS Louis (editors). Handbook of Research on Educational Administration, Second Edition. San Francisco (CA): Jossey-Bass Publishing; 1999.
16. Slaughter S, Leslie LL. Academic Capitalism: Politics, Policies, and the Entrepreneurial University. Baltimore (MD): The Johns Hopkins University Press; 1997.
17. Emanuel EJ, Steiner D. Institutional Conflict of Interest. N Engl J Med 1995; 332 (4): 262-267.
18. Boyer CM, Lewis DR. Maintaining Faculty Vitality Through Outside Professional Consulting. In SM Clark, DR Lewis (editors). Faculty Vitality and Institutional Productivity: Critical Perspectives for Higher Education. 1985.

5. Understanding Misconduct

Preventing Scientific Misconduct: Insights from "Convicted Offenders"

Mark S. Davis, Justice Research & Advocacy, Inc., USA

Michelle L. Riske, Justice Research & Advocacy, Inc., USA

Keywords: Equity theory, Prevention, Rationalization techniques, Research misconduct, Responsible conduct of research, Scientific misconduct

The mere seriousness of certain social behaviors implies the need to prevent them. In the case of conventional crime, for example, survivors of homicide victims or the victims of physical or sexual assault, when asked what they want most, often will say they wish the incident had never happened. For them, a successful homicide prosecution does not bring back the lost loved one. A long prison term for the rapist will not restore the victim to the state she enjoyed prior to the crime. As a result, we strive to identify and implement various ways of reducing opportunities for both offending and victimization.

Although the perceived harm in research misconduct may not be as great as in violent crime, its consequences nevertheless can have disastrous and far-reaching effects. After-the-fact measures such as the investigation of allegations and the sanctioning of the guilty, while necessary for justice and the vindication of the moral order, seldom can undo the harm caused by each instance of fabrication, falsification, plagiarism, or other serious departure from the norms of science. The retraction of a published paper cannot restore the time wasted by other investigators pursuing pointless lines of research or by editors and referees reviewing meaningless results. An apology and a signed voluntary consent agreement by one found guilty of research misconduct does not automatically lift the taint from the supervisor and colleagues in whose lab the misconduct occurred. And for those who suffer from life-threatening diseases and consequently hold out hope for a cure, the broken trust of falsified clinical trials has far more devastating effects. To be sure, the shock waves emanating from a single incident of research misconduct can create untold collateral damage, including the tarnishing of reputations of scientists, institutions, and of the enterprise of science itself.

In view of our collective inability to undo the damage and effect restoration to all parties in these cases, the prevention of research misconduct is a desirable end. The question then becomes, what can the scientific community do to keep research misconduct from occurring in the first place? The purpose of this preliminary analysis is to explore largely untapped data sources in order not only to advance theoretical work in this area, but also to glean information of practical import.

In order to tackle the challenge posed by prevention, we must acknowledge that prevention can occur at more than one level. Douglas Weed, employing public health's notions of primary and secondary prevention, suggests that we first need to know something about etiology, and he argues that there are causal factors both internal and external to the scientist who engages in research misconduct (1). Examples of internal causal factors would include psychological problems, financial motivations, or perhaps the desire to hurt others. Causes external to the scientist, on the other hand,

Corresponding author: Michelle L. Riske, J.D., M.A., 849 Cleveland Ave., Amherst, OH 44001, 440-988-8455 (voice), 440-988-8455 (fax), riske@lor.net.

are factors such as the pressure to publish, inadequate training or supervision, or the fierce competition for research grants.

In either case, successful prevention requires that we somehow interrupt one or more processes that lead to an instance of research misconduct. For example, if we knew that individual psychopathology was responsible for research misconduct, we perhaps could administer the Minnesota Multi-phasic Personality Inventory (MMPI), the Narcissistic Personality Inventory, the Psychopathy Checklist, or other psychometric tools to help us screen out applicants who were predisposed to engaging in unethical research practice. In an effort to address an external cause such as inadequate supervision, we might institute regular meetings between lab supervisors and their staff members.

Objectives

This pilot study focuses on two individual-level explanations for research misconduct. First, Cressey's research on embezzlement in financial institutions was examined (2). Cressey's subjects, who largely perceived themselves to be respectable people, had three characteristics in common:

1. A non-shareable financial problem, for example, one the individual could not discuss without suffering humiliation;
2. An awareness the problem could be solved by violating the position of financial trust; and
3. Suitable rationalizations for the embezzlement of funds to resolve their self-conception as a trusted person.

Applying Cressey's work to scientific researchers, is it possible that some have non-shareable problems, not necessarily financially-based, which motivate them to engage in research misconduct? The possibilities could include the inability to produce replicable work under pressure, a perceived lack of talent for research, or personal problems such as marital or emotional difficulties. For example, William Summerlin, the protagonist in one of the best-known cases of research misconduct, intimated that he had been under a lot of pressure from the head of the lab to produce results. Could the inability to withstand this sort of pressure constitute a non-shareable problem?

In addition to possibly having such non-shareable problems, how do researchers who engage in misconduct formulate rationalizations for their behavior? And what form might these rationalizations take? Sykes and Matza, in their research on juvenile delinquency, discuss several of what they refer to as "techniques of neutralization" including (3) :

- *Denial of a victim* (Who am I really hurting by fudging these data?)
- *Denial of an injury* (What is the harm?)
- *Condemnation of the condemners* (They're out to get me.)
- *Denial of negative intent* (I never meant to hurt anyone.)
- *Metaphor of the ledger* (For most of my time here in the lab I've been a hard-working, loyal employee. I'm entitled to a slip or two. All in all, I've done more good than bad.)

Is it possible that individuals who commit research misconduct may employ one or more of these techniques in order to justify their conduct?

The second perspective employed for this study was social psychology's equity theory, which speaks to perceived fairness in dyadic relationships (4). Equity theory is exemplified in the common phrases "You scratch my back and I'll scratch yours" and "One good turn deserves another." Social beings have come to expect reciprocity when dealing with others. If people perceive they are getting less from a relationship than they are given, they may suffer distress. It is common, then, for the ostensibly exploited person to take measures to relieve this distress and restore a sense of equity. In the case of research misconduct, scientists may be more likely to engage in misconduct if they believe they were deprived of what was rightfully theirs, such as the co-authorship on a publication or a coveted promotion. Accordingly, individuals may engage in scientific misconduct as a form of retaliation against a coworker or supervisor if they believe that they have been slighted or exploited.

Design

Two sources of data were gathered for this study. The first was information from the case files of individuals against whom a finding of scientific misconduct was made by the Office of Research Integrity (ORI). A standard data collection form was used to record data including the institution, type of alleged misconduct, information from the respondent, response of the institution, and finding by the ORI. A member of the research

team read each case file and wrote narrative responses to the items on the data collection form summarizing information primarily pulled from the investigative reports by the universities and from the investigative reports of ORI and its predecessors. These narrative responses were analyzed for this part of the study. A total of 21 case files were reviewed for the initial pilot study. These case files included 16 cases reviewed as part of a pretest, as well 5 additional cases that included cases closed prior to the formation of the ORI, i.e., these cases were handled by the Office of Scientific Integrity (OSI), ORI's predecessor.

The second source of data consists of interviews with scientists against whom a finding of scientific misconduct was made by the ORI. Subjects who were included in the first nine case files used as part of the pretest comprised the sample for this portion of the data collection process. Because some scientists approached could not be located or were unwilling to participate in the interviews, only three out of the nine contacted were interviewed. It is possible that the experience of having been accused and found guilty of research misconduct was so unpleasant that some subjects have little interest in dredging up the past. One scientist who declined to participate in the study summed up his feelings in an e-mail to the senior author:

> "I am very sorry to disappoint you but after more then ten years I have no inclination to discuss this issue with anybody. With my very poor English I found it useless to talk about the inquisition. I have no idea what is a (sic) subject and goal of your research, but I wish you a (sic) success in your work in the name of justice, science and humanity."

One of the interviewees summed up his feelings more bluntly when thanked for his time:

> "The time is not the problem; it's the pain of having to relive this crap."

The researchers signed a confidentiality agreement with ORI to protect sensitive case file information. The researchers also took additional steps to ensure confidentiality during the data collection process, by excluding the subjects' name and case file number from the data collection instruments. Subjects were identified by the assignment of a subject number. To match files with subjects being interviewed, a list including the subject name, institution, ORI case number, and subject number was created. The information was only used to link interview subjects with the case file reviews. Upon completion of the interviews, the subject list was given to ORI. Both data collection instruments were approved by an Institutional Review Board and by the U.S. Department of Health and Human Services, Office for Protection from Research Risks.

Methods of Analysis

Because theoretical work on scientific misconduct is relatively meager, we chose to use a qualitative approach borrowed from phenomenological psychology. Rather than first searching for evidence of specific theories or propositions, the investigator examines the data more for "explication" than explanation (5). This results in the listing and preliminary grouping of terms or phrases revelatory of, in this case, etiology. As a check against possible bias created by prior knowledge or other factors, the analyst extracts exact phrases rather than interpreted concepts. Another analyst approaches the data in the same way, identifying exact wording to convey possible sources of causation. The second step involves the two analysts coming together to compare and reconcile their lists. In the third step, the analysts group the phrases into common themes or constructs. Finally, the constructs are examined to see if they relate back to the selected theoretical approaches in order to help us interpret and discuss the relevance of these constructs or central themes in explaining the etiology of research misconduct. For example, in looking at Cressey's notion of the non-shareable problem (6), the analyst would group together those extracted phrases suggesting such themes as psychological issues, marital difficulties, financial pressure, lack of knowledge, difficulty with expectations of a supervisor, lack of supervision, or other problems an individual might reasonably be uncomfortable sharing with others.

Data obtained from the case file reviews and from the interviews eventually will be content analyzed using the QSR-NUDIST software. Content analysis is a means of systematically analyzing textual information to find recurring themes, issues, and motifs, which can then be isolated, counted, and interpreted (7, 8). If the appropriate statistical criteria are met, the data will also be analyzed to examine relationships among variables in order to assess, for example, if a certain type of misconduct or rank is

associated with the existence of a non-shareable problem.

The Sample

The data collected was part of a pilot study to test the efficacy of the data collection instruments developed, which were then used as part of a larger study examining all individuals against whom a finding of scientific misconduct was made by the ORI as of December 2000. A total of 21 case files were reviewed for the pilot study. Many of the respondents held academic positions as either Senior Faculty or Junior Faculty (each category included 8 out of the 21 subjects). Senior Faculty included professors, associate professors, and directors/heads of departments, institutions or clinics. Junior Faculty is defined as assistant professors, postdoctoral students, research fellows and residents. Other researchers, including research associates, predoctoral students, and administrative assistants, made up the remaining positions (5 out of 21). It should be noted that tenure status could not be gleaned from the case files.

With respect to the types of research misconduct committed by these 21 respondents, 38% of the cases were for plagiarism, 19% were for fabrication, and 19% were for falsification. Fabrication/falsification made up 14% of the cases, and the remaining 10% were for a combination of falsification, fabrication, and plagiarism.

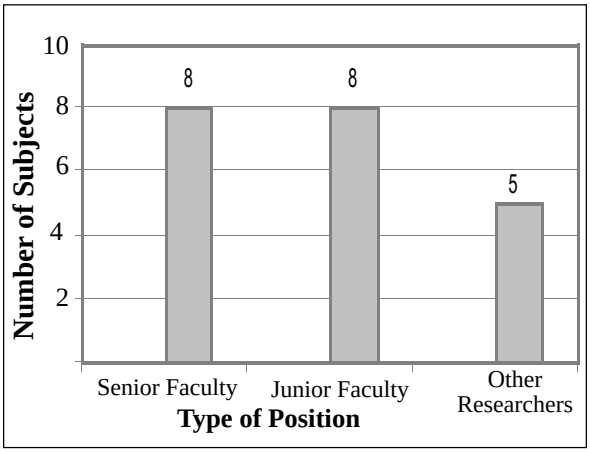


Fig. 1. Researcher's Academic Position

Results

Data from the case files reviewed were analyzed using the qualitative phenomenological approach.

Etiology

The systematic search for possible etiological factors related to our two theoretical perspectives yielded data in support of both theories. Phrases or elements extracted from the case files showed evidence of non-shareable problems such as publish-or-perish pressure, lack of knowledge or experience, difficulty with supervisor's expectations/lack of supervision, and personal problems. These phrases were usually extracted from information contained in the University investigative reports or the

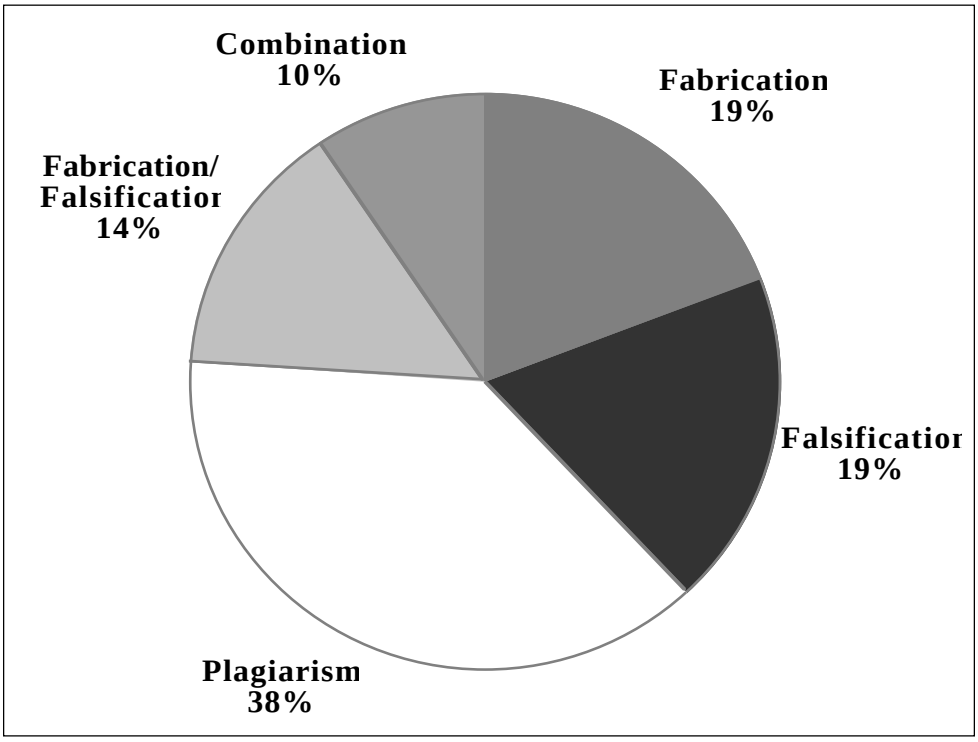


Fig. 2. Classification of Research Misconduct

Job Pressure	Lack of Subject Matter Knowledge	Personal Problems	Problems with Supervision
Enormous pressure to produce/ Pressure to produce by supervisor	Understanding of grant application process/First proposal	Personal insecurity	Could not fully satisfy the expectations of the supervisor/ If supervisor had more realistic expectations this incident might never had occurred
Time factors - short deadlines/ Short cut to save time	Different interpretation of the normal modes of responsible authorship	Personal/Family difficulties	Supervisor was demanding in research results
Pressure to keep the system working	Understanding of the principles of attribution in review articles	Medical illness	Lacked proper scientific guidance from supervisor/ Unsupervised
Insecure position	Not able to handle position/ Saddled with responsibilities which in hindsight were out of proportion to subject's training and experience		Under personal pressure from supervisor to publish data in order to secure a research position.
Isolated laboratory with few peers to discuss situation or possible career alternatives	Never trained in appropriate record keeping		Negligent oversight/ Deficiencies in oversight/ Supervisor's oversight was inadequate
Difficult job situation/ Stressful job situation			

Table 1. Etiology - Non-shareable problem

investigative reports from ORI; therefore, the information is hearsay, as the actual statements made by the respondent or other interested parties were usually not contained within the case files.

Information obtained from the interviews also provided evidence in support of a non-shareable problem by the respondent, which may have contributed to his misconduct. For example, one interviewee stated:

> "How am I going to get a position where I don't have to worry every 2-3 years about if I don't get my grant I'm gonna be out on the street. This feeling of being a second, kind of a second class citizen. Um, the pressures to produce papers. And, you know, it was, I knew I was doing something wrong, but I reached a point where I didn't care."

The data also contained summarized statements from respondents indicating rationalization techniques of denial of an injury, condemnation of condemners, and denial of negative intent.

Although information extracted from the case files did not definitively point to instances where the subject engaged in conduct in order to restore a perceived loss of equity in a dyadic relationship, some of the phrases taken from the case files suggest possible motivation by the subjects that could indicate retaliatory conduct in response to perceived exploitation. For example, some of the subjects said that they falsified data in order to annoy colleagues or that they were not recognized for their scientific expertise. Other subjects discussed competition in relation to positions within the university or institution and competition for submitting papers for publication.

Implications for Prevention

If we look at the preliminary evidence for our theoretical questions, we can infer some tentative implications for prevention. Information pertaining to lack of proper supervision or training suggests that it might be prudent for universities to implement better procedures and guidelines for supervisors with respect to employee oversight and monitoring responsibilities. We found some support that periodic reviews or audits of research notebooks, as well as the original data collected for all experiments, by the supervisor may help to reduce research misconduct. Ensuring that

Denial of an Injury	Condemnation of the Condemners	Denial of Negative Intent
No harm done because the experiments were preliminary, unimportant, and had not been published	Subject had opposite and competing opinions to research performed by colleagues of the complainant	Fabricated sampling times were preliminary and never intended to be published
Worked on several of the articles which were used as references for the proposal and therefore permitted to incorporate these materials into the proposal	Allegations by complainant were an attempt to "get rid of" the subject from the University	Going to tell supervisor the truth after the subject had a chance to obtain valid counts, but the subject didn't have the chance
If there was some faulty reporting of findings, that it was minimal since it was not the central issue of the study		

Table 2. Etiology - Neutralization Techniques

employees are properly trained on all experimental techniques prior to performing such experiments could also help reduce the researcher's lack of knowledge on the subject matter, as well as apprehension about acknowledging that as a problem. Similarly, discussing the serious ramifications of research misconduct can also discourage some of the denial its perpetrators use to rationalize their actions with such conduct; for example, that there indeed is harm associated with these actions that affects a variety of actors and institutions, including, most importantly, the patient population.

The three interviews conducted to date have also provided some insights for prevention. One subject credited the careful handling of data for his own demise:

> "...when the technician did the, you know, do the random stuff, yes, there would be a copy on the computer, but he would also print out the data, you know, a paper copy and put that into their books. So, it was, you know, like, it was also like a guarantee that I would eventually be found out and that it could all be traced back."

So upon his returning to the lab from an extended trip:

> "...basically they sat me down and confronted me with the fact that these data sets don't fit. And, it was a situation of, uh, what do you say if you're caught red-handed? You know all the original data was there. It was very easy for them to go back to the original sets and see that there were discrepancies."

This same interviewee briefly contemplated trying to cover up the misconduct, but again realized:

> "...it was truly a situation where the record keeping system that I had set up was such that there was no way I could possibly go back through all the computer files and alter those. There was, you know, everything, the techs had always printed out paper copies, so there was shelves of three ring binders with all the data. It was a situation of, it can't be done."

One interviewee felt that training might help prevent some research misconduct:

> "I think that there should be more training, study in just the basics of the scientific method and, you know, what is appropriate, you know,

Evidence of possible motivation to retaliate	Evidence of possible motivation to exploit
Made up data to annoy a colleague	Future dependent on rapid success in the laboratory
Some friction between subject and others in the lab	Laboratory placed too much emphasis on short-term productivity
Bitter relationship between subject and supervisor	Competitive pressure for tenure-track positions
Failed to make changes because upset with others	Insecure position
Attempt to get rid of subject	
Personal animosity against the subject/Prejudice against the subject	

Table 3. Etiology - Equity Theory

> what is not appropriate in terms of experimental methodology, in terms of statistics, in terms of, if you're going to discard data, you know, what are the other experimental reasons for discarding the data? For example, oh yeah, I had a sudden sneeze and I sneezed and I botched test tubes or I knocked over this particular test tube, or I tested this particular agent and found that, oh my gosh, I actually added 10 times the amount of a particular component, you know, those are valid reasons for discarding data. You know, I don't think there's enough emphasis placed on teaching people the proper scientific method."

Another subject offered what he referred to as an "easy" solution to the problem of fabrication and falsification:

> "What you do, is you have, uh, open laboratory meetings where everyone in the laboratory knows what everyone else is doing. Uh, you say you did an experiment that took a hundred rats, but only five rats came into the, into the lab, it's pretty clear that you didn't do a hundred rats. Uh, if you're not there doing the work, uh, that people think you're doing or know that you're supposed to be doing, uh, so I think, uh, open laboratories, with regular, uh, presentations of data prevent that."

Conclusions

We used a qualitative approach to explore selected aspects of individual-level etiology of research misconduct. These preliminary data offer some tentative support for our theoretical perspectives. More definitive conclusions will have to await the collection and analysis of the data from the larger study.

This research-in-progress also offers support for certain forms of prevention. These suggestions, rather than the product of well-meaning, but less-than-well-informed commentators, come from those most intimately involved in actual cases. Returning to the analogy of crime, learning from those who have engaged in research misconduct is not unlike debriefing convicted burglars on what would have kept them from choosing a particular dwelling as a target. Who should know better than those who have done it?

Bibliography

1. Weed, DL. Preventing scientific misconduct. American J Public Health 1998; 88(1): 125-129.
2. Cressey, D. Other people's money: The social psychology of embezzlement. New York: The Free Press; 1953.
3. Sykes, G, Matza, D. Techniques of neutralization: A theory of delinquency. American Sociological Review 1957; 22: 664-670.
4. Walster, E, Walster, GW. Equity: theory and research. Boston (MA): Allyn & Bacon; 1978.
5. von Eckartsberg, R. Existential-phenomenological research. In von Eckartsberg, R, editor. Phenomenological inquiry in psychology: Existential and transpersonal dimensions. New York: Plenum Press; 1996, p. 3-61.
6. Cressey, D. Other people's money: The social psychology of embezzlement. New York: The Free Press; 1953.
7. Denzin, NK, Lincoln, YS. Part I: Method of collecting and analyzing empirical materials. In Denzin, NK, Lincoln, YS, editors. Collecting and interpreting qualitative methods. Thousand Oaks (CA): Sage; 1998, p. 35-45.
8. Maxfield, MG, Babbie, E. Research methods for criminal justice and criminology. 2nd ed. Belmont (CA): West/Wadsworth Publishing Co; 1998.

The Relative Efficiency of Research Misconduct Investigations Involving Personal Injury vs. Injury to the Scientific Record

Andrew J. Hogan, Department of Medicine, Michigan State University , USA

Ronald J. Patterson, Department of Microbiology and Molecular Genetics, Michigan State University, USA

Robert L. Sprague, Department of Kinesiology , University of Illinois-Urbana/Champaign, USA

Keywords: Efficiency, Research misconduct, Type of injury

Research misconduct investigations conducted by universities and other research institutions are sometimes highly contentious affairs whose findings are disputed both internally and externally. The central question of the research reported in this paper is whether certain features of the typical internal research misconduct investigation contribute to the likelihood of closure or to continued contention.

Most research misconduct investigations undertaken in institutions that receive Federal research contracts and grants follow the investigational model proposed by the National Institutes of Health (NIH), described here as the tribunal model. In civil law, similar types of disputes (civil fraud, misappropriation of property, etc.) are dealt with in adversarial proceedings. One measure of the efficiency of the typical model for conducting a research misconduct investigation is to determine how often that model produces a definitive finding, or alternatively how often it leads to further proceedings.

The objective of this study was to test whether the presence of personal injury associated with a research misconduct allegation influences the likelihood of a post-investigation proceeding (lawsuit, grievance, legislative hearing, administrative inquiry, etc.), in the context of the use of the tribunal model of investigation. We hypothesized that the standard tribunal model, which was designed principally to protect the integrity of the scientific record, might not be very efficient in addressing misconduct allegations in which a personal injury was the central feature.

Materials and Methods

Data. Cases were identified in the files of Dr. Robert Sprague of the University of Illinois-Urbana/Champaign, which contain 1,100 references on the 231 research misconduct cases (hereafter referred to as the "Sprague files"). The Sprague files consist primarily of copies of news stories in scientific journals, such as Science and Nature, or academic trade journals, such as the *Chronicle of Higher Education* and *Lingua Franca*.

Sixty-three cases were identified as having adequate documentation of alleged misconduct

Corresponding author: Andrew J. Hogan, Ph.D., Health Services Research Division, Department of Medicine, Michigan State University, B210 Clinical Center, East Lansing, MI 48824, 517-353-4361 (voice), 517-432-1326 (fax), hogana@msu.edu

involving either a personal injury or an injury to the scientific record. A personal injury case was one in which a person directly involved in the misconduct allegation identified some kind of personal loss, usually misappropriation of intellectual property — plagiarism or the unauthorized use of confidential information from grants or articles under peer review. A scientific record case was one involving some form of contamination of the scientific record. Scientific record cases usually involved falsification/fabrication, but sometimes involved misappropriation of the intellectual property of non-parties to the allegation.

Post-investigation proceedings included grievances filed within the institutions, lawsuits, complaints to regulatory or funding agencies, and requests to legislative or administrative bodies. A post-investigation proceeding was classified as a due process case if one or more of the parties raised due process issues (hearing notification, right to call or cross-examine witnesses, impartial decision-makers, etc.) related to the research misconduct investigation.

In the tribunal model of a research misconduct investigation, an individual files an allegation with an institution, and the institution forms a panel to investigate the allegation. The panel is responsible to gather evidence, call and examine witnesses, and make a finding; in common parlance, the tribunal is prosecutor, judge and jury. The standard NIH-tribunal model often attenuates some due process rights commonly found in adversarial proceedings, in particular rights to call or cross-examine witnesses and to present evidence. Current NIH policy suggests that the complainant in such an investigation be treated as a witness, rather than as a party.

In an adversarial proceeding, one party (complainant) accuses the other party (respondent) of misconduct. The parties gather and present evidence, call and examine and cross-examine witnesses. The institution provides an adjudicator to process the allegation, hold hearings and render a decision. We were able to identify no unambiguous cases in which the adversarial model was employed in a research misconduct investigation.

Data Collection and Reliability. We reviewed 221 documents related to the 63 identified cases. For each document, a form was completed (see Appendix A) identifying the case name and the document number in the Sprague files. The abstractor (Hogan or Patterson) identified the type of misconduct alleged (fabrication/falsification, misappropriation of intellectual property, other serious deviations, retaliation, or other). The abstractor then determined the nature of injury based on whether there was an injured party known to the individual alleged to have committed misconduct; if so, the case was classified as one involving personal injury, otherwise as injury to the scientific record. Next the abstractor coded for the type of institutional investigation (tribunal or adversarial), based principally on whether the complainant was a witness or a prosecutor.

The abstractor then determined whether there were other proceedings consequent to the institutional research misconduct investigation, such as:

- Internal grievances, discrimination complaints, etc.
- Lawsuits, complaints/appeals to administrative agencies, complaints/appeals to legislative bodies.

In those cases where there was some sort of post-investigation proceeding, the abstractor determined whether due process issues were raised.

Finally, the abstractor examined each document regarding the role of the institutional legal counsel as being supportive, neutral, or obstructive of the procedural fairness of the institutional investigation. The abstractor looked for any references to the role of institutional legal counsel regarding the selecting or preparing witnesses, selecting or preparing panelists, selecting or preparing administrators, handling administrative problems/complaints, issues of attorney-client privilege, providing or withholding information, applying legal indemnification, deliberating or making findings, the preparing or editing of reports, the protecting of parties' due process rights.

To assure the reliability of the abstraction process, the first 20 cases were reviewed by both abstractors to establish interrater reliability using a data collection tool. Review of the reliability of the initial cases indicated a 94% agreement on which documents were relevant to each case, a 70% agreement regarding the type of misconduct, and a 91% agreement on whether the injury was personal or to the scientific record. There was a 60% agreement on which documents indicated the type of institutional investigation,

but 100% agreement on the type of institutional investigation. There was also 100% agreement regarding the existence of post-investigation proceedings. The reasons for the discrepancies in the classification of misconduct allegations were discussed and resolved before finishing the abstraction of the remaining cases.

Results

No unambiguous cases where the original research misconduct investigation was administered using the adversarial model were found. All of the results related to research misconduct investigations which were conducted under the standard tribunal model.

Of the 63 cases described in the 221 documents reviewed, 41% of cases resulted in a post-investigation proceeding, and 69% of these involved a due process issue. Of the 63 cases, 41% of cases involved personal injury, and 70% of personal injury cases resulted in a post-investigation proceeding. Of the personal injury cases resulting in a post-investigation proceeding, 61% of these proceedings involved a due process issue.

Ten percent of the 63 cases involved some controversy regarding the role of the institutional attorney. Although we looked for instances where the role of the institutional attorney was supportive of procedural fairness, only negative statements appeared in the literature examined. Twenty-one percent of cases arose in the context of a funded grant.

Multivariate logistic regression analysis was performed to determine the likelihood of post-investigation proceedings. The results are presented in Table 1. Personal injury cases are at least 10 times more likely to result in a post-investigation proceeding than cases involving injury to the scientific record. When allegations are made in the context of a funded grant, the likelihood of a post-investigation proceeding is reduced, although this effect is only marginally statistically significant.

Parameter	Odds Ratio	95% Bounds	
		Upper	Lower
Personal injury	10.34**	36.46	2.94
Attorney controversy	3.71	33.39	0.41
Grant context	0.22*	1.12	0.04

*Table 1. Logistic Regression Analysis: Likelihood of Post-Investigation Proceeding. n=63, ** = p < 0.05; * = p < 0.10*

In the subset of cases where due process issues were raised, any controversies regarding the role of the institutional attorney in the research misconduct case tended to increase the likelihood of a post-investigation proceeding by more than six-fold (see Table 2). However, this result was only marginally statistically significant.

Parameter	Odds Ratio	95% Bounds	
		Upper	Lower
Personal Injury	3.39**	11.28	1.028
Attorney controversy	6.50*	46.16	0.92
Grant Context	0.35	1.88	0.07

*Table 2. Logistic Regression Analysis: Likelihood of Post-Investigation Proceeding Involving Due Process. n=63 / ** = p < 0.05 / * = p < 0.10*

Conclusions

Because we were able to identify only two ambiguous cases of research misconduct investigations possibly employing an adversarial model, we were not able to determine whether the adversarial model would result in fewer post-investigation proceedings than the tribunal model arising out of misconduct investigations involving personal injury.

Under the standard tribunal approach to research misconduct investigations, cases involving personal injury are much more likely to produce a post-investigation proceeding. We speculate that the tribunal approach frustrates the ability of personally injured complainants to seek redress. From the lofty perspective of protecting the integrity of the scientific record, personal injury cases may often appear trivial or unimportant and clouded by interpersonal bickering that borders on the unprofessional.

Very often personal injury cases involved intellectual misappropriation disputes between students or junior faculty and senior faculty members. In such cases, the administrators and the members of the tribunal conducting the investigation tend to be more the peers of the respondent than the complainant. Complainants, rightly or wrongly, often believe that the investigation is biased toward the respondent and that the tribunal procedures prevent them from making the most effective cases against the respondent.

ORI's recent policy statement about treating whistleblowers as witnesses will probably

increase the likelihood of a post-investigation proceeding by giving complainants even less standing than they previously held.

In some cases the external funder offers a post-investigation appeals process including a full due process hearing, for example, the Departmental Appeals Board of the Department of Health and Human Services. The existence of this appeal mechanism may alter the conduct of the original investigation, leading to fewer post-investigation proceedings. The existence of an external appeal mechanism may discourage some institutions that might be tempted to bias a research misconduct investigation toward an outcome most favorable to the institution's reputation or financial interests; the possibility of disclosure and/or reversal at an appeals hearing could act as a check on such institutional behavior.

Institutional attorneys may face conflicts of interest when fair treatment of the parties to an investigation is not perceived to be in the institution's interest. Legal representation of an organization presents many potential ethical pitfalls for attorneys, especially when conflicts arise within an organization, as is the case when a university must investigate a research misconduct allegation against a faculty member or student.

While most judges are attorneys, most attorneys are not judges and most attorneys are trained to act as advocates for their clients. Some institutional attorneys may see their roles as advocates for procedural fairness, but they also understand that a finding of misconduct can carry heavy financial and reputational consequences for the university as well as the individual respondent.

Moreover, any of the parties to a misconduct investigation could become a potential litigant against the university because of decisions made during the case by university administrators. Therefore there may be a strong tendency to act as legal advisor to university administrators as opposed to advocates for a fair and impartial investigation.

In this research, it is difficult to determine whether controversial actions by institutional attorneys was a cause or consequence of post-investigation proceedings, since the timelines necessary to distinguish cause from effect are often missing in the kinds of documents reviewed. Also the frequency of such reports are low, but this could arise from the confidentiality of attorney-client communications as well as from lack of incidents to report.

Caveats. Most reports of research misconduct are from news stories in scientific or trade magazines (*Science*, *Nature*, *Chronicle of Higher Education*). Reliance on these sources could introduce a possible reporting bias, since only the most disputatious cases would be considered news worthy. This reporting bias could significantly affect the prevalence data presented earlier, but probably would not have a major effect on the results of the multivariate analysis.

NIH/ORI reports on the outcomes of research misconduct investigations were also a major source of cases. NIH/ORI reports also contain relatively few plagiarism/ownership cases, which might tend to underestimate the number of personal injury cases.

Some observers believe that the handling of research misconduct cases has improved over time. The results of this study found a slight and statistically insignificant temporal decline in the number of cases resulting in post-investigation proceedings. However, this decline was confounded by a concurrent decline in the number of cases reported over time. Because the cases presented here were identified from the scientific news literature, this latter decline could be a function of either fewer cases (better management) or less reporting (declining newsworthiness) or both. A separate study based on a fixed baseline of research misconduct allegations in the institutions in which they arose has been proposed to disentangle these confounded effects.

Case References (Reference Identifier in Sprague Database, citation)

5310. Marshall E. EPA faults classic lead poisoning study. *Science* 1983; 222: 906-7.

6411. Weissman v. Freeman. D. New York 1989 Feb 23; Nos. 225, 353.

6438 . Investigations, S. o. O. a. Fraud in NIH grant programs. 1988.

6439 . Crewdson J. *Science* on the hot seat. *Chicago Tribune* 1989 Mar 19; Sect. 5, 9, 13.

6483 . Dakins D. Copyright lawsuit illuminates debate over academic standards. *Diagnostic Imaging* 1989 May; 54-60.

6484. Mervis J. Bitter suit over research work asks “Who deserves the credit?’. *The Scientist* 1989; 1, 4.
6486. Blum D. Yeshiva professor wins copyright lawsuit. *Chronicle of Higher Education* 1989 Apr 5; A14.
6494. Raymond C. Institutes of Health find that Purdue professor plagiarized experiments. *Chronicle of Higher Education* 1989 Jul 12; A2.
6496. Marshall E. USGS reports a fraud. *Science* 1989; 244: 1436.
6512. Barinaga M. NIMH assigns blame for tainted studies. *Science* 1989; 245: 812.
6519 . Gupta’s defense. *Science* 1989; 245: 1192.
6521. Erben H. Carelessness, or good faith? *Science* 1989; 245: 1165-6.
6523. Coughlin EK. Geologist accused of misidentifying fossils calls charges ‘misleading’. *Chronicle of Higher Education* 1989 Sep 13; A9.
6532. Raymond C. Allegations of plagiarism of scientific manuscript raise concerns about ‘intellectual theft.’. *Chronicle of Higher Education* 1989 Jul 19; A4, A7.
6551. Water T, Oliwenstein L. *Science* duds. Discover 1990: 42-44.
6565. Zurer P. NIH panel strips research of funding after plagiarism review. *Chemical & Engineering News* 1989; 24-5.
6603. Greenberg DS. Squalor in academe Plagiarist gets promoted, victim is out of her job. *Science* & *Government* 1990, May 1; 1, 5-7. Gualtieri, C. T. “Letters.” .
6622. Henderson J. When scientists fake it. *American Way* 1990; 56-62, 100-1.
6623. Heidi S. Weissmann v. Leonard M. Freeman. 684 Fed Sup: 1248.
6629. Kaplan, R.J. (1989). Letter.
6636. Scandal upon scandal. *Nature* 1990; 343: 396.
6661. Gordon G. Lab aide sues U-M, ex boss. *Detroit News* 1990 May 6: 1A, 9A.
6689 . Culliton BJ. NIH misconduct probes draw legal complaints. *Science* 1990; 249: 240-2.
6728. Hamilton DP. NIH sued over misconduct case. *Science* 1990; 2419: 471.
6738. Twedt S. Federal study raps Pitt medical school. *The Pittsburgh Press* 1990 Sep 9: Sect. A:1,7.
6744. Hallum J, Hadley S. NIH Office of Scientific Integrity: Policies and procedures (Letter to the Editor). *Science* 1990; 249: 1227-8.
6745. Weismer G. Orofacial impairment in Parkinson’s disease (Letter to the Editor). *Neurology* 1990; 40: 191-3.
6747. Hamilton D. Science misconduct legalese thickens. *Science* 1990; 249: 1102.
6812. Bradshaw RA, Adler AJ, et al. Bridges committee procedures (Letter to the Editor). *Science* 1990; 250: 611.
6831. Pinney GW. ‘Whistle-blowers’ are encouraged to come forward at ‘U,’ dean says. *Minneapolis Star Tribune* 1990 Oct 9: 4B.
6832. FDA investigating doctor at ‘U’ who led drug study. *Minneapolis Star Tribune* 1990 Dec 9; 4B.
6842. Hamilton DP. NIH misconduct procedures derailed. *Science* 1991; 251: 152-3.
6874. Cordes C. Professor asserts Stanford lied to U.S. about misconduct probe. *Chronicle of Higher Education* 1991 Mar 13; A21, A24-A25.
6920. Hallum JV, Hadley SW. NIH misconduct procedures: Effect of new ruling (Letter to the Editor). *Science* 1991; 251: 1296.
6945. Even misconduct trials should be fair. *Nature* 1991; 350: 259-260.
6990. Rooney P. UI researcher steps down after admitting plagiarism. *News Gazette* 1991 Apr 11; A1, A10.
6991. Rooney P. Colleagues see pressure in plagiarism case. News Gazette 1991 Apr 12; A3. Ames, Rusell University of Illinois 19913
7057. London R. Judge strikes rules on scientific fraud inquiries. *New York Times* 1991 May 3; Section 2, 11.
7079. Weiss T. Too many scientists who ‘blow the whistle’ end up losing their jobs and careers. *Chronicle of Higher Education* 1991 Jun 26: A36.
7090. Monson N. Misconduct in the labs. *Health Watch* 1991: 24-33.
7097. Wheeler D. U.S. has barred grants to 6 scientists in past 2 years. *Chronicle of Higher Education* 1991; A1, A6-A7.
7216. Wiernik J. Suits charge plagiarism in research. *Ann Arbor News* 1991 Sep 18; B1- B2.
7279. Woolf P. Fraud in science: How much, how serious? 1989 *Hastings Center Report* 11, No. 5: 9-14.
7284. Blakely E, Poling A, et al. *Fraud, fakery, and fudging.* : 313-330.
7327. Rennie D. The Cantekin affair. *JAMA* 1991; 266: 3333-7.
7328. Winslow R. Earache studies lead to painful dispute. *Wall Street Journal* 1991 Dec 18; B1, B3.
7329. Lead researcher under investigation. *Science* 1991; 254: 1575.
7486. Ernhart CB, Scarr S. Lead challenge (Letter to the Editor). *Science* 1992; 255: 783.
7501. Moore JA. Cases of misconduct related to speech and hearing science. Paper for HSS 1991; 494/Psych 493.
7551. Roberts L. Misconduct: Caltech’s trial by fire. Science 1991; 253: 1344-7.
7584. Wheeler DL. Scientists begin to question confidentially of fraud investigations. *Chronicle of Higher Education* 1992; A1, A8, A11.
7632. Woolf PK. Deception in scientific research. *Jurimetrics Journal* 1988; 29: 67-95.
7673. Adler T. Psychologists criticize study of low lead levels. *Monitor* 1992 Mar; 40-1.
7689. Woolf PK. Accountability and responsibility in research. *Journal of Business Ethics* 1991; 10: 595-600.
7757. Rigert J, Lerner M. ‘U’ researcher could face federal charge. *Star Tribune* 1992 Aug 22; 1A, 4A-1A, 5A.
8020. Anderson M. *Impostors in the temple*. New York: Simon & Schuster 1992.

8070. Rigert J, Lerner M. In courting industry, 'U' has invited trouble." *Minneapolis Star Tribune* 1992 Dec 31; 1A, 14A.

8186. Scarr S, Ernhart C. Of whistleblowers, investigators, and judges. *Ethics & Behavior* 1993; 3: 73-93.

8187. Ernhart CB, Scarr S. On being a whistleblower. *Ethics & Behavior*. 1993; 3: 199-206.

8197. Minnesota psychiatrist accused of fraud in study. *News Gazette* 1993 Feb 17: D-2.

8217. Rigert J. Top child psychiatrist at 'U' indicted on charges of fraud. *Minneapolis Star Tribune* 1993 Feb 17; 1A, 11A.

8218. Sweeney P. U professor faces charges on drug study irregularities. *Saint Paul Pioneer Press* 1993 Feb 17; 1A, 6A.

8259. Cordes C. Research project on ear infection dramatizes challenge of conflicts. *Chronicle of Higher Education* 1993 Mar 3: A23, A28.

8262. McCuen RH. Institutional oversight to enhance integrity in research. *Journal of Professional Issues in Engineering* 1990; 116: 27-37.

8298. Penn S. Misconduct in the modern university. *Thistle* 1992; 1, 7.

8299. Loh P. Vest sued for coverup of scientific misconduct. *Thistle* 1992; 1, 8-1,11.

8323. Zorza M. "Letter." 1993 May 10.

8419. Cotton P. Harassment linked to fraud, lawsuit alleges. *JAMA* 1992; 267: 783-784.

8439. Burd S. Shalala's criticism of U.S. polices on scientific misconduct troubles some lawmakers and watchdog groups. *Chronicle of Higher Education* 1993; A31, A34.

8478. Lueders B. Hunger strike feeds UW controversy. *Isthmus* 1993; 5, 7.

8479. Kerr S. Feds freeze fraud-finders A Wisconsin connection? *Shepherd Express* 1993 Jun 17-24;. 3.

8501 . Leo A. Paquette, Ph.D., Ohio State University. *NIH Guide for Grants and Contracts* 1993; 2-3.

8502 . James H. Freisheim, Ph.D., Medical College of Ohio. *NIH Guide* for Grants and Contracts 1993; 1-2.

8506 . Mark M. Kowalski, M.D., Ph.D., Dana Farber Cancer Institute and Harvard University. *NIH Guide for Grants and Contracts* 1993; 2.

8523 . Burd S. New policy of naming scientists who are found guilty of fraud renews debate on federal role. *Chronicle of Higher Education* 1993 Jun 30; A24, A25.

8549 . Final findings of scientific misconduct. *NIH Guide for Grants and Contracts* 1993; 2-3.

8585 . Holland E. Committee finds validity in allegations against Paquette. *On Campus* 1993 Aug 19.

8586 . Lore D. OSU hunts core of plagiarism. *Columbus Dispatch* 1993 Aug 23; C1.

8587 . Cleverley M. Chemistry professor has student support. *The Lantern* 1993 Aug 26.

8604 . Polsby M. NIH spasmodic torticollis. Bethesda, MD: National Institutes of Health. 1993 Jul 30.

8606 . Polsby M. "Letter." 1993 Aug 20.

8607 . Rosen JE. "Letter." 1993 Jun 26.

8608 . Rosen JE. Statement of Dr. Jane Rosen. Undated.

8609 . McCutchen CW. "Letter." 1993 Aug 24.

8611 . Rosen JE. "Letter." 1993 Aug 24.

8612 . Rosen JE. "Letter." 1993 Aug 12.

8627 . Taubes G. Fraud busters. *Review of Academic Life* 1993; 47-53.

8648 . Grassley CE, Cohen WS. "Letter." 1993 Sep 23.

8791 . Hilts PJ. U.S. scientist says superiors stole credit for her vaccine. *New York Times* 1993 Dec 18; 1010.

8841 . Nimmons D. Sex and the brain. *Discover* 1994; 64-71.

8902 . Butler K, Ohland G. Making the case against discrimination. *LA Weekly* 1991 Jan 18.

8915 . Anderson D, Attrup L, Axelsen N, Riis P. *Scientific dishonesty & good scientific practice*. Copenhagen: Danish Medical Research Council 1992 .

9097 . Lock S, Wells F, et al., Eds. *Fraud and misconduct in medical research*. 1993.

9185 . Immunologist fined $5,000 in lab misdeed. *New York Times* 1994 Aug 31; A16.

9186 . Final findings of scientific misconduct. *NIH Guide for Grants and Contracts* 1994; 23: 2.

9417 . Findings of scientific misconduct. NIH Guide 1995.

9542 . Hilts PJ. A university and 4 scientists must pay for pilfered work. *New York Times* 1995 May 19; A20.

9568 . Bauer HH. Ethics in science. [Internet] Discussion of Fraud in Science; 1995 May 1.

9603 . Taubes G. Plagiarism suit wins: Experts hope it won't set a trend. *Science* 1995; 268: 1125.

9614. Herling J. Embarrassment in Hong Kong. *Chronicle of Higher Education* 1995 Mar 24; A43-A44.

9616 . Parrish CJ. The Lumigen Project - or how Wayne State University went into business and lost its shirt and a lot more! *Newsbriefs* Wayne State University 1994 Sep 7;. 1-7.

9617 . Shellum B. Chemical conspiracy. *Detroit Free Press* 1995 Apr 7; 1E-2E.

9618 . Wilson M. Judge: WSU chemist stole major discovery. *The Detroit News* 1994 Aug 17; 1A, 6A.

9636 . Lytton H. Frequency of FFP. [Internet] 1995 Jun 11.

9642 . Kytle R. HHS release. [Internet] 1995 Jun 23.

9649 . Holden C. More on genes and homosexuality. *Science* 1995; 268: 1571.

9654 . Crewdson J. 'Gay gene' study comes under fire. *Chicago Tribune* 1995 Jun 25, 1, 6-1, 7.

9662 . Benditt J. Conduct in science. *Science* 1995; 268: 1705-1718.

9684 . Marshall E. NIH's "Gay gene" study questioned. *Science* 1995 Jun 30; 268: 1841.

9707 . Findings of scientific misconduct. *NIH Guide* 1195; 24.

9720 . Integrity, O. o. R. *Annual Report*. Department of Health and Human Services 1994 Sep.

9727 . Researcher assessed damages for distribution of a cell line. *ORI Newsletter* 1994; 3: 1-2.

9739. Burd S. Nursing professor found guilty of scientific misconduct. *Chronicle of Higher Education* 1995 July 7; A27.

9757. Bivens LW. Findings of scientific misconduct. Office of Research Integrity. Notice 1993 June 21.

9784 . Wilchek M. [Letter to the Editor]. *Science* 1995; 269: 1034.
9812 . Office of Research Integrity, *Annual Report* 1994. 1995.
9826 . Findings of scientific misconduct. *NIH Guide* 1995; 24, No. 33.
9854 . Findings of scientific misconduct. NIH Guide 1995; 24, No. 34.
9890 . ORI/AAAS Conference plagiarism and theft of ideas. 1993.
9959 . Case summaries. *ORI Newsletter* 1996; 4, No. 2: 5.
9963 . Findings of scientific misconduct. NIH Guide 1996; 25, No. 15.
9972 . Findings of scientific misconduct. NIH Guide 1996; 25, No. 13.
9974 . Miller JP. AIDS research scientific fraud. [Internet] 1996 Apr 30.
9975. Findings of scientific misconduct. NIH Guide 1996; 25, No. 13.
9976. Findings of scientific misconduct. NIH Guide 1996; 25, No. 10.
9983 . Jayaraman KS. Court allows fossil fraudster to return. *Nature* 1996; 380: 570.
10003 . Cherry M. Black scientist faces inquiry at South African university. *Nature* 378: 324.
10079. Findings of scientific misconduct. *NIH Guide* 1996; 25, No. 22.
10084. Alan L. Landag. *ORI Newsletter* 1995; 4, No. 1: 5.
10110 . King RT. How a drug firm paid for University study, then undermined it. *Wall Street Journal* 1996 Apr 25.
10116 . Jock F. MIT torn by bitter dispute over missile. *Science* 1996; 271: 1050-2.
10153 . Findings of scientific misconduct. *NIH Guide* 1996; 25, No. 22.
10154 . Findings of scientific misconduct. *NIH Guide* 1993; 25, No. 22.
10163 . Gail L. Daubert. *ORI Newsletter* 1996; 4: 3-4.
10164 . Jamal Z. Farooqui. *ORI Newsletter* 1996; 4: 4.
10165 . Andrew Friedman. *ORI Newsletter* 1996; 4: 4-5.
10166 . Joan Gans. *ORI Newsletter* 1996; 4: 5.
10212 . Hilts PJ. Dispute on gulf war missile continues to fester at MIT. *New York Times* 1996 Feb 24; 11.
10229 . Zinberg D. A cautionary tale. *Science* 1996; 273: 411.
10260 . Findings of scientific misconduct. *NIH Guide* 1996; 25, No. 27.
10358 . Garber P. MCG looking at changes admid research investigation. *Augusta Chronicle* 1996, October 5.
10463 . Rex D. International recruitment highlights need to track scientific behavior. *Nature* 1996, September 12; 383: 107-108.
10605 . Vipin Kumar. *ORI Newsletter* 1996; 6-7.
10693 . Findings of scientific misconduct. *NIH Guide* 1997; 26, No. 15. Findings of scientific misconduct. *NIH Guide* 1997, May 9; 26, No. 15.
10719 . David F. Eierman. *ORI Newsletter* 1995; 3.
10724 . Yahya Abdlahi. *ORI Newsletter* 1996; 5.
10772 . Ann Marie Huelskamp, M.H.S. Johns Hopkins University School of Medicine. *ORI Newsletter* 1997; 5.
10789 . Zalewski D. Do corporate dollars strangle scientific research? *Lingua Franca* 1997; 51-59q.
10802. Findings of scientific misconduct. *NIH Guide* 1997; 26, No. 22.
12345 . Mazur A. The experience of universities in handling allegations of fraud or misconduct in research. Project on Scientific Fraud and Misconduct: Report on Workshop Two, AAAS 1989.

Appendix A

DATA COLLECTION SHEET FOR ORI ABSTRACT RESEARCH PROJECTS

CASE NAME_____ DOCUMENT NO._____

TYPE OF MISCONDUCT ALLEGED:
(check all that apply)

Fabrication/Falsification _____
Misappropriation of Intellectual Property _____
Other Serious Deviations _____
Retaliation _____
Other:_____

NATURE OF INJURY
(Is there an injured party known to the alleged misconductor?)

Personal Injury _____Injury to the Scientific Record _____

TYPE OF INSTITUTIONAL INVESTIGATION
(Is the complainant a witness or a prosecutor?)

Tribunal _____ Adversarial_____

OTHER PROCEEDINGS CONSEQUENT TO THE INSTITUTIONAL INVESTIGATION

Internal (grievances, discrimination complaints, etc.) ____
If yes, was due process an issue? ____
External:Lawsuits ____
If yes, was due process an issue? ____
Complaints/Appeals to administrative agencies ____
If yes, was due process an issue? ____
Complaints/Appeals to legislative bodies ____
If yes, was due process an issue? ____

ROLE OF INSTITUTIONAL LEGAL COUNSEL

As regards the following, was there any evidence as regards the role of institutional legal counsel as being (S)upportive, (N)eutral, (O)bstructive or (U)nknown of the procedural fairness of the institutional investigation? (circle one in each line)

Selection or preparation of witnesses:	S	N	O	U
Selection or preparation of panelists:	S	N	O	U
Selection or preparation of administrators:	S	N	O	U
Handling administrative problems/complaints:	S	N	O	U
Issues of attorney-client privilege:	S	N	O	U
Providing or withholding information:	S	N	O	U
Application of legal indemnification:	S	N	O	U
Deliberation or making findings:	S	N	O	U
Preparation or editing of reports:	S	N	O	U
Protection of parties' due process rights:	S	N	O	U

Ethical Evaluation of Misconduct Cases

Doric Little, School of Medicine, University of Hawaii, Honolulu, USA.

Martin D. Rayner, Pacific Biomedical Research Center, University of Hawaii, Honolulu, USA

Key Words: Ethical evaluation, Misconduct, Scientific divorce

The policies governing the actions of the Ethics Committee at the University of Hawaii were developed during the late 80's when the dominant paradigm for Ethics investigations was the "whistleblower" model. In this model a person of relatively low power in the academic hierarchy complains of scientific or ethical misconduct perpetrated by a person of higher rank and/or power, typically within their own academic unit.

For such cases to be handled in an appropriate manner (and to ensure that whistleblowers feel free to come forward) the confidentiality of the complainant must be carefully protected. Administrative procedures should minimise the chances that the accused person can use his/her academic power: a) to have the complaint disregarded without adequate investigation and/or, b) to instigate reprisals against the whistleblower. However, innocent faculty also need to be protected from frivolous or malicious complaints. Thus, an initial Inquiry (Phase 1) was required, during which the existence of the complaint is withheld from the accused, with the accused being informed and interviewed only after the complainant has convinced the Review Panel that a thorough investigation is justified. At that point, a full Investigation (Phase 2) is initiated, the accused is informed of the complaint while his/her lab notebooks, computer files and other pertinent sources of information are immediately sequestered. The accused then has the opportunity to present detailed rebuttal. If the evidence in support of this rebuttal seems inadequate, then the committee so reports to the Administration and a more formal Phase 3 Hearing is set up. It is only after the innocence of the accused has been reasonably established (typically following the completion of Phase 2) that more difficult issues may be considered, such as the possibility that the complaint was motivated by envy or by malice. Furthermore, to conclude that the complaint is malicious requires the committee to assess the motivations of the accuser at the time the accusation was made. Thus, even if strong suspicions exist, it is not likely that sufficient evidence will be uncovered to confirm suspicions of malicious intent.

Despite the even-handed principles involved in this approach, the Inquiry Phase of such investigations is necessarily limited to evidence provided by the complainant. And, more significantly, both Phase 1 and Phase 2 primarily address the guilt or innocence of the accused. While we understand that this sharp focus is appropriate in some situations, our experience suggests that this is *not* necessarily a "one size fits all" model. This committee has experienced scientific misconduct cases in which this approach prevented a fair and balanced Inquiry. We suggest that specific circumstances exist in which policies based on this model may need to be modified to ensure an appropriately ethical analysis of the complainant's case.

Corresponding author: Martin D. Rayner, Interim Director, Pacific Biomedical Research Center, 1993 East-West Rd., University of Hawaii, Honolulu, HI 96822, 808-956-5184 (voice), 808-956-9574 (fax), martin@pbrc.hawaii.edu.

Year	Complaint	Outcome	Whistleblower protections	Collaboration Breakdown	$$ issues
92	Intel. Prop. Theft	Sustained	Required	No	No
92	Plagiarism	Sustained	Required	No	No
93	Plagiarism	Dismissed	Required	No	No
93	Admin. Miscon.	Dismissed	No	No	No
94	Plagiarism	Sustained	Required	No	No
95	Authorship	Dismissed	No	Yes	No
96	Intel. Prop. Theft	Dismissed	No	Yes	No
96	Intel. Prop. Theft	Dismissed	No	Yes	No
97	Intel. Prop. Theft	Negotiated	No	Yes	No
98	Misapp. of funds	Reimbursed	Required	No	Yes
99	Theft/fabrication	Dismissed	No	Yes	Yes
99	Intel. Prop. Theft	Dismissed	No	Yes	Yes
99	Intel. Prop. Theft	Sustained	No	Yes	Yes
99	Sci. Miscond	Sustained	No	Yes	Yes
00	Hum Subj. issue	Sustained	No	No	No

Table 1. Analysis of cases presented to the University Ethics Committee from 1992 to 2000

Results

Despite the many high-profile cases, nationally, which seemed to fit the whistleblower model during the 80's and early 90's, we have noted significant changes in the nature of the complaints coming before our committee over the last five years (see Table 1). As shown in this Table, six of the nine cases occurring after 1995 involved issues of intellectual property. Before this time, however, only one case out of six involved a clear intellectual property dispute. Seven out of the nine cases since 1995, but only one out of the six earlier cases, involved breakdowns in scientific collaborations. Similarly, five out of the nine post-1995 cases involved high financial stakes, whereas none of the earlier cases seem to have been primarily motivated by financial considerations. Finally, whereas four out of the six early cases required whistleblower protections to protect the identity of a junior complainant, only one complaint out of nine cases since 1995 benefited from such protections. Thus, whistleblower protections are still needed, although cases that fit that specific model are no longer a major part of our workload.

Discussion

Ethics Evaluations in A Changing World

Two nation-wide trends may well have been responsible for these changing patterns. First, changes in funding patterns have increased the payoff for collaborations between potentially competing laboratories. Second, as scientific information has become increasingly regarded as potentially marketable intellectual property, it is inevitable that disputes will arise as to the ownership of that property. The stakes are further raised when University Administrators suggest that returns to research units from the marketing of such intellectual property should become a significant component of the budgets of academic research units. In apparent response to these trends, our recent cases have been motivated primarily by disputes over the ownership of potentially valuable intellectual property. These situations are not consistent with the whistleblower model on which our Ethics policies and procedures are based - making them difficult to evaluate. However, these cases cannot be dismissed as being merely "authorship disputes" beneath the level of interest of those whose duty it is to evaluate true scientific misconduct issues, in view of the very high stakes which may be involved. Finally, we have seen such cases start at the level of an authorship dispute, only to later expand into full-scale accusations of data fabrication.

Nevertheless, our university's policies as well as the general awareness of the scientific community remain tuned to the whistleblower

model. So, as one might well expect, our cases continue to be presented in the "approved" whistleblower format, promising to reveal significant instances of scientific misconduct.

If one fails to understand their origins, such cases can be difficult to evaluate. In one such instance we were unable even to conclude that a valid case existed under our restrictive rules for Phase 1 Inquiries. What does one do when Phase 1 of a "denial of authorship" complaint leads to the complainant eventually submitting letters from the accused in which the accused pleads with the complainant to accept authorship on the paper in question? Should the accused have been interviewed during Phase 1, in this case, so as to gain additional understanding of the background against which the complaint was made? The initial decision that there was no case to pursue, precipitated a seemingly endless series of requests for external intervention and/or re-evaluation of our committee's policies. We need to do better than that.

Similarly, other recent cases before our committee have seemed to involve inherent conflicts between the superficial appearance and the underlying realities of each case. The stage now seems set for continuing problems arising, in part, from our evaluative approaches. Perhaps, significant changes should be proposed in both the published procedures and investigative approaches so as to permit effective evaluation of cases that do not fit the whistleblower paradigm. However, these cases raise arguments for modifications of our procedures that might, if implemented, remove key protections for more classic whistleblowers.

This seems a potentially dangerous situation in which it would be all too easy for university faculties and administrations to make serious mistakes while acting from the highest ethical motivations. To address these concerns recent cases have been re-evaluated to search for potentially generalizable patterns within what had seemed to be "property disputes". Such a pattern could provide the theoretical grounding from which a more systematic approach could be developed towards this different class of misconduct complaints.

Excluding situations involving "priority of discovery" issues, or situations of outright theft (none of which we have yet seen), when two groups feel that they both have valid claims to some piece of the same pie this is probably a pie they baked together. In other words, the majority of such disputes seem to arise from the breakdown of formerly effective collaborations. And, since most collaborations collapse from personality conflicts, it is hardly surprising that such breakdowns lead to disputes over the custody of intellectual property. The comparison with that other graveyard of failed collaborations, the divorce courts, is inescapable. The level of acrimony over rights to intellectual property seems fuelled by these underlying personal issues, just as rights to child custody may become the focus of a parent's sense of violation in a divorce situation. An Ethics Committee that must stick its head into a "scientific divorce" needs to be well aware just how high the emotional stakes may have become for the individual contestants regardless of the monetary worth of the objective data.

The committee will need to remember that not all fights are about money. Some fights are incomprehensible from any other motive than to humiliate the opponent. And they will need to recognise that when it takes at least two people to bake such a pie, it often takes two to spill it on the floor. Of course, the participants in this "divorce" may not have behaved equally badly, but the party most wronged is not necessarily the one who complains the most loudly. This is dangerous territory for an investigative committee, where the most fundamental assumptions of the whistleblower model may no longer be valid.

Formulating a working hypothesis

The essence of the issue is this: whereas the whistleblower model appropriately evaluates the validity of the complaint, in a "scientific divorce" it cannot be assumed that the substance of the complaint is valid. Furthermore, it was clear that our case load in Hawaii would not be sufficient to permit even a minimally rigorous prospective study of such cases - which is why we are presenting our ideas to this meeting. If analysis of our experience resonates with the experience of other similar committees, perhaps they will also take up this issue.

"Scientific divorces" may need to be evaluated by different procedures. In these cases one should not focus on the guilt or innocence of the accused, but rather survey the ethical landscape in which the breakdown of collaboration occurred. Specifically, it is not appropriate to assume that the complaint is valid or that the complainant is not a material

contributor to the situation under investigation. To support this approach, the preliminary instructions given to our Review Panels were changed. When the initial complaint indicated that either an intellectual property dispute, or a breakdown in collaboration, was involved, it was suggested that both the complainant and the "accused" needed to be interviewed during Phase 1. In other words, it may be impossible to determine whether or not misconduct is likely to have occurred unless both parties are interviewed. In a situation of this kind, however, the committee needs to be aware that the complainant will have had time to purge any files that might prove embarrassing, although the accused may well have been taken by surprise.

Additionally, even in Phase 2 of the investigation, we suggested that the Review Panel delay considering whether the accused might be guilty or innocent of misconduct. First, they should focus their attention on a different question: "What happened to create the present conflict?". However, they should be prepared to take as much detailed testimony as necessary to answer that very simple question. Only when the committee has reached a clear consensus as to "what happened", should they attempt to consider which actions taken by each participant might rise to the level of scientific misconduct. The danger here is that such open-ended investigation can get out of hand – the Chair of the Review Panel may need to remind its members that focus should be maintained on immediately relevant events.

These instructions appear to have substantially facilitated the appropriate ethical evaluation of difficult cases. Our Review Panels have been models of good committee interactions where all decisions have been unanimous following considerable discussion but without significant disputes. This surprising degree of agreement resulted from a comprehensive consensus as to "what really happened" – committee members have all felt comfortable that "blame", where blame has been needed, was fairly assigned. Finally, shared understanding of the underlying issues allowed them to make some very tough calls in potentially explosive cases. Even in these hard cases, committees appear to have appropriately surveyed each situation without bias and to have resolved the issue appropriately.

Next steps

The most effective method needs to be explored by which to merge this "Ethical Landscape model" into policies written to protect whistleblowers. We would like to avoid a triaging mechanism which would separate cases into, for example: intellectual property cases, misconduct cases and "harm/rights" cases with different guidelines (as in the the separate courts of our legal system). Instead, we have hoped to find some way to treat all our cases from an ethical perspective, while at the same time preserving our protections for whistleblowers. We now believe that ALL cases can be addressed from this ethical approach in which we do not ask "is the accused guilty?" but instead ask "what really happened?" Once the Panel can answer that question, then they can consider the extent to which each participant has behaved in an ethical or unethical manner - and we are ready to ask whether any of these behaviors rise to the level of scientific misconduct. By contrast, Phase 3 of the investigation (when this is necessary), should be the point at which standard legal models are introduced.

Fortunately, only one small change in our policies is required to implement this approach. The Review Panel **needs the discretion to interview the accused during Phase 1,** should they conclude that this can be carried out without threat to the complainant. Given that freedom, the Panel can then adopt either the "standard" approach to Phase 1, or the "ethical landscape" approach, as seems most fitting to the case under investigation.

Nevertheless, the open-ended investigational approach advocated here can lead to unusual situations. For example, in one recent case the Committee's final report to the University Administration recommended censure not only for the accused but also for the complainant (whose actions contributed to the wrongdoing), as well as for a third party who facilitated the situation to his own benefit. To have reported only on the guilt of the accused would have seemed a violation of our Committee's ethical duty in this instance.

Acknowledgements

We acknowledge receipt of travel funds from Alan Teramura, Senior Vice-President for Research at the University of Hawaii.

Potential Cultural Factors In Scientific Misconduct Allegations

Walter M. Meyer, III, Department of Psychiatry and Behavioral Sciences, The University of Texas Medical Branch, Galveston, USA

George M. Bernier, Jr., Department of Internal Medicine, The University of Texas Medical Branch, Galveston, USA

Key words: Authorship disputes, Ethnicity, Gender, Scientific misconduct

Since 1993, The University of Texas Medical Branch has had 16 allegations of scientific misconduct. They were each examined carefully during an inquiry by a faculty committee and the scientific integrity officer for evidence of fabrication, falsification or plagiarism. Only one of them was judged to be scientific misconduct. It involved plagiarism, which was acknowledged by the respondent, and this case will not be discussed further in this document. The remaining 15 allegations did not reach the stage of investigation. They involved a variety of other types of complaints: an authorship dispute in 4 cases, inadequate sharing of data in 3 cases or allegations of questionable research practices in the remainder. Since many of these disputes involved individuals who were not born in North America and were raised in different cultural settings, the authors hypothesized that cultural factors underlie many of these allegations. In order to examine this question, they have done a retrospective review of the 15 allegations.

Methods

A retrospective review of these 15 allegations was done to detect the possible involvement of gender, academic status, ethnic factors or cultural concerns. To determine whether any ethnic or cultural group appeared to be overly represented as complainant or respondent, the cultural/ethnic background status of the entire faculty, post-doctoral fellows and research technical personnel was compared to those involved in these allegations.

Results

The 15 complaints involved 29 people; 13 White (10 European descent, 3 Middle Eastern descent), one African American and 15 Asians (9 Indians and 6 Chinese). See Table I for ethnic distribution of the complainants and respondents. One of the Indians was involved in two separate instances, once as a respondent and once as a complainant. All the Asians were born and raised outside of the United States. Six of the complainants were White (4 European descent, 2 Middle Eastern descent) and 3 of these were born and raised outside of North America. Seven of the respondents were White (5 European descent, 2 Middle Eastern) and two were born outside of North America. The one African American individual, born in the United States, was a respondent. Nine Asians (4 Chinese and 5 Indians) were complainants and 7 Asians (2 Chinese and 5 Indians) were respondents.

Corresponding author: Walter J. Meyer III, Department of Psychiatry and Behavioral Sciences, University of Texas Medical Branch, Galveston, TX, 77555, 409-747-8355 (voice), 409-747-8351 (fax), wmeyere@email.utmb.edu.

		Complainants					
		White, US	White, Foreign	Asian, Indian	Asian, Chinese	African American	Total
Respondent	White US born	1	2	1*	1	0	5
	White Foreign born	0	1	0	1	0	2
	Asian, Indian	2	0	3*	0	0	5
	Asian, Chinese	0	0	0	2	0	2
	African American	0	0	1	0	0	1
	Total	3	3	5	4	0	15

Table I: Number of complainants and respondents by ethnic group
** One person was a complainant and a respondent*

Three subjects involved in these allegations were technicians, seven were post-doctoral fellows and the remaining 19 individuals were faculty. One faculty was involved in two allegations, once as a complainant and once as a respondent. The complainants and the respondents were of very similar ages, mean ages of 45.7 and 44.0 years, respectively. In ten cases, the complainants were older than the respondents and in five they were younger. Ten of the complainants were of lower rank in the university than their respective respondents. Only five of the 29 individuals were female (two Whites, two Indians and one Chinese). These 5 women were involved in a total of 3 allegations.

Six of the allegations involved individuals from different ethnic groups. The remainder involved individuals from the same ethnic or cultural background. Of the six disputes involving more than one ethnic group, three involved White of European origin and Indians; two, a White and Chinese; one, a African American and an Indian. Nine disputes involved individuals from the same ethnic group: two involved Chinese; three involved Indians; and four involved White. Among the disputes involving White as both complainant and respondent, one involved both parties being from the Middle East; one involved both parties born in the USA and of European descent; one involved a complainant born in an eastern block country and a respondent born in the USA; and the last involved a foreign born middle eastern (Egyptian) complainant and an American born Israeli respondent. Two of the allegations involving Asians referred to deep-seated distrust of individuals from similar backgrounds in their country of origins. In one instance, the complainant stated that he knew that the misconduct had occurred because people from the village of the respondent were evil. In the other instance, the complainant referred to the political leanings of the respondent as they related to their country of origin, i.e., brands of communism.

To determine whether any ethnic or cultural group appeared to be overly represented as complainant or respondent, the cultural/ethnic background status of the entire group of university employees (faculty, bachelor level technicians or post-doctoral fellow) was compared to those involved in complaints. All but one female professor was or had been employees of the university. Only five of the individuals were female (two Whites and three Asians). The faculty is 24% female and 17% of these allegations involve females.

There is a great difference in the ethnic distribution of the total faculty compared to those individuals involved in scientific misconduct allegations. The medical school has a faculty of 750 individuals (550 White, 39 Hispanic, 24 African American and 136 Asian). Of the 136 Asian, at least 55 are from India and 43 are from China. Table II illustrates the differences in ethnic distributions between the faculty, bachelor level research technicians and post-doctoral fellows at large and those individuals involved in scientific misconduct disputes. There is a significant difference between the individuals involved in scientific misconduct allegations and the total group of individuals in the same category for the faculty ($p < .0001$ by chi-square), the technicians ($p < .0001$ by chi-square) and the post-doctoral fellows ($p < .001$ by chi-square). The country of origin was not discerned for the faculty. But there does seem to be among the White individuals an unexpectedly large number of individuals born in the Middle East.

Discussion

In the early 1990's many universities started

		White	Nat. Am.	Hispanic	Indian	Asian	Total
Total	Faculty*	73.0	3.2	5.2	0.5	18.1	100
Total	Technicians**	56.6	4.6	9.8	0.0	29.0	100
Total	Postdoctoral ***	40.0	3.0	4.0	1.0	52.0	100
Involved in Scientific Misconduct Disputes	Faculty *	52.6	5.3	0.0	0.0	42.1	100
	Technicians**	33.0	0.0	0.0	0.0	67.0	100
	Postdoctoral ***	28.6	0.0	0.0	0.0	71.4	100

Table II: Differences expressed as percent of total in ethnic distributions between the faculty and postdoctoral fellows at large and those individuals inovlved in scientific misconduct disputes
**Significantly different p <.0001 by Chi Square, **Significantly different p <.0001 by Chi Square*
****Significantly different p <.001 by Chi Square*

establishing a very formal process to consider scientific misconduct charges. The initial definitions were focused on fabrication, falsification and plagiarism but did leave an opening for examining 'other practices that seriously deviate from those that are commonly accepted within the scientific community for proposing, conducting or reporting research' (so called unusual or questionable practices) (1-3). The allegations or complaints were usually none of these; rather they reflected personal disputes between the complainant and respondent. Questionable research practices were particularly difficult to define and often the scientific integrity officer and/or relevant faculty committee were called upon to make a judgment of intent. Therefore these disputes were almost always impossible to discern with any assurance for fairness. In order to gain insight into these types of complaints, a fairly large amount of work has been done nationally to examine the nature of the complaint. In fact, certain types of complaints such as authorship complaints were rejected as scientific misconduct. Also the Office of Science and Technology Policy has established, but not formally implemented, a more narrowed definition to exclude questionable research practices and to include with fabrication, falsification, and plagiarism only the inappropriate use of documents which might be seen as part of the review process (4). Even with this narrower definition the complaints about authorship, data ownership and access and questionable or sloppy research practices will continue to plague the university committees and scientific integrity officers.

In contrast to open discussion about the nature of the complaints and allegations, almost nothing has been written about the nature of those who made the complaints or those who were the target of the complaints. The little we do know refers only to the respondents who have been determined to have committed scientific misconduct. We know little about those who brought the complaint forward because of the appropriate concern about damaging the whistleblower. Also almost nothing has been written about those allegations, which did not meet the definition of scientific misconduct as defined by falsification, fabrication, and plagiarism. One study of authorship disputes received at the Ombuds office of Harvard Schools and affiliated hospitals reported that the number of disputes has greatly increased between 1991-2 to 1996-7 (5). Women were involved in the majority (53%) of the complaints and non-US citizens were involved in 21% of them (5). The current study seems to be the only other venture into this area. This study identifies a higher than expected number of individuals who were born, raised and partially educated outside of the United States. In addition, the complaints are often against individuals from the same ethnic background and gender as the complainant. This data is provocative. If substantiated in other universities, it indicates a need to reexamine our education of faculty and post-doctoral fellows concerning the proper use of the scientific misconduct complaint process. Also other mechanisms need to be identified to help settle these misunderstandings among scientific colleagues.

There are significant hazards to doing this type of retrospective review. This type of endeavor invites accusations of racism, gender bias, and other un-American activities, such as racial profiling. In order to get different perspectives on this issue, the authors had the Director of our Affirmative Action Office and a member of our Institute of Medical Humanities

review this manuscript. We are attempting only to describe as a group the complainants and respondents, not to speculate why one group rather than another might utilize the scientific misconduct complaint process to address other related issues in the research group setting. One speaker at the recent ORI conference on research (6) suggested that misconduct complaints are increasing because of the increased collaborative nature of research and increased difficulty in obtaining funding. Only three of our allegations involved collaborations outside of the complainant's research group. Four of our allegations could be linked to some financial factors but they did not seem to be the main issue. Usually the complaint involved very poor communication between the respective parties. Some ground rules for working together need to be taught as part of the research curriculum.

Conclusions

The vast majority of complaints did not involve scientific misconduct as currently defined. This retrospective review suggests that cultural concerns may contribute to the complaints to the scientific integrity office. Proportionally the Asian group is over represented in the scientific misconduct complaint process. This report documents for one university the magnitude of the apparent influence of cultural differences in the scientific misconduct complaint process. On the surface, this retrospective review suggests that cultural differences account for many of the authorship and other scientific misconduct disputes. Since the vast majority of complaints in this retrospective review did not involve scientific misconduct as currently defined, we believe there is a need for an increased educational effort on the part of the university to orient faculty, bachelor level research technicians and post-doctoral fellows on the appropriate use of the scientific misconduct process and to develop other mechanisms to help them resolve conflicts with fellow scientists. Guidelines for data ownership and management (7), authorship of grants, and authorship of papers (8) have been recently established on our campus to aid in this process.

Acknowledgment

The authors wish to express their gratitude to Deborah Reynolds for her help in preparing this manuscript.

Bibliography

1. U.S. Public Health Service. Responsibility of PHS awardee and applicant institutions for dealing with and reporting possible misconduct in science. 54 Federal Register 151; 42 CRF Part 50 (8 August 1989), 32446-51.
2. Panel on Scientific Responsibility and the Conduct of Research. *Responsible Science: Ensuring the Integrity of the Research Process, I.* Washington DC: National Academy Press; 1992.
3. Dresser R. Defining scientific misconduct: the relevance of mental state. JAMA 1993; 269: 895-97.
4. Office of Science and Technology Policy. Federal policy on research misconduct. 65 Federal Register 235; (6 December 2000), 76260-64.
5. Wilcox LJ. Authorship: the coin of the realm, the source of complaints. JAMA 1998; 280: 216-7.
6. Little D, Rayner M. Ethical evaluation of misconduct cases. Research Conference on Research Integrity. November 18-20, 2000.
7. Meyer WJ. Data Management: recording, retention, access and ownership. In: David TP, editor. *Management of Biomedical Research Laboratories: A National Conference Proceedings*. Tucson, AZ: University of Arizona Press; 1999. p. 173-78.
8. The University of Texas Medical Branch Web Page accessed at *http://www.utmb.edu/avpr/integrity/memo.htm.*

*Whistleblowers in Environmental Science, Prevention of Suppression Bias, and the Need for a Code of Protection**

Elihu Richter, Occupational and Environmental Medicine, Hebrew University School of Public Health and Community Medicine, Jerusalem, Israel

Colin L. Soskolne, Department of Public Health Sciences, University of Alberta, Edmonton, Canada

Joseph LaDou, International Center for Occupational Medicine. University of California School of Medicine, San Francisco, USA

Tamar Berman, Occupational and Environmental Medicine, Hebrew University School of Public Health and Community Medicine, Jerusalem, Israel

Keywords: Code of protection, Environmental science, Suppression bias, Whistleblowers

Suppression bias is the distortion in the estimate of findings on hazard and risk inimical to special or national interests, and is well known (1-4). The direct and indirect repercussions of suppression bias are issues of direct importance not only to environmental scientists and health and safety professionals, but also to the public itself. These repercussions raise questions as to the adequacy and degree of protection provided by professional organizations, research institutions, and the legal system against such suppression bias.

Suppression bias is rooted in the way societies react to troublesome information, as we know from the tradition of shooting the messenger of bad news. The trial of Socrates served as the classic case study of the risks to messengers. The jurors of Athens, a city besieged from without and insecure from within, convicted Socrates and sentenced him to death for corrupting the morals of the youths of Athens (5-6). Legal scholars have pointed out that Socrates would be convicted by a modern jury for the same reasons that he was convicted by the jury in Athens: his teachings undermined order, stability, and state security. For Athenians, there was a Benthamite rationale for putting Socrates to death: silencing him was necessary to preserve the greatest good for the greatest number in a society weakened by external wars and internal divisions (7).

Environmental scientists and occupational health and safety professionals measure and report health risks from exposures to toxic and physical agents so that preventive measures can be put into

* An earlier version of this paper was published in the Int Journ of Occ and Env Health, 7:68-71, 2001, based on a presentation at the Collegium Ramazzini Annual Meeting in Carpi Italy on October 29 2000. We thank the Councilors and members of the Committee on Ethics and Philosophy of the International Society for Environmental Epidemiology for advice and encouragement, the International Student Foundation and the Collegium Ramazzini for financial support, and Dr Herbert Levine (Bethesda MD), Mr Ron Dror (Tel Aviv) and Dr Richard Laster (Jerusalem) for their helpful insights.

Corresponding author: Elihu D. Richter, Hebrew University-Hadassah School of Public Health and Community Medicine, POB 12272, Jerusalem, Israel, 972-2-6758147 (voice), 972-2-6784010 (fax), elir@cc.huji.ac.il.

effect. We define epidemiologic messengers, or whistleblowers, as persons who are subjected to harassment, lawsuits, ostracism, job loss, loss of funding, intimidation, abuse, threats, or even force after reporting such risks, or are prevented from investigating or reporting risks altogether.

In most scientific fields, the rewards go to investigators who report «positive findings». But in the environmental sciences, the situation is the opposite. In environmental and occupational medicine, and in epidemiology and related disciplines, "positive" findings about hazards and risks are threatening to powerful interests. Investigators who study or report these risks are therefore at increased risk for harassment by the very nature of their work.

Ultimately, suppression of information about hazards and their health risks may itself become hazardous to public health. There has not been sufficient recognition of the possibility that such pressures may serve to deter investigation or assessment of health risks from exposures, and thereby delay or block the implementation of preventive measures. So far, there have been few systematic efforts to examine the impact of such pressures on the direction, content, and work output of environmental epidemiologists, physicians in occupational medicine, and other scientists. Nor has there been sufficient attention as to how to respond to these pressures.

Methods

This paper reviews past reports and summarizes work now being carried out by the ISEE Committee on Philosophy and Ethics and the Collegium Ramazzini. This work documents episodes of harassment of environmental scientists and episodes of responding to requests for assistance from environmental messengers subject to harassment. We also make recommendations for future action by governmental organizations, which define standards for research policy.

Findings

In the 1980's, the United States Environmental Protection Agency (EPA) published a document which described the hazards unique to environmental scientists and the forms of harassment to which they may be subject. It made the point that harassment is most likely directed at younger or less well-known scientists, employees of government or industry, or members of the exposed population itself in settings where protection of human rights is weak. However, information is not readily available on the degree to which this or other Federal agencies defined institutional responsibilities to protect investigators from external or internal harassment.

The context and content of the problem

Martin (8) has listed the five methods of suppression bias. These are: (a) preventing creation of data (b) controlling, (c) blocking, (d) distorting data, and (e) attacking researchers. This simple list shows that using harassment to block dissemination of data on hazard and risk and attacking researchers who report such findings are only part of a syndrome of suppression bias, leading to what is known as lamppost toxicology or epidemiology. Martin and Deyo have reviewed the driving forces, context and methods of harassment of epidemiologic messengers or whistleblowers, and have provided case studies (1, 2, 8).

The reported distribution of the problem: sentinel episodes

Does suppression bias deter the prompt detection, reporting and prevention of hazard and risk? If so, is this bias systematic, episodic, or sporadic and what are its distributions and determinants? The details of whistleblower harassment are not frequently publicized (9), but below we present a list of episodes that have come to light in the past years from reports gleaned from the professional and lay literature, and from our own direct contacts.

Cases of suppression by a governmental institution

- Cate Jenkins, an environmental scientist with the US EPA, claimed that chemical industry studies had consciously minimized the hazard of dioxin (10-11). She received a written reprimand for writing down what she knew about the history of the dioxin incinerator regulations (12-13), and was transferred from her position.
- Omar Shafey, an epidemiologist in the Florida State of Health, was forced to leave his position after publishing an epidemiologic report on complaints of acute illness in residents exposed to drift from aerial spraying of malathion, used to control the Medfly

(14).

- Desi Mendoza Rivero, a Cuban physician, was imprisoned after he issued statements regarding an epidemic of dengue fever (15).
- Grigory Pasko and Alexander Nikitin, government scientists in Eastern Europe, were accused of treason and subjected to physical abuse after they reported dangers from nuclear waste in Murmansk (16-17). From newspaper reports, it appears that Pasko's subsequent acquittal was reversed. (17)
- Melvin Reuber, a toxicologist at the Frederick Cancer Research Facility in Maryland (which is part of US National Cancer Institute) studied links between pesticides and cancer. As a result of his studies, he is one of the world's leading critics of pesticides. In 1981, he was subjected to an attack on his work and his credibility that shattered his career (18-19).
- In the United Kingdom, a Health and Safety Executive (HSE) memo indicates that several researchers and health and safety activists who exposed poor health and safety practices were targeted for special surveillance (20).

Cases of suppression by an academic institution

- John Coulter, a medical researcher at the Institute of Medical and Veterinary Science in Adelaide, South Australia was dismissed from his post after releasing a report that ethylene oxide was mutagenic (21).
- Robert van den Bosch of the University of California, Charles Lincoln of the University of Texas, and Robert Fleet of Texas A&M University all suffered abuse because of their research on the hazards of pesticides (22).
- David Kern, an occupational physician and epidemiologist at Brown University Medical School, received notice that his tenure would not be renewed and his clinic closed after he reported numerous cases of interstitial lung disease in nylon flockers at Microfibres (23).
- In Israel, Dr Jerome Westin was greylisted for any governmental or academic appointments after publishing findings on massive contamination of the nationwide milk supply with organochlorines (24).

Cases of suppression by industry

- In the 1940's, Randolph Byers, the Harvard pediatrician, was sued for defamation and damages by the Lead Industries Association for publishing findings on brain damage from acute lead poisoning in children from nibbling paint chips (25-26).
- Doug Johnson, a safety specialist for Tatitlek, Chugach, and Chenega Corporation in Alaska was fired after raising environmental concerns regarding Alyeska's oil spill response program in Prince William Sound (27).
- Myron Mehlman, a Mobil Oil Corporation toxicologist, was fired after advising a Mobil subsidiary in Japan to stop selling gasoline with hazardous levels of benzene, a known carcinogen (28).
- Alexandra De Blas of Australia was threatened with a suit for defamation by a mining company when she attempted to publish a thesis about environmental impact of its operations. (29).
- Dr Yoram Finkelstein, an Israeli neurotoxicologist with important publications on organophosphates and lead, is currently the target of a SLAPP (Strategic Lawsuit against Public Protestors) lawsuit for libel after writing a medical opinion on the health risks from emissions of hexavalent chromium, Cd, lead, Ni, and other pollutants from an aluminum foundry (30).

Survey Results

At the Annual Conference of the International Society for Environmental Epidemiology (ISEE) held in 1999 in Greece, the Committee on Ethics and Philosophy distributed a questionnaire to the delegates. Out of 10 individuals who completed the questionnaire, five reported harassment following publication of research findings on health risks from environmental exposures. The following is a brief description of these cases:

- Male MD, age 47, a scientist in a major Cancer Institute in Italy, experienced ostracism after publishing findings on asbestos exposure in a petroleum refinery and lung

cancer.

- Female MD, MPH, age 60, was threatened with loss of her job after publishing findings on TCDD exposure and cancer.
- Male MPH, PhD., age 53, experienced ostracism and the threat of job loss after publishing findings on cancer mortality in Vietnam veterans exposed to Agent Orange.
- Two Female MD, investigators age 59 and 47, experienced both ostracism and confiscation of data after publishing findings on ethylene oxide exposure and breast cancer.

Pressures on institutions

Deyo et al have reviewed Congressional harassment of the CDC Injury Prevention Unit following its epidemiologic work on impact of gun control laws on violent deaths (2).

Actions to date:

The International Society for Environmental Epidemiology (ISEE) Committee on Ethics and Philosophy and the Collegium Ramazzini Committee to Protect Whistleblowers are working in parallel to provide moral and professional support to whistleblowers (31). The ISEE has already developed procedures designed to provide an international service of advice, referral, and support for environmental whistleblowers which was first presented in a special workshop at the ISEE International Conference in Athens in 1999 (not far from the site where Socrates was convicted.) The Collegium Ramazzini is now doing the same, and is planning to expand media reporting of whistleblower harassment, with particular attention to occupational medicine professionals in developing countries. The aim of both professional societies is to establish systems for monitoring and reporting harassment and abuse of whistleblowers, and to offer support and assistance should it be requested.

In 1996-97, before ISEE developed these procedures, it reacted to two situations in which investigators were subject to political pressures resulting from the publication of their findings. In the case of Dr. Herbert Needleman, ISSE sent a petition signed by many of its members to the University of Pittsburgh asking that its review of the validity of his findings on the effects of low level lead exposure on intelligence scores, behavior, and mood status be insulated from outside pressures and be governed by the criteria used for peer review. In the second case, Professor Viel from France reported to the Ethics and Philosophy Committee being the target of job threats following publication of papers in the British Medical Journal on risks for cancer among seashore residents living near nuclear power plants. This investigator also reported pressures from the nuclear industry to reveal the identity of individuals whose health records were part of an epidemiologic study. The Ethics and Philosophy Committee convened an ad hoc subgroup, under the late Professor John Goldsmith, one of its founding members, which communicated with Professor Viel, and offered to provide moral support for the issues raised. In both the Needleman and Viel cases, the issues of concern were resolved, but it is not known whether and to what degree ISEE's response played a role. Both Needleman and Viel are well-known senior investigators who published their work in prestigious journals. Their situations are exceptions to the rule that most whistleblowers do not have the protection of status and seniority, their findings or warnings may not be particularly original, and they may be prevented from either from publishing their findings or completing investigations in progress.

Through 2001, ISEE has responded to two cases, that of Yoram Finkelstein and Omar Shafey, and is working on a third, that of a pathologist sent to prison in Belarus.

Discussion

The case studies above provide support for the hypothesis that powerful governmental, military, economic, and political interests are often the driving forces and the sources of legal and illegal harassment of environmental messengers and, at times, the institutions they work for. But most of the case reports are from Western countries with developed research cultures and codes for the protection of human rights. The high-risk settings for exposure to pressures against environmental scientists are those where research is most needed, i.e., where exposures and risks are severe, where there are few environmental scientists, and occupational safety and health is not properly regulated and enforced by law. The risks are increased where legal safeguards for human rights are weak, and where access to a free press is blocked.

Yet, data are not readily available to examine

the working hypothesis that the exposure settings in which scientists are at greatest risk for threats, harassment, and legal pressure are those in which they are most needed. Africa, Latin America, Asia, the Mid-east and Eastern Europe are the regions of the world with the worst environmental and occupational health problems, the fewest environmental scientists, and the weakest safeguards to protect the rights of investigators. The situation is probably the worst for physicians working in occupational medicine who serve remote populations, given their relatively low status on the professional totem pole. In many of these countries, the situation for environmental scientists parallels the situation with regard to human rights, and suppression bias, like poor working conditions, is accepted as part of the normal research environment. It therefore stands to reason that in these regions, the absence of information on harassment of researchers can almost be said to be evidence of the effectiveness of suppression bias as a deterrent to investigation of environmental hazards. So far, neither the ISEE nor the Collegium Ramazzini have received requests for help from these settings.

In the developed countries, we need to ask whether a more subtle institutional form of suppression bias could be taking hold. Academic institutions are entering into strategic business alliances, most often with biotechnology and pharmaceutical firms (32). The close ties between university and business are a frontal assault on the last vestiges of "academic freedom" of the faculty members. Moreover, the diminishing role of governments in funding public health research causes academic institutions to pursue corporate funding. This trend furthers the alliance of university and business, and increases the likelihood of suppression bias.

We suggest that suppression bias and the occurrence of environmental hazards circularly reinforce each other. Alibek has pointed out that in the former Soviet Union, suppression of information on health hazards to personnel and the environment from activities related to weaponizing bacterial and viral organisms for bioterrorism led to a scenario in which safety was jeopardized over and over again in the name of national security. He described a scenario in which suppression bias *resulting* from the harassment of epidemiologic messengers endangered public health (33).

Institutional safeguards against harassment in environmental science

Until now, research on ethics in environmental epidemiology has focused on the obligations of individual researchers to comply with norms of truth and not engage in scientific misconduct (34-35). But there has been insufficient discussion of the obligations of institutions to protect their workers and their findings from external harassment when their findings are embarrassing, costly, or threatening to powerful interests. Such harassment serves as a deterrent to investigating and reporting information about hazards and risks.

Measures to protect messengers in environmental and occupational epidemiology should be required of grant recipients of research contracts around the world and should become a worldwide institutional norm.

Messengers can be wrong

The statements made by epidemiologic messengers on the presence of a hazard or risk may be right or they may be wrong. We suggest that pressures, harassment, and abuse are no substitute for access to the peer-review process. At the same time, there is the need to be concerned about pressures on this peer review process by new trends in the academic world to forge alliances between industrial or technological interests and the research community.

What Next?

Professional societies derive their legitimacy from their mission in promoting the public good. Investigation and reporting environmental hazards and their risks are essential to prevent damage to public welfare. As we noted at the outset, the protection of epidemiologic messengers derives from the primacy of protecting public health. Ironically, Benthamite rationales —stretched somewhat—could have served to acquit Socrates were it to have been shown that his teachings were necessary for protection of the greatest good for the greatest number, or more fundamentally, for the health and welfare of all individuals, in keeping with traditions of the sanctity of preserving individual human life.

Organizations concerned with ethics in

science in recent years rightfully called attention to the need to establish rigid standards for preventing scientific misconduct by individuals. The first generation of work on ethics in research focused on setting standards, procedures and codes of practices which defined responsibilities of individual scientists at all levels, to work according to codes of truth, quality assurance and quality control, precision and accountability (36-37). This first generation of work addressed issues raised by whistleblower scientists who drew attention to scientific misconduct in the laboratories of their superiors. These episodes of misconduct led to the distortion of findings, failures in quality assurance and quality control, and lapses in precision and accountability. The issue at hand now is standards for preventing institutional misconduct. There has been no parallel effort of equivalent force to enact standards that prevent misconduct by institutions—be they the scientist's employer or other bodies—which results in harassment of epidemiologic messengers.

We suggest that failure to ensure proper access to independent peer review insulated from internal and external pressures is a form of institutional misconduct. The same statement applies to failure to provide protection against legal harassment, such as occurs with SLAPP lawsuits. Therefore, the second generation of work in ethics and scientific integrity has to deal with a new and different set of problems. These pertain to the need for standards, procedures, and codes of practice that define the responsibilities of institutions and organizations to prevent the harassment of individual environmental scientists who either attempt to investigate or report findings on hazard and risk which go against powerful interests that could be damaged by such information.

The issues at hand here are not quite the same as those having to do with investigations of scientific misconduct, i.e., the falsification or fabrication of research results. In investigations of scientific misconduct, there is a more or less level playing-field for right and wrong: the peer reviewed literature and its well elaborated codes and norms for evaluating scientific evidence. In the case of whistleblowing in environmental and occupational epidemiology, the problem is to promote access to this level playing field, and to ensure that the playing-field is indeed level. There is a need to ensure that outside interests, often commercial, economic or political, do not obstruct access to or distort the process of peer review.

There is a need to recognize a dissonance between the emphasis of the first generation of ethics on promotion of research integrity and that of the second on prevention of suppression bias. Often there is a two-stage scenario in which investigators—or officials in need of a rapid estimation of hazard or risk— are first blocked from access to the exposed populations and relevant databases, and then their reports are disqualified because they are incomplete, imperfect or imprecise. In short, the very criteria used to define the quality of investigation may serve as barriers to reporting its substance. This situation —in which being precisely wrong is considered preferable to being approximately right—is the classic scenario of delay.

One form of harassment of environmental epidemiologists and other investigators is to subject their databases and records to a battery of legal subpoenas. If transparency is our norm, it is hard to fault such challenges. However, such subpoenas pose potential challenges to the privacy of research on individuals, and may serve as a deterrent to their giving permission to use data on individual exposure and risk. But, in the case of environmental epidemiology and related fields, the ultimate form of harassment is to deny the investigator access to databases, so as to prevent a complete investigation. In epidemiology, in particular, barriers to accessing databases on denominators can be particularly devastating, because they effectively forestall precise calculations of risk. Such barriers, by delaying or impeding investigations, may not only block research, but they permit the persistence of situations hazardous to the health and safety of the populations themselves. We see use of the term "sound science" to disparage attempts to make do with limitations of estimates of risk based on studies not meeting "gold standard" requirements because data sets may not be complete. (38).

A second form of harassment is lawsuits for libel. To address this hazard to environmental scientists, there is a need to explore the use of insurance policies modeled after those available to writers. Grants to environmental scientists should include budgeting for such insurance.

Conclusions

Until now, there has been no watchdog address

for environmental and occupational epidemiologists to which to turn for assistance.

We suggest that major granting agencies follow the lead of ISEE and the Collegium Ramazzini in protecting environmental scientists from harassment. We call for studies on the impact of harassment of research scientists on the detection and prevention of health risk. We call for the development and use of codes for protecting environmental scientists from harassment when they are engaged in this mission. We recommend that measures to protect messengers in environmental and occupational epidemiology be required of recipients of research grants or contracts around the world. These codes should become a worldwide institutional norm. Codes that protect epidemiologic messengers in environmental and occupational medicine will serve also to protect the public health.

Acknowledgements

Councillors of ISEE and Collegium Ramazzini, which provided a grant for EDR's work, and an anonymous donor for TB's work.

Bibliography

1. Martin, B (1981) The scientific straightjacket: the power structure of science and the suppression of environmental scholarship. The Ecologist. Vol 11(1): 33-43.
2. Deyo, RA, Psaty BM, Simon G, Wagner EH, Omenn GS (1997) The messenger under attack- intimidation of researchers by special-interest groups. New England Journal of Medicine. 366 (16): 1176- 1180.
3. Relyea, HC (1994) Silencing Science: National Security Controls and Scientific Communication. Norwood, New Jersey: Ablex Publishing.
4. Hess, DJ (1999) Suppression, bias, and selection in science: the case of cancer research. Accountability in Research. 6: 245-257.
5. Plato (1963) The trial of and death of Socrates. New York: Heritage Press.
6. Dror R. Notes on the trial of Socrates. Based on conversations with the late Professor Yeshayahu Liebowitz (personal correspondence, June 1999).
7. Bentham, J (1907) An introduction to the principles of morals and legislation. Oxford: Clarendon Press.
8. Martin, B (1999) Suppressing research data: methods, context, accountability, and responses. Accountability in Research, Vol. 6: 333-372.
9. Bahr, MH. "Violations of academic freedom: official statistics and personal reports", Social problems, 14, 1967, pp. 310-320.
10. Cate Jenkins, "Newly Revealed Fraud by Monsanto in an Epidemiological Study Used by EPA to Assess Human Health Effects from Dioxins", EPA memorandum to Raymond Loehr, Chairman, Executive Committee, EPA Science Advisory Board, February 23, 1990. As cited in Rachel's Hazardous Waste News #171. Dioxins and Cancer–Fraudulent Stories. March 7, 1990.
11. Jenkins C. (September 3, 1991) Affidavit: Recent Scientific Evidence Developed after 1984 Supporting a Causal Relationship between Dioxin and Human Health Effects. Full text of the affidavit is available online from Agent Orange Victims, Families, and Friends in Touch. http://softvision.netfirms.com/ao_vets/.
12. Memo from Gerald H. Yamada, Office of [EPA's] General Counsel, to Catherine L. Jenkins, dated February 26, 1993. As cited in Rachel's Hazardous Waste News #328. An update on two key incinerator battles. March 11, 1993.
13. Jenkins v. U.S. Environmental Protection Agency, 92-CAA-6 (ALJ Dec. 14, 1992). Available online, http://www.oalj.dol.gov/public/wblower/decsn/92caa06a.htm.
14. http://www.pmac.net/pesticide_wars.html. Pesticide wars: the troubling story of Omar Shafey.
15. http://www.web.amnesty.org/ai.nsf/Index/AMR250281997?OpenDocument&of=COUNTRIES%5CCUBA. Dr Desi MENDOZA Rivero, Cuba.
16. http://www.web.amnesty.org/ai.nsf/index/EUR460421996 (1996) Federal Security Service (FSB) versus Prisoner of Conscience Aleksandr Nikitin: Return to Soviet Practices.
17. www.rsf.fr/uk/html/europe/cplp01/lp01/261201.html - Reporters wtihout Borders, Dec 2001 http://www.internews.ru/ngo/pasko/indexe.html (1997) Grigory Pasko sentenced to four years in prison.
18. Martin B (1993) Stamping Out Dissent. Newsweek. 26 April: 49-50.
19. Pesticide & Toxic Chemical News (1981) Dr. Mel Reuber, Pathologist, Gets Sharp Censure, Warning from his Supervisor (25 April): 22-23.
20. Watterson A (1998). Occupational health and safety policy in the UK: who is driving the agenda with the "greylisting" of health and safety activists? Int J Occup Environ Health 4(4):275-77.
21. Coulter, J (1980) Memorandum re use of ethylene oxide as a sterilant in the S.P.F. laboratory at Northfield, 16 April.
22. van den Bosch, R. (1978) The Pesticide Conspiracy. Garden City, NY: Doubleday.
23. http://home.cwru.edu/activism/kern.html. Conflicts of interest: the case of David Kern.
24. Westin, JB and Richter, ED (1990) The Israeli breast-cancer anomaly. Ann NY Acad Sci 609:269-279.
25. Byers, RK and Lord, EE (1943) Late Effects of Lead Poisoning on Mental Development, American Journal of Diseases of Children. Vol. 6: 471-494.
26. Smythe, L (1980) letter, New Scientist, 85: 775-776. As cited in reference Martin, B (reference 1).
27. http://www.whistleblower.org/www/998news.htm#Oil. Big oil and the battle to protect the environment. Government Accountability Project, Fall 1998.
28. Sternberg, S (1996) An Unusual Case: Law Overseas

Myron Mehlman's case against Mobil Corp. attempts to apply New Jersey statute to incident in Japan. The Scientist 10(1).
29. De Blas, A. (1994) Environmental Effects of Mount Lyell Operations on Macquarie Harbour and Strahan. Sydney: Australian Centre for Independent Journalism, University of Technology, Sydney.
30. http://www.jpost.com/Editions/2000/03/20/News/News.4270.html. Siegel-Itzkovich, J (2000). Environmental group aiding expert in libel suit. Jerusalem Post Online Edition. March 20, 2000
31. http://www.iseepi.org/ethguide.htm#victimized. ISEE support for victimized colleagues.
32. Abraham, J (1995). Science, Politics and the Pharmaceutical Industry: Controversy and Bias in Drug Regulation. London: ICL Press.
33. Alibek, KW and Handelman S (2000). Biohazard. Dell Publishing Company.
34. Nigg, HN, Radulesc G (1994). Scientific Misconduct in Environmental Science and Toxicology. JAMA 272:168-170.
35. LaFollette MC. Stealing Into Print: Fraud, Plagiarism, and Misconduct in Scientific Publishing. Berkeley: University of California Press; 1992:294.
36. American College of Epidemiology (2000). Ethics Guidelines. Ann Epidemiol. 10: 487-497.
37. http://ori.dhhs.gov/html/policies/oripolicies.asp. Office of Research Integrity.
38. Ong, EK and Glantz, SA (2001). Constructing "sound science" and "good epidemiology": tobacco, lawyers, and public relations firms. Am J Public Health 91(11):1749-57

ISEE Ethics Committee Epidemiologist Whistleblower/Messenger Questionnaire:

1. Personal status

- ISEE Member? Y/N _____; ISEA Member? Y/N_____; Age _____
- Gender M/F _____
- Personal Status: M, S, D, W _____
- Children (Give no___)

Education	**From**	**To**	**Where* (see Code)**	**Code:**
Undergrad	____	____	_____	America: NA, CA, LA
MD	____	____	_____	Europe: WestE, Med,
MPH/MSc/MS/MA	____	____	_____	EastE Mideast: ME
PhD/DPH	____	____	_____	Africa: WA, EA, SA Asia:
Post Doc	____	____	_____	Ind, CentAsia, Jp, Ch,
Residency Spec	____	____	_____	SEA Oceania: Aus, PI

2. Currently Employed

Where? _____ see code above
By: Govt Civilian Military Police (Circle one)
Level: National Regional/Province/District/Municipal (Circle one)
University/College _____
Independent research institute
Foundation _____
Trade Union NGO Self Employed
Industry/Corporation: If yes? _____
Multinational Y/N _____
Other_____

3. Tenured or permanent? Y/N
Rank (Univ): Prof Sr Lect/Lect__ Asst__Other_____

4. Research/salary funded by: (Circle correct answer)
Government
Industry
Foundation Other
No funding

5. Harassment: Following publication of research findings on health risks from environmental exposures, have **you** ever experienced:

Ostracism Y/N	**Demotion** Y/N	**Criminal investigation /Prosecution/Trial** Y/N
Confiscation of data Y/N	**Loss of job** Y/N	**Physical threats** Y/N
Threat of loss of job Y/N	**Threats of lawsuits** Y/N	**Physical Attack** Y/N
Transfer Y/**N**	**Lawsuits** Y/N	**Imprisonment** Y/N
Other		

How many **other** co-researchers were there? ____ Did **they** experience any of the responses? Y/N

6. Research on specific problem which lead to episode(s) of harassment or threat or abuse: Years during which research carried out: From_____To_____

Was this research on the hazard/risk published in:
Peer reviewed journal (sited in SCI CIT INDEX)___
Masters thesis___
Doctorate___
Internal document of organization in which you were then employed/studied?
Professional society___
Non peer-reviewed journal____
Other____
Date of publication?_____ Would you be able to provide the Citation? *Leave blank if you wish*
_____

7. Response

7a. Did you receive **assistance** after being subject to any of the above problems? Yes___
No____
7b. If yes, from: Individual colleagues_____ Superiors____ Professional societies ___ NGO's inside country_____ Journalists/Media_____ Lawyers or legal aid groups____Colleagues outside country____ NGO's outside country_____Family_____Other_____

8. Publication If findings were not published, were you **prevented** from submitting findings on health risks on a hazardous exposure/risk for publication in a peer reviewed journal? Yes___ No____
OPTIONAL_____

9. Findings: Could you summarize the findings you discovered/reported for which you were harassed?

Study design (Cohort, CC, Prev, TS, Other)	**Pop(s) / N**	**Exposure(s)**	**Outcome**	**RR/OR**	**Reference**
_____	_____	_____	_____	_____	_____
_____	_____	_____	_____	_____	_____

10. In retrospect, were your findings: understated?____a proper assessment?____overstated?_____ **For further information**: *http//:www.iseepi.org/ethguide.htm*

II. Teaching

6. Training in the Responsible Conduct of Research

Influencing the Moral Dimensions of Professional Practice: Implications for Teaching and Assessing for Research Integrity

Muriel J. Bebeau, Center for the Study of Ethical Development, University of Minnesota, USA

Key Words: Assessment, Moral development, Professional ethics, Research ethics

This paper will present implications for teaching and assessing for research integrity from 20 years of experience designing and assessing ethical development in the dental profession. Data sources for the implications include: 1) pretest/posttest data for 18 cohorts of dental students who completed a well-validated ethics program; 2) pre/post assessments of 28 practitioners referred by a licensing Board[1] for individualized ethics instruction because they violated the State Dental Practice Act; and 3) efforts in several professions to influence moral judgment development.

After pointing out some of the features of the Minnesota ethics program, the program's theoretical foundations (e.g., the processes of morality) are described. Each process suggests research questions that motivate inquiry and assessment methods that were developed or used to investigate the research questions and to gather evidence on program effectiveness. The paper continues with a summary of data supporting the independence of the component processes and a discussion of the ongoing search for behavioral indicators that could provide the "acid test" for the model. The paper concludes with a discussion of the implications for the teaching and assessing for research integrity.

Special features[2] of the curriculum include: 1) 43 contact hours distributed over four years; 2) required attendance and participation; 3) small group instruction—using dilemma discussion and role-play; 4) an emphasis on student performance, self-assessment and personalized feedback; 5) use of validated assessment methods that are checked for reliability; 6) involvement of high status professionals (in measurement development and feedback); and 7) involvement of faculty in the teaching. Thus, the curriculum isn't a one-shot intervention, nor is it the isolated property of one instructor.

Theoretical Foundations

The ethics curriculum, for students and referred practitioners, is designed to promote functional processes that give rise to morality: 1) ethical sensitivity; 2) moral reasoning; 3) moral motivation and commitment; and 4) ethical implementation (1). Moral failing is conceptualized as the result of deficiencies in one or more of the processes. Rest's Four Component Model of Morality, operationally defined below, is a substantial departure from much of the work in psychology that arbitrarily divides moral functioning into affects, cognitions, and behaviors (2).

The Four Component Model of Morality

Early in the cognitive developmental research program initiated by Kohlberg, he noted that, in

Corresponding author: Muriel J. Bebeau, Department of Preventive Sciences, School of Dentistry, University of Minnesota, 515 Delaware S.E., Minneapolis, MN 55455, 612-625-4633 (voice), 612-626-6096 (fax), bebea001@umn.edu.

addition to moral judgments, other processes were important to the production of moral behavior (3). Rest made these processes more explicit in what he called the Four Component Model of Morality (1). Starting from the question: how does moral behavior come about, Rest suggested that the literature supports at least four component processes, all of which must be activated in order for moral behavior to occur. These include:

1. Moral sensitivity (interpreting the situation as moral)

This process highlights the idea that moral behavior can only occur if the individual codes the situation as moral. Specifically, Component 1 focuses on the various actions that are available and how each action might affect the self and others.

2. Moral judgment (judging which of the available actions are most justified)

This is the process that Kohlberg emphasized. Here the focus is on judging which of the various options are the most ethically justified. Further, the job of a psychologist and educator is in sketching out how the justification process develops and under what conditions these processes inform real-world choices.

3. Moral motivation (prioritizing the moral over other significant concerns)

Less understood than the other processes, the main concern of Component 3 is, "why be moral." The model acknowledges that individuals have a number of legitimate concerns that may not be compatible with the moral choice: for instance, career pressures, established relationships, idiosyncratic personal concerns, among many others. Some of the most notable lapses of ethical behavior in the professions can be attributed to low priority placed on the moral, even when the moral choice is very well understood.

4. Moral character (being able to construct and implement actions that service the moral choice)

Component 4 represents the processes by which one constructs an appropriate course of action, avoids distractions, and maintains the courage to continue.

It is important to notice that the model is not conceived as a linear problem-solving model. For example, moral motivation may impact moral sensitivity, and moral character may constrain moral motivation. In fact, Rest (1) makes clear the interactive nature of the components. Further, the Four Component Model assumes that cognition and affect co-occur in all areas of moral functioning. Thus, moral action is not simply the result of separate affective and cognitive processes operating in interaction, as suggested by traditional models of moral function that focus on three domains—cognitions, affects and behavior (4, 5). Instead, each of the four components are mixes of affective and cognitive processes that contribute to the component's primary function (e.g., identifying a situation as moral). Bebeau, Rest, & Narvaez suggest that researchers focus attention on identifying processes as they contribute to moral action, rather than attempting to understand moral actions from a starting point defined by arbitrarily dividing moral functioning into affect, cognitions, and behavior (2).

The debate on the usefulness of a psychological theory of morality, that has its foundation in the work of Lawrence Kohlberg, is addressed in "Postconventional Moral Thinking" (6). This paper presents a theory of moral judgment development that is not grounded in a particularistic moral theory—as was Kohlberg's—but is grounded in empirical evidence illustrating that as individuals develop, so do the basic understandings they bring to resolving complex moral problems. Such findings are of importance to ethics education in general, as the goal of ethics education is, simply put, to promote ethical development. The authors contend that their findings will be of particular importance to research ethics educators because of their interest in promoting critical thinking about responsible research conduct (6). In the past, ethicists working in the professions questioned the usefulness of a moral development theory (and related measures) that favored a particular moral theory, observing that practitioners working on real problems often developed well-reasoned solutions without regard to a particular theory or even to principlism as a way of arriving at moral judgments (7).

By amending a theory of moral judgment development to make it congruent with advances in moral philosophy, the authors hope to counter current views of the obsolescence of moral psychology and support more interdisciplinary collaboration in the design and evaluation of moral education programs. Further, a more enlightened view of the role of tests of moral judgment development should enable educators

to put such tests to more appropriate use.

Besides drawing attention to a broader conception of postconventional moral thinking, the authors direct the reader's attention to a broader conception of morality, one that encompasses moral judgment, but that also addresses other aspects of moral functioning, including moral sensitivity, motivation, character, and competence. The Four Component Model of Morality has been a centerpiece for research activities at the Center for the Study of Ethical Development for nearly 20 years.

Educational Interventions Assessed in Terms of the Four Components

A program of research and educational development to investigate the usefulness of the model was initiated by Jim Rest and the author in the early 80s. Variations on these research questions motivated the inquiry: Can ethical sensitivity (or any of the other components) be reliably assessed? Do students differ in ethical sensitivity (or other components)? Can sensitivity (or other components) be enhanced? And, is ethical sensitivity distinct from other components?

The Four Component Model offers unique information and direction for educational development. First, it suggests profitable areas for measurement development. To claim that a program is effective in a broad sense, it seems reasonable to expect changes within each of the four components. For the dental curriculum, measures of each component were designed and validated, and data from them helped identify deficiencies to consider as the curriculum was designed. There are measurement models and methods for assessing each of the components (2, 8). These can be used as templates for assessment in various contexts.

Second, the model provided direction for instructional design for groups, as well as for individual referrals. For referred practitioners, deficiencies were noted in various components and were associated with particular moral weaknesses (9). Targeting specific deficiencies in an individualized instructional program proved to be an effective intervention strategy, resulting in substantially enhanced posttest performance.

Measures for the Components of Morality

Five measures are used to assess performance in the Dental Ethics Curriculum. A brief description of each measure and the findings are summarized as follows:

Component I: Ethical Sensitivity

The Dental Ethical Sensitivity Test (DEST)

The DEST (Form A or B) (10, 11) assesses the ability to recognize the ethical issues hidden within the professional problems dentists encounter in practice. Students' verbal responses to four audio-taped dramas are recorded and transcribed, and provided to the student and to a practicing dentist, who each apply the DEST coding scheme, then meet for personalized feedback. The validity and reliability of the DEST are reported in several studies, summarized in Bebeau (8) and Fravel and Bebeau (12). Briefly, the results support these conclusions: 1) Ethical sensitivity can be reliably assessed. Calibrated raters achieved item agreement ranging from 84.7 percent to 88 percent. Reliability estimates for individual cases ranged from .83 to .92; 2) Students and practitioners vary in sensitivity to ethical issues. Students at different levels of education in medicine and dentistry (physicians vs. technicians or dentists vs. hygienists) differed significantly, such that those with longer preparation showed higher levels of sensitivity. Further, the DEST is sensitive to institutional differences; 3) Women have a slight edge over men in recognizing ethical issues, but differences were not attributed to differential recognition of the care and justice issues; 4) Ethical sensitivity can be enhanced through instruction; 5) Ethical sensitivity is distinct from moral reasoning abilities. Correlations between the DEST and Defining Issues Test (DIT) posttest are consistently low (see later section for more detail); 6) Despite the stressful nature of the DEST assessment—responding on the spot to complex cases, having responses taped, transcribed, and sent to a practicing professional is a high-stakes examination—students value the assessment and feedback experience.

Component II: Moral Reasoning and Judgment

In this section, two measures are described: a well-established measure (DIT) and a newly-devised, context-specific test of ethical reasoning and judgment (Dental Ethical Reasoning and Judgment Test [DERJT]). In the case of the DIT,

the discussion will include findings from new analyses with new indices for three of the recent cohorts of dental students.

The Defining Issues Test

The DIT measures life-span development of moral reasoning and judgment (13). The DIT is the most widely used test of moral judgment development and is often used as an outcome measure for intervention studies, because it has an exceptional validation history.[3] Students read dilemmas, and then rate and rank the importance of each of 12 arguments to support their position. Confirmatory factor analysis of a mega-sample of over 44,000 subjects shows that items (arguments) cluster around three general moral schemas: Personal Interest, Maintaining Norms, and Postconventional schemas (14). Typically, researchers have reported scores in terms of the P score—the proportion of items selected that appeal to Postconventional moral frameworks for making decisions. The average adult selects postconventional moral arguments about 40 percent of the time, the average Ph.D. candidate in moral philosophy or political science about 65.2 percent of the time, the average graduate student 53.5, with the average college graduate at 42, and the average high school student at 31.8 percent.

Progress in moral judgment is developmental, and development proceeds as long as an individual is in an environment that stimulates moral thinking. College has a powerful effect on moral judgment development. McNeel's meta analysis of 22 longitudinal studies of liberal arts students estimates first year college students at 36, seniors at 46, estimating an effect size of .80 (15). Effect sizes of about 0.80 are among the largest effect sizes for many college impact variables that have been studied. In fact, effect sizes are higher for moral judgment than for the many cognitive and affective college outcome variables that have been studied (16). Yet professional schools (e.g., Veterinary Medicine, Medicine, Dentistry, and Accounting) are programs where one does not typically see gains associated with the educational program, unless the program has a specially-designed ethics curriculum (17). Further, for some students and some professions, programs actually seem to inhibit growth (18, 19).

Change in moral judgment can be attributed to the ethics curriculum (18). The average entering Minnesota dental student scores 46 (with cohorts ranging from 42 to 49 across the 15 classes tested). The average graduate selects postconventional arguments 51 percent of the time (with cohorts ranging from 47 to 55). Effect sizes vary across classes, with a range of .12 to .78, with an average of .43. For each cohort, scores tend to be normally distributed. For entering students, as many as 35 percent are not using postconventional moral schemas as often as the average adult, with about seven percent above the mean of philosophy and political science graduate students. Although we see an upward shift in the distribution at posttest, with 16 percent lower than the mean of the average adult, and 20 percent above the mean of philosophy and political science graduates; of particular interest are the proportion of students who showed no change or regressed from pretest to posttest. By classifying students' change scores into categories defined by the standard error of measurement (18), Bebeau reported that 44 percent of the 1,229 students who participated in the curriculum made moderate to highly significant gains, 40 percent showed no change, and 16 percent regressed on the P score (20).

New Indices and New Analyses of DIT Scores

Observations of what appeared to be regression in postconventional reasoning in our intervention studies prompted the validation studies, including development of an alternate form of the DIT and a reanalysis of moral education interventions that attended to several moral cognition variables derived from DIT scores (6, 14, 21, 22, 23, 24).

Moral Schema Profiles. Instead of relying only on the P score as a measure of pretest to posttest change, a profile showing the proportion of times a student rates was constructed to illustrate important items for each of three general schema: a Personal Interests schema (Kohlbergian Stage 2 and 3 items); a Maintaining Norms schema (Stage 4 items): and a Postconventional schema (Stage 5 and 6 items). Figure 1 illustrates how two profiles with similar P scores can reflect differing levels of moral judgment development. Examining profiles from students who did not show gains in DIT P scores from pretest to posttest (20) illustrates a substantial reduction on the Personal Interest schema coupled with an increase on the Maintaining Norms schema, without significant change on the Postconventional schema score. In fact, when the statistically significant pretest/posttest change for the 18 cohorts of students that participated in the dental curriculum was

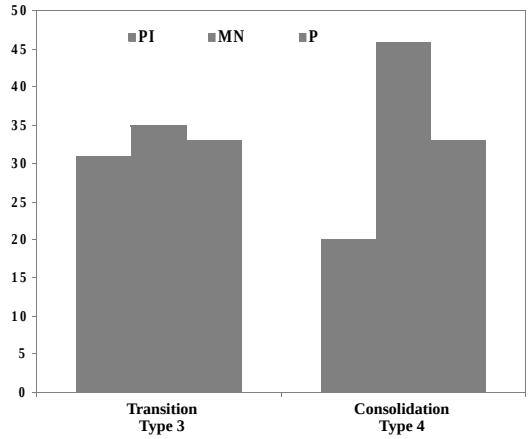


Figure 1. Moral judgment profiles illustrating similar P scores, but differences on other moral cognition variables.
PI = Personal Interests Schema
MN = Maintaining Norms Schema
P = Postconventional Moral Schema

reanalyzed, the reduction in the Personal Interests schema score appeared much greater and more consistent across cohorts than changes in P score. By focusing only on the P score, researchers may be missing change that is quite significant.

Consolidation/Transition. Figure 1 illustrates another variable to consider in describing change. When there is little evidence of discrimination among the schema-typed items, students are classified as transitional. A flat profile is viewed as a marker of developmental disequilibrium, or transition, since there is no evidence of a schema preference. A further discussion of this topic is addressed by Thoma and Rest (22). A pretest/posttest analysis of consolidation/transition status was conducted for 222 dental students (20), showing that nearly half the students (46.9%) were in a transitional status at pretest, whereas only 27.1 percent exhibited the transitional status at posttest.

Type. Profiles can further be classified by type (22), where type reflects both the predominant schema and the extent of its use. By reexamining several intervention studies reported in the literature, Yeap showed that Type provided a more illuminating description of change that occurred as a result of an intervention than relying simply on the P Score (24). A pretest/posttest analysis of six Types was also conducted for the 222 students reported above. Whereas the pretest responses were distributed among Types 3, 4, 5, and 6, 61.2 percent were classified at Types 5 and 6 (postconventional types), with the distribution peaking at Type 6. For the posttest responses, 75.8 percent were classified at Types 5 and 6, with 59.9 percent at Type 6. By way of comparison, Yeap reported college student samples peaked at Type 3.

These new analytical procedures may help to unravel some of the puzzles researchers have cited, where professional groups like Accounting and Auditing (19) seem to regress on moral judgment as a result of an educational program. Such analysis may clarify McNeel's findings that programs that are too careerist (focus narrowly on technicalities of beginning job performance) or too dogmatic (in closing off questioning and inquiry) inhibit growth in reasoning (15). Such findings would have implications for developing research integrity. Courses that focus narrowly on the rules of research conduct may focus attention on the minimal (legal) standards, rather than on aspirational standards for research integrity.

Tests like the DIT are valuable for assessing general reasoning that is a critical element of professional ethical development, but they may not be sensitive to the specific concepts taught in a professional ethics course—or indeed, in a research ethics course. The question (for educators) is often whether to teach specifically to the codes or policy manuals, or to teach concepts particular to a discipline—informed consent, intellectual property, conflict of interest, etc.

The Dental Ethical Reasoning and Judgment Test (DERJT)

The DERJT is a first effort to test application of context-specific concepts (taught in ethics courses) to real cases (25). The test is similar to the DIT, in that cases are presented followed by lists of action choices and justifications. The action and justification choices for each problem were generated by a group of Minnesota dental faculty and residents. The scoring key was developed by a group of "dental ethical experts." When taking the test, a respondent rates each action or justification, then selects the two best and two worst action choices, and the three best and two worst justifications. Scores are determined by calculating the proportion of times a respondent selects action choices and justifications consistent with "expert judgment." In validation studies, Bebeau and Thoma have seen clear expert novice differences (25). Further, scores for students, practitioners, and referrals appear to be normally distributed. In a

study comparing our graduates' responses to familiar vs. unfamiliar problems presented on the test, it appears that a good grasp of postconventional moral schemas is a necessary condition for transfer to new problems.

Component III: Motivation and Commitment

The Professional Role Orientation Inventory (PROI)

The PROI assesses commitment to privilege professional values over personal values (26, 27). Likert scales assess dimensions of professionalism that are theoretically linked to models of professionalism described in the professional ethics literature. The PROI scales, in particular the responsibility and authority scales, have been shown to consistently differentiate beginning and advanced student groups and practitioner groups expected to differ in role concept. By plotting responses of a cohort group on a two dimensional grid, four distinctly different views of professionalism are observed (26) and, if applied, would favor different decisions about the extent of responsibility to others. In comparing practicing dentists with entering students and graduates, our graduates consistently express a significantly greater sense of responsibility to others than entering students and practicing dentists from the region. This finding has been replicated for five cohorts of graduates ($\underline{n} = 379$). Additionally, the graduates' mean score was not significantly different from a group of 48 dentists, who demonstrated special commitment to professionalism by volunteering to participate in a national seminar to train ethics seminar leaders. A recent comparison of pretest/posttest scores for the Classes of 1997-1999 (20) indicates significant change ($p < .0001$) from pretest to posttest. Cross-sectional studies of differences between pre and posttest scores for a comparable dental program suggests that ethics instruction accounts for change.

To provide students or practitioners with individualized feedback on their role concept, an interpretive guide is provided enabling a respondent to sum his or her own scores on each scale, plot them on the two dimensional grid (one grid is provided for the authority and responsibility scales, one for the agency and autonomy scales), and then compare responses to their cohort. Descriptions of each of the models of professionalism are included to stimulate thinking about the model of professionalism that appears to be dominant for the individual. When the scales and interpretive guide are used in an educational setting, participants can compare and discuss items and challenge each other's thinking.

Developing a concept of role appears to require instruction and opportunities for reflection. At entry to professional school, Minnesota dental students do not illustrate a good understanding of key concepts of professionalism like service to society, or the priority of patient well-being, or the duty to self-regulation (8). But, even after participation in an instructional program in which students write an essay describing their perception of their professional role (the program is of demonstrated effectiveness and includes generous amounts of practice and feedback on performance), key concepts like self-regulation, service to society, and the basic duty to place patient's rights before self-interest are still frequently omitted or miscommunicated by as many as 20 percent of the students. The literature on concept learning has helped us see that when students have no functional schema for a particular concept, several educational experiences are required to instill a clear concept of the professional's role.

Whether instilling a clear idea of the professional's role will motivate students to place moral values over personal ones is a key question. The most direct evidence of a relationship between role concept and professionalism comes from the study of performance of the 28 members of the practicing community, referred for courses in dental ethics because of violations of the dental practice act. Although the practitioners varied considerably on measures of ethical sensitivity, reasoning, and ethical implementation, 27 of 28 were unable to clearly articulate role expectations for a professional (9).

Component IV: Moral Implementation (character and competence)

Shifting to the last component, character and competence, the authors have observed that guided practice changes the expectation of efficacy that is likely to change behavior. Role-playing builds competence and confidence in resolving thorny ethical problems, and skills in

communication and negotiation are necessary requisites of this competence.

A Professional Problem Solving Index

Problem-solving and role-playing performance scores are calculated for eight complex cases that present difficult human interaction problems (8, 20). Students are directed to prepare 1) an interpretation of the facts that must be addressed if the problem is to be resolved efficiently; 2) an action plan; and 3) a verbatim dialog to illustrate the implementation of the action plan. A checklist, prepared for each case, assures some uniformity in judging responses. Each response is reviewed by a peer and by the course instructor who provide written comments identifying the strengths and shortcomings of the assignment. As with other measures, scores are normally distributed and cohort differences are observed.

Independence of the Components of Morality

Rest's Four Component Model predicts the independence of the components (1). Prior studies have typically reported low to very low correlations between ethical sensitivity and moral judgment, but correlations among the other components have varied from very low to an occasional moderate correlation. Often sample sizes have been low, challenging the reliability of the estimates. Recently, Bebeau reported correlations between components for a larger sample (230 students) (20). Except for the expected moderate correlations (.46) between the DIT Pretest and Posttest and between the PROI Pretest and Posttest scales (.38), each measure appears to provide unique information about ethical decision making competence. Consistent with earlier studies, correlations are consistently very low between the DEST and the DIT, and between the DEST and other component measures (8). The exception is between the DEST and the DERJT justification score, where there appears to be some overlap between the two tests ($r = .28$). Also consistent with earlier reports (27), there appears to be some low to moderately-low relationship between the PROI Responsibility Scales and the DEST and DIT.

The Continuing Search for Behavioral Indicators

Several attempts have been made to show the contributions of each of the components to meaningful behavioral indicators. Although moral judgment is linked to a wide range of pro-social behaviors (28), including clinical performance ratings for nurses (29, 30), physicians (31) and dentists (8), and to preferences for the more altruistic law disciplines for law students (32), the search for behavioral measures to examine the relative contribution of each component to the behavioral outcomes has been a frustrating one. The author's most recent effort (20) has been to calculate a productivity index that reflects students' success in interacting effectively with patients to achieve acceptance and completion of treatment recommendations. To meet competency requirements, the student must achieve an average monthly index (over all months of clinical practice) of .75 or above. Although there was considerable range in productivity from .67 to 1.19, since students must meet a .75 overall average in order to graduate, the productivity index, while identifying highly effective students, also produces a highly skewed distribution (Mean = .80, S.D. = .08). In the analysis, productivity, like Grade Point Average, was not related to any of the measures of morality.

The explanatory power of the Four Component Model is observed, taking a somewhat different approach, i.e., working backward from disciplinary action to examining deficiencies in the components. Baldwin observed a relationship between the number of malpractice claims and moral judgment scores, noting that a high DIT score had a kind of protective effect, insulating one from claims (33). For dental practitioners referred for ethics instruction, disciplinary actions were directly tied to significant deficits in one or more of the components (8, 9). Further, one consistent observation, in addition to a deficiency in either sensitivity, reasoning or implementation, is the difficulty 27 of the 28 referrals had in articulating the expectations of the profession. After targeted instruction, directed toward role concept development and remediation of one or more other deficiencies, we observed measurable improvements in performance, coupled with documented changes in the behaviors that gave rise to the disciplinary action. Further, to date, there have been no cases of recidivism.[4] Examining case studies bolsters the understanding of the connection between the components and behavior, and provides direction for education.

Conclusions

Analyzing data from the sources cited indicates: 1) striking individual differences among students and practicing professionals on each of the measures; 2) that competence on one of the processes does not predict competence on another; 3) that curricula of rather modest duration can influence performance in measurable ways (our curriculum consists of 43 contact hours); and 4) that strengths and weaknesses in each of the processes are linked to real-life ethical behavior. The findings described in this paper support Rest's contention that moral failings can result from deficiencies in one or more of the processes. Findings also support the importance of attending to each when designing curriculum. Further, whether a curriculum promotes ethical development depends on whether that curriculum incorporates the elements of effective instruction.

Implications for Teaching and Assessing for Research Integrity

If the objective is to develop thoughtful and responsible scientists who act with integrity and have broad understanding of their role and a commitment to integrity in science, it is important to do more than teach the rules and policies that apply to the conduct of research. Before engaging in case discussions, research ethics teachers need to address the expectations of a scientist. Students cannot be expected to intuit the norms and values that undergird the research enterprise. And, it is not clear that they can "pick them up" from role models. The expectations need to be explicitly taught and formally assessed, preferably in writing. By asking students to express the concepts in their own words, and in writing, misperceptions can be identified and addressed before they become an issue. Once the expectations of the scientist are clear, engage students in active learning (using cases, if possible) to facilitate ethical sensitivity, reasoning and problem solving. When designing case materials, careful thought should be given to the particular process that is of concern. Too often, cases are written and participants are asked: What should the protagonist do? Such a question focuses on problem solving, rather than problem identification or moral reasoning. Certainly a skilled facilitator can redirect attention to reasoning or problem identification, but it is sometimes much more difficult.

The author's experience suggests that for novice ethics teachers (which most of us are) focusing on sensitivity, reasoning, and role concept independently of one another will more efficiently develop the skill needed for effective problem solving. Ethics teachers should not expect that carefully targeted courses will develop the more advanced skills in ethical reasoning that might result from courses in moral philosophy. Yet, problem-based practice (using cases) can be especially effective in helping students recognize and subsequently avoid personal interest arguments while strengthening awareness and adherence to the rules of responsible research conduct.

Notes

1. The referrals from the State Board came about because some of the Board members have been involved in the undergraduate curriculum for students. They wondered whether violations of the Dental Practice Act reflected ethical deficiencies that could be remediated by the kinds of experiences we provided for students.
2. For a detailed account of the undergraduate dental ethics curriculum, see Bebeau (1994).
3. There is extensive literature on the construct validity of the DIT. See Rest, Narvaez, Bebeau, & Thoma (1999) for a summary and references to the 400 published studies using the DIT.
4. It is important to note that none of the practitioners referred for remediation involved problems with impulse control, substance abuse, mental illness, or significant personality disorders.

Bibliography

1. Rest J. Morality. In: Mussen PH, series editor, Flavell J, Markman E, volume editors. Handbook of child psychology: vol. 3, cognitive development. 4th ed. New York: Wiley; 1983. p. 556-629.
2. Bebeau MJ, Rest JR, Narváez DF. Beyond the promise: a perspective for research in moral education. Educational Researcher 1999;28 (4):18-26.
3. Kohlberg L. Stage and sequence: the cognitive developmental approach to socialization. In: Goslin DA, editor. Handbook of socialization theory. Chicago: Rand McNally; 1969. p. 347-80.
4. Eisenberg N. Altruistic emotion, cognition and behavior. Hillsdale (NJ): Erlbaum; 1986.
5. Lickona T. Educating for character. New York: Bantam; 1991.
6. Rest J, Narvaez D, Bebeau MJ, Thoma SJ. Postconventional moral thinking: a neo-Kohlbergian approach. Mahwah (NJ): Erlbaum; 1999.
7. Beauchamp TL, Childress JF. Principles of biomedical ethics. 4th ed. New York: Oxford University Press; 1994.

8. Bebeau MJ. Influencing the moral dimensions of dental practice. In: Rest JR, Narvaez D, editors. Moral development in the professions: psychology and applied ethics. Hillsdale (NJ): Erlbaum; 1994. p. 121-46.
9. Bebeau MJ. Continuing ethics education for remediation. Paper presented at the meeting of the American Association of Dental Schools; Orlando (FL); 1997.
10. Bebeau MJ, Rest JR, Yamoor CM. Measuring the ethical sensitivity of dental students. J Dent Educ 1985;49 (4):225-35.
11. Bebeau MJ, Rest JR. The Dental Ethical Sensitivity Test. Minneapolis (MN): Center for the Study of Ethical Development; 1990.
12. Fravel BJ, Bebeau MJ. The role of the American College of Dentists in assessing ethical sensitivity. Poster session presented at the 25th annual meeting of the Association of Moral Education; Minneapolis (MN); 1999 Nov.
13. Rest J. Development in judging moral issues. Minneapolis (MN): University of Minnesota Press; 1979.
14. Rest J, Thoma S, Edwards L. Designing and validating a measure of moral judgment: stage preference and stage consistency approaches. Journal of Educational Psychology 1997;89 (1):5-28.
15. McNeel SP. College teaching and student moral development. In: Rest JR, Narvaez DF, editors. Moral development in the professions: psychology and applied ethics. Hillsdale (NJ): Erlbaum; 1994. p. 27-50.
16. Pascarella ET, Terenzini PT. Moral development. In: How college affects students: findings and insights from twenty years of research. San Francisco (CA): Jossey-Bass; 1991. p. 335-68.
17. Rest JR, Narvaez D, editors. Moral development in the professions: psychology and applied ethics. Hillsdale (NJ): Erlbaum; 1994.
18. Bebeau MJ, Thoma SJ. The impact of a dental ethics curriculum on moral reasoning. J Dent Educ 1994;58 (9):684-92.
19. Ponemon L, Gabhart D. Ethical reasoning research in the accounting and auditing professions. In: Rest JR, Narvaez D, editor. Moral development in the professions: psychology and applied ethics. Hillsdale (NJ): Erlbaum; 1994. p. 101-19.
20. Bebeau MJ. Influencing the moral dimensions of dental practice: implications for the professions. Invited address. American Educational Research Association; New Orleans (LA); 2000 Apr 24-28.
21. Rest JR, Thoma SJ, Narvaez D, Bebeau MJ. Alchemy and beyond: indexing the Defining Issues Test. Journal of Educational Psychology 1997;89 (3), 498-507.
22. Thoma SJ, Rest JR. The relationship between moral decision-making and patterns of consolidation and transition in moral judgment development. Developmental Psychology 1999;35:323-33.
23. Rest J, Narvaez D, Thoma SJ, Bebeau MJ. DIT2: devising and testing a revised instrument of moral judgment. Journal of Educational Psychology 1999;91 (4):644-59.
24. Yeap CH. An analysis of the effects of moral education interventions on the development of moral cognition [dissertation]. Minneapolis (MN): University of Minnesota; 1999.
25. Bebeau MJ, Thoma SJ. "Intermediate" concepts and the connection to moral education. Educational Psychology Review 1999;11 (4):343-60.
26. Bebeau MJ, Born DO, Ozar DT. The development of a Professional Role Orientation Inventory. Journal of the American College of Dentists 1993;60 (2):27-33.
27. Thoma SJ, Bebeau MJ, Born DO. Further analysis of the Professional Role Orientation Inventory [abstract 116]. J Dent Res 1998 Mar;77 (Special Issue):120.
28. Thoma S. Moral judgments and moral action. In: Rest JR, Narvaez D, editors. Moral development in the professions: psychology and applied ethics. Hillsdale (NJ): Erlbaum; 1994. p. 199-212.
29. Duckett LJ, Ryden MB. Education for ethical nursing practice. In: Rest JR, Narvaez D, editors. Moral development in the professions: psychology and applied ethics. Hillsdale (NJ): Erlbaum; 1994. p. 51-70.
30. Krichbaum K, Rowan M, Duckett L, Ryden MB, Savik K. The clinical evaluation tool: a measure of the quality of clinical performance of baccalaureate nursing students. J Nurs Educ 1994;33 (9):395-404.
31. Sheehan TJ, Husted SD, Candee D, Cook CD, Bargen M. Moral judgment as a predictor of clinicial performance. Evaluation and the Health Professions 1980;3:393-404.
32. Landsman M, McNeel SP. Moral judgment and preference for public interest law practices among beginning law students. Paper presented at the Annual Meeting of the Association for Moral Education; Glasgow, Scotland; 2000 July 7-11.
33. Self DJ, Baldwin DC Jr. Moral reasoning in medicine. In: Rest JR, Narvaez D, editors. Moral development in the professions: psychology and applied ethics. Hillsdale (NJ): Erlbaum; 1994. p. 147-62.

Research Ethics in US Medical Education: An Analysis of Ethics Course Syllabi

James M. DuBois, Center for Health Care Ethics, Saint Louis University, USA

Jill E. Ciesla, Center for Health Care Ethics, Saint Louis University, USA

Kevin E. Voss, Center for Health Care Ethics, Saint Louis University, USA

Keywords: Curriculum, Ethics, Medical education, Medical ethics, Medical schools

Medical education trains future physicians as medical practitioners. For this reason ethics education for medical students has traditionally focused on themes revolving around the patient-physician relationship: veracity, informed consent, fidelity, confidentiality, non-maleficence, and the like (1-3). While many of these themes overlap with themes in research ethics, these ethics courses may be inadequate for those future physicians who will engage in research of any kind – including clinical trials, patient surveys, or program assessments (4-7). Research ethics introduces new and important themes related to experimental design, interaction with communities, and the dissemination of information (8,9). The well being of patients, physicians, and research institutions is at stake when physicians fail to abide by rules for ethical research (9,10).

Recent, highly publicized failures to follow protocol at major medical centers reinforce the idea that Institutional Review Boards (IRBs) are inadequate to ensure ethical research behavior. These facts give rise to an important research question: To what extent is research ethics incorporated into the ethics curriculum at medical schools in the United States (US), where future clinical researchers are trained? This question takes on additional significance when one considers that medical students may be engaged in clinical research in various forms even before completing undergraduate medical studies (5,11,12).

This study builds upon a larger study that the first two authors of this paper conducted on the ethics curriculum in US medical schools. DuBois and Ciesla analyzed syllabi from required ethics courses in US medical schools with the aim of identifying and rank-ordering course objectives, teaching methods, course content, and methods of student assessment (13). (The term “ethics course” is used here to refer broadly either to a self-standing course or to a formal educational unit within a larger course.) The present study analyzes in detail the content of the research ethics portion of required ethics courses in the 4-year medical doctor (MD) curriculum at US medical schools. It makes no attempt to describe responsible conduct of research (RCR) education at medical schools as a whole, which frequently house graduate and postgraduate programs in the biomedical sciences, and accordingly offer more extensive RCR courses outside of their MD programs.

Methods

This study was presented to the Institutional Review Board of Saint Louis University. It was approved

Corresponding author: James M. DuBois, PhD, DSc, Center for Health Care Ethics, Saint Louis University, Salus Center, 4th Floor, St. Louis MO 63104, 314-977-6660 (voice), 314-268 5150 (fax), duboisjm@slu.edu.

as an exempt study given guarantees that participation would be voluntary, subjects would be adults, and confidentiality would be maintained by publishing only aggregated data.

Instrument and Participants

The American Association of Medical Colleges (AAMC) provided mailing labels for all curriculum directors of 4-year medical colleges in the US (N=121). A 1-page survey was sent to all curriculum directors asking whether ethics is taught as a formal required component, as an elective, or not at all. It also inquired into the year or years in which ethics is taught. The survey further requested course syllabi for all formal ethics components in the 4-year medical curriculum.

Analysis

In the larger study, two researchers read all syllabi using an open coding method to produce a comprehensive list of all elements found in the syllabi that fell into one of four generic categories: (1) course objectives, (2) teaching methods, (3) course content, and (4) student assessment methods. All other statements (e.g., pertaining to class times, locations, and instructors) were ignored. The specific elements of the syllabi were then placed into categories. These categories were used to create variables in a SPSS database. Schools, rather than syllabi, constituted cases in the database: if a school had more than one required ethics component, data from all required course syllabi were entered into that case. Data from 10 syllabi (17%) were entered by two researchers to establish interrater reliability.

The present study identified those syllabi that included content on research ethics.

The research ethics sections of syllabi were read using an open-coding method to generate a comprehensive list of research ethics content. The results of this open-coding process were then placed into general categories. These categories were entered into an expanded SPSS database. Statistical analysis aimed above all to provide descriptive data on the frequency of various research ethics content. Pearson's r was used to test whether the mean number of content areas covered was significantly correlated with either class size or tuition cost.

Results

Surveys were returned by 72% of the schools (n=87). Seventy-nine percent (n=69) of these schools claimed to require a formal ethics course. Of these schools, 84% (n=58) provided ethics course syllabi. The two raters categorized items the same in 90% of the cases. In the predecessor study, analysis and codification of all syllabi identified 10 course objectives, 8 teaching methods, 39 content areas, and 6 methods of student assessment. The mean for individual schools was 3 objectives, 4 teaching methods, 13 content areas, and 2 methods of assessment.

Among the 39 different content areas, research ethics ranked 11th. Twenty-three of the 58 syllabi (39.6%) addressed research ethics in some fashion. Analysis of the research ethics sections of these syllabi revealed 82 specific themes that fall under 17 different general categories.

Table I (below) presents these 17 general categories in rank order, along with the specific themes that fall under each category. It further indicates where the categories and specific themes overlap with the US Public Health Service's (PHS) "Core Instruction Areas" for courses on the Responsible Conduct of Research (RCR) (14). (This policy of December 1, 2000 was suspended by the Bush administration in February 2001 pending further study. This paper refers to the policy because it continues to serve as a model for many institutions and it remains under discussion among legislators and policy makers.)

The average number of general research ethics topics addressed in these 23 syllabi is 6, with individual schools covering anywhere from 1 to 11 topics. Only six topics were covered by more than half of those syllabi that address research ethics. In rank order these are: clinical trials; informed consent; general ethics of human subject research; government committees and regulations; history and background to research ethics; and protecting vulnerable populations. No research ethics topic was covered by more than 21% of the 87 participating schools. The number of research ethics topics covered did not correlate significantly with either school enrollment ($r=.10$, $p<.45$) or tuition costs ($r=.10$, $p<.43$).

Discussion

While Mastroianni and Kahn conducted a useful and informative pilot study of NIH grantee institutions' training efforts in RCR, this study is the first to examine comprehensively the RCR curriculum in US medical programs. Our study

exposes two possible causes for concern. First, too few medical schools teach research ethics in any fashion within their MD program. No topic in research ethics – including clinical trials – is covered by more than 21% of all medical schools. The topic of Institutional Review Boards is covered by less than 13% of medical schools, despite the fact that medical researchers are most likely to work precisely with human subjects. Second, it appears that important topics are wholly missing even in those programs that teach research ethics. This becomes clear when comparing the specific research ethics topics covered within medical ethics syllabi to the "Core Instruction Areas" PHS identified for RCR education (14). For example, the first five of nine core areas PHS identifies (data acquisition, management, sharing, and ownership; mentor / trainee responsibilities; publication practices and responsible authorship; peer review; and collaborative science) seem wholly missing from these syllabi. (The only possible exception is one syllabus that mentions industry/university relationships.)

It is, of course, possible that some of these topics are covered under other general headings (e.g. 'collaborative research' might be discussed under 'clinical trials'). This is one limitation of the method used: a topic is identified only if it explicitly appears on the course syllabus. This means that syllabi using only very general headings will be shortchanged. Nevertheless, a course syllabus should be a reliable statement of the objectives and content of a course, and most syllabi were quite detailed (as the larger study demonstrated). Thus, it seems safe to conclude both that very few MD programs discuss research ethics and that those that do ignore at least half of the topics PHS wants to see addressed.

However, the significance of these findings cannot be firmly established until other questions are answered:

- To what extent are medical students participating in clinical research?
- Are current requirements for RCR instruction likely to be successful in targeting future physicians who are funded by private industry?
- To what extent do clinical researchers encounter special ethical topics that are not covered in general RCR courses?

These questions remain unanswered. Literature in academic medicine has addressed the roles of undergraduate medical students in research (5,11,12). However, the prevalence and extent of students' roles and whether they are specifically listed in study protocols remains unknown. Thus, it is difficult to know whether education in RCR is a pressing need for medical students, or whether these years might be viewed simply as a convenient time to introduce education in RCR.

Research has shown that private industry is now funding more research than is the government (15). Government requirements regarding RCR instruction pertain only to government-funded research, and according to at least one study, two-thirds of NIH grantee institutions require RCR instruction only to the extent that the government mandates it (16). These facts suggest that a "blanket" approach to educating future physicians would be the safest route to ensuring RCR instruction for clinical researchers. However, given the scope of recent government requirements, such a blanket approach would have to be initiated by a professional institution like the AAMC.

Finally, it is difficult to anticipate how well the RCR programs that are currently being mandated will address the specific ethical concerns that arise in clinical, medical research. This study has shown that 13 of our 17 categories could easily be subsumed under just one PHS Core Area: #6, Human Subjects. This suggests that typical RCR instruction aims to cover a broad range of issues that arise in research (such as authorship, peer review and the treatment of animals), whereas physicians feel the need for a highly focused and intensive treatment of human subject research. The years of medical school may be the best or only time to provide this sort of special-tailored education in RCR.

While this study has provided new answers to questions about the current educational training of medical students in RCR, it has also managed to bring new questions to the fore. Only after these questions are answered, will the significance of this study's findings be properly understood.

Acknowledgements

The authors thank Doris Rubio, Ph.D., for assistance in study design. They thank all participating medical schools for their cooperation. They thank Gerard Magill, Ph.D. and Dennis Daly, SJ for securing funding for this project. This study was funded by the Marchetti Fund at Saint Louis University.

Bibliography

1. Miles SH, Lane LW, Bickel J, Walker RM, Cassel CK. Medical ethics education. Coming of age. Acad Med 1989;64:705-14.
2. Fox E, Arnold RM, Brody B. Medical ethics education: past, present and future. Acad Med 1995;70:761-9.
3. Musick DW. Teaching medical ethics. A review of the literature from North American medical schools with emphasis on education. Medicine, Health Care, and Philosophy 1999;2:239-54.
4. Mastroianni AC, Kahn JP. The importance of expanding current training in the responsible conduct of research. Acad Med 1998;73:1249-54.
5. Taylor I, Baum M, Cooper A, Johnston IDA. Dilemmas facing surgical research in the '90s. Ann R Coll Surg Engl 1991;73 Suppl 4:70-2.
6. Friedman PJ. Research ethics: a teaching agenda for academic medicine. Acad Med 1990;65:32-3.
7. Jonsen AR. Clinical research ethics. Journal of Medical Education 1975;50: 918-9.
8. Brody B. The ethics of biomedical research. An international perspective. New York: Oxford University Press; 1998.
9. Levine RJ. Ethics and regulation of clinical research. Baltimore: Urban & Schwarzenberg; 1986.
10. Spicker SF. The use of human beings in research: with special reference to clinical trials. Boston: Kluwer; 1988.
11. Christakis N. Do medical student research subjects need special protection? IRB: A Review of Human Subject Research 1985;7:1-4.
12. Makarushka JL, Lally J. Medical schools, clinical research, and ethical leadership. Journal of Medical Education 1974;49:411-8.
13. DuBois, JM, Ciesla, JE, Voss, KE. Ethics education in US medical schools. A study of syllabi. Academic Medicine. In press.
14. Office of Research Integrity and Public Health Service. PHS policy on instruction in the responsible conduct of research. Adopted and published on December 1, 2000. Available from: URL: http://www.ori.dhhs.gov
15. Cho MK, Shohara R, Schissel A. Policies on faculty conflicts of interest at US universities. JAMA 2000; 284: 2203-8.
16. Mastroianni AC, Kahn JP. Encouraging accountability in research: a pilot assessment of training efforts. Accountability in Research 1999; 7: 85-100.

Table I: Rank Order and Content of the Research Ethics Categories

*An **asterisk *** followed by a number indicates that the general category or specific topic overlaps with a PHS Policy "Core Instructional Area." The number indicates which of nine instructional areas it overlaps with.*
***'Percent valid'** indicates how often a research ethics topic is included in those syllabi from the 23 schools that actually teach research ethics.*
***'Percent all'** indicates how often a research ethics topic is included among all participating schools (i.e., the 87 schools that returned a survey).*

1. CLINICAL TRIALS (*6) – 78% of valid / 21% of all
   - Therapeutic vs. non-therapeutic research
   - Person as patient vs. research subject
   - Physician as clinician vs. physician as scientist
   - Selection of subjects for clinical trials
   - Randomization
   - Patient as research subject vs. health research subject
   - Ethics of medical students' roles in clinical research
   - Drug testing and the role of the FDA
   - Whether scientific methods provides sole criterion for treatment efficacy
   - Industry / university relationships (*possibly 5 & 9)
   - Types of clinical trials
2. INFORMED CONSENT (*6) – 70% of valid / 18% of all
   - Informed consent in clinical vs. research setting
   - Sample consent form for adults
   - Emergency waiver of informed consent
   - Coercion
   - Deception – active and passive
   - Placebos
3. GENERAL ETHICS OF HUMAN SUBJECT RESEARCH (*6) – 65% of valid / 17% of all
   - Ethics of human experimentation
   - Justification of research involving human subject
   - Challenges to human subject protections
4. GOVERNMENT COMMITTEES & REGULATIONS (*6 & others) – 61% of valid / 16% of all
   - Belmont report
   - President's Commission for the Study of Ethical Problems in Medicine and Biomedical and Behavioral Research (1979-83)
   - National Commission for the Protection of Human Subjects of Biomedical and Behavioral Research (1974-78) [Published Belmont Report]
   - Federal regulations
   - National Bioethics Advisory Committee
   - Declaration of Helsinki
   - Practice and regulations
   - OPRR reports, Protection of Human Subjects
   - Title 45, Code of Federal Regulations, part 46 (1994)
   - Nuremberg Code (as living document)
5. HISTORY AND BACKGROUND OF RESEARCH ETHICS – 57% of valid / 15% of all
   - Nazi experimentation / Holocaust (awareness of attitudes toward)
   - Nuremberg Code (as historical document)
   - Tuskegee study of syphilis (awareness and attitudes toward)
   - Abuses and errors of early eugenics
   - "Frankenstein"
   - Sloan-Kettering experiments
   - Willowbrook experiments
   - Henry Beecher revisited (article by DJ Rothman)
   - Introduction to sulfonamides revisited (articles by BH Lerner)
   - Research in the Hippocratic Oath (i.e., the fact that it is not addressed therein)

6. PROTECTING VULNERABLE POPULATIONS (*6) – 52% of valid / 14% of all
   - Minorities
   - Newborns, Infants, Children
   - Soldiers
   - Prisoners
   - Mentally ill
   - AIDS patients
7. IRB (*6) – 48% of valid / 13% of all
   - IRB issues
   - Definition of research / Novel therapy vs. research
8. RESEARCH INTEGRITY & MISCONDUCT (*8 & 9) – 39% of valid / 10% of all
   - Accuracy of published data
   - Research fraud (*8)
   - Appearance of impropriety
   - Scientific misconduct (*8)
   - Scientific integrity
   - Appropriate credentials
   - Research quality guidelines for both academic and non-academic environments
   - Conflicts of interest (*9)
9. ETHICAL PRINCIPLES IN HUMAN SUBJECT RESEARCH (*6) – 39% of valid / 10% of all
   - Respect autonomy
   - Do good (beneficence)
   - Fairness / justice
   - Avoid harm to subjects (non-maleficence)
   - Justify level of risk
   - Apply process of ethical decision making to research ethics
10. ANIMAL EXPERIMENTATION (*7) – 30% of valid / 8% of all
    - Animal rights
    - Use of animals for research
    - Poor living conditions for research animals
11. GENETIC RESEARCH AND THERAPY (*6) – 26% of valid / 7% of all
    - Genetic research
    - Germ-line therapy
    - Somatic cell genetic therapy
    - National Human Genome Research Institute
    - Genetic information and privacy
    - Cystic fibrosis research
12. RESEARCH AND THE SOCIAL GOOD (*6) – 22% of valid / 6% of all
    - Medicine and the goals of society
    - Research in the international context
    - Social utility of research
    - Relationship between ethics, science, and technology
    - Balancing society's mandates, competing pressures to innovate
13. MINIMIZING RISKS (*6) – 22% of valid / 6% of all
    - Establishing gold standard
    - Asking whether risk is proportionate to benefit
14. SUBJECT SELECTION (*6) – 13% of valid / 3% of all
    - Ensuring the inclusion of women, children and minorities (a concern of justice, rather than protection)
15. EMBRYO AND FETAL RESEARCH (*6) – 9% of valid / 2% of all
    - Stem cell research
    - Research on live-born fetuses
16. EPIDEMIOLOGY (*6) – 4% of valid / 1% of all
    - Ethics of epidemiology
17. MILITARY RESEARCH ETHICS (*6) – 4% of valid / 1% of all
    - Experiments related to weaponry
    - Using compounds not fully tested in a wartime situation

Teaching Ethics in Biomedical Science: Effects on Moral Reasoning Skills

Elizabeth Heitman, University of Texas School of Public Health, Houston, Texas, USA

Patricia J. Salis, University of Texas School of Public Health, Houston, Texas, USA

Ruth Ellen Bulger, Uniformed Services University of the Health Sciences, Bethesda, Maryland, USA

Key words: Defining Issues Test, DIT, Ethics education, Evaluation, Responsible conduct of research

Academic institutions that train professionals play an important role in ensuring that trainees learn the ethical norms of their respective disciplines, and that they learn to behave ethically from the start of their professional lives. The National Institutes of Health requirement that funded research training programs include education in scientific integrity has made formal courses on the responsible conduct of research increasingly common in academic medical centers and research universities.

There is still no consensus on what constitutes the most appropriate subject matter, format, methods, or faculty for teaching the responsible conduct of research. The objectives of general courses on the responsible conduct of research and scientific integrity typically include increasing students' understanding of the norms of scientific practice, their recognition of ethically problematic situations in science, and their ability to analyze and respond to such situations in a morally mature manner. Courses vary in the specific content, the number of contact hours, the format (lecture, small-group discussion, video or web-based tutorials), and the instructors' professional background and ethical expertise. The effectiveness of available courses probably also varies. Studies of how students are affected by formal ethics courses in such disciplines as engineering, law, dentistry, medicine, nursing, journalism, accounting, veterinary medicine, and social work have found that course design influences the extent to which students' ethical reasoning skills change during the courses (1-3). Such evaluation in the area of scientific integrity, however, is still in its infancy.

The syllabi of courses on the responsible conduct of research in several institutions suggest that such courses present at least three different kinds of instruction to students. The first is the "how-to" of science, in which the practical, procedural dimensions of science, rather than its ethical dimensions, are the focus: how to devise an experiment, give a talk, or write a manuscript. The second kind of instruction relates to the rules, regulations, and professional norms articulated by the organizations in which scientists work, their professional societies, and/or the government: how to make experimental data available for use, how to address suspected research misconduct, and how to deal ethically with animal and human subjects. Ethical considerations are often addressed as an aspect of these practical issues. Lecture and individual reading assignments are effective mechanisms for teaching both of these traditional types of subject matter, and students' understanding and retention can be evaluated by an objective written (including computerized) or oral exam.

Corresponding author: Elizabeth Heitman, PhD, University of Mississippi Medical Center, 2500 North State Street, Jackson, MS 39216, 601-815-1134 (voice), 601-366-3262 (fax), eheitman@surgery.umsmed.edu.

The third type of instruction presented by these courses relates to students' ability to recognize the ethical aspects of problems that they encounter in their research, and their ability to address these issues in a considered way. This instruction involves their developing moral reasoning skills rather than simply comprehending information, and it frequently uses case discussion or problem-based learning. Two decades ago the Hastings Center Project on the Teaching of Ethics proposed three criteria for evaluating the effectiveness of such instruction: 1) whether the student understands the central concepts; 2) whether the student can make cogent oral and written ethical arguments; and 3) whether the student can recognize ethical problems and examine them rationally (4). This evaluation is typically conducted through a more subjective examination using actual case analysis, possibly in a written or oral exam, but ideally in a more interactive setting.

The Hastings Center Project emphasized that helping students develop skills to recognize and analyze ethical issues and stimulating their moral imagination are fundamental to the effective teaching of ethics. The Association of American Medical Colleges handbook, *Teaching the Responsible Conduct of Research through a Case Study Approach* (5), has also stressed the need to enhance students' ethical awareness and problem-solving skills in formal education on the responsible conduct of research. Ideally, the courses should have a positive effect on students' actual and future behavior, helping individuals avoid ethically problematic behavior and enhancing their ability to resolve unfamiliar ethical conflict appropriately.

After several years of teaching a formal course on the responsible conduct of research at the University of Texas Health Science Center at Houston, the course's organizers sought to assess its effects and to determine what outcomes could be evaluated formally. The course, The Ethical Dimensions of the Biomedical Sciences, originated in 1984 as an institutional response to an incident with a foreign graduate student that would have been considered plagiarism for a student schooled in the United States (6, 7). Consideration of the case highlighted the administration's and faculty's need to articulate the university's ethical expectations and to teach U.S. academic and professional standards to all students. The primary objectives of the course subsequently developed by Dr. Ruth Bulger, and later continued by Drs. Stanley Reiser and Elizabeth Heitman, have been to encourage students' interest in the ethical development and goals of science, and to teach students to prevent, recognize, analyze, and resolve ethical conflicts in the daily conduct of their work (8).

From the beginning, the course has used a combination of formal reading assignments, didactic lecture, and small-group case discussion to address a wide variety of issues in the responsible conduct of research. Its faculty have always included both ethicists and bench and clinical researchers from various disciplines, both as lecturers and as discussion leaders. Most are senior faculty. Since 1988, the course has been a requirement for graduation from the Graduate School of Biomedical Sciences, and it is an elective for graduate students in the School of Public Health. For the past four years, approximately 120 students have enrolled in the course each fall, including 90+ from the Graduate School of Biomedical Sciences' 22 degree programs, 10-15 students from the School of Public Health's 11 degree programs, and several post-doctoral fellows from the UT Medical School and MD Anderson Cancer Center. Students in biomedical sciences typically take the course in their first semester, while others often enroll in the second half of their formal graduate study.

Objective written examinations demonstrated that the course effectively enhanced students' knowledge and understanding of both the practical how-to of science and the rules, regulations, and professional norms of research that the course addressed. Written analysis in the final exam demonstrated students' ability to identify and consider ethical issues. Students' course evaluations also confirmed that most of them found the course valuable to their professional development. However, the faculty wanted to assess the more comprehensive effects of the course on students' professional attitudes and behaviors.

To affect students' current behavior and shape their future action, instructors of courses in the responsible conduct of research must have three things: 1) an effective way to teach desired behaviors; 2) an effective way to motivate students to adopt these behaviors; and 3) a reliable way to measure behavior change. In a broad literature review, we found no clearly identifiable, successful method for teaching ethical behavior or motivating students to act

ethically. While there has been work on how best to evaluate students' comprehension and retention of information related to ethical conduct, we found no generally accepted way to measure the presumed beneficial effect of ethics courses on behavior.

In the absence of accepted measures of behavior change and future practice, surrogate measures of the effectiveness of courses on the responsible conduct of research are needed. Bebeau (9) and her colleagues have developed a set of teaching materials for education in the responsible conduct of research that considers four psychological processes in the decision to act ethically: moral sensitivity (the ability to interpret a moral situation and the effects of various courses of action on the parties involved); moral reasoning (judgment about which course of action is right); moral commitment (intention to do what is right) and moral perseverance (the ability to follow through with ethical behavior). Their method of evaluating the effectiveness of courses that use the group's instructional materials assesses the essential components of students' moral discernment and moral reasoning.

Efforts to define, implement, and assess education in the responsible conduct of research in graduate science programs have parallels in medical education, where considerable work has been done on the teaching of professional ethics and the evaluation of such teaching. The effects of ethics courses on medical students' moral reasoning skills have been studied since the late 1970s (10). Such evaluations have linked different types of ethics education with changes in students' moral reasoning, and have suggested that case-based discussion can significantly increase students' moral reasoning ability.

The Defining Issues Test (DIT) is the instrument used most frequently to measure moral reasoning skills and the effects of education on moral reasoning. The DIT was developed by James Rest and colleagues at the University of Minnesota Center for the Study of Ethical Development (11). The test is a standardized, computer-scored test that is easily administered to groups. It is based on Kohlberg's theory of cognitive moral development, which considers the principle of justice as the highest moral good. The DIT presents six morally problematic scenarios; the subject ranks the importance of various moral criteria for judging how to act, then chooses a course of action. Scores are reported in terms of a P%, which measures the extent of "principled" reasoning behind the individual's assessment of the cases. Cross-cultural applications have found that DIT scores increase consistently with subjects' age and education level.

This study explored whether two offerings of our course on The Ethical Dimensions of the Biomedical Sciences had an effect on students' principled moral reasoning, as measured by the DIT.

Methods

Following an IRB-approved protocol, a total of 215 graduate students who were enrolled in The Ethical Dimensions of the Biomedical Sciences course were asked to complete the DIT at the beginning (before-course) and the end (after-course) of the 1997 and 1998 classes. Use of individual codes protected students' confidentiality. Computerized scoring by the University of Minnesota Center for the Study of Ethical Development generated P% scores.* The analyses used students' change scores — the after-course test score minus the before-course test score — as the data. A preliminary analysis of differences in change scores between the 1997 and 1998 classes (t-test, independent samples) was performed to determine whether it was possible to combine the data from the two classes. Next the effectiveness of the course in improving students' principled judgment by was tested directly analyzing whether their change scores differed significantly from zero (t-test, matched pairs). Finally, an analysis of variance (ANOVA) test was run to determine whether students' gender or country of undergraduate education (US or non-US) was related to differential change scores.

Results

One hundred seventy-two students (80% of the original 215 students) completed both a before-course and an after-course test, 95 students in 1997 (87% of 109) and 77 in 1998 (73% of 106) (Table 1). One or both tests from 14 of these 172 subjects were excluded from analysis based on scoring criteria used by the University of Minnesota Center for the Study of Ethical Development. The final sample therefore contained 158 students who had valid scores for both the before-course and the after-course tests. Change scores did not differ significantly between the 1997 and 1998 classes (t=-0.88,

p=0.38), so a combined analysis of the two classes was possible.

The primary analysis assessed the course's effect on principled judgment: It revealed that the students showed no significant after-course improvement in principled judgment, as measured by the DIT P% score (Figure 1, Table 2). Indeed, the pattern in six of the eight sub-groups (Figure 2) was for after-course scores to drop slightly.

Follow-up analyses of the influence on change scores of students' gender and location of undergraduate schooling indicated that neither gender nor location of education had a significant effect for the combined 1997 and 1998 courses (Table 3), for the 1997 students alone (Table 4), or for the 1998 students alone (Table 5). For the combined group and the 1997 group, there was no significant interaction between the gender factor and the location-of-schooling factor, but this interaction was significant in the 1998 group (Table 5). The 1998 data in Figure 2 suggest that this result arose from the distinctive pattern among men educated in the U.S. Their after-course scores declined somewhat, while those of both groups of women and of men not educated in the U.S. either improved very slightly or stayed essentially the same.

Conclusions

The finding that no significant change had occurred in P% scores after the course on the

	Combined 1997 & 1998 Classes	1997 Class	1998 Class
No. people who took at least 1 test	215	109	106
No. people who took 2 tests	172 (80% of 215)	95 (87% of 109)	77 (73% of 106)
No. test pairs sent for scoring	172	95	77
No. test pairs not used	14	11	3
Final no. people or test pairs	158 (92% of 172) (73% of 215)	84 (88% of 95) (77% of 109)	74 (96% of 77) (70% of 106)

Table 1. Composition of final study sample.
* one or both tests in pair purged by scorers for invalidity; one of pair purged by us due to absence of valid pair-mate; scorers failed to process test pair

Group	Change Score Mean (SD)	t Value *	p Value
1997 & 1998 Combined	-1.27 (11.75)	-1.36	0.17
1997	-2.05 (11.64)	-1.61	0.11
1998	-0.40 (11.89)	-0.29	0.78

Table 2. Statistical evaluation of course's effect on DIT P% scores: t-tests (matched pairs) of change scores (after-course minus before-course) in combined classes and in each class alone.

Source of Variance	Degrees of Freedom	F Value	p Value
Gender	1	0.21	0.65
Country of education	1	0.54	0.46
Gender X country interaction	1	0.90	0.34
Error	154		
Total	157		

Table 3. Statistical evaluation of effect of gender and location-of-schooling on DIT P% scores: analysis of variance of change scores (after-course minus before-course) in combined classes.

Source of Variance	Degrees of Freedom	F Value	p Value
Gender	1	0.16	0.69
Country of education	1	0.09	0.77
Gender X country interaction	1	0.25	0.62
Error	80		
Total	83		

Table 4. Statistical evaluation of effect of gender and location-of-schooling on DIT P% scores: Analysis of variance of change scores (after-course minus before-course) in 1997 class.

Source of Variance	Degrees of Freedom	F Value	p Value
Gender	1	1.75	0.19
County of education	1	0.97	0.33
Gender X country interaction	1	4.86	0.03
Error	70		
Total	73		

Table 5. Statistical evaluation of effect of gender and location-of-schooling on DIT P% scores: Analysis of variance of change scores (after-course minus before-course) in 1998 class.

responsible conduct of research was a surprising and frustrating outcome, given the course's perceived value within the university and the number of studies that report significant changes in students' moral reasoning skills after similar courses in professional ethics. Even more perplexing was that students in most sub-groups actually showed slight declines in P% scores after the course.

Upon reflection, the authors concluded that principled moral reasoning is only one of a number of skills and concepts that we hope to teach and foster in our course. Much of the material and related discussion in the course focuses on common conflicts and practical ethical strategies in research and collegial interaction. Rest and colleagues (12) noted in 1999 that Kohlberg's theories, and thus the DIT, address formal ethical structures of society, what they call macromorality, and do not illuminate the micromoral phenomena of personal, face-to-face interactions in everyday life. Thus these null findings suggest that it is essential to ask different questions or use different methods to evaluate the complex issue of the outcomes of the course.

The establishment and ultimate success of education in the responsible conduct of research will require effective means of assessing the impact of such programs on students' knowledge, awareness, and moral reasoning. Under the most recent proposal requiring such education in all Public Health Service-funded institutions, a wide variety of formats appear to satisfy the new credentialing standards. Suggested options range from semester-long academic courses to day-long workshops to hour-long web-based tutorials, to self-study reading programs. As academic research institutions develop the expertise needed to provide education in the responsible conduct of research, mechanisms must also be developed to assess the extent to which these different formats are effective in enhancing participants' moral reasoning skills. Recent observations reported by Bebeau and colleagues suggest that some apparently unchanged DIT scores may mask important differences in moral sensitivity and reasoning (13). Expanded use of the DIT should strive to uncover all significant changes in moral reasoning in order that academic courses can target their educational intervention appropriately.

However, if the objective of education in the

responsible conduct of research is to shape the behavior of researchers and to reform the culture of research, methods for evaluating such change must be developed, and instructors must learn how to present the rules, regulations, and professional norms of science in a way that motivates researchers to adhere to them.

Note

***** The Center generated the P% scores using its new system of validity checks, which should be considered when comparing these results to those of older studies.

Bibliography

1. Schlaefli A, Rest JR, Thoma SJ. Does moral education improve moral judgment? A meta-analysis of intervention studies using the Defining Issues Test. Review of Educational Research 1985; 55: 319-20.
2. Self DJ, Baldwin, DC Jr., Wolinsky FD. Evaluation of teaching medical ethics by an assessment of moral reasoning. Medical Education 1992; 26: 178-184.
3. Self DJ, Wolinsky FD, Baldwin, DC Jr. The effect of teaching medical ethics on medical students' moral reasoning. Academic Medicine 1989; 64: 755-759.
4. Callahan D, Bok S. Hastings Center project on the teaching of ethics. In: Callahan D, Bok S, eds., Ethics teaching in higher education. New York: Plenum Press, 1980.
5. Association of American Medical Colleges. Teaching the responsible conduct of research through a case study approach. Washington, DC: AAMC, 1994.
6. Bulger RE, Reiser SJ, Studying science in the context of ethics. Academic Medicine 1993; (Suppl. 3): 55-59.
7. Reiser SJ, Heitman E. Creating a course on ethics in the biological sciences. Academic Medicine 1993, 68: 876-879.
8. Bulger RE, Heitman E, Reiser, SJ. The ethical dimensions of the biological sciences. New York: Cambridge University Press, 1993.
9. Bebeau M. Moral reasoning in scientific research. Cases for teaching and assessment. Bloomington, IN: Poynter Center for the Study of Ethics and Assessment, 1995.
10. Self DJ, Baldwin DC Jr. Moral reasoning in medicine. In Moral development in the professions: Psychology and applied ethics. In: J Rest & D. Navares, eds., Hillsdale, NH: Lawrence Erlbaum Associates, 1994.
11. Rest JR. Development in judging moral issues. Minneapolis, MN: University of Minnesota Press, 1979.
12. Rest J, Narvaez D, Bebeau MJ, Thoma SJ. Postconventional moral thinking: A neo-Kohlbergian approach. Mahwah, NJ: Lawrence Earlbaum Associates, 1999: 15.
13. Bebeau MJ. Influencing the moral dimension of professional practice: Implications for teaching and assessing for research integrity (abstract). A Research Conference on Research Integrity, Bethesda. MD, November 19, 2000.

Figure 1. Mean DIT P% Scores Before and After C Combined 1997 and 1998 Classes

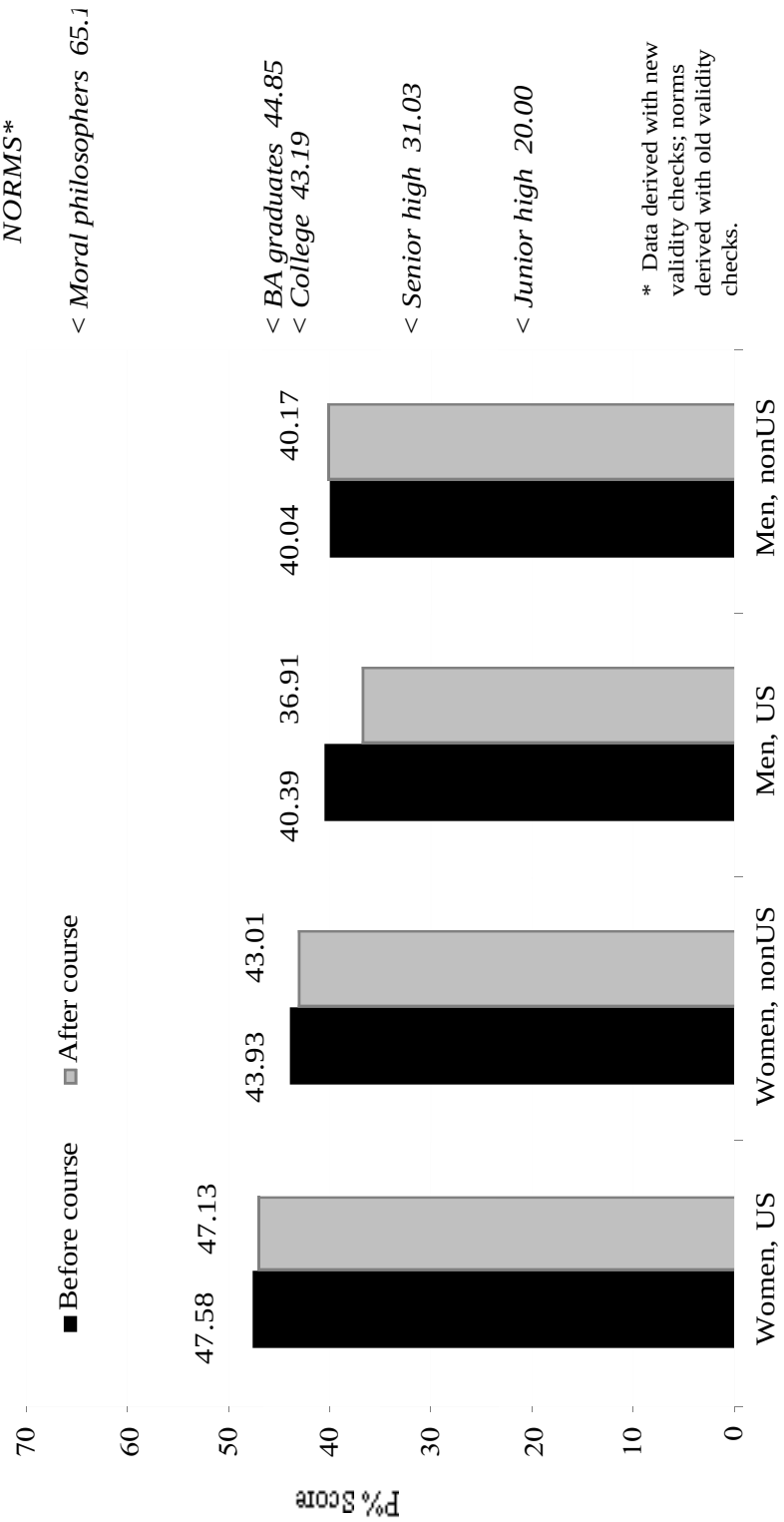
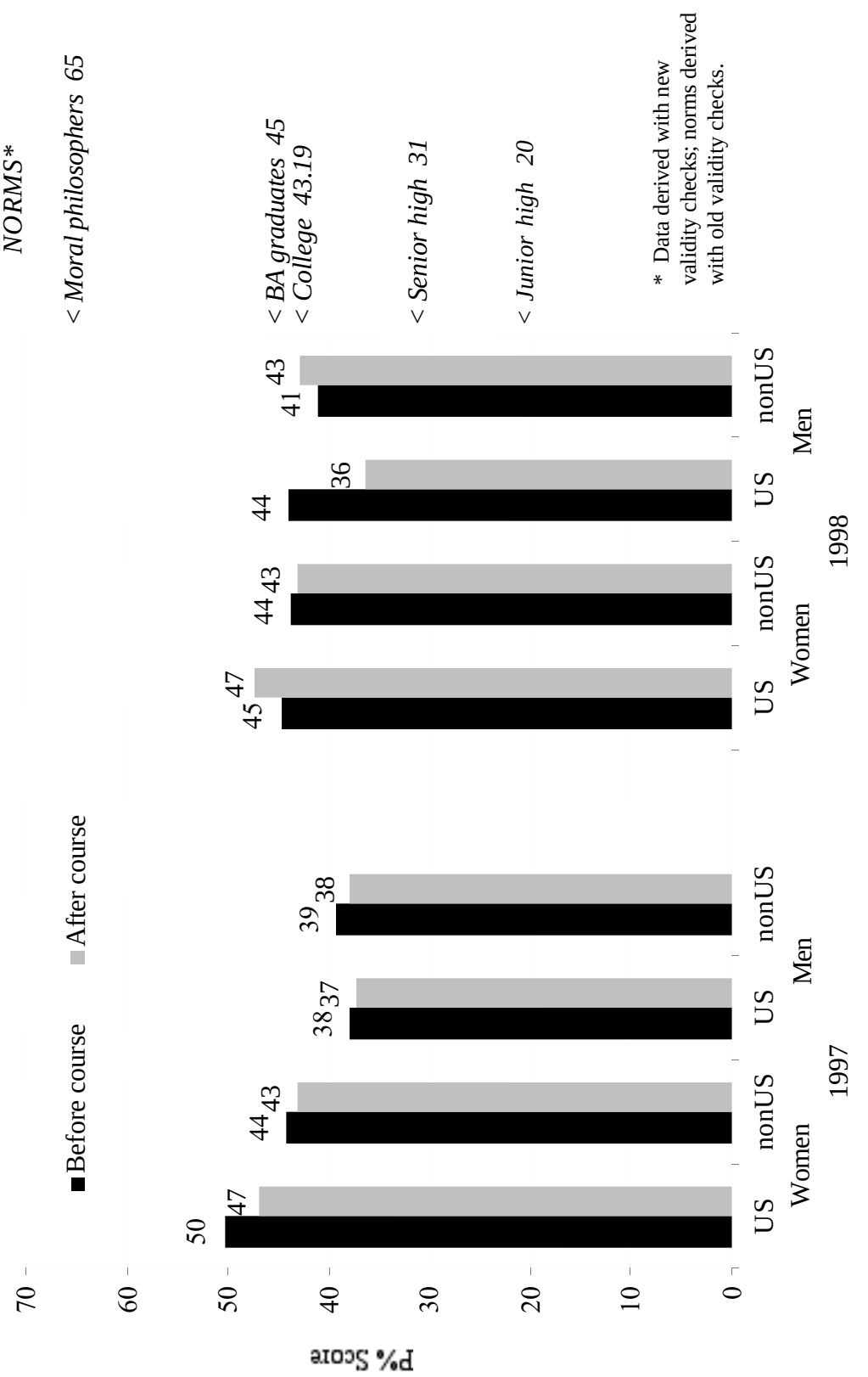


Figure 2. Mean DIT P% Scores Before and After C
1997 Class and 1998 Class



Fostering Research Integrity through Educational Programs: Lessons Learned at the University of Minnesota

Jeffrey P. Kahn, Center for Bioethics, University of Minnesota, USA

Peggy A. Sundermeyer, Office of the Vice President for Research, University of Minnesota, USA

Muriel J. Bebeau, Department of Preventive Sciences, University of Minnesota, USA

Virginia S. Seybold, Department of Neuroscience, University of Minnesota, USA

Keywords: RCR curriculum development, Research integrity education programs, Responsible conduct of research education

The implementation of a Public Health Service (PHS) policy on Instruction in the Responsible Conduct of Research (RCR) would be a significant challenge to universities because of its broad inclusion of personnel involved in research. The University of Minnesota is already meeting this challenge with the delivery of a comprehensive educational program to over 2,000 faculty and principal investigators (PIs) in calendar year 2000.

The University of Minnesota is a large, land-grant institution. The intellectual diversity of the institution is reflected in its 21 collegiate units, 3,000 tenure and tenure-track faculty, and 10,000 graduate students enrolled in 150 masters and doctoral programs. The foundation of our educational programming in RCR developed centrally, early in the 1990's, to support the educational requirement of training grants. These programs were expanded to faculty in the mid-90's in response to growing institutional and national concern about misconduct in research. The current curriculum is the result of an institutional corrective action plan initiated by National Institutes of Health (NIH) in 1997. Therefore, a unique set of circumstances required the University of Minnesota to implement a comprehensive educational program in RCR before announcement of the PHS policy on Instruction in RCR.

Our goal is to share the experience of our institution in order to aid others in the development of programs to meet the requirements of the PHS policy. Points of discussion within the context of the evolution of the educational program at Minnesota include 1) policy as framework for education, 2) development and delivery of the curriculum, 3) resources and financial investment, and 4) evaluation.

Policy as Framework in Education

One strength of the educational initiative at the University of Minnesota is that the importance of RCR is reflected in institutional policies. The Board of Regents, the administrative authority of the

Corresponding author: Jeffrey Kahn, Center for Bioethics, University of Minnesota, N504 Boynton, 410 Church St. SE, Minneapolis, MN 55455-0346, 612-624-9440 (voice), 612-624-9108 (fax), kahnx009@tc.umn.edu.

University, passed the Code of Conduct in 1996. This policy pertains to all members of the University community and states that we will "adhere to the highest ethical standards of professional conduct and integrity." While affirming the common values of research and scholarship, it is a clear demonstration of institutional ownership of these values. In 1999, the Board of Regents passed a revised policy on Principal Investigator Eligibility on Sponsored Projects. This policy requires PIs to complete a required education in RCR before any awarded funds are released for spending. The policy was implemented March 1, 2001, preceding the PHS policy by approximately two years and providing the motivation for compliance with the educational requirement. Both policies can be viewed at http://www.ospa.umn.edu/policy/respolcy.htm.

The University of Minnesota has a strong tradition in faculty governance, so it is not surprising that the faculty senate has also promoted RCR. In 1999, the faculty senate passed the policy on Education in Responsible Conduct of Sponsored Research and Grants Management (see http://www1.umn.edu/usenate/policies/grantsmgmt.html). Whereas this policy reiterates the expectation that PIs and project personnel have the responsibility to behave in accordance with the highest ethical standards, it also defines the responsibility of the University to provide individuals involved in research with information and resources that support responsible conduct. The policy describes the framework for implementing educational programs under the leadership of the Vice President for Research and Dean of the Graduate School. It outlines the formation of three advisory committees, one for each major constituency: Academic personnel (including faculty and academic administrators), research staff (including graduate and postdoctoral trainees as well as project staff), and administrative staff (including accounting and secretarial support). The charge to each of these committees is to define the educational needs of the constituency, develop the curriculum, recommend delivery formats for the curriculum, propose appropriate recognition/accreditation, and establish appropriate continuing education requirements. The Vice President for Research and Dean of the Graduate School is also charged with the responsibility of maintaining a database on meeting the educational requirements.

Development and Delivery of the Curriculum

The development and delivery of the educational program in RCR for investigators has been led by the Faculty Education Advisory (FEA) Committee. The FEA Committee is in its third year of existence and is made up of faculty, with senior administrators serving in ex officio capacity. The Committee is staffed by personnel from the Office of the Vice President for Research. The Committee meets monthly and has had remarkably consistent participation over the three years. Members were added recently to increase representation of disciplines within the University.

Members of the FEA Committee are senior and respected faculty and broadly represent the diversity of the University's colleges, departments, and programs. The commitment of faculty leaders, coupled with resources and commitment from high-level University administration, has been crucial to the success of the FEA Committee's effort. The Committee has focused on three areas in RCR education: (1) defining and identifying the target populations; (2) identifying topic areas; and (3) implementation.

Defining and identifying target populations for RCR education and training

The initial focus of RCR educational programming has been PIs, both because it represents the largest group of faculty and staff responsible for the performance of research, and because the University has a system for certification of PI status. This cohort represented nearly 2,000 individuals, from across every college and a diverse range of departments and research areas.

This diversity led to a recognition that education in RCR could not be successful as a "one size fits all" program, and that we needed to speak to the needs and interests of researchers from outside biomedical research areas. But in spite of the diversity of researchers' needs, the FEA Committee agreed on a need to achieve a shared basic level of understanding for all researchers on a core set of RCR issues. This is based on the view that all researchers belong to the University's research community, and that

such membership brings with it certain responsibilities, including basic familiarity with the rule and issues in areas such as research that involves human or animal subjects. So while many researchers may never engage in human or animal research, it is unacceptable for them to pass it off as someone else's problem. For those researchers engaged in research involving human or animal subjects, more in-depth education and training in those areas is required. In addition to both basic training for all and in-depth training when appropriate, the FEA Committee is developing recommendations for continuing education in RCR.

Identifying topic areas

The FEA Committee's second task was to identify topic areas for curriculum development. Since our efforts pre-dated the PHS/Office of Research Intetrity (ORI) draft of final guidelines, an initial list of topics was drawn from the list of suggested topic areas in the ethics requirement for NIH Training Grants (T32). The FEA Committee then worked to make the list of topics relevant to PIs._The current list of topics includes:

Social Responsibility and Misconduct
Authorship and Peer Review
Data Management
Intellectual property
Conflict of Interest
Fiscal Responsibility
Human Subjects
Animal Subjects
Environmental health and Safety

After the PHS/ORI guidelines were issued, we compared our list of topics to the guidelines in an effort to assess what changes, if any, are needed, and determined that we need to add content on both collaborative science and mentoring.

Implementation

After identifying the target population, and the topic areas that would be covered, the FEA Committee's last task was to develop strategies for implementation. Key components in our effort include recruiting instructors with appropriate expertise and experience, drawing mostly from the ranks of the faculty; and a commitment that face-to-face interaction be part of the educational experience.

We have employed three separate formats for instruction—classroom sessions totaling six hours; web-based instruction for some financial and grants management topics, followed by a 1.5 hour classroom session; and in-depth special topic instruction involving a 1.5 hour classroom session, web resources, and case studies.

Because of the number of hours of instruction required and the diversity of investigators who need to participate, a large and diverse pool of instructors was recruited. We have between four and six faculty who are prepared to deliver one topic area; faculty are paired with relevant professional staff for some topics. These 37 instructors represent 13 colleges and 3 administrative units, and include 4 department heads, and 2 associate deans. While all of the faculty agreed to teach in our RCR efforts on a volunteer basis, the FEA recommended and the University's Vice President for Research agreed that formal and material acknowledgement of their efforts is appropriate. To that end, funds were committed to provide small professional development awards to all faculty participating as instructors in the RCR programs.

Resources & Financial Investment

A cornerstone of our program is faculty involvement in the delivery of the curriculum. Faculty are presenters or facilitators of discussion for each topic. For some topics they are partnered with staff who are available to answer more technical questions. For example, faculty who deliver the module on Intellectual Property are paired with a staff member from the University office of Patents and Technology Marketing. Faculty are also involved in revising instructional materials used in workshops and on the web, as well as the curriculum itself.

The commitment of respected, senior faculty, demonstrated by their leadership on committees or their development of the curriculum, enabled us to recruit other faculty for the delivery of the curriculum. Another critical element for recruitment was a detailed syllabus for each topic of the curriculum. The syllabus includes learning objectives, relevant policies, principles, issues for discussion, reference materials, and case studies for some topics.

One limitation of the curriculum was its bio-medical flavor, particularly in case studies, largely because of the disciplines represented on the initial faculty advisory committee. Recognizing this, we targeted faculty in

underrepresented disciplines to achieve greater balance for delivery of the curriculum. Over 50 faculty from 34 departments are currently involved in RCR curriculum development or delivery. Besides enriching the curriculum, we believed that faculty involvement throughout the University would increase ownership and spread commitment to the RCR. An unexpected outcome of the diversity of disciplines has been the high level of interest maintained by the faculty as they see the issues in their topic take on new dimensions and challenges from one discipline to another.

Besides the demonstrated commitment of faculty, a successful educational program in RCR requires strong support services. Instructional materials are revised and shared amongst presenters. When the faculty audience asks previously unanswered questions, the experts are consulted. The answers are incorporated into future workshops, and the curriculum and instructional materials are revised as appropriate.

There are also numerous administrative tasks associated with scheduling presenters, rooms, and equipment; preparation of printed and web based materials; registration and documentation of attendance; tabulation of evaluations; and feedback to and coaching of faculty presenters. Although these activities happen mostly behind the scenes, they are critical to the program.

Finally, communication is a critical support service. Requirements and rationale must be conveyed to the faculty and other research personnel; progress must be reported to the advisory committee (FEA), faculty senate committees, administrative offices, and academic administrators. All available communications vehicles are used, including monthly newsletters of the sponsored projects office and of colleges as well as the University's multiple publications; printed brochures and flyers; web based home pages and events calendars; meeting in person with faculty committees, academic deans, and special constituencies (IRB); and e-mailings from the Vice President of Research, Deans, and Department heads or chairs.

So what does all of this cost? The direct expenses of the 62 workshops for 2,400 investigators over a 12-month period is the most straight forward. Based on actual cost to date for printing of materials, rental of rooms and equipment, and similar expenses, these direct expenses are projected to be $48,600, or $15.20 per person per session. This amount does not include any compensation for the faculty involved in the delivery. Based on the average actual salaries of faculty involved in the workshops, with an average of 1 – 2 hours depending upon the topic, the value for delivery would be an additional $32,300. This does not include any estimate of faculty time for preparation or involvement in discussions, via e-mail or in person, of improvements or additions to the materials, sharing of additional references, or similar and recurring work. Although faculty were recruited without any hint of monetary reward, we were able to give those most involved small professional development grants of $1,000 – 2,000, for an expense of $24,000.

Direct administrative costs include the salary and fringe benefits of 1.75 staff years: one full time program coordinator, a 50% administrative appointment of a faculty member acting as program director; and an administrative fellow (graduate student). However, the direct cost of additional support services including design and maintenance of web-based tutorials as well as registration and recording keeping activities are nearly impossible to tally since they are provided by a number of centralized offices from the graduate school administration to the human resources offices.

Hardest yet to calculate are the cost of faculty hours spent in participation. Since the University of Minnesota has no formula for faculty productivity or opportunity costs, one simple estimate was based on salary. Applying the composite faculty salaries for Category 1 universities in our region from the March 4, 2000, issue of *Academe* and the University of Minnesota fringe benefits rate against the estimate of 9,600 hours spent by faculty in workshops or reading materials, we estimate the cost of faculty participation at $425,000. However, the benefit side of this equation is even harder to estimate. Certainly the potential liabilities exceed the total cost of the program, including loss of faculty time.

Evaluation

Assessing for continuous course improvement
The RCR curriculum is currently offered in 2 parts of 3 hours each. At the end of each session, participants are asked to complete a one-page course evaluation form which asks 1) whether the level of content for each topic is appropriate, 2) whether the information on each topic is

useful, 3) whether the session increased your understanding, and 4) whether the materials and resources were helpful. Finally, there is a place for comments. Data from these forms have been summarized and used to make course improvements.

During the first 6 month period, 66% of the participants (N=522) returned the evaluation for part 1; 43% (N=1162) for part 2. In general, 80% of the participants judged the material presented to be appropriate. Lists of resources and websites were considered the most useful resources. Early on, criticisms outpaced satisfactory remarks 3 to 1. Constructive comments included: make the course more interactive, provide readings ahead of time, incorporate web based materials, and shorten the length of time. Subsequent iterations of the course adopted these suggestions. As a result, the overall rating of usefulness improved, from 2.7 to 3.0 on a 4 point scale (with 4 being very useful) for part 1 and from 2.5 to 2.9 for part 2. In addition, there were fewer critical comments, and the number of statements of praise increased.

Reflecting on the course evaluation data and our efforts at course improvements, we have identified the following contributors to participant satisfaction:

Interactive programming. The more interactive the program, the more it is viewed as useful.

Group size. Smaller groups are better received than larger groups.

Presenters from diverse disciplines. Participants have been less satisfied when the presenters are all from the same discipline.

Topic. Some topics seem to be inherently more interesting than others. For example, authorship seems to be rated as most interesting irrespective of who presents the material. Other topics, like intellectual property and conflict of interest typically get lower ratings for usefulness. However, when we have broadened the topic of intellectual property to include more on copyright, there were some improvements in rating. Staff have speculated that in areas like intellectual property and conflict of interest may be inherently dissatisfying as it is seldom possible for the presenter to give definitive answers to questions.

Assessing promotion of responsible conduct

Documenting faculty participation in an initial and on-going educational program in RCR demonstrates compliance with a federally mandated corrective action plan (e.g., the NIH plan currently in effect for the University of Minnesota). It does not, however, provide evidence that the attitudes, values, and behaviors that gave rise to the disciplinary action have changed. Likewise, installing a model system for financial accountability, such as the Electronic Grants Management System (EGMS), can alert an individual faculty member and his/her unit head when a proposed action is not within the bounds of sanctioned behavior. It does not, however, assure that the moral climate in which research is conducted is enhanced, or will it necessarily improve the ability of investigators to interpret ambiguous situations and identify better choices. If we hope to provide evidence that we have improved the integrity of the researcher and climate of the institution, we need measures that assess the more elusive outcomes of the research ethics enterprise and that can be used to examine the effectiveness of our educational programs and compliance systems.

In Fall of 1999, a faculty group was convened to identify opportunities for assessment of outcomes. The following were identified:

Self-assessment questions in web-based modules. Self assessment items have been included in several topics: Fiscal Responsibility, Intellectual Property, Conflict of Interest, Informed Consent, Protecting Human Subjects. Although self assessment items are included, we have decided not to invest resources to assess knowledge level outcomes.

University-wide climate surveys to track perceptions of ethical research practices. The last Faculty and Staff Climate Survey of the University of Minnesota was conducted in 1997, with a summary reported in 1999. Questions are being prepared for the next survey. The purpose will be to track perceptions of the extent to which the University climate supports ethical conduct generally. Questions would be directed toward ethical research practices as well as issues of academic integrity.

Narrative interviews of unit administrators. In addition to eliciting their perceptions of the norms of research conduct, interviews with unit administrators is a way of identifying areas needing attention.

Graduate student perceptions of the doctoral experience. Melissa Anderson directs the

Academic Life Project, funded by NSF, which studies the normative experiences of doctoral students (see paper by M. Anderson in these proceedings for additional information on this study).

Adaptation of measures of ethical reasoning and role concept. One reason for the paucity of information on assessment of instructional effects in this area is the lack of well-validated outcome measures. Measures must be grounded in a well-established theory of ethical development and be sufficiently user friendly to enable their use for a variety of purposes. We propose to develop two outcome measures: (1) a measure of ethical reasoning and judgment about common problems arising in the research setting, and (2) a measure of role concept, i.e., how the researcher understands his/her role relative to other researchers. The measures will assess two of the four dimensions of competence described by Rest's Four Component Model of Morality (Rest, 1983). The areas are chosen because prior studies support the usefulness of the methods for outcome assessment and for demonstrating the links between performance and day-to-day ethical behavior. The two measures will be modeled after existing measures designed for assessing the outcomes of ethics education in dentistry. (See paper by M. Bebeau in these proceedings for additional information on these approaches).

In summary, a national effort is required to design outcome measures that can be used to assess the effectiveness of institutional education programs in RCR. Measures must well-grounded theoretically, well validated, and sufficiently user friendly to enable their use for a variety of purposes. Such purposes may include: 1) determining the range of criteria that define competence in topic areas among different disciplines, 2) conducting a needs assessment to identify areas where instructional resources should placed, 3) identifying individual differences or problems that require intervention or remediation, 4) providing feedback to individuals, departments, and institutions on research ethics competence, 5) determining the impact of current programs, and 7) studying the relationship between competence and ethical behavior.

Bibliography

1. Rest, J. Moraliaty. In Mussen PH (ser. ed.), Flavell J, Markman E (vol. eds.), Handbook of Child Psychology, Cognitive Development, Vol. 3, (4th ed). New York: Wiley; 1983; p. 556-629.

Being a Scientist: Educating for Ethical Conduct

Chloe Doby Little, Department of Physical Therapy, Western Carolina University, USA

Katherine Le Guin White, Department of Physical Therapy (retired), Western Carolina University, USA

Key Words: Ethical conduct, Problem-based learning curriculum, Reflection-in-action skills, Reflection-in-experimentation, RCR

This project is predicated on a reflective way of life for being a scientist as the epistemological foundation for educating health professions students in the ethical conduct essential for scientific integrity and progress. Thus, being a scientist exemplifies a reflective way of life; and educating health professions students for ethical conduct embodies the reflective practitioner epistemology explicated by Schon in his books, *The Reflective Practitioner* and *Educating the Reflective Practitioner* (1, 2). Schon (1) challenges traditional professional curricula and educators that continue to implement course content based on the positivist, technical-rational epistemology of the nineteenth and twentieth centuries. The reflection-in-action epistemology Schon (2) pioneered offers health professions educators and practitioners a theoretical system of knowledge for helping faculty in science-based professions education update curricula.

The thesis of this project is that a transitional problem-based learning (PBL) curriculum in the allied health professions provides an excellent framework for education of reflective practitioners. Reflective practitioners are problem solvers and ethical scientists. Faculties who are themselves exemplary reflective researchers and teachers can teach ethics through successful PBL experiences that guide health professions students in development of ethical conduct as the foundation for their way of life as science-based, reflective practitioners.

A transitional PBL curriculum in the health professions is structured to guide students from acquisition of new information and knowledge through application of that knowledge in solving clinically-based problems to reflection-in-action as practitioners. Put another way, the transitional PBL curriculum helps health professions students progress from information gathering and knowledge warehousing to practitioners who know through reflection-in-action and are therefore wise clinicians rather than master technicians.

Faculties, who are science-based, reflective practitioners and instructors, integrate scientific research, scholarship, and teaching. Successful implementation of reflection-in-action epistemology in health professions curricula depends in large measure on the participation of wise, dedicated faculty whose ethical conduct as scholars and as teachers is manifested in their successful participation in those reflective dimensions of problem-based learning experiences.

Introduction

Keith-Spiegel, et al., (3) report that scientific misconduct is socialized during undergraduate years

Corresponding author: Chloe D. Little, PW Box 2145, Cullowhee, NC 28723, 828-227-2288 (voice), 828-227-7071 (fax), littlec@rabun.net

with students believing that significant results will earn them better grades. Recent research by Davidson, et al., (4) lends additional support to these findings. One also can speculate that scientific misconduct reflects the attitudes of society. Dishonesty and misrepresentations have become commonplace and acceptable in the absence of social sanctions against these behaviors and also as a result of increased professional competition and increased pressure to produce. Since the 1940's the incidence of self-reported academic misconduct by college students has risen 55-78 percent. (5) Other examples of misconduct include medical school faculty applicants misrepresenting research citations, (6) ethics committees endorsing unnecessary research, (7) peer-reviewed journals editors misappropriating authorship, (8) and researchers faking data in experiments or failing to report unfavorable results (9). Some researchers suggest that there has been a "reorientation away from traditional values," especially in scientific inquiry (10) . Others speculate that fraud and dishonesty in scientific research are the inception rather than the rule (11).

Regardless, scientists and institutions must maintain quality and integrity in scientific research if progress and public support are to be sustained. To promote responsible research, college and university faculties must sensitize future scientists to the critical issues in research ethics and guidelines. Also, the National Institutes of Health requirements mandate all institutions participating in training grants show they provide instructions to faculty and students in the principles of scientific integrity (12). Additionally, the final report of the Commission on Research Integrity noted the importance of providing "formal and informal educational opportunities to sensitize both junior and senior scientists to critical issues in research ethics and their institution's guidelines" (13, p.16). Although expecting college and university faculties to single-handedly prevent research misconduct is unrealistic, faculties can create informal learning environments to promote high standards by engaging students in open discussions of ethical and unethical research practices, carefully supervising and mentoring student research, encouraging responsible data management, and modeling ethical behaviors. Faculties also can create formal methods for integrating the study of scientific values and responsible conduct in the academic courses.

This project presents informal and formal methodologies to encourage health professions graduate students to develop reflection-in-action skills and values that foster ethical practice in health professions services and clinical research. The ultimate goal is to describe a curriculum for promoting active student learning throughout a series of scientific research courses.

Implementing Problem-Based Learning Curriculum in Scientific Research for Graduate Health Professions Students

First semester course content includes three case-based problems for students to study and discuss: university-specific guidelines for conduct of scientific research, how to construct a research project, and the virtues of ethical research. Second semester course content is focused on student implementation of the research project constructed during the first semester. In subsequent semesters, students reflectively examine with faculty mentors their completed student projects for ethical integrity.

Learning issues in the first case-based problem explored in semester one focused on defining scientific misconduct through differentiating negligence from deliberate dishonesty and examining institutional research policies, especially distinguishing human and non-human research, confidentiality, and the obligations of scientific researchers. Students complete an institutional review board proposal for their subsequent projects. The second problem progresses students to application of those skills and behaviors learned in the first case-based problem on the rudiments of responsible scientific conduct. Learning issues for this case include practicing ethical data management and examining the ethical content of published research studies. The third problem is structured to concentrate student learning on management of conflicting interests, determination of criteria for multiple authorship, reporting scientific misconduct, and the process by which research grants are awarded.

Second semester learning issues arise from reflection on students' performances as they begin to conduct their research projects, structured during the first semester. Throughout this course faculty and student reflection-in-action and faculty mentoring become critically important. Learning experiences during this semester are more informal than those structured

for the first course. Students complete their projects in subsequent semesters, varying from one to three. Equally critical throughout these times is informal student-faculty discussions, supervision, and reflection that occurs during regularly scheduled small group or individual meetings.

Benefits of Problem-Based Curriculum and Learning Experiencesfor Faculty and Students

Students and faculty alike are beneficiaries of PBL experiences and curricula. Students develop problem-solving skills through student-directed discussions and information gathering assignments. They also learn to become self-directed and independent learners, habits that equip them for lifelong learning in practice communities, even in those remote settings where colleagues and library resources may be scarce. As they become more independent learners, students begin to actively demonstrate increasingly critical, creative thinking.

Assessment of one's peers during PBL experience is an essential dimension of PBL that requires active participation of all students in a learning group. To that end, students must learn to assess themselves and their colleagues in honest, thorough, deep, and sincere ways. Learning to work critically in this manner helps students reach greater depths of understanding the importance of frequently and realistically evaluating their performance as team members and learners; they also become skilled in applying the same sensitivities to evaluating the participation and performance of their peers in learning groups. These assessment skills and values also relate to other aspects of PBL: information management, creation of measurable knowledge bases for solving problems, and assessing peers, social and ethical skills, communication effectiveness, and the ability to work effectively as a team member.

Finally, development of leadership skills is fostered through revolving, shared group leadership. For each problem-solving session, students select a group leader, facilitator, and recorder. All group members serve in each capacity throughout a semester.

If PBL is to be successful, faculties must become models and coaches, relinquishing their traditional roles as lecturers and purveyors of information. In this role, faculties develop skills that monitor student learning during a problem-solving session and throughout the curriculum. To properly monitor student learning, faculties must become proficient in classroom reflective behaviors that probe and challenge students' thinking conclusions and processes, keep students involved throughout exploration of the problem, adjust levels of challenge to students, and manage group dynamics so that processes move toward constructive resolution of the problem. Development of learning materials and writing comprehensive clinical problems that challenge students demand faculty creativity and planning that exceed those faculty demands imposed by a curriculum predicated on traditional technical-rational epistemology. Faculties relinquish the resident expert status to become guides for student learning that is independent and self-directed. Faculty expertise in asking rather than telling, planning and guiding rather than showing is essential for successful discussions and problem solving sessions.

Formal and Informal Methodology Designs

Problem-based learning methodologies presented here are designed to encourage first-semester health professions graduate students to develop reflection-in-action skills and values for ethical practice as clinicians and as researchers. The ultimate goal of the methodology is to promote active student learning in the education of future scientists who will consistently demonstrate ethical scientific research behaviors.

As with the previously discussed benefits of PBL for students and faculty alike, effective PBL methodology design occurs only when faculties and students participate successfully in the process. At a minimum, faculties must openly discuss with students during learning group sessions those ethical and unethical behaviors in scientific research reported in the literature and in the faculty member's experience as a scholar-researcher. Faculties also must carefully and continuously supervise student research activities while mentoring student development as novice researchers. To be credible leaders for development of ethical behaviors in students, faculties must be personally engaged in ongoing and successful scientific research and scholarship.

Student involvement in design of PBL methodology requires full participation of all group members in researching the literature

available on ethical and unethical practices in scientific research. Students also must learn to engage faculty and student peers in reflective discussions throughout the problem solving group experiences.

Finally, students must demonstrate learned ethical behaviors in their own student research projects completed after their first semester.

Formal faculty and student responsibilities for methodology design and successful implementation are focused on scientifically rigorous planning and participation guidelines. Faculties are charged with responsibility for developing curriculum materials that include a series of complex, real world, "ill-structured" problems to stimulate learning, integration and organization of learned information that ensure application of past learning to future problems. Curricular materials include learning objectives for each PBL problem, definition of PBL techniques, appointment of small groups of 5-7 student learners, identification and instruction of tutors, guidelines for student leadership process and responsibilities during group learning sessions, and development of assessment tools. Beyond these process design matters, the essential faculty responsibility is creating multiple cases that form the bases for student learning. Without solid, reality-based clinical cases, the process cannot proceed as a valid or effective learning experience. As stated earlier, faculty also must model the values promoted as ethical conduct for scientists. They must consistently demonstrate their ability to reflect-in-action as they participate in the group learning experiences.

Students likewise have many formal responsibilities for achieving successful PBL. Students must learn to formulate hypotheses as individuals and as learning team members. They must learn to participate effectively and responsibly as group members for many outcomes, including designing a plan to solve the problem, researching available and pertinent information, justifying individual and group decisions and conclusions, recognizing multiple acceptable solutions to a given problem, evaluating the performance of themselves, their peers, and their tutors, and demonstrating novice reflection-in-action skills and values.

Discussion and Conclusion

Problem-based learning, based on small group discussion and clinically-based problems, encourages independent learning during which students develop depth of understanding of content (14). Through PBL students become more involved in and responsible for their own learning. The objectives of PBL are to assist the process of active learning by students as they develop effective clinical reasoning skills, such as critical appraisal, decision making, collaboration, and self-directed learning habits in order to participate effectively and actively in the small group discussions during the problem solving of cases. (15, 16) Each problem should be designed to provoke critical inquiry, to encourage independent access to multiple and diverse learning resources, and to generate lively, focused, and pertinent small group discussions. Reflection-in-action during and after completion of a problem promotes transfer of learning as well as generation of new concepts (16). Recent research findings suggest PBL curricula are effective methods of learning and that students successfully transfer knowledge and skills in timely and meaningful ways (17, 18, 19).

Researchers have shown PBL promotes higher order thinking skills (16). PBL is a curriculum approach that places students in the active role of problem solver during the process of constructing meaning from case-based problems that mirror real-world situations. Throughout the process students develop problem-solving and information gathering strategies, reflection skills, and discipline-specific knowledge bases. In the absence of actual clinical experiences during problem solving discussions, students learn to make judgments based on facts, information, logic, and rationalization alone, they must use higher thinking orders to justify decisions based on application of learned principles. Nevertheless, the defining measurement of learning during an academic course is the quality of research produced by the student, an outcome that may not be evident throughout the span of the course. Therefore, continued supervision and mentoring of a student's future research activities beyond the first semester is essential for facilitating ethical development. The authors believe that through PBL students will exhibit reflection-in-experiment skills that will culminate ultimately in reflection-in-action skills1 as they complete their student research projects and move toward mastery as scientific researchers.

Bibliography

1. Schön DA. *The reflective practitioner: How professionals think in action*. NY, NY: Basic Book, A Division of Harper Collins Publishing; 1984.
2. Schön DA. *Educating the reflective practitioner: Toward a new design for teaching and learning in the professions*. San Francisco: Jossey-Bass Publishers; 1987.
3. Keith-Spiegel P, Lee B, Spiegel D, Zinn-Monroe M. Research described in Scientific Misconduct: An annotated bibliography by Keith-Spiegel, P., Aronson, K., & Bowman, M. 1994. Available from: URL: http://www.lemyne.edu/OTRP.
4. Davidson EW, Cate HE, Lewis CM, Hunter M. Data manipulation in the undergraduate laboratory: What are we teaching? First ORI Research Conference of Research Integrity; 2000 November 19-21; Bethesda, MD. 2000.
5. Educational Testing Service. Cheating is a personal foul: Academic cheating background. 1999. Available from: URL: http://www.nocheating.org/adcouncil/research/cheatingbackgrounder.
6. Goe LC, Herrera AM, Mower WR. Misrepresentation of research citations among medical school faculty applicants. *Acad Med*, 1998; *73*: 1183-1186.
7. Savulescu J, Chalmers I, Blunt J. Are research ethics committees behaving unethically? Some suggestions for improving performance and accountability. *BMJ*, 1996; *313*(7069): 1390-1393.
8. Flanagin A, Carey LA, Fontanarosa PB, Phillips SG, Pace BP, Lundberg GD, Rennie D. Prevalence of articles with honorary authors and ghost authors in peer-reviewed medical journals. *JAMA*, 1998; *280*(3): 222-224.
9. Friedberg M, Saffran B, Stinson T, Nelson W, Bennett C. Evaluation of conflict of interest analysis of new drugs used in oncology. *JAMA*, 1999; *282*(15): 1453-1457.
10. Bechtel HK, Pearson W. Deviant scientists and scientific deviance. *Deviant Behav*, 1985; *6*: 237-252.
11. Knight JA. Exploring the compromise of ethical principles in science. *Perspectives Biol Med*, 1984; *27*: 432-441.
12. National Institute of Health. Requirements for programs on the responsible conduct of research in National Research Service Award Institutional Training Programs. *NIH Guide for Grants and Contracts*, 1989; *19*: 1-35.
13. U. S. Department of Health and Human Services, Commission on Research Integrity. *Integrity and misconduct in research*. Washington, DC: Office of Research Integrity; 1995.
14. Schmidt HG. Problem-based learning: Rationale and description. *Med Educ*, 1983; *17*: 11-16.
15. Dolmans DH, Schmidt H. What drives the student in problem-based learning? *Med Educ*, 1994; *28*: 372-380.
16. Barrows HS. (1998). The essentials of problem-based learning. *J Dent Educ*, 1998; *62*(9): 630-633.
17. Borduas F, Carrier R, Drouin D, Deslauriers D, Trenblay G. An interactive workshop: An effective means of integrating the Canadian Cardiovascular Society clinical practice guidelines on congestive heart failure into Canadian family physicians' practice. *Can J Cardiol*, 1998; *14*: 911-916.
18. Dolmans DH, Schmidt H. The advantages of problem-based curricula. *Postgrad Med J*, 1996; *72*: 535-538.
19. Dolmans DH, Gijselaers WH, Schmidt HG, van der Meers SB. Problem effectiveness in a course using problem-based learning. *Acad Med*, 1993; *68*: 207-213.

*Encouraging Accountability in Research: A Pilot Assessment of Training Efforts**

Anna C. Mastroianni, School of Law and Institute for Public Health Genetics, University of Washington, USA

Jeffrey P. Kahn, Center for Bioethics, University of Minnesota, USA

Keywords: Pilot assessment, RCR education and training, Training grants

The objective of this pilot assessment was to describe the response of a sample of grantee institutions to the federally-mandated training requirement in the responsible conduct of research that is part of NIH Training Grant (T32) funding. Materials collected by the Department of Health and Human Services (DHHS) were reviewed and described with the following five research goals:

- describe the target audience for training programs
- describe the locus of instructional responsibility for training programs
- describe whether all trainees at an institution participate in the same training program
- describe the program approaches, materials used and program contents
- create a source of baseline information for planning evaluations of future training programs
- identify areas for further research and analysis

Methods

The sample consisted of a collection of materials assembled by DHHS. These included syllabi, course outlines, case studies, reading lists, institutional research policies, and other information provided by training grant recipient institutions about their research ethics programs. In June 1996, the Office of Science Policy, Office of the Assistant Secretary for Planning and Evaluation, DHHS, sought to create "a library of course materials that are being used by T32 grantees." A letter was sent to a stratified sample of T32 grantees requesting "any training materials currently used to instruct trainees in research integrity and misconduct" (1). The stated goal of collecting this library of information was to provide an understanding of training programs in the responsible conduct of research, including the range of institutional approaches for meeting the training grant requirement. This information was not collected as part of assessing regulatory compliance or as part of any oversight effort, but to create a resource and a source of baselines information for planning evaluations of future training programs (2).[1] This sample served as a convenient and best available sample for this review.

Excerpted from: Mastroianni A, Kahn JP. Encouraging accountability in research: A pilot assessment of training efforts. *Accountability in Research* 1999;7:85-100. Some policy implications of the results presented here are also discussed in Mastroianni, A.C. and Kahn, J.P. The importance of expanding current training in the responsibleconduct of research. *Acad Med* 1998; 73(12):1249-1254.

Corresponding author: Anna Mastroianni, University of Washington School of Law and Institute for Public Health Genetics 1100 NE Campus Parkway, Seattle, WA 98105-6617, 206-616-3482 (voice), 206-616-3480 (fax), amastroi@u.washington.edu.

DHHS contacted awardees at 50 of the 210 institutions that held training grants as of October 1995. (3) DHHS selected these 50 based on number of training grants, geographical location, status as public or private institution, and number of T32 trainees at the institution. For those institutions with multiple training grants, individual grants were selected for inclusion in the sample in order to obtain diverse representation. Selection factors included: the number of trainees, the distribution of pre- and post-doctoral students, and representation of clinical and basic research.

DHHS contacted Principal Investigators by telephone and follow-up letter, and requested that they provide "any training materials currently used to instruct trainees in research integrity and misconduct [including] materials such as the syllabi, course outlines, case studies, reading lists, institutional codes of conduct in research, etc., [and] any information [that] readily . . . describes the context in which such materials are introduced to students and the method of training" (4). Respondents from 45 of the 50 institutions contacted provided information concerning a total of 75 training grants.

Access to and copying of these publicly available materials was provided by the Office of Science Policy, Office of the Assistant Secretary for Planning and Evaluation, DHHS, in November 1996.

Approach

A coding form was developed as a method to collect and summarize information from the sample. Descriptive statistics were calculated using standard statistical software.

The characteristics of the sample were described at the level of either the institution (n=45), or the responsible conduct of research training program (n=75). In order to understand whether institutions shared characteristics based on number of training grants, the sample of institutions was stratified into thirds by number of training grants. For this purpose, these groupings were categorized as: "low-density" institutions (14/45 [31.1%] of the institutions) which held four or fewer training grants; "medium-density" institutions (15/45 [33.3%] of the institutions) which held from five through nine training grants; and, "high-density" institutions (16/45 [35.6%] of the institutions) which held ten or more training grants. Institutions also could have been grouped by total number of trainees. In examining total number of trainees and number of T32s against other variables, each was found to be a proxy for the other. Variables, where appropriate, are grouped by numbers of T32s only.

Results

There were 45 institutions in the sample representing 660 T32s (number of T32s at each institution ranges from 1 to 60, with a median of 6) and 4,883 trainees (number of T32 trainees at each institution ranges from 3 to 507, with a median of 38). Responses concerning 75 training grants were represented in the sample.

Of the 45 institutions, 25 [55.6%] were public educational institutions, 17 [37.8%] were private educational institutions, and 3 [6.7%] were non-academic institutions (i.e., a professional organization, a non-profit service provider, and an independent research organization).

Institutional Characteristics

The sample was reviewed to determine the target audience for the training programs. Two-thirds of institutions represented in the sample required that only T32 trainees receive training in the responsible conduct of research. In this sample, this result was not affected by the number of training grants held by the institution: 9/14 [64.3%] of low-density, 10/15 [66.7%] of medium-density, and 11/16 [68.8%] of high-density institutions required training only for T32 trainees. Over one-quarter of all of the institutions, however, required much broader participation of either all trainees in the school or college, all graduate students or all trainees in the institution.

In half (23/45 [51.2%]) of the institutions represented in the sample, the responsibility for the responsible conduct of research training program was located at the departmental or Principal Investigator level. Another quarter located the responsibility at the institutional level. In the materials submitted, 4 [8.9%] of the institutions placed responsibility for the program in their ethics faculty. The institutions that placed responsibility for the program in their ethics faculty were among the highest-density institutions in the sample. They each had 18 or more training grants, and represented the top quarter of the sample by number of training grants. The majority of low-density and medium-density institutions had the locus of

program responsibility at the department level [64% and 66%, respectively], while the majority of high-density institutions had the locus of program responsibility above the department level [75%].

For those 41 institutions with more than one NIH training grant, 24 [58.5%] used the same responsible conduct of research program for all those required to receive training in the responsible conduct of research. As the number of training grants at an institution increased, the proportion of institutions utilizing the same responsible conduct of research training program decreased. Seven of the 10 [70%] low-density, 9 of the 15 [60%] medium-density, and 8 of the 16 [50%] high-density institutions used the same program for all trainees.

Program Characteristics

The material from the 45 institutions in the sample included information from 75 training grants. Depending on the characteristic being examined, the following analyses were based on either the number of institutions (n=45) or the number of programs (n=75). The denominator is noted in each case.

Program approach

Submitted materials indicated that one-quarter of the programs specifically tailored training to the trainee population, with either discipline-specific focus or both general and discipline-specific material.

Of the 45 institutions, 28 [62.2%] had a formal course in place to satisfy the training grant requirement. A greater proportion of medium-density and high-density institutions utilized a formal course than did low-density institutions: 5 of the 14 [35.7%] low-density institutions, 13 of the 15 [86.6%] medium-density institutions, and 10 of the 16 [62.5%] high-density institutions had a formal training course in place.

Fourteen [31.1%] of the institutions represented in the sample had programs that indicated the availability of ethics training that could be taken to supplement the course or training offered to satisfy the training grant requirement.

Only two institutions indicated that formal training was provided to faculty who then carried out the required responsible conduct of research training—a "train the trainer" approach. These two institutions were among the highest-density institutions.

Lecture was the most popular method of instruction represented in the sample (53/75 [70.7%]). (Table I) To examine whether programs relied solely on lectures to satisfy the requirement, the frequency of lecture format in combination with other methods of instruction was determined. (Table II) For those programs that used lectures as a method of instruction, only a small proportion (4/53 [7.5%]) did not supplement lectures with some less didactic method or methods of instruction that provide opportunities for greater interaction. It is interesting to note that the materials indicated that there was very little use of "brown bag" discussions to satisfy the requirement.

Contact hours could be determined for 42 of the 75 [56%] programs for which information was received. The median number of contact hours for these programs was 10 hours. The range was from 4 to 72 contact hours.

Method of Instruction*	**# [%]**
Lecture	53 [70.7]
Case study	42 [56.0]
Small group	36 [48.0]
Seminar	21 [28.0]
Student presentation	11 [14.7]
Mentor	9 [12.0]
Brown bag	1 [1.3]
Computer	0 [0]

Table 1. Method of program instruction. n=75
** programs could have more than one method of instruction*

Methods of Instruction	**# [%]**
Lecture only	4 [7.5]
Lecture + seminar	3 [5.7]
Lecture + small group	11 [20.8]
Lecture + case studies	16 [30.2]
Lecture + small group + case studies	14 [26.4]
Lecture + seminar + small group	3 [5.7]
Lecture + seminar + small group + case studies	1 [1.9]
Lecture + brown bag + small group	1 [1.9]

Table 2.Combination of methods of program instruction with lectures. Fifty-three programs used lecture as part of their instructional format. n= 53

Program Contents

Material from the 75 training grants was reviewed to determine whether course content included the five topic areas recommended by NIH in the NRSA policy—conflict of interest, responsible authorship (including issues of peer-review, plagiarism, and research reporting), policies for handling misconduct (including institutional policies, federal policies, whistleblowing and reporting misconduct), policies regarding the use of human and animal subjects, and data management (including fabrication, falsification, handling research data, materials and information, and data and objectivity).

Fifty-one [68%] of the T32 programs covered four or five of the NIH recommended program content areas while 24 [32%] of the T32 programs covered three or fewer of the categories. The top five ranked categories fell within the five NIH recommended program content areas, and the top ten ranked categories were addressed by at least half of the T32 programs. (Table 3)

Rank	Content Area	# [%]
1	Authorship	65 [86.7]
2	Data Management	56 [74.7]
3	Human Subjects	53 [70.7]
4	Animal Use	51 [68.0]
5	Conflict of Interest	49 [65.3]
6	Institutional Policy	45 [60.0]
7	Skills-Writing	44 [58.7]
7	Confidentiality	44 [58.7]
9	Skills-Publishing	38 [50.7]
10	Intellectual Property	37 [49.3]
11	Mentor/Mentee	35 [46.7]
12	Information Sharing	24 [32.0]
13	Whistleblowing and Reporting Misconduct	22 [29.3]
14	Moral Reasoning	21 [28.0]
15	Other Content	20 [26.7]
16	Federal Policies	16 [21.3]
16	Grants Management	16 [21.3]
18	Skills-Grant Preparation	15 [20.0]
19	Organizational Structure	14 [18.7]
20	Skills-Oral Presentation	11 [14.7]
21	Science and Society	10 [13.3]
22	Laboratory Safety	9 [12.0]
23	Skills-Teaching	6 [8.0]
24	Skills-Tenure	4 [5.3]
24	Skills-Funding	4 [5.3]
26	Skills-Jobs	3 [4.0]

Table 3. Ranking of program content areas. N = 75; programs can have more than one content category.

Content issues that were identified by fewer than half the programs include:

- whistleblowing and reporting misconduct (22 of 75 programs)
- the more theoretical issues encompassed in a category we labeled "moral reasoning" (21 of 75 programs)
- social issues encompassed in a category we labeled "science and society" (10 of 75 programs)
- development of certain skills necessary for becoming a productive scientist, e.g. grants preparation and funding, job hunting, oral communication, tenure, teaching, etc., (3 to 15 programs).

General skills related to publishing and writing received greater attention, with 38 and 44 programs addressing them, respectively.

Thirty-six of the 75 [48%] programs provided syllabi or other similar program materials in the information sent in response to the DHHS request. Of those, 6 [16.7%] identified goals and objectives for the responsible conduct of research training program. Based on this limited information, few programs set forth traditional goals and objectives for their educational efforts.

Training Materials

The information submitted was reviewed to identify the most frequently noted training materials used by programs. The top three referenced training materials were: 1) institutional policies concerning the responsible conduct of research (45/75 [60%]); 2) Korenman et al., *Teaching the Responsible Conduct of Research through a Case Study Approach: A Handbook for Instructors* (5) (30/75 [40%]); and, 3) the National Academy of Science's On Being a Scientist (6) (24/75 [32%]). While the institutional policies are specific to

each institution, Korenman et al. (5) and NAS (6) are prepared and widely distributed by professional societies. Of the materials referenced in the sample, the following four each are marketed as offering complete training materials in the responsible conduct of research without need to supplement: Korenman et al. (5); Macrina, *Scientific Integrity: An Introductory Text with Cases* (7); the American Association for the Advancement of Science's *Integrity in Scientific Research: Five Video Vignettes* (8); and Bulger et al., *The Ethical Dimensions of the Biological Sciences*. (9) Forty-three [57.3%] of the programs used one or more of these four materials. A greater proportion of high-density institutions (12/16 [75%]) used at least one of these four "ready-to-use" training materials, than did low- or medium-density institutions (7/14 [50%]; and 7/15 [46.6%] respectively).

Discussion

In this sample, training in the responsible conduct of research in response to the NIH requirement was most often directed at T32 trainees. While the NIH policy encourages expanding training to others, it requires that only T32 trainees receive such training. If this result is representative of institutional commitment to training in the responsible conduct of research, future scientists' exposure to responsible conduct of research will largely depend on their source of funding. The characteristics of the minority of institutions that make a broader commitment to responsible conduct of research education and training for its trainees deserve further exploration.

The T32 recipient institutions in the sample employed a diversity of approaches to satisfying the training grant requirement. Approaches varied both among and within institutions. Further, the number of T32s held at the institution had some impact on how the training grant requirement was met.

Locating program responsibility at the departmental or Principal Investigator level, as did about half of the institutions in the sample, may offer ethics training that is more tailored to the trainees' disciplines. In the materials reviewed, a quarter of the programs offered some discipline-specific training. Further research is necessary to determine whether a relationship exists between discipline-specific training and location of program responsibility within an institution.

The finding that a greater proportion of high-density institutions placed program responsibility above the departmental level may indicate that as institutional demand for responsible conduct of research training programs increases, more shared institutional resources are sought. However, based on T32 density, those institutions with the highest density had the smallest proportion that utilized the same responsible conduct of research training program for all trainees. This finding may be attributable to more diverse training programs for which different approaches are used, even if some institutional resources are shared. Perhaps the administrative level at which the ethics training decision is made affects the institutional approach. Future research might focus on examining this question, and the sharing of institutional resources regardless of any differences in program approach.

The small number of institutions that placed responsibility for teaching in their ethics faculty may be a reflection of the fact that institutions with greater numbers of training grants are more likely to have ethics faculty—it would be interesting to compare the characteristics of institutions that have ethics faculty and place program responsibility in them.

Contrary to the expectations of the authors, lecture format alone was rarely used; nearly two-thirds of the programs employed lectures plus additional instructional approaches. Also contrary to popular belief among teachers of responsible conduct of research, brown bag discussions were rarely identified as an approach used to satisfy the training grant requirement. The wide range of contact hours offered by programs underscores the great diversity in the implementation of the requirement.

The majority of programs (51/75 [68%]) specifically addressed four or five of the NIH-recommended subject categories. Either the recommendations in the NIH policy have influenced program content or the subject categories are well-chosen and represent the commonly accepted basic issues in the responsible conduct of research.

Some variation in the subject matter covered by programs may result from differences in the needs of trainees in basic versus clinical research. However, four of the five NIH-recommended categories are relevant to all scientific research, i.e, one category, human and animal research issues, may not be relevant to all researchers.

Therefore, one would expect a higher proportion of programs than was observed to address at least four of the categories.

Most educational efforts in other areas typically identify goals and learning objectives as a way of focusing teaching approaches and assessing their success. In this sample, few of the T32 programs (6/36) identified goals and objectives. This would seem to imply that programs do not approach training in the same manner they would typically approach other educational efforts.

"Ready made" materials and materials sanctioned and made available by professional organizations were a popular source of training materials. This underscores the need to ensure that these materials, which are so heavily relied upon, are of high quality, complete, appropriately tailored for the target audiences, and widely available.

The most popularly used material, institutional ethics policies, is critical for trainees' basic understanding of responsible conduct of research. The proportion in the sample who used these policies as a educational tool could be viewed as unexpectedly low (45/75 [60%]).

Future Research

In addition to the findings discussed, this review indicates the need for further research on institutional approaches to education and training in the responsible conduct of research. First, additional research is needed on the characteristics of training programs. A description of primary and participating instructors in training would be instructive, particularly knowing the extent to which an institution's ethics faculty are involved in program development, administration and teaching. In addition, it would be useful to understand the differences in approach and content of training provided for trainees in different disciplines, particularly in the clinical sciences as compared to the basic sciences. This information would point to differences in the perceived needs for subgroups of trainees, and could aid development of appropriate materials and programs, for example, the use of core programs with tailored components.

Second, research is needed on the effectiveness of training initiatives. Evaluation of a variety of programs and their approaches would be particularly useful. Some target areas for program evaluation include:

- the use of a core program plus tailored discipline- and skill-specific components
- resource utilization-sharing by multiple programs within and among institutions
- skill-based training programs, with assessment of trainee competencies
- the importance of goals and objectives of programs as away to focus the educational effort
- resource needs for "train the trainer" approaches
- the effectiveness of stand-alone training materials
- the effectiveness of one-day programs compared to series of sessions

There is also a need to identify how to broaden current training efforts to ensure that all scientists-in-training are prepared to address ethical dilemmas in their professional careers, regardless of the source of funding for their training. Such initiatives might include education of institutional administrators about the importance of responsible conduct of research training beyond T32 trainees and the enlisting of institutional commitments for broadened training efforts. In addition, there is a need for improved dissemination of effective approaches to responsible conduct of research training in the relevant professional literature.

The results of this review should not be viewed as representative of responses to the NIH mandate at either the programmatic or institutional level because of the sample's limitations. The way the sample was selected and the generality of the government's request for materials may have had some impact on the results. Since the materials were collected independently from this review, a targeted questionnaire would provide more detailed information. However, the results of this review are a valuable first step in describing how institutions and investigators meet the mandate for training in responsible conduct of research

Conclusion

The intent of this pilot assessment was to describe for the first time how institutions and investigators are responding to the NIH mandate for training in the responsible conduct of research that is part of NIH Training Grant (T32) funding. The results provide a snapshot of the variety of

approaches used in programs across the country.

Understanding the range of approaches taken in the education and training in the responsible conduct of research is a crucial part of any effort to encourage accountability in research, on the part of trainees, researchers, institutions, and funders. Those engaged in training and education can gain important insights for further study given the diversity of approaches seen in this review, while at the same time pointing to the need for some consistency of training content. Further, education and training in the responsible conduct of research should be part of all the training of all scientists and not a function of the source of funding for training. Only by assuring the highest standard of research conduct, can we be confident that the trust the American people continue to place in biomedical research is truly deserved.

Acknowledgements

Thanks to William Raub and Lily Engstrom at the U.S. Department of Health and Human Services for providing access to the sample of materials, and to Wayne Hartley of the Center for Bioethics at the University of Minnesota for his assistance in statistical compilation.

Notes

1. DHHS staff selected this approach to the collection of resources because their primary purpose was to gain insights into the scope and character of the materials being used to teach responsible conduct of research, and in a way that minimized the reporting burden for the cooperating institutions. They recognized from the outset that this approach would enable only qualitative characterization at best, and unlike a formal survey, would not yield readily analyzable data. (DHHS, 1997)

Bibliography

1. DHHS. Raub WF. Office Communication. Director of the Office of Science Policy, Office of the Assistant Secretary for Planning and Evaluation, DHHS, to a sample of T32 recipient institutions, June 25 [on file with authors]; 1996a.
2. DHHS. Research integrity project. [Unpublished memorandum prepared by Preya Sharma on file with authors]. 1996b Aug 15.
3. DHHS. Raub WF. Personal Communication, DHHS; 1997 Oct 24.
4. National Institutes of Health. NIH program announcement PAR-98-006: Mentored scientist development award in research ethics. NIH Guide Grants Contracts 1997a; 26(37): Nov 7.
5. Korenman SG, Shipp AC, AAMC Ad Hoc Committee on Misconduct and Conflict of Interest in Research Subcommittee on Teaching Research Ethics. *Teaching the Responsible Conduct of Research through a Case Study Approach: A Handbook for Instructors*. Washington, D.C.: Association of American Medical Colleges, Division of Biomedical Research; 1994.
6. National Academy of Sciences. *On Being a Scientist* (2nd Edition). Washington, D.C.: National Academy Press; 1995.
7. Macrina FL. *Scientific Integrity: An Introductory Text with Cases*. Washington, DC: ASM Press; 1995.
8. American Association for the Advancement of Science. *Integrity in Scientific Research: Five Video Vignettes. Washington, D.C.: AAAS; 1996.*
9. Bulger RE, Heitman E, Reiser SJ, editors. *The Ethical Dimensions of the Biological Sciences.* Cambridge: Cambridge University Press; 1993.

A Plea for Pursuing New Dimensions of Assessment in the Teaching and Learning of Research Integrity

Carl Mitcham, Liberal Arts & International Studies, Colorado School of Mines, USA

Barbara M. Olds, Liberal Arts & International Studies, Colorado School of Mines, USA

Ronald L. Miller, Chemical Engineering & Petroleum Refining, Colorado School of Mines, USA

Keywords: Engineering ethics, Ethics in science, Learning assessment, Responsible conduct of research education, RCR, Scientific integrity, Science policy, Teaching assessment

Our basic thesis is simple: There are abundant research opportunities involved with the need to assess the teaching and learning of research integrity. In one sense this thesis is a cliché. Research opportunities are abundant everywhere; more research can be conducted on almost anything and everything–even in quite narrowly defined areas such as the quantitative assessment of teaching and learning about research integrity.

It is nevertheless possible to interpret our thesis in a broader and more provocative sense and to argue for breaking out of a restricting if well established, four-sided system of constraints. The teaching and learning of research integrity is, after all, concerned with integrity–from the Latin integritas, which signifies not only purity or correctness but also and more fundamentally soundness or completeness, the undiminished or unimpaired wholeness of a thing. Integrity is related to *integritas,* bringing together. There is more to ethics than what has been going on in research ethics, and research ethics will profit from more extensive connections than heretofore pursued.

Before making an effort to move beyond the constraints, it will be useful to describe in slightly greater detail the two-dimensional box in which this issue of assessing the teaching and learning of research integrity is currently confined.

Narrow Interpretations of RCR Education

It is increasingly common at research universities to teach courses or modules on research integrity or the responsible conduct of research (RCR)–as is now required by National Institutes of Health and Public Health Service grant award guidelines, and as has been reported more generally in Michael Davis (1). To date, however, efforts to measure the effectiveness of RCR curricula have been limited if not anecdotal. Nicholas Steneck's bibliographic background report for the present proceedings volume begins to identify such limits (2), although he is not as critical as we are of the present state of affairs.

Constituting a first restriction, the whole literature on research integrity is highly concentrated in the biomedical field. There are modest exceptions, but the most prominent instances of teaching and

Corresponding author: Carl Mitcham, Liberal Arts & International Studies, Stratton Hall 301, 1005 14th Street, Colorado School of Mines, Golden, CO 80401, 303-273-3648 (voice), 303-273-3751 (fax), cmitcham@mines.edu.

learning about RCR–and thus of possibilities for RCR assessment–are found in the health care fields, from general medicine to dentistry and diverse medical research specialities. Given the emphasis on informed consent issues in both research and clinical practice, and the public profile of regulations related to the treatment of animals in research, this is perhaps to be expected. It need not, however, be accepted without question.

A second restriction is that research ethics teaching focuses heavily on what may be termed internalist over externalist issues. Issues concerned with doing things right crowd out all discussions about what might be the right things to do; process overshadows substance. Questions of precisely how to handle data management, treat human and animal subjects, pursue publication, deal with conflicts of interest, and mentoring protocols dominate, at the expense of critical reflection on the proper ends to pursue with these methods (see the NIH Bioethics Resources on the Web at nih.gov/sigs/bioethics/researchethics.html, especially the NIH supported link to Resources for Teaching Research Ethics at medicine.ucsd.edu/research/ethics/resources).

Still a third restriction is that although formal RCR instruction obviously raises questions about whether such teaching makes a difference–whether it reduces research misconduct–confirming evidence remains slight. In fact, there is scant agreement even on the immediate goals of RCR teaching and learning, thus making it difficult to decide what would count as evidence for or against short- or long-term success. In consequence, many assessments of RCR education have produced ambiguous results.

Finally, a fourth restriction is that what unambiguous assessment results do exist have relied almost exclusively on the utilization and adaptation of two specific instruments, the Defining Issues Test (DIT) developed by James Rest (3) and the Sociomoral Reflection Measure (SRM) developed by John Gibbs (4), both of whom had studied with, and in their work attempted to more readily operationalize moral development theorist Lawrence Kohlberg's Moral Judgment Interview (MJI). A clutch of studies generated by Muriel Beabeau at the University of Minnesota and her colleagues (5-7) and Donnie Self at Texas A&M University and his colleagues (8-10) all observe measurable if modest correlations between ethics education and moral reasoning skills, and some possible implications for attitudes or behaviors. Michael Kalichman and colleagues at the University of California at San Diego (11, 12) have developed an independent instrument that shows similar correlations, although other studies (13) raise doubts about the full significance of such correlations.

No doubt partly as a result of the restrictions in, if not the inconclusiveness of, existing assessments, it has been argued that the goals of ethics education should not be attitudes or behaviors at all but simply skills and knowledge (14). Indeed, the most common classroom assessments of research ethics teaching emphasize solely the learning of ethical reasoning skills, with little attempt to gauge the potential for long-term changes in behavior. Arguments have even been made to the effect that much more effective than RCR teaching in the promotion of scientific integrity would be the establishment of clear behavioral guidelines followed by some form of monitoring such as data audits (15). When education fails, try social control.

Broader Interpretations of RCR Education

Quantitative assessment of teaching and learning about research integrity in the academic classroom is thus boxed in on four sides. Such constraint reflects the analytic and reductionist strategy of modern scientific methodology, which is based on the demand for and promise of metrical results; this is a strategy that must continue to be pursued. At the same time, there is no need to completely restrict approaches to such a flat plane. Indeed, especially given the wealth of issues associated with moral education, there are grounds for stepping beyond such constraints–that is, for expanding our horizons in the assessment of the teaching and learning of research integrity.

First, boundaries may be extendend slightly by recognizing the limits of particular instruments such as the DIT and SRM. One modest movement in this direction would be to consider the relevance of other instruments for assessing cognitive or intellectual development such as the Reflective Judgment (RJ) scale developed by Patricia King and Karen Kitchener (16) on the basis of the work of William G. Perry, Jr. (17). It may be noted, for instance, that

the RJ instrument has been pioneered at the Colorado School of Mines (18) and repeated at Pennsylvania State University (19) as a tool to assess the intellectual development of engineering students. Although not focused on moral development, RJ has potential implications for ethics learning that deserve exploration.

Second, precisely because RCR education raises research questions about long- as well as short-term effectiveness, the goals of the teaching and learning about research integrity should themselves become themes for research. This would constitute, as it were, an initial step off the flat place of quantitative research. Research into goals, as opposed to research on the effective implementation of goals, calls for more than quantitative or empirical study. It calls for historical and philosophical analysis and reflection. It may be noted, for instance, that current assessment strategies tend to carry forward, more or less uncritically, the applied ethics movement that arose during the 1980s.

At the very beginning of this revival Daniel Callahan (20) proposed five goals for the teaching of ethics in higher education: (a) stimulating the moral imagination, (b) recognizing ethical issues, (c) eliciting a sense of moral obligation, (d) developing analytic skills, and (e) tolerating and reducing disagreement and ambiguity. Viewed against the background of the analytic meta-ethics dominant at that time, these were all worthy and even modestly revolutionary goals. Historically, however, the focus has increasingly narrowed to simply developing analytic skills. The teaching and assessment of research ethics has largely accepted this narrow inheritance, as is reflected in the very terminological emphasis on "responsible conduct of research." Philosophically, there are even deeper historical issues to be raised if RCR education is examined in the light of such classic reflections on the moral life as those present in the works of Plato, Aristotle, and Augustine, not to mention the Upanishads, the Sutras, the Torah, or the Gospels.

Third, reflective reconsideration of the goals of teaching and learning about research integrity may stimulate recognition that as much if not more pertinent teaching and learning goes on outside the classroom as well as within it. This recognition may, in turn, promote a search for ways to assess meta-classroom learning. One meta-classroom context is the professional association. Yet lack of assessment is also common among scientific professional societies. Although most societies have codes of ethics that clearly bear on research integrity, Mark Frankel, director of the Scientific Freedom, Responsibility and Law Program at the American Association for the Advancement of Science (AAAS), has concluded that few scientific societies are able to tell whether these codes are working (21, 22).

Finally, extending such reflection even further, it reasonably may be argued that such internalist issues as data management, the treatment of human and animal subjects, publication protocols, conflicts of interest, and mentoring standards cannot in reality be separated from the focused externalist issues of science and technology policy. Indeed, international recognition of the immoral behavior of some members of the medical research establishment during World War II stimulated adoption of the Nuremburg Code for free and informed consent in human subjects research; political concern in the United States during the 1980s about the improper behavior of scientists using public funds has been one of the primary drivers to promote RCR education. Surely both of these historical points deserve to be taught along with the norms of data management and peer review.

Three (Intentionally Provocative) Suggestions

Without attempting to draw definitive conclusions from this four-fold unsystematic expansion of the RCR educational context, we would like to pose three summary pleas for the pursuit of new dimensions in assessing the teaching and learning of research integrity. In this way we seek to make common cause with others such as J. Andre (23) who have also called for not limiting professional ethics courses to moral reasoning analyses.

First, in light of the public policy roots of RCR education and the larger philosophical and religious traditions of ethics, is it appropriate to focus on reasoning or analytic skills in ways that slight attitudes and behavior? Would it not be possible to develop, for instance, an instrument for assessing cynicism and idealism among students, and indeed to attempt to counteract a too common passive cynicism? Social idealism is an honorable heritage of the scientific tradition, as exhibited by scientific leaders from Francis Bacon to Albert Einstein. In a talk to

scientists and engineers at the California Institute of Technology in 1931, for instance, Einstein argued that

> Concern for man himself and his fate must always form the chief interest of all technical endeavors . . . in order that the creations of our mind shall be a blessing and not a curse to mankind. Never forget this in the midst of your diagrams and equations (24).

Contemporary witnesses to this tradition of idealistic science can be found in the public interest activism of International Pugwash founding member and Nobel Peace Prize winner Joseph Rotblat (25) as well as SunMicrosystems co-founder Bill Joy (26). Introduction to such moral heros of what may be termed scientific social idealism should not be slighted to carve out time for parsing moral dilemmas in conflict of interest or authorship adjudication, as important as these may well be.

Second, does research ethics need to be conceptualized as distinct from engineering ethics, as it has been so far? Does the engineering/science separation not perpetuate stereotypes of the pure scientist versus the applied engineer–images at odds with reality in a world in which virtually all science is dependent on complex technological instrumentation? Moreover, is it not the case that scientists have something to learn from engineers regarding ethics? Long before scientists, engineers formulated ethics codes at the beginning of the 20th century; they also began taking them into the classroom well before scientists (26).

In the engineering education community today, considerable attention currently is being given to ABET Criteria 2000, the new set of accreditation guidelines developed by the Accreditation Board for Engineering and Technology (available at www.abet.org). Criterion 3, for instance, contains 11 attributes that graduates should possess, including "understanding of professional and ethical responsibility." Many engineering programs are developing methods to assess student progress in this area, including the use of such instruments as the DIT. There are also unexplored possibilities for assessing teaching and learning in engineering ethics by correlating results from the Fundamentals of Engineering (FE) and Professional Engineering exams required of all professional engineers.

Research integrity should not be separated from academic integrity in the research university setting. The practical RCR educational potential of student honor codes–some specific to schools of engineering–perhaps deserves as much attention as relations to engineering ethics codes.

Finally, does the assessment of teaching and learning itself not also deserve some assessment. An assessment of teaching and learning assessment requires both community engagement and critical analysis. The practice of any assessment should be guided by the principles developed by the Assessment Forum of the American Association for Higher Education (28), which include the following:

- Assessment is most effective when it reflects an understanding of learning as multidimensional, integrated, and revealed in performance over time.
- Assessment works best when the programs it seeks to improve have clear, explicitly stated purposes.
- Assessment works best when it is ongoing.

It is our contention that assessing of the teaching and learning of research integrity has only begun. This is true not only in the narrow senses associated with quantitative investigation of RCR, but also in the much broader senses of attempts to develop relations between RCR and idealistic science activism, engineering ethics and academic codes, and the reiterative assessment of assessment itself.

Bibliography

1. Davis, M. Ethics and the university. New York: Routledge; 1999.
2. Steneck, N. Assessing the integrity of publicly funded research. Background report. A Research Conference on Research Integrity, Bethesda, MD, Nov. 18-20, 2000.
3. Rest JR., Narváez D, Bebeau MJ, Thoma S. Postconventional moral thinking: A neo-Kohlbergian approach. Mahwah, NJ: Erlbaum; 1999.
4. Gibbs, John C., Karen S. Basinger, and Dick Fuller. Moral maturity: Measuring the development of sociomoral reflection. Hillsdale, NJ: Lawrence Erlbaum; 1992.
5. Bebeau MJ. Designing an outcome-based ethics curriculum for professional education: Strategies and evidence of effectiveness. Journal of Moral Education 1993; 22(3): 313-26.
6. Bebeau MJ, Thoma SJ The impact of a dental ethics curriculum on moral reasoning. J Dent Educ 1994; 58(9): 684-92.
7. Bebeau MJ, Rest JR, Narvaez D. Beyond the promise: A perspective on research in moral education. Educational Researcher 1999 May; 28(4): 18-26.

8. Self DJ, Wolinsky FD, Baldwin Jr DC. The effect of teaching medical ethics on medical students= moral reasoning. Acad Med 1989; 64: 755-9.
9. Self DJ, Schrader DE, Baldwin Jr DC, Root SK, Wolinsky FD, Shadduck JA. Study of the influence of veterinary medical education on the moral development of veterinary students. J Am Vet Med Assoc 1991; 198(5): 782-7.
10. Baldwin Jr DC, Daugherty SR, Self DJ. Changes in moral reasoning during medical school. Acad Med 1991; 66(9) Suppl: S1-S3.
11. Kalichman MW, Friedman PJ. A pilot study of biomedical trainees= perceptions concerning research ethics. Acad Med 1992; 67(11): 769-75.
12. Brown S, Kalichman MW. Effects of training in the responsible conduct of research: A survey of graduate students in experimental sciences. Science and Engineering Ethics 1998; 4(4): 487-98.
13. Eastwood S, Derish P, Leash E, Ordway S. Ethical issues in biomedical research: Perceptions and practices of postdoctoral research fellows responding to a survey. Science and Engineering Ethics 1996; 2(1): 89-114.
14. Elliott, D, Stern JE. Evaluating teaching and students=learning of academic research ethics. Science and Engineering Ethics 1996; 2(3): 345-66.
15. Shamoo AE, Dunigan CD. Ethics in research. Exp Bio Med 2000; 224(4): 205-10.
16. King PM, Kitchener KS. Developing Reflective Judgment. San Francisco: Jossey-Bass; 1994.
17. Perry Jr WG. Forms of Intellectual and Ethical Development in the College Years: A Scheme. New York: Holt Rinehard and Winston; 1970.
18. Pavelich, MJ, Moore WS. Measuring the effect of experiential education using the Perry model. Journal of Engineering Education October 1996 85(4): 287-92.
19. Palmer B, Marra RM, Wise JC, Litzinger TA. A longitudinal study of intellectual development of engineering students: What really counts in our curriculum? 30th ASEE/IEEE Frontiers in Education Conference, Proceedings Oct 18-21 2000: S3A-2-6.
20. Callahan D. Goals in the teaching of ethics. In: Callahan D, Bok S, editors. Ethics Teaching in Higher Education. New York: Plenum Press; 1980. p.61-80.
21. Brainard, J. 2000. Societies urged to fight research misconduct. Chronicle of Higher Education April 21, 2000; 46(33): A38.
22. DuMez E. The role and activities of scientific societies in promoting research integrity. Professional Ethics Report 2000; 13(3): 1 and 7-8.
23. Andre J. Beyond moral reasoning: A wider view of the professional ethics course. Teaching Philosophy 1991 Dec; 14(4): 359-73.
24. Einstein A. Einstein on Peace. Nathan O, Norden N, editors. New York: Simon and Schuster; 1960.
25. Rotblat, J. Taking responsibility. Science 2000 Aug 4; 289: 729.
26. Joy, B. Why the future doesn't need us. Wired 2000 April; 8.04: 238-62.
27. Mitcham, C. The achievement of "Technology and Ethics": A perspective from the United States. In: Technology and Ethics. Brussels: In press.
28. Assessment Principles. ASEE Prism 1997 Oct; 7(2): 10.

7. Responsible Conduct of Research Courses

The Responsible Conduct of Animal Research

Lida Anestidou, University of Texas-Graduate School of Biomedical Sciences at Houston, USA

Elizabeth Heitman, University of Texas-School of Public Health at Houston, USA

Key Words: Animal, Animal research, Curriculum, Responsible conduct of research, RCR

Students in graduate education in the basic sciences have a high probability of using live animals at some point in their research training. Although animal rights are a volatile issue for public debate, the use of animals in graduate science education raises little controversy among research trainees. Due to a National Institutes of Health (NIH) mandate, most graduate science programs today offer instruction in the responsible conduct of research that may include the ethics of experimentation with animal subjects[1]. Similarly, federal requirements for animal research review committees include provisions for the technical training of students and others conducting procedures with live animals[2].

As part of their responsibilities for overseeing the housing and care of research animals and the safe conduct of research, the veterinary staff of the University of Texas-Health Science Center at Houston offers formal training sessions in the safe and humane handling of laboratory animals and proper techniques for a variety of procedures. These sessions are offered regularly and are often filled well in advance.

The University's Institutional Animal Care and Use Committee (IACUC) and the veterinarians of the Center for Laboratory Animal Medicine and Care (CLAMC) are justly proud of their record of concern for animal welfare and the institution's humane research practices. Nonetheless, faculty involved in the required research ethics course at the University of Texas-Graduate School of Biomedical Sciences at Houston routinely hear comments from first- and second-year students who feel uncomfortable in their animal work, particularly in mastering routine procedures after the formal training has ended. Often these comments, made in small group discussions, are about the value of biomedical research with animals and questions about animal suffering. The same students typically express unwillingness to ask for help or further instruction for fear of criticism from their faculty and/or older peers. Nonetheless, many agree that more direct training in the handling and use of specific research animals would improve their skills, confidence, and attitude toward the work, as well as improve the quality of their research.

Research in medical education has demonstrated that trainees who ignore or discount their emotional responses to patients and the pain that medical procedures may cause are at risk of becoming emotionally stifled, cynical, and even punitive in response to the suffering of others. In contrast, by including formal attention to the emotional dimensions of patient care, medical educators have been shown to foster trainees' compassion and personal satisfaction in their work[3]. Moreover, by learning to identify and address their emotional responses directly, medical trainees have been

Corresponding author: Lida Anestidou, UT-Houston Mdical School, Deptartment of Integrative Biology, Pharmacology & Physiology, 6431 Fannin, Houston, TX 77030, 713-500-5082 (voice), 713-500-7444 (fax), anestido@girch1.med.uth.tmc.edu.

found to improve the accuracy of their diagnosis and treatment. Parallel risks and opportunities exist for researchers who use animals, and efforts to address the emotional dimension of animal use make a valuable addition to the institution's efforts to enhance the integrity of scientific research.

In response to the perceived need for more focused education and hands-on training for graduate students in the biomedical sciences, the authors organized a new intensive course entitled "*The Humane Use of Animals in Biomedical Research.*" The course offers a highly structured and multidisciplinary approach to responsible animal research. Its goal is to provide instruction in the ethics and regulatory aspects of animal research, approaches to the reduction of the numbers of animals used in specific protocols, including alternative research methods, and extensive practical training tailored to the individual animal model that each participant expects to use. Using a combination of didactic sessions, case discussions, and direct, hands-on laboratory instruction under the close supervision of institutional veterinarians, the course faculty seek to enhance students' theoretical knowledge base, technical skills, practical compassion, and professional confidence.

An aspect unique to this course is the inclusion of structured group discussion intended to help students address their personal experiences, concerns, values, and attitudes regarding their interaction with animals and the demands of animal research. Faculty facilitators help students recognize and prepare for the personal and ethical challenges of live animal experimentation using a modified version of the Balint method, which has been used in medical education to promote personal awareness and effective, compassionate patient care[4].

The course was offered to graduate students, post-doctoral fellows, research associates and technicians across the University for the first time in July 2000. The course schedule, including topics, instructors, and format appears in Table 1. The list of assigned readings for the course appears in the Appendix.

Evaluation (Students', Instructors', Course Coordinators')

As part of the wrap-up on the last day of class, students were encouraged to provide a comprehensive evaluation of the course, with particular attention to the aspects of reading assignments, class structure and timing, and the integration of theoretical material and practical skills. One week following the end of the course, the instructors and course-coordinators held a similar debriefing and evaluation session with a special focus on potential changes for subsequent course offerings. The following constructive suggestions were made by course attendees:

Positive points

1. The readings were comprehensive and challenging.
2. The practical aspects and methodologic training were invaluable even to students not working in laboratories.
3. Learning about regulations and IACUC activities from IACUC members was very enlightening about the practicalities of researchers' obligations and institutional review.
4. The information on alternative methods to animal research was important to new researchers considering a variety of techniques.
5. The presence, knowledge, and guidance of veterinarians were a tremendous intellectual and practical asset.
6. The variety of viewpoints presented by interdisciplinary faculty and guest lectures was useful in understanding the scope of animal research and its ethical gray areas.
7. Discussion of the personal demands of research was valuable for integrating interdisciplinary issues and helpful for students seeking to come to terms with the demands of their work.
8. The intensive class format enhanced rapport among students and faculty.

Drawbacks and obstacles

1. The time commitment in an intensive 2-week format was extremely hard for students to manage along with their regular daily schedules.
2. The summer offering made scheduling faculty assignments difficult because of their travel schedules and other special commitments.
3. The logistical complexity of organizing multiple faculty in both classroom and laboratory was very time consuming for the course organizers.
4. More practical discussion of alternative methodologies by practicing researchers was needed.
5. Students in science are often uncomfortable with ethical ambiguity and like

clear answers.

6. Faculty need to focus more on the links between ethical debate, science policy, and practical demands of research.

7. The costs of laboratory materials for a larger enrollment are likely to be considerable

8. Students' perception of the need for such a course is variable. Faculty need to identify and address the multiple goals of different students in different backgrounds throughout the class.

Conclusion

Evaluation by the student and faculty participants and a critique of the course by the course

DATE	CLASS	Topic	INSTRUCTOR
Monday 07/17	Lecture	• Historical uses of animals in biomedical research • Ethical and regulatory perspectives on animals in biomedical research	Heitman Anestidou
Tuesday 07/18	Lecture	• Scientific approaches to refining animal research (the three Rs) • Balint group discussion	Heitman Anestidou
Wednesday 07/19	Lecture	• IACUC: its function and responsibilities • How to fill out animal protocol forms	Smith Heitman Anestidou
Thursday 07/20	Lecture	• Alternatives to animal models	Heitman Anestidou Bjerckey
Friday 07/21	Lecture	• AAALAC and the Guide • Housing and standards of care for laboratory animals- Facility tour • Balint group discussion	Goodwin Blasdel Heitman Anestidou
Monday 07/24	Lecture Lab	• Mouse biology, care, and management	Head
Tuesday 07/25	Lecture Lab	• General anesthesia and pain control; rodent-specific protocols; • Anesthesia matters (video) • Rodent anesthesia practicum	Smith
Wednesday 07/26	Lecture Lab	• Monkey retirement facility speaker • Balint group discussion	Griffin Heitman Anestidou
Thursday 07/27	Lab	• Disposition of animals after research • Euthanasia	Blasdel Head
Friday 07/28	Lecture Discussion	• Wrap up course material • Evaluation	Heitman Anestidou

*Table **1. The Human Use of Animals in Biomedical Research-Course Outline and Schedule***

coordinators resulted in significant enthusiasm to repeat it. The course will be offered again in the summer 2001 term, using mostly the same didactic methods and material, but in a less intensive format. The course coordinators, CLAMC veterinarians, IACUC members, and the University's administration hope that in the next few years the course will be developed into both an integrated part of many students' education at the Graduate School and a continuing education course available to researchers and others from outside our institution.

Bibliography

1. U.S. Department of Health and Human Services. *Guide for the care and use of laboratory animals (revised).* NIH Publication #86-23. Washington, DC: DHHS, 1985.
2. National Institutes of Health, Alcohol, Drug Abuse, and Mental Health Administration. Reminder and update: requirement for programs on the responsible conduct of research in National Research Service Award institutional training programs. *NIH Guide* 1990; 19 (30): 1.
3. Gorlin R, Zucker, HD. Physicians' reactions to patients: a key to teaching humanistic medicine. New England Journal of *Medicine* 1983; 308: 1059-63.
4. Novack DH, Suchman AL, Clark, W Epstein RM, Najberg E, Kaplan C. Calibrating the physician. Personal awareness and effective patient care. *JAMA* 1997; 278: 502-09.

Appendix.

The Human Use of Animals in Biomedical Research-Course Readings

(by topic)

History of animals in biomedical research; ethical & regulatory perspectives on animals in biomedical research

- F. Barbara Orlans, "The Beginnings of Institutionalized Animal Experimentation" and "Current Attitudes and Ethical Arguments" in *In the Name of Science: Issues in Responsible Animal Experimentation,* New York: Oxford University Press, 1993: 3-34.
- Caplan, Arthur, "Beastly Conduct: Ethical Issues in Animal Experimentation", Science, 1983, 406: 159-169
- Brody, Baruch "The Use of Animals in Research" in *the Ethics of Biomedical Research: An International Perspective*, New York: Oxford University Press, 1998: 11-30.
- National Association for Biomedical Research, "The Strict Regulations that Govern Research" *Animal Research Facts,* http://www.fbresearch.org/research98.htm

Procurement of animals for research and education; Scientific approaches to refining animal research (the three Rs)

- F. Barbara Orlans, "The Source of Laboratory dogs and Cats: Pound versus Purpose-Bred Animals", in *In the Name of Science: Issues in Responsible Animal Experimentation,* New York: Oxford University Press, 1993, 209-220.
- "Shelter Intake and Euthanasia Trends", *Animal Policy Report* 2000, 14 (2): 2.
- Judith Reitman, "From the Leash to the Laboratory", *Atlantic Monthly* 2000, 286(1): 17-21.
- "Pet Theft: Urban Myth Makes Useful Propaganda", *FBR Facts* (Foundation for Biomedical Research), 2000, 7(2), 2 pages. http://www.fbresearch.org
- Joanna Weiss, "Squid's Fate: Science of Seafood", *Houston Chronicle (from Boston Globe),* June 27, 2000, 3D.
- W.M.S. Russell & R.L. Burch, "Introduction" in *The Principles of Humane Experimental Technique*. London: Methuen & Co., 1959, 3-13.
- Alan M. Goldberg, Joanne Zurlow, & Deborah Rudacille, "The Three Rs and Biomedical Research" (editorial), *Science*, 1996, 272: 1403.
- Ruth Ellen Bulger, "Use of Animals in Experimental Research: A Scientist's Perspective", A*natomical Record*, 1987, 219: 215-220.
- National Association for Biomedical Research, "Animals in Research 1998", *Animal Research Facts,* http://www.fbresearch.org/research98.htm
- Michael F.W. Festing, et al., "Reducing the Use of Laboratory Animals in Biomedical Research: Problems and Possible Solutions" — The Report and Recommendations of ECVAM Workshop 29, A*TLA* 1998, 26: 283-301.

The function and responsibilities of the IACUC

- Robert A. Whitney, Jr., "Animal Care and Use Committees: History and Current National Policies in the United States", *Laboratory Animal Science* 1987, Special Issue, 18-21.
- Henry J. Baker, "Essential Functions of Animal Care and Use Committees", *Laboratory Animal Science* 1987, Special Issue, 30-33.
- "Consensus Recommendations on Effective Institutional Animal Care and Use Committees" *Laboratory Animal Science* 1987, Special Issue, 11-13
- F. Barbara Orlans, "What Does the Public Have a Right to Know?", in *The Human Use of Animals Biomedical Research*. New York: Oxford University Press: 103-117.
- UT -Houston Health Science Center, Animal Welfare Committee, "Animal Protocol Review Form", revised 11/1/96.

Alternatives to animal models

- Michael Ballis, "Why is it Proving to be So Difficult to Replace Animal Tests?" *Lab Animal* 1998, 27 (5): 44-47.
- Richard N. Hill & William S. Stokes, "Validation and RegulatoryAcceptance of Alternatives", *Cambridge Quarterly of Healthcare Ethics* 1999, 8, 73-79.
- Jacques LeClaire & Odile De Silva, "Industry Experience with Alternative Methods", *Toxicology Letters* 1998, 102-103: 575-579.
- Seymour Levine & Arthur Saltzman, "An Alternative to Overnight Withholding of Food from Rats", *Contemporary Topics (American Assn. for Laboratory Animal Science)* 1998, 37: 59-60.
- Sharron Kirchain & Robert P. Marini, "A Tissue Harvesting Program as a Method for Implementing the 3Rs of Biomedical Research", *Lab Animal* 1998, 27 (8): 37-39.
- Adrian Smith, Richard Fosse, David Dewhurst, & Karina Smith, "Educational Simulation Models in the Biomedical Sciences", *ILAR Journal* 1997, 38 (2), 82-88.

Standards of care and housing for laboratory animals

- National Research Council, Institute of Laboratory Animal Resources, *Guide for the Care and Use of Laboratory Animals*, Washington, DC: National Academy Press, 1996.

Anesthesia and pain control

- Lawrence R. Soma, "Assessment of Animal Pain in Experimental Animals", *Laboratory Animal Science* 1987, Special Issue, 71-74.
- American College of Veterinary Anesthesiologists, "Position Paper on the Treatment of Pain in Animals", *JAVMA* 1998, 213(5), 628-630.
- American Association for Laboratory Animal Science, "Position Statement: Recognition and Alleviation of Pain and Distress in Laboratory Animals", http://aalas.org
- American Association for Laboratory Animal Science, "Policy #12 — Consideration of Alternatives to Painful/ Distressful Procedures — June 21, 2000. http://aalas.org

Disposition of animals after research

- American Veterinary Medical Association Panel on Euthanasia, "1993 Report of the AVMA Panel on Euthanasia", *JAVMA* 1993, 202: 229-249.
- UT -Houston Health Science Center, Center for Laboratory Animal Medicine and Care", "Animal Adoption Release form", revised 11/1/96.

Wrap up and evaluation

- Diane J. Gaertner, Lele K. Riley, & Dale G. Martin, "Reflections on Future Needs in Research with Animals", *ILAR Journal* 1998, 39: 306-310.

An Effective Short Course on Research Integrity

Bernard S. Gerstman, Department of Physics, Florida International University, USA

Key words: Course, Ethics, Integrity, Interactive, Research, Responsible conduct of research, RCR, Training

Ethical conduct in research has always been considered of utmost importance within the research community. Historically, it was assumed that scientific ethics did not require special training. Instead, the ethical manner in which to carry out research was presumed to be learned by new scientists automatically and unconsciously, as if by osmosis, as the technical aspects of the research were carefully taught by their superiors. This was of course, never true. Mendel and Millikan may have fudged their data, along with numerous others of less renown.

More recently, consideration has been given to developing methods for training scientists in research ethics, rather than relying on osmosis (1). Part of the impetus for this change is that the problems associated with unethical procedures in research have become especially visible to the public when they occur in research in the health sciences (2). This paper reports on a course of short duration that is designed to train students efficiently and effectively in the ethical conduct of research.

Design

The course is designed for graduate students and undergraduates who have shown an interest in a career in science. There is no obvious reason why the course design would not be applicable to students outside the sciences. At this time, all science majors at the home institution do not take the course. The science undergraduates who are required to take the course are affiliated with special programs such as the Research Experience for Undergraduates funded by the NSF as well as NIH funded programs.

The course is designed to meet for one hour each week and to contain a maximum of 15 students. If necessary, such as in summer sessions, the course can be compressed into a two-week period, but some of its effectiveness is lost. This will be discussed later in this section when the reason for this loss in effectiveness will be clear.

The initial course meetings are organized like a traditional class with the faculty member explaining various aspects of research integrity and unethical behavior. This is best introduced by a short (one hour) summary of the general principles of ethics in western society, which can then be used as the basis for the principles of research integrity and ethics. It is important that this explanation of ethics in general be presented as a summary. If it is presented in another form, such as an “Introduction to Western Ethics” or any other form that does not convey immediate *de facto* credibility, the course runs the danger of degenerating into a philosophy discussion on ethics in general. Valuable time will then be taken from the specific goal of training the students in scientific integrity and the course is likely to be neither short nor effective.

In addition to explaining the principles of research integrity, it also is important to be explicit

Corresponding author: Bernard S. Gerstman, Department of Physics, Florida International University, University Park, Miami, FL 33199, 305-348-3115 (voice), 305-348-6700 (fax), gerstman@biophys.fiu.edu.

about the importance of adhering to these principles. Thus, the first few lectures of the course should cover the following topics:

1) the general principles of Research Integrity (1);
2) how scientific progress is enhanced by adherence to integrity by all researchers;
3) how scientific progress is slowed by unethical behavior, or even the perception thereof; and
4) the direct impact of ethical misconduct in research:
   i) wasted money by universities and funding agencies,
   ii) wasted time by researchers who trust the results of others, and
   iii) injury or death to patients (biomedical research).

The middle part of the course shifts to a preceptorial structure with faculty led discussions of selected reading material on recent cases concerning violations of research integrity. These case studies summarize the accusations, how they were investigated, the decisions that were reached, and penalties imposed, if any. These case studies can be found in the Annual Report from the Office of Research Integrity of the Department of Health and Human Services (3).

These case studies supply concrete examples of the topics discussed in the first part of the course. The vast majority of cases involve data fabrication and falsification. This also presents the opportunity to discuss types of research misconduct that are common but insidious: sloppy data taking and self-deception (4). In these instances, the researcher is not consciously violating the principles of ethical behavior. Unfortunately, because the misconduct is unconscious, there is no chance for self-correction (5). The case studies are useful in training the students against sloppy data taking and self-deception, which can appear to be, or easily become, data fabrication or falsification.

The case studies also present concrete examples of a new topic — the penalties suffered by researchers who are found to violate the principles of research integrity. The usual penalties(3) are disbarring from receiving federal funding for 3 to 5 years, monitoring of a researcher by the home institution, mandatory retraction or correction of publications, and occasionally dismissal. Students initially consider these penalties too light and suggest criminal prosecution. The faculty member at this point can explain the severe ramifications of these penalties for the researcher's career.

The third, and last, part of the course is the most important for successfully conveying the principles of research integrity and the necessity of adhering to these principles. It requires each student to make a half-hour presentation to the class about a case of suspected unethical behavior in research that they have investigated through a literature search. The students are expected to use what they have learned in the earlier parts of the course in discussing the following points:

1) an explanation of what actions constituted unethical behavior, entailing enough of an explanation of the scientific research so that other students can understand why the behavior was unethical;
2) how the unethical behavior was uncovered;
3) what the motivation might have been for the unethical behavior;
4) what, if any, penalties (real or intangible) were suffered by the perpetrators; and
5) what penalties the student thinks would have been appropriate.

Information for these presentations can be obtained from books(6,7,8) on the subject, science magazines such as *Scientific American*, and with especially well-known and recent cases, newspapers and general readership magazines. Students are informed early in the course about the presentation and are told to choose a case as soon as possible. It is hoped that by giving the students several weeks to prepare for their presentation, they will use the time to follow a meandering path in their literature search and learn about several different cases. If two students choose the same case, the second student to notify the faculty member is instructed to pick another case.

Results

The first two parts of the course give the students a customary introduction to the issues of research integrity. The third part of the course is crucially important for consolidating these issues. The students are enthusiastic about making their presentation and peer pressure motivates them to do a thorough job. The presentation forces the

students to “step into the mind” of a scientist who is behaving unethically. This obliges them to confront the temptations to behave unethically and solidifies the need for self-vigilance.

Conclusion

A short course can be effective in conveying the necessity of integrity in research and in training the students on how to perform research in an ethical manner. For the course to be effective, the students must be required to take an active role. A class presentation by each student is of crucial importance and the most important element of the course.

Bibliography

1. Marshall Thomsen, Editor, Ethical Issues in Physics: Workshop Proceedings, July 17-18, 1993, Eastern Michigan University, Ypsilanti, Michigan.
2. United States Congress House Committee on Government Operations: nineteenth report. Are scientific misconduct and conflicts of interest hazardous to our health? Washington: U.S. G.P.O., House report / 101st Congress, 2d session ; 101-688, September 10, 1990.
3. Office of Research Integrity (US), Annual Report 1998, Office of the Secretary, Office of Public Health and Science, Department of Health and Human Services.
4. Lederberg J, Commentary: Sloppy Research Extracts a Greater Toll Than Misconduct, The Scientist, February 20, 1995, page 13.
5. Thomsen M and Resnick D. The Effectiveness of the erratum in avoiding error propagation in physics, Science and Engineering Ethics, 1995, 1:3.
6. William Broad and Nicholas Wade. Betrayers of the Truth, Simon and Schuster, New York, 1983.
7. Alexander Kohn, False Prophets: Fraud and Error on Science and Medicine, Basil Blackwell, New York, 1986.
8. Robert Park, Voodoo Science: The Road from Foolishness to Fraud, Oxford University Press, New York, 2000.

Resources for Instruction in Responsible Conduct of Research

Michael W. Kalichman, Department of Pathology, University of California San Diego, USA

Francis L. Macrina, Philips Institute of Oral and Craniofacial Molecular Biology,Virginia Commonwealth University, USA

Jeffrey P. Kahn, Center for Bioethics. University of Minnesota, USA

Key Words: Instruction, Internet, Research ethics, Responsible conduct of research, RCR, World Wide Web

In recent years it has become clear that, despite its importance, training in ethics, standards, and responsible conduct is too frequently minimal or absent in academic science. This deficit is being corrected in part by the requirement that fellows funded by National Institutes of Health (NIH) training grants receive such instruction. This requirement has been important to the development of a variety of outstanding texts now available (1-8) and a number of very effective, thoughtful programs developed across the country. However, no network provides ready communication about the goals, resources, tools, or methods for such programs. As a result, the design and implementation of a new program in responsible conduct of research (RCR) training can be frustrating if not intimidating.

It can be difficult to pull together material for a new RCR program. Unfortunately, such effort is frequently duplicated even within the same institution and the resulting RCR instruction is uneven in quality, topics covered, and audience reached. In addition, it appears that the most likely audience for these programs has been limited to only those NIH trainees required to take part. This is contrary to the goal that such training is best met by a program that reaches the broad spectrum of the academic community including staff, undergraduates, medical students, pre- and post-doctoral fellows, and both junior and senior faculty. However, with the rapid changes in access to the Internet, the technology is now available to make formats, examples, contacts, and resources immediately available to any institution interested in providing effective RCR instruction.

The Internet is now being used for a variety of purposes relevant to RCR instruction (9-17). In just the last couple of years, these resources have evolved rapidly in both form and content. Many institutions have created web sites that provide considerable content as well as lists of links to other sites (9-10), typically in the area of ethics. In addition, many universities now have course materials posted on the web (11-13) and in some cases Internet-based courses, designed to be run without traditional classroom meetings (14,15). Finally, web-based information is available on programs such as the “Survival Skills and Ethics” (16) and “Teaching Research Ethics” (17) workshops for teaching about the teaching of responsible conduct of research. All of these resources provide important contributions, but diverse audiences, differences between disciplines, and the frequency of significant new developments, all minimize the value of any one approach to RCR instruction. The proposed alternative is a continually evolving web site.

Corresponding author: Michael Kalichman, Ph.D., Director, Research Ethics Program, 0003, University of California, San Diego, La Jolla, CA 92093-0003, 858-822-2027 (voice), 858-534-4772 (fax), kalichman@ucsd.edu.

A web site dedicated to resources on instruction in the responsible conduct of research could provide national access to the most effective programs, materials, and methods for such training. The long-term goal would be to improve the quality and extent of RCR instruction. Such a site would not only make it possible for virtually any institution to develop an RCR program, but would also increase general awareness about what is being done, and what can be done, to enhance instruction in RCR. It is intended that this site would complement, not replace, other tools for RCR programs (1-17). Given the ongoing NIH requirement for training grants to include instruction in RCR and the proposed extension of this requirement to all research staff working on PHS-supported projects, many institutions need help to either extend limited existing programs or to develop new programs. However, even in the absence of federal requirements, it should be enough to know that access to proven materials and methods for RCR instruction can only help to foster responsibility in the conduct of research.

Methods

The core of the web site was first assembled from materials already available for courses taught at the University of California in San Diego, Francis Macrina's course at Virginia Commonwealth University and his book on "Scientific Integrity," and course materials under development at the University of Minnesota.

The site was initially designed to cover nine topic areas: (1) Getting started; (2) Defining the goals of an RCR program; (3) Elements of an RCR program; (4) Guidelines, requirements, and procedures; (5) Resources; (6) Case studies; (7) RCR programs; (8) Contacts; and (9) Evaluating an RCR program. The plan was that these primary divisions would be subdivided into topics generally considered to be relevant to responsible conduct of research (e.g., conflict of interest, use of animals in research, and criteria for authorship). Using this framework for the content available in the authors' institutions, the initial goals were to design and implement a framework for the web site, insert materials from the authors' institutions, and annotate the resources.

After completion of the first steps of the project, the web site was to be improved through an iterative process, including three phases of external reviews, plus soliciting of suggestions for additional materials. For this review phase, the primary goals were to solicit new materials from other institutions, modify the framework of the site as needed to accommodate the new resources and reviewer suggestions, annotate the resources, and publicize the site.

For evaluation of the web site, reviewers were asked to rank various aspects of the site's form and content in a brief online form. Numerical rankings were to be scored using a scale of 1 to 5 (1=very low, 2=low, 3=average, 4=high, 5=very high). Additional questions asked for specific suggestions to improve the web site, including recommendations of material to be added.

Results

The first phase of this project was to develop a web site framework for presenting resources on instruction in the responsible conduct of research. Beginning in September of 1999, work on the web site began at the University of California, San Diego with ongoing assistance from collaborators at Virginia Commonwealth University and the University of Minnesota. During the initial months, the web site evolved through several different formats until a version was considered ready for external review. In July of 2000, the first phase of external review was begun. The three planned phases of review were completed by November 1, 2000.

The first external review was based on a limited release of the web site to four reviewers (two from government agencies and two from non-governmental organizations). In a series of questions about web site form and content, scores averaged between 3.25 and 4.75 with medians between 3 and 5. The lowest scores were generally assigned to the appearance and navigability of the web site. Several valuable suggestions were made for future improvements, but one–ease of navigation–was sufficiently important to address before the next phase of review. Based on this concern, the structure of the web site was considerably modified to provide the user with a linear arrangement of topics. This and other changes were completed by the beginning of August 2000.

For a second external review, 13 people were asked to participate. One of the 13 did not respond to the invitation, three declined because of conflicting commitments, but two recommended other choices for reviewers. Ultimately, of nine who agreed to review the site,

QUESTIONS	AVERAGE	MEDIAN
1. CONTENT		
A. How would you rate the choices of topics covered?	5.0	5.0
B. How would you rate the quality of information provided?	4.6	5.0
2. NAVIGATION		
How would you rate the ease for navigating within the Web site?	4.2	5.0
3. APPEARANCE		
How would you rate the appearance of the Web site?	3.8	4.0
4. OVERALL		
A. How would you rate the likelihood you would recommend this resource for someone developing a new training program?	4.6	5.0
B. How would you rate the likelihood you would recommend this resource for someone improving an existing program?	4.8	5.0

Table 1. Second phase of external review (6 reviewers). Using a scale of 1-5 (1 = very low, 5 = very high), the reviewers answered the following six questions.

three failed to meet the deadline. The six reviewers who responded were from two public universities, one private university, two government agencies, and one non-governmental organization.

A summary of the average and median of the second phase of reviewer evaluations is provided in Table 1. The reviewers were extremely positive about the content of the web site (averages of 4.6 to 5.0). Compared to the previous round of review, these reviewers were also more positive about navigation (4.2 vs. 3.25-3.75). Although considered acceptable, no reviewer scored appearance of the web site as a 5. In addition, the reviewers offered many practical suggestions for improvements in content, navigation, and appearance.

A third external review was begun in September of 2000. A total of 48 people were asked to review the web site by early October; 31 responded that they had the time and would be willing to do so. Of those, reviews were completed by 23 reviewers (16 public institutions, 4 private institutions, 2 government agencies, 1 Canadian government agency).

A summary of the average and median of the third phase of reviewer evaluations is provided in Table 2. Evaluation rankings were generally in the range of 4 to 5. Lowest scores were for the appearance of the web site (average=3.8) and highest scores were for the likelihood that the reviewers would recommend this web site as a resource for someone developing a new training program (average=4.8). The reviewers were again generally positive, but several made

QUESTIONS	AVERAGE	MEDIAN
1. CONTENT		
A. How would you rate the choices of topics covered?	4.4	4.0
B. How would you rate the quality of information provided?	4.1	4.0
2. NAVIGATION		
How would you rate the ease for navigating within the Web site?	4.0	4.0
3. APPEARANCE		
How would you rate the appearance of the Web site?	3.8	4.0
4. OVERALL		
A. How would you rate the likelihood you would recommend this resource for someone developing a new training program?	4.8	5.0
B. How would you rate the likelihood you would recommend this resource for someone improving an existing program?	4.5	5.0

Table 2. Third phase of external review (23 reviewers). Using a scale of 1-5 (1 = very low, 5 = very high), the reviewers answered the following six questions.

excellent suggestions for changes in structure and content to the site. Few comments were repeated across reviewers. Major areas of criticism included:

1. Content: One reviewer was looking for a pre-packaged RCR course along the lines of the web-based tutorials for training researchers working with human subjects. As this reviewer observed, this web site does not provide such a course.
2. Format: The most frequently voiced concern was that the background (a grid similar to a lab notebook page) was distracting.
3. Audience: It wasn't clear to some reviewers who the audience (i.e., instructors of RCR courses) was for this web site.
4. Structure: Several reviewers failed to find key elements of the web site (e.g., the examples of courses) and some pointed out confusion about the structure of some of the sections (esp. resources and cases). Related to this problem, a couple of the links did not work, or did not work as expected.

Several of the reviewers were quite supportive of the web site, for example:

> "The choice of topics to be covered in teaching research ethics is excellent. I particularly think it is useful that 'minimal instruction' is defined for each of the topics and that more advanced versions of the units are also suggested. This will be quite helpful to faculty who are just beginning to teach RCR, and who want to know what is the minimum level of instruction they need to meet."

> "I think the site looks great. It is very well organized. It will be especially useful for newcomers."

> "Best collection of materials related to RCR I have found. The logical progression of steps should make it easy to develop or improve courses without becoming overwhelmed by the task at hand. Linked pages were relevant and provide materials for inspiration and contrast."

> "This is a very strong site and I learned a lot just skimming. The links for case studies and analysis of instructional delivery options were quite good."

> "This is a great program. I think its strongest feature is the way it brings together a wealth of material in a useful and usable form."

Based on the reviewer comments, further significant changes were made to the structure of the web site. As of its release, the structure of the web site was designed around five topic areas: Goals (Goals for RCR instruction), Content (Suggested RCR topics: Descriptions and reading lists), Format (Formats for RCR instruction: Descriptions and examples), Tools (Tools for RCR instructors: Texts, cases, and contacts), and Evaluation (Evaluation of RCR instruction: Overview and examples). After checking that the structure of the web site was consistent and that all links were active and accurate, the web site was released for public use on November 1, 2000.

Discussion

As proposed, a new web site was developed to facilitate access to resources for instruction in the responsible conduct of research. With the support of constructive comments from external reviewers, an initial version of the web site was made available to the research community beginning on November 1, 2000. Based on reviewer comments, this web site will be of value both to those first developing programs of RCR instruction and also to those seeking to improve on existing programs of instruction.

To achieve the long-term goals for this web site, it will be necessary for the site to evolve both in terms of content and format. For this purpose, the authors intend to solicit the latest information about content and format of existing RCR programs nationwide. Further, it will be important to include mechanisms for ongoing evaluation of the merits of the resources listed on the web site and the web site itself. During this next phase, the primary goals will be to survey existing programs in RCR, solicit new materials from these institutions, continue modifying the framework of the site as needed to accommodate the new resources, and implement mechanisms for evaluating the effectiveness of the web site and the resources listed.

Acknowledgments

The Office of Research Integrity (ORI) and particularly Alicia Dustira, Lawrence Rhoades, and Chris Pascal, all of ORI, are thanked for their encouragement and support for the initial

development of this web site. The University of California, San Diego is thanked for its support of the Project Director (MWK), helping to make this project possible. In addition, the generous contributions of time and suggestions of the 33 external reviewers are gratefully acknowledged. This work was supported by a contract from the ORI to UCSD (contract #99T07035001D)

Bibliography

1. Bulger RE, Heitman E, Reiser SJ. The Ethical Dimensions of the Biological Sciences. New York: Cambridge University Press; 1993.
2. Elliott D, Stern JE. Research Ethics: A Reader. Hanover, New Hampshire: University Press of New England; 1997.
3. Grinnell F. The Scientific Attitude. Second edition. New York: Guilford Press; 1992.
4. Korenman SG, Shipp AC. Teaching the Responsible Conduct of Research through a Case Study Approach: A Handbook for Instructors. Washington, D.C.: Association of American Medical Colleges; 1994.
5. Macrina FL. Scientific Integrity: An Introductory Text with Cases. 2nd edition, Washington, D.C.: American Society for Microbiology Press; 2000.
6. National Academy of Sciences, National Academy of Engineering, Institute of Medicine. On Being a Scientist: Responsible Conduct in Research. Washington, D.C.: National Academy Press; 1995 [also available online at: http://www.nap.edu/readingroom/books/obas]
7. Penslar RL. Research Ethics: Cases & Materials. Bloomington, Indiana: Indiana University Press; 1995.
8. Stern JE, Elliott D. The Ethics of Scientific Research: A Guidebook for Course Development. Hanover, New Hampshire: University Press of New England; 1997.
9. Center for Bioethics, University of Pennsylvania: http://bioethics.net
10. Online Ethics Center for Engineering and Science: http://onlineethics.org
11. Research Ethics Program, University of California, San Diego: http://ethics.ucsd.edu
12. Scientific Integrity, a one-semester course, Virginia Commonwealth University: http://www.vcu.edu/courses/rcr
13. Teaching Ethics for Research, Scholarship, & Practice, University of Minnesota: http://www.research.umn.edu/ethics
14. Research Ethics, an Internet-based course, University of Nebraska Medical Center: http://www.unmc.edu/ethics
15. Scientific Integrity, an Internet-based course, University of California, San Diego: http://ethics.ucsd.edu/courses/integrity
16. Survival Skills and Ethics, University of Pittsburgh: http://www.edc.gsph.pitt.edu/survival
17. Teaching Research Ethics, Indiana University: http://www.indiana.edu/~poynter/tre.html

An Interactive Web Site for Ethics Training: http://storiesandquestions.com

Rudolph J. Marcus, Stories and Questions, Occidental, CA, USA

***Keywords**: Case studies, Computer-based instruction, Ethics training instrument, Moral advisor, Self-directed learning, Work environment*

This paper reports on the construction and use of a web site for ethics training. The paper is divided into three parts. Various uses of the web site as a delivery vehicle for ethics training are outlined in the first part. Practical advantages of an ethics training instrument usable by individuals at their own pace, place, and time are discussed in the second part. The web site itself and its operation are described in the third part located after the references. The paper ends with suggestions for further work in adding more seminars to the web site, further measuring the web site's effectiveness, and developing guidelines for facilitators.

Ethics Training With The Web Site

The computer-based ethics training instrument is conceived as a delivery technique. What might such a training instrument contain?

After the apprentice model lost its effectiveness for ethics training it was replaced by a recital of the appropriate ethics codes (Table I). Discussion of case studies is used for ethics training that is more specialized and pointed toward particular disciplines and tasks. Both training in ethics codes and in case studies can be delivered by this sequenced text-and-question technique. The present web site adds a third category to the ethics training instrument, awareness training.

Types
Codes
Case studies
Awareness

Table I: Types of ethics training

Experience with the web site has shown different results from bulk presentation and from sequenced presentation. Bulk presentation, where the whole story and all of the questions are presented at one time, usually draws either no response or a "head trip" response. The sequenced presentation of a part of the story at a time or an exercise with the story, accompanied by a single question, appears to encourage the thoughtfulness and inner work that lead to real attitude change. It is that experience that leads to the statement of the previous paragraph that the necessary ethics training of the ethics codes themselves, and of applicable case studies, can be delivered by this sequenced text-and-question

Training	**Message**
Codes	What are the limits? What can I get away with?
Case studies	What are the limits? What can I get away with?
Awareness	What is going on here? What can I do about it? What might be the right thing to do?

Table II Messages from various types of ethics training

Corresponding author: Rudolph Marcus, 17010 Burl Lane, Occidental, CA 95465, 707-876-1959 (voice), catnrudy@monitor.net

technique.

The implication that trainees get from recital of appropriate ethics codes or from discussion of case studies is "What are the limits?" The message conveyed by such training is "What can I get away with?" (Table II).

By contrast, awareness deals with questions such as "What is going on here?" and "What can I do about it?" The message conveyed by awareness training is "What might be the right thing to do?"

The reader may sense that I find the "What might be the right thing to do?" more significant and necessary in ethics training because it goes beyond the given guidelines. However, knowledge of applicable ethics codes and of their application is case studies is an essential and equally necessary part of ethics training. I have discussed the multiplicity of the overlapping ethics codes that researchers are subject to and some of the conflicts between those codes in an encyclopedia chapter (1).

Results of ethics training by recital of appropriate ethics codes usually are measured by attendance at the training, or by recall of the contents (Table III).

Training	Results reported as
Codes	Attendance or recall
Case studies	Discussion of case
Awareness	Attitude changes

Table III: Results from various types of ethics training reported as

Results of ethics training with case studies are measured by qualities of participants' discussion of the cases. Results of awareness training can be assessed by noting attitude changes. The same question is asked near the beginning and near the end of each seminar. In experience so far with the web site, and in the talks and seminars on which the web site is modeled, the response after the second time the question is asked is much better elaborated, is more grounded in the story, and shows a better grasp of how the story applies to the participant's own life, problems, and actions. Changes become apparent in the perceived self image of the seminar participant. The changes are reinforced by the facilitator calling attention to them or asking whether the participant has noticed the changes.

Practical Advantages Of Web Site Ethics Training

Self-directed learning: No travel or time off from work

The web site is an alternative learning mode for ethics training which permits study and testing at each individual's time, place, and pace. It reduces or eliminates trying to get everyone together at a given time and place, with resultant schedule changes and resistance to "a waste of time."

"Valuing of self-directed learning" (2) applies to academic institutions and to needs for ethics training there as well as in industrial and governmental laboratories. Gunzburger writes. "The survey results indicate that most schools have not established systems for the valuing of curriculum time that include balancing values for lecture time and values for self-directed learning."

One surprise of experience with the web site was the amount of motivation which is generated by the sequenced self-study. No coaxing or arguments by the facilitator were needed to get clients to complete a seminar. If there was no response to a session within two weeks, re-presentation of the segment was sufficient to elicit a response, often with an explanation of what the holdup had been.

Engendering an ethical environment in self and in work place

An unspoken assumption seems to be that infractions of an ethics code are deliberate; if not deliberate, then the incident was an "accident" (3). Based on that assumption, an infraction can be stopped and the perpetrator exposed and/or punished.

Inherent conflicts and inconsistencies between the different kinds of ethics codes, as well as in individual codes themselves, are discussed in the encyclopedia article previously referred to (1). There are few, if any, places where those caught in a conflict between different kinds of codes can get help and advice. The encyclopedia article concludes with a training program that I have designed for situations like that, and this training program has now been put on the web site described in this paper. In brief, the seminar participant who uses the web site ethics training learns where to look for the advice needed to cope with overlapping

and conflicting ethics codes

A different way of stating the problem is found in a letter to a college student who asked for advice on ethical behavior In the letter, C. G. Jung describes codes of moral values as "general and not specific," and therefore "they don't exactly apply to individual situations." Jung's advice to the student is "try to live as consciously, as conscientiously, and as completely as *possible* (italics are Jung's) and learn who you are and who or what it is that ultimately decides" (4).

The problem also has been stated in a law review article about ethics in government, which applies equally well to ethics in science (5): "Our current obsession with investigating and prosecuting individual wrongdoing may actually prove counterproductive in our efforts to promote ethical [science], promoting instead public cynicism about [science]. To counteract this emphasis on individual wrongdoing, we need to pay more attention to ensuring that [science] institutions are designed to engender an ethical environment."

All three ways of stating the problem contradict the beginning assumptin—the old vocabulary—that infractions of ethics codes are deliberate or an "accident." All three statements indicate that infractions of ethics codes are NOT always deliberate and that the perpetrators may not even be aware of their inadvertent and often avoidable errors affecting research integrity.

Jung's advice to "learn who you are" to behave ethically is exactly the aim of the training program that I described in the encyclopedia article. The training is to live with the opposites in resolving conflicts of various ethics codes within which researchers have to work (1, 6). It is that kind of training program that I have now brought out of the lecture and workshop stage and put into the web site for one-on-one work with an experienced ethics consultant. It is a self-contained course that meets a previously unrecognized need.

Present Status And Future Work

At this time the web site contains four such seminars. One of the seminars deals with collegiality and civility in the work place or, to see it from the other side, conflict resolution in highly polarized situations (7). Another deals with how to find a moral advisor in a hierarchically structured work environment (8). Both describe work environments that often lead to alleged scientific misconduct and how to deal with them creatively.

Material is at hand for expansion of the web site to about 12 seminars during the coming year. The immediate next additions to the web site will be four seminars dealing with the origins of science (9). Together, they show four successive stages of scientists working "consciously, conscientiously and as fully as possible." I have used that material about twenty times in national tour lectures for the American Chemical Society under the title of "Nature and Gods, Science and Scientists."

Further research aspects of this work consist of:

1. Adding more seminars to the web site.
2. Assessing its effectiveness. Effectiveness can be gauged by looking at changed opinions, feelings, or assessments of problem situations by seminar participants as the seminar progresses. Records of such changes are already being kept while maintaining seminar participants' analytical confidentiality, which is a hallmark of the seminars in workshop and web site modes.
3. Developing guidelines for facilitators. As more people use this method of self-study for ethics training, they too may want to become facilitators and learn more by helping others to start self-study in scientific ethics.

Bibliography

1. Marcus RJ. Government funding of research. In: Encyclopedia of Applied Ethics. San Diego: Academic Press; 1998. Vol. 2, p. 471-80.
2. Gunzburger LK. U. S. Medical Schools' Valuing of Curriculum Time: Self-Directed Learning versus Lectures. Academic Medicine 1993; 68: 700-702.
3. Marcus RJ. Ethical considerations. The New Yorker 1996 March 11; p. 14.
4. Jung CG. Letter to William Kinney (1956 May 26). In: Letters. Princeton NJ: Princeton University Press; 1973. Vol. 2, p. 300-1. .
5. Clark K. Toward More Ethical Government: An Inspector General for the White House. *Mercer Law Review* 1998; 49: 553-64.
6. Marcus RJ. Ethics in research and government funding. Chemtec*h* 1997 Nov, p. 8-13.
7. Marcus RJ. Ethics workshops when collegiality fails. In: Abstracts of Annual Meeting, Association for Practical and Professional Ethics. 1998; Paper V H.

8. Marcus RJ. Sequenced self-study for constellating a moral advisor. In: Abstracts ofJoint Services Conference on Professional Ethics (JSCOPE 2000).
9. Marcus RJ. Mythological stories as case studies for teaching scientific ethics. In: Abstracts of 1996 Annual Meeting, Association for Practical and Professional Ethics. 1996; Paper II H,

The Web Site

Introduction

Welcome to this web site, the home of *"Stories and Questions,"* a personal journey in self enrichment. Here is a method of exploring who one is by reading short stories and responding in the moment to simple questions. These stories allow one to stop and feel, and the question permits feelings about one's life and its direction. *Stories and Questions* is a series of individual seminars facilitated by Rudy Marcus. Rudy has done stories and questions for 16 years (and if one counts his research experience in the sciences, for 50 years) and has experienced for himself their ability to effect personal change.

On this web site, you can start your own journey of exploration. Please read the Introduction and follow its suggestion for "How Do I Start?" Feel free to contact me at: rudy@storiesandquestions.com

This is an introduction to seminars designed for self-study. Each session or envelope contains a story or direction for an exercise with the material of the story. You, the participant in this seminar, encounter the story or do the exercise, and then respond to one or more question(s) on this web site. The response can be in writing or any other form of expression, and can be telephoned or sent by e-mail or post to: response@storiesandquestions.com. Rudy will then send you the next session (e-mail) or envelope (paper) of the self-study seminar. The method is adapted for self-study from group workshops using different stories and questions.

Method

If you were using this material at a group workshop or seminar, you would be sitting in a circle. Each person in the circle would hear from a facilitator what is on the web site as Session I, or what is in envelope I in the paper version — a short story, and then a question to which each member of the circle responds. It is not a discussion group, and there is not a consensus to be reached. Rather, each response is respected as that person's truth at that particular time and place. In such a workshop, there would be a long break after the discussion of the material in Session I. That break might even take the form of lunch, a nap, a walk in the woods, and/or a swim. More thoughts about the story, and additional responses to the questions occur, and those might be written in a journal or one's workshop notes.

Stories

In a group seminar using this material, the facilitator would have warned participants NOT to identify with any of the characters in the story. That is important and it applies as well to the self-study.

The seminar participant encounters the story as if the participant were seeing it on a stage. The participant is not on the stage with the story characters. The participant is in the audience watching the actions of all the characters, being privy to the knowledge, habits, and actions of all the characters at that point in the story.

In the language of psychology, the participant brings one's ego to the story, one's own awareness, rather than identifying with, taking the part of, one or the other character in the story The more cross-cultural the story is—for example, all cultures are likely to have creation stories, and stories about the origin of science—the more universally valid or typical do those characters seem, and the easier it is for the hearer of the story to say, "Hey, that character IS me, and that is MY story." Try NOT to do that.

The comparison of story with stage is quite apt because as action on stage involves feelings and emotions of onlookers, so encounter with story can activate an individual participant's inner knowledge and experience analogous to the story character(s)' knowledge and experience. That can happen whether or not the participant had previously been aware of any feeling or actions corresponding to those of one or more characters in the story. A shorthand phrase for such activation is that one or another of those story characters is constellated in a participant by the participant's work with story and questions.

Another way of saying this is that no one character in the story is or describes the whole of me, but it often describes a part of me. I may not have been aware of that part of me prior to my work with that story.

Questions

If you encountered this material in a group workshop, the facilitator would ask one or more questions after each story. Responses to the question are addressed not to the facilitator, and not to other members of the group but to the center of the circle. The story is "myth," a technical term that has been defined by Joseph Campbell as something that happened and continues to happen to many people at many times and places, in contrast to history which happened at one time and place to one person or a group of persons (and, a scientist would add, in contrast to science which is something that happens and is replicable under specified conditions). As a myth, the story is considered to bubble out of the center of the group circle rather than from the facilitator. Similarly, the question asked is considered to be coming from the center of the circle, and therefore responses are addressed to the center of the circle rather than to a person at the perimeter of the circle or to the facilitator. That should minimize any onus on, or defensiveness by, a respondent about one's response. Similarly, others in the circle are expected to respect any response since it is addressed to the center and not to them. Responses are not to be discussed or argued with, particularly not by the facilitator.

The discussion circle consists of center and perimeter. Without participation from the perimeter, in the form of responses to questions, the circle and therefore the seminar would not exist. It is important for each participant to hear oneself verbalize a response, even if that response sounds similar to one that has already been given.

The questions are designed to evoke choice and feelings rather than to test knowledge or recall. Reasons for, or explanations of, choices may be asked for. Respondents will be asked to stick to the subject matter of a question because one of the easiest ways of escaping a choice or a feeling is to talk about something else. Similarly, the question asked at a session is about the story told in that particular session, not about the end of the story or about another story. Each question is a universe of discourse, embedded in the universe of discourse of the story of that particular session. (By definition, a universe of discourse is a separate world surrounded by an impermeable barrier.) A participant is free to state different choices, feelings, and opinions at any time. Such changes are seen as fresh insights rather than as error or shame for earlier choices.

"I've heard that story before"

Of course you have heard that story before. It is part of our cultural heritage. It is a myth. It is universal truth. You and I have that story in our bones.

But in the group workshop I am not listening to, and in the self-study I am not reading, that story to find out how it ends, "who done it," or what the facts are. I am called on to take a stand vis-a-vis that story at this particular moment. I am listening to it, or reading it, to see whether I still have the same answers to questions evoked by the story or to see whether the story evokes new questions.

Bruno Bettelheim has noted that a fairy tale asked for most frequently by a child most likely describes what that child feels to be its most vital problem at that stage of its life. (For example, the Grimm Brothers' fairy tale that gripped my emotions for many decades was "The Goose Girl." Like that protagonist, I also lost home, country, and native tongue as a pre-pubescent.)

Ritual has been described as a process of activating a myth. In that sense repeated exposure to a story—by rereading it, by noting whether responses to story and questions change, and by asking new questions about it—reconstellates powerful, often numinous, characters within myself.

Instructions

You are not at a group workshop now, so you have an opportunity to create your own pace and place. Find a private space and time for 45 minutes, turn off the telephone, put out the cat, and open one of the stories. Read the story, consider it, and respond to the question(s) in writing or other art form. Send your response to: response@storiesandquestions.com. Stay with the story for a day or more—preferably a week. Look at the story, questions, and your responses occasionally, and write down any additional thoughts. Note any additional insights.

Your response will be acknowledged and the next Session will be sent to you. Repeat the process with the next Session. Continue in that manner until the final Session.

Because this kind of work is an ongoing process and new insights keep popping up, it is well to keep the Session materials, your

responses, and the facilitator's comments in a notebook. You will find that collection a growing resource as new insights arise. You will also find that it becomes a valued friend and adviser in dark times.

Which Story?

Four stories are available as Sequenced Self Studies at this time. They are:

- Cracked Pot (11 sessions)
- Moses Mendelssohn's Dream (5 sessions)
- Rainmaker (5 sessions)
- Becket (8 sessions)

Any of those Sequenced Self-Studies is worth doing in its own right in the same way that one goes to a movie or takes a trip for adventure, enjoyment, or enrichment.

Just as movies or trips also may be taken with specific purposes in mind, such as information or education, these stories can be used for specific purposes as well as in their own right. For example, *Cracked Pot* and *Moses Mendelssohn's Dream* have been used for working with self-worth problems. *Rainmaker* has been useful for conflict resolution in highly polarized situations. *Becket* is a good practicum for finding moral advisers in hierarchically structured organizations. Both *Rainmaker* and *Becket* are excellent self-studies for ethics training.

How Do I Start?

On the following pages [of the web site] you will find the first Session of each of the available self-studies. Choose one, follow the instructions, and send your response to: response@storiesandquestions.com.

Rudy will then comment on your response and activate the next session of your self-study.

III. Research Theory and Methods

8. Detection Methods

The Misuse of Statistics: Concepts, Tools, and a Research Agenda

John S. Gardenier, National Center for Health Statistics, Centers for Disease Control and Prevention, USA

David B. Resnik, Brody School of Medicine, East Carolina University, USA

Key Words. Competence, Ethics, Journals, Research misconduct, Statistics

While it is widely recognized that the proper use of statistics is a key element of research integrity, there has been considerable debate about how to understand or respond to the *misuse* of statistics in research. To understand what is meant by "misusing statistics," it is important to describe the role of statistics in the scientific method and relate the concept of "misuse" to other ethical concepts, such as "misconduct" or "incompetence" or "negligence." We believe that some misuses of statistics can be considered misconduct, although most misuses should be viewed as negligence or deficits of competence.

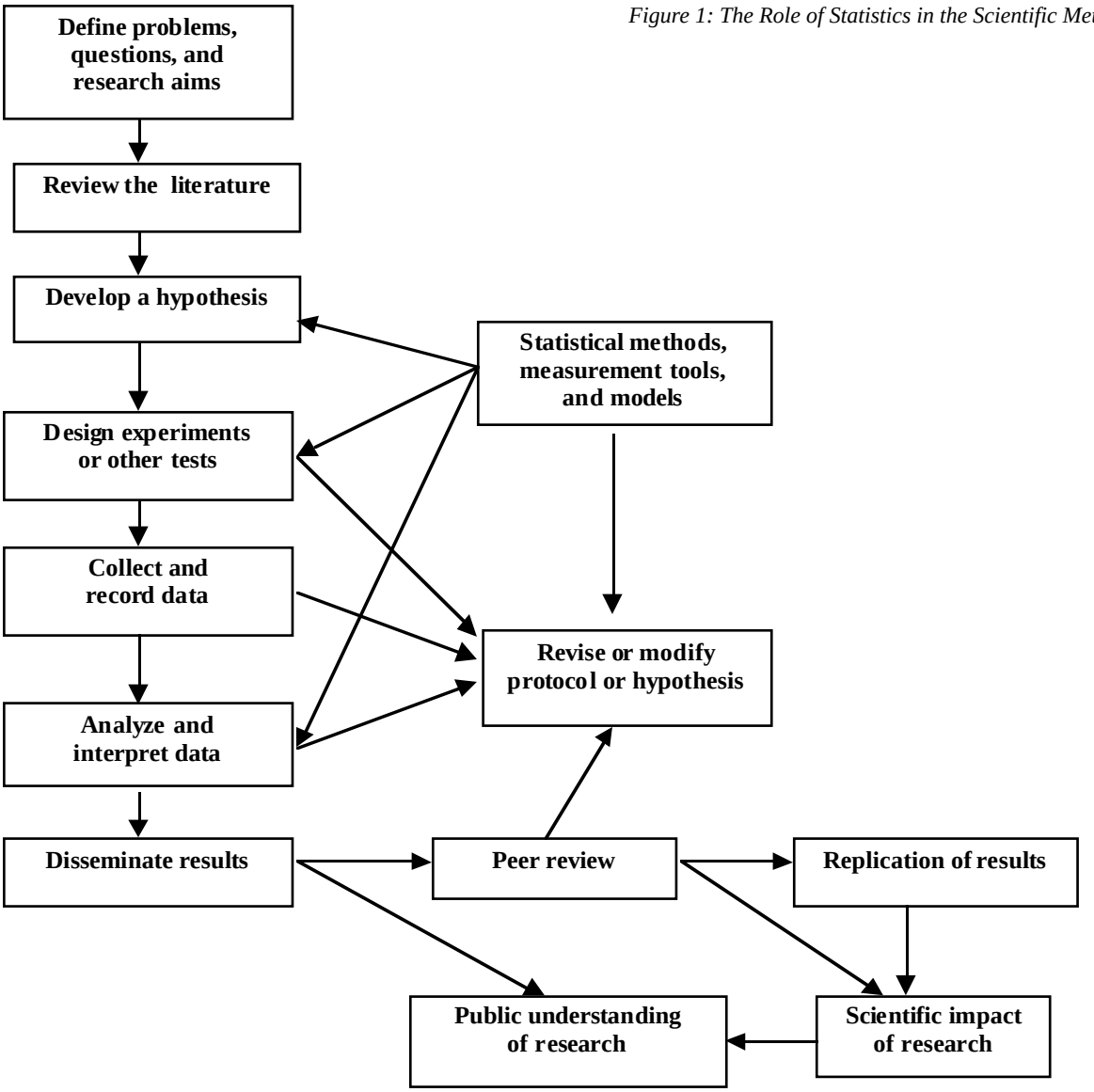
Statistical methods, theory, techniques, and models play an important role in several stages of the scientific method, but we will focus here on just two stages (See Figure 1). First, statistics is essential to good *experimental design* as in randomized clinical trials, for example. In order to obtain a rigorous test of a hypothesis, it important to obtain data that can provide evidence for or against the hypothesis. If the hypothesis is a comparative or quantitative statement, such as "drug x is more effective than drug y" or "less than five percent of patients suffer serious side effects from drug x," then the conclusions must be based on statistically significant results. For example, an experiment that compares the effects of two drugs on only ten patients is very unlikely to produce statistically significant results. If some or all of those patients are subjected to health risks in the experiment, this creates two additional ethical problems. First, it is unethical to expose a human subject to an unnecessary experimental risk, unless the potential benefits (to the individual or to society) of exposure to the risk outweigh the potential harms. If the experiment is not well designed such that no meaningful conclusions can be drawn, then the potential benefits will not outweigh the potential harms. Second, when patients give informed consent to participate in research, they usually believe that the research is valuable and may advance science. Encouraging or even allowing subjects to participate in an experiment that is highly unlikely to yield valid results is implicitly deceptive. It is important to address the statistical issues before conducting experiments or tests, because once one has gathered and recorded data, it may be too late to correct statistical (or ethical) flaws in the design of the experiment (1). The expression "garbage in, garbage out" applies here.

Second, statistics is important in an *analyzing and interpreting data*. There are many different statistical tools that one may use to analyze data, ranging from simple procedures, such as t-tests and

Corresponding author: John S. Gardenier, CDC/NCHS, 6525 Belcrest Road, Hyattsville, MD 20782, 301-458-4229 (voice), 301-458-4021 (fax), jsg6@cdc.gov.

* Disclaimer: This paper represents the personal views of the authors. It does not necessarily represent policies or positions of their institutions. The research effort reported here was not supported by grant funds.

Figure 1: The Role of Statistics in the Scientific Method



linear regression, to more complex ones, such as analysis of covariance and statistical modeling. It is not our aim to discuss these methods here, but we would like to point out that it is relatively easy to misuse these methods. To apply any statistical method correctly, one must have information about the variables used (continuous or discrete, gaussian or bimodal, etc.), information about the sampling process used (sample size, independence, randomness, representativeness, etc.), and a sound understanding of the theory and assumptions underlying that method. If a researcher does not use a method correctly, then conclusions may overestimate or underestimate an important relationship or effect. If we think of statistics as a tool for distinguishing between random "noise" in the data and the real signal, then someone who incorrectly uses statistics may produce a result that is distorted or even artificial. A person who correctly uses statistics will amplify and clarify the signal without distorting it (2).

With this understanding of the role of statistics in research in mind, we can clarify what we mean by "misuse" of statistics. Not all misuses have equivalent ethical implications, as we discuss later. A "misuse," for our purposes, is an *incorrect* use, i.e., a use of statistics that is not appropriate, given the research question, the experimental design, and the methods being used. For example, it may be appropriate to exclude outliers if there is credible evidence that such points are not part of the statistical population represented by the sample. It may

also be appropriate to use statistical methods to fill in (or impute) missing data for the purposes of statistical analysis. What's the difference between appropriate and inappropriate exclusion of outliers or appropriate and inappropriate imputation of data? Many books on statistical methods discuss these topics, but from an ethical viewpoint they boil down to the following: an appropriate exclusion (or imputation) is one that dampens the noise without altering the signal that describes the relationship or effect.

Misuses of statistics can also occur in the absence of erroneous or distorted results. Misuse can also arise from a failure to provide the research community with *important information* about the methods used or the experimental design. Researchers need to address such statistical issues as excluding outliers, imputing data, editing data, "cleaning" data, or "mining data."[2] These practices are often practical, or even necessary, but it is important to discuss them honestly and openly when reporting research results (3).

Thus, there are two types of misuses in statistics: (1) using statistical methods, techniques, or models in ways that produce distorted or artificial results; (2) failing to disclose important information about statistical methodology to researchers. Misuses of statistics may (or may not) violate several ethical obligations, such as the duty to be honest, the duty to be objective, the duty to avoid error, and possibly the duty to be open (4). There has been considerable debate about whether "misuse of statistical methods" should be classified as misconduct (5). The federal government and the scientific community have moved toward a narrow definition of misconduct that focuses on *fabrication, falsification, and plagiarism* (6, 7). The new federal policy implies that the misuse of statistics could be classified as a form of misconduct when it involves *intentional deception*. Some misuses could be classified as "fabrication" if they involve making up data or results, or "falsification" if they involve manipulating, changing, or omitting data or results. Misuses of statistics that do not involve intentional deception could be viewed as honest error, incompetence, bias, or "serious deviations" from acceptable practice (8). A person who makes excessive errors due to haste, ignorance, or sloppiness may be considered to be negligent or lacking the needed degree of competence, statistical or otherwise (9). Professionalism requires adequate application of both statistical and subject matter expertise to analyses. There might be varying degrees of culpability in a failure to meet this criterion. Clearly, honest error is never misconduct. Neither is it misconduct when two or more well qualified statisticians or other researchers disagree about technical issues in a given research protocol. Still, some misuses of statistics in research do fit the definition of misconduct used by the federal government. That may be hard to establish by a preponderance of the evidence. When a person changes or fabricates data, one at least has some kind of record that one can use to imply intent. When a person manipulates analyses of data, there may be no record to prove the manipulation was deliberate or even culpably negligent. Thus, as a purely practical matter, it may be very difficult investigate or prosecute such cases (10).

The Importance of Correcting Statistical Misuse

Statistics play vital roles in most aspects of modern post-industrial societies. Although statistics are sometimes dismissed as trivia or fuzzy math, distrusted as biased, or directly equated with lying, the truth is that they are inescapably important (11). As noted in the Preamble to the Ethical Guidelines for Statistical Practice:

> The professional performance of statistical analyses is essential to many aspects of society. The use of statistics in medical diagnoses and biomedical research may affect whether individuals live or die, whether their health is protected or jeopardized, and whether medical science advances or gets sidetracked. Life, death, and health, as well as efficiency, may be at stake in statistical analyses of occupational, environmental, or transportation safety. Early detection and control of new or recurrent infectious diseases depend on sound epidemiological statistics. Mental and social health may be at stake in psychological and sociological applications of statistical analysis.
>
> Effective functioning of the economy depends on the availability of reliable, timely, and properly interpreted economic data. The profitability of individual firms depends in part on their quality control and their market research, both of which should rely on statistical methods. Agricultural productivity benefits greatly from statistically sound applications to research and output reporting. Governmental policy decisions regarding public health, criminal justice, social equity,

> education, the environment, the siting of critical facilities, and other matters depend in part on sound statistics.
>
> Scientific and engineering research in all disciplines requires the careful design and analysis of experiments and observations. To the extent that uncertainty and measurement error are involved – as they are in most research – research design, data quality management, analysis, and interpretation are all crucially dependent on statistical concepts and methods. Even in theory, much of science and engineering involves natural variability. Variability, whether great or small, must be carefully examined both for random error and for possible researcher bias or wishful thinking.
> . . .
>
> Because society depends on sound statistical practice, all practitioners of statistics, whatever their training and occupation, have social obligations to perform their work in a professional, competent, and ethical manner. (12)

If researchers are careless or deceptive in their use of statistics, harms and costs to society will result. Poor statistics in science leads to poor science. The research record can be corrupted or polluted, wasting the time and energy of other researchers. At the very least, research funds lost in bad research represent an opportunity cost in that those funds could have been allocated to more deserving projects.

For all of these reasons, it is important that scientists and science administrators pay careful attention to the quality of statistics in science as funded, performed, and reported in their areas of jurisdiction and of competence. Good statistical work should be defended when it is attacked inappropriately. Bad statistical work should be detected and corrected as appropriate.

What are the Contributing Factors to Misuse?

There is not a great deal of evidence that has a direct bearing on the misuse of statistics in research. However, if one assumes that many of the factors that contribute to other ethical problems in research, such as misconduct, probably also play a role in the misuse of statistics, then one could cite the following factors, i.e., the "usual suspects" (13, 14).

- Pressures to publish, produce results, or obtain grants
- Career ambitions or aspirations
- Conflicts of interest and economic motives
- Inadequate supervision, education, or training

We believe that all of these factors probably play a role in misuses of statistics, but our conclusions are merely speculative. More research is needed on this topic. However, we would like to discuss two other possible factors in the misuse of statistics that are not on the above list of "usual suspects."

First, there are now many computer programs that analyze data. These programs are very user-friendly; all you need to do is load your data set and choose your statistical test in order to get results. One may even run several different tests in an attempt to increase the significance level (or p-value), although this can invalidate the testing. While these programs save a great deal of time and effort, they may contribute to statistical misuse in that it is possible to plug some numbers into one of these programs without knowing how the analysis works, or why a certain test would (or would not) be an appropriate test. We think this problem has a fairly obvious solution: teach more statistics in research. If students and researchers understand how to use statistics properly, then they should have fewer problems using statistical computer programs. Indeed, we believe that education is the key to improving statistical practice.

Second, it has become standard practice in some areas of research to only publish results that have a p-value of 0.05 or less. The best journals use more comprehensive criteria enforced by competent statistical peer review. We here address only those journals that place excessive reliance on the p-value. The value of 0.05 is an arbitrarily chosen number; there is no sound statistical or philosophical reason why a p-value of 0.06 is fundamentally different from a p-value of 0.05. However, under pressure to publish, researchers may decide to massage or manipulate data in order to obtain "significant" results. Furthermore, there is now a growing body of literature on publication bias in research (15-17). Publication bias occurs when there are discrepancies between the published research record and the complete research record. The discrepancies occur because journals tend to publish only "significant" results. There are some good potential solutions to the p-value problem. First, researchers should realize that p-values are merely conventional, not sacrosanct.

Second, they are also often sensitive to various theoretical assumptions and may give erroneous results due to mere artifacts of a data sample. Third, not all statistical computer packages compute all tests correctly. Fourth, journals should be willing to publish results that are substantial contributions to the literature of the field, not just those that appear to have met a conventional statistical test. The test result reported may not be correct, and even a correct conclusion that a certain hypothesis was not statistically supported by data from a well-designed study may be useful in limiting future fruitless research by others. Finally, researchers and research organizations should create databases for unpublished data of archival value and make those data publicly available (18).

Statistical Ethics, a Powerful Tool for Research Integrity

Statistical ethics is a relatively recent development. The seminal work, by W. Edwards Deming, was first published in 1965 (19). The American Statistical Association developed a series of statistical ethics codes or guidelines starting in 1979. Their current official Ethical Guidelines for Statistical Practice was promulgated in 1999 (12). The International Statistical Institute instituted its Declaration on Professional Ethics in 1985 (20). The United Nations has published Fundamental Principles of Official Statistics in the early 1990s, the current official version being dated 1994 (21).

The pattern that emerges from this brief history is that initial efforts to approach the issue tend to be optimistically simple. Corrections over time add to the scope and complexity of the documents. The most recent document breaks out areas of ethical responsibility for all people using statistical methods professionally (12). It covers separately, for example, responsibilities in publications and testimony, responsibilities to funders or employers, to research subjects, to research team colleagues, and responsibilities regarding allegations of misconduct. Beyond addressing responsibilities of the individuals, moreover; it also addresses the responsibility of those employing practitioners of statistical methods to provide a suitable moral climate for that work.

Such statistical ethics documents become tools for research integrity when they are integral to actual practice. For example, if a federal research funding agency were to adopt a policy of stating in grant announcements that all grant proposals received for projects employing statistical methods would be expected to be performed in accordance with the Ethical Guidelines for Statistical Practice, that would put real moral pressure on both proposers and grantees to avoid misuse of statistics. If journal editors were to state in notices to authors that any papers containing statistical methods submitted to that journal would be implicitly subject to those same guidelines, some of the authors would be more highly motivated to avoid misuse of statistics.

If all scientists and engineers who are competent in statistical methods would note published examples of misuse of statistics and report those to the funding agencies or journal editors involved, then the recipients would become more motivated to enforce sound statistical practice. In short, we should not let ethics documents sit unused on shelves or in unvisited cyberspace. Ethical considerations have practical consequences for good or evil. The failure of good people to use them effectively contributes to the likelihood that other people may perpetuate statistical misuse either through intent to deceive or simply through deficits of statistical competence.

A Proposed Research Agenda

While we believe that there are still many important conceptual and theoretical issues relating to the use/misuse of statistics in research, it should be clear from this brief discussion that more empirical research is required on the incidence of statistical misuse, its causes and effects¸ and on the efficacy of using ethics education and ethics documents as tools for improvement. The following are some of the empirical research questions we think are important to study:

1. How many (or what percentage of) published studies make statistical mistakes?
2. How many allegations of research misconduct involve misuses of statistics?
3. How many researchers believe that the misuse of statistics is an important ethical issue in research?
4. Do different fields have different statistical practices or take different approaches to the misuse of statistics?

5. What is the incidence of publication bias in various fields?
6. What do researchers and students know about statistics?
7. Where, when, and how do students learn about misuses of statistics in research or other ethical issues in statistics?
8. How often do researchers use statisticians or other statistical consultants?
9. Are editors and reviewers able to catch statistical misuses?
10. Can data audits detect misuses of statistics?
11. Do research ethics codes or policies address misuses of statistics?
12. When ethics education or ethics documents are used as tools to improve research integrity, how effective are they at promoting the proper use of statistics?
13. How often do institutional review boards (IRBs) discuss statistical issues in human subjects research? Do IRBs use statisticians?
14. How do misuses of statistics affect the public? Do such misuses ever cause harm to the public or threaten public health or safety?
15. How often do statistical issues arise in public policy debates?
16. What does the public know (or not know) about statistics?
17. How do lay people interpret important statistical concepts, such as "probability," and "risk"?

Bibliography

1. Johnson R. and Bhattacharyya G. *Statistics: Principles and Methods*. New York: John Wiley and Sons, 1985.
2. DeMets D. Statistics and ethics in medical research. *Science and Engineering Ethics* 1999*;* 5: 97-117.
3. Resnik D. Statistics, ethics, and research: an agenda for education and reform. *Accountability in Research* 2000; 8: 163-88.
4. Bailar J. Science, statistics, deception. *Annals of Internal Medicine* 1986; 104: 259-60.
5. Guston D. Changing explanatory frameworks in the U.S. government's Attempt to define research misconduct. *Science and Engineering Ethics* 1999; 5: 137-54.
6. Office of Research Integrity. Federal policy on research misconduct. Federal Register: October 14, 1999; 64, 198: 55722-55725.
7. Office of Science and Technology Policy. Federal policy on research misconduct. *http://www.ostp.gov/html/001207_3.html* , 2000.
8. Panel on Scientific Responsibility and the Conduct of Research. *Responsible Science: Ensuring the Integrity of the Research Process.* Washington: National Academy Press, 1992.
9. Grinnell F. *The Scientific Attitude*. New York: The Guilford Press, 1992.
10. Dahlberg, J. Personal communication. Office of Research Integrity, Rockville, MD. 20 November 2000.
11. Huff D. *How to Lie with Statistics*. New York: WW Norton, 1954.
12. American Statistical Association. *Ethical Guidelines for Statistical Practice*. Alexandria, VA: American Statistical Association, 1999. Available at *http://amstat.org/profession/ethicalstatistics.html.*
13. Macrina, F. *Scientific Integrity*. Washington: American Society for Microbiology Press, 2000.
14. Commission on Research Integrity. *Integrity and Misconduct in Research*. Washington: Public Health Service, 1995.
15. Ioannidis J. Effect of the statistical significance of results on the time to completion and publication of randomized efficacy trials. *JAMA* 1998*;* 79: 281-86.
16. Stern J and Simes R. 1997. Publication bias: evidence of delayed publication in a cohort study of clinical research projects. *BMJ* 1997; 315: 640-45.
17. Easterbrook P, et al. Publication bias in clinical research. *Lancet* 1991; 337: 867-72
18. Rennie D. Fair conduct and fair reporting of clinical trials. *JAMA* 1999; 282: 1766-68.
19. W. Edwards Deming. Principles of Professional Statistical Practice. *The Annals of Mathematical Statistics* 1965; 36: 1883-1900.
20. International Statistical Institute. Declaration on Professional Ethics. 1985. Available at *http://www.cbs.nl/isi/ethics.htm.*
21. United Nations. Fundamental Principles of Official Statistics. 1994. Available at *http://www.cbs.nl/isi/fundamental.htm* .

Images as Evidence: Forensic Examination of Scientific Images[1]

John W. Krueger, Division of Investigative Oversight, Office of Research Integrity, DHHS, USA

Keywords: Autoradiograms, Blots, Image processing, Manipulation and falsification, Scientific misconduct

A "questioned" scientific image, i.e., suspicions of falsification (or plagiarism) of image data, such as photographs of PAGE gels, autoradiograms, and blots (Western, Northern, and Southern) can give rise to an allegation of misconduct in science. Pursuing oversight review of institutional investigations and reviewing allegations that ORI receives directly, ORI commonly examines the evidence through image processing. Typically, the examination can extend beyond merely asking "what is the evidence the image is/isn't authentic?" and/or "are two contested images really the same?" Examples from these cases illustrate the general principles in forensic image processing and several methods that ORI has found useful in resolving the questions at hand. They provide an opportunity for further instruction as to what constitutes data falsification in an image.

Design/Methods

Source of Material: The material for this presentation was taken from a survey of 19 ORI cases that involved allegations of falsification or plagiarism of the images of gels, blots, auto-radiograms, and micrographs. The cases span a period from 1990 to 2000. The number of such questioned image allegations has generally increased, as has their incidence relative to other ORI cases. (Figure 9) A compilation from this review is discussed below.

Software: Most of ORI's image analysis was done on a Macintosh® computer. The reason is both historical and practical; files transfer easily from the Windows® platform to the Macintosh®; but the opposite is not always true.

ORI has found several different image processing programs that are readily available and well documented so that the results can be easily shared with all parties in a potentially adversarial dispute. (1, 2) Each separately —or in combination with the others— offers distinct advantages. The image processing was conducted using either NIH Image (3) and/or Adobe Photoshop® (4), both of which were equipped with the Image Processing Tool Kit® (IPTK) plugins. (5) NIH Image, developed at the National Institutes of Health, is in the public domain and is ideal for analytical treatment of 8 bit (256 shades) monochromatic images. Photoshop is better suited for conducting overlay comparisons of two images and for working with color, but it requires the IPTK's plugins for analytical work. Finally, ImageJ (6) is an update of the NIH public domain software that is compatible across computer platforms and will process images at 8, 16, and 32 bit depth; thus, it can detect vastly fainter features that might be hidden in the image.

Other Resources: Articles that can serve as guidance to issues involved in the forensic examination of contested documents can be obtained on the Internet. (1, 2) Those sites can serve as

Corresponding author: John Krueger, Office of Research Integrity, 5515 Security Lane, Suite 700, Rockville, MD, 20852, 301-443-2263 (voice), 301-594-0039 (fax), jkrueger@osophs.dhhs.gov.

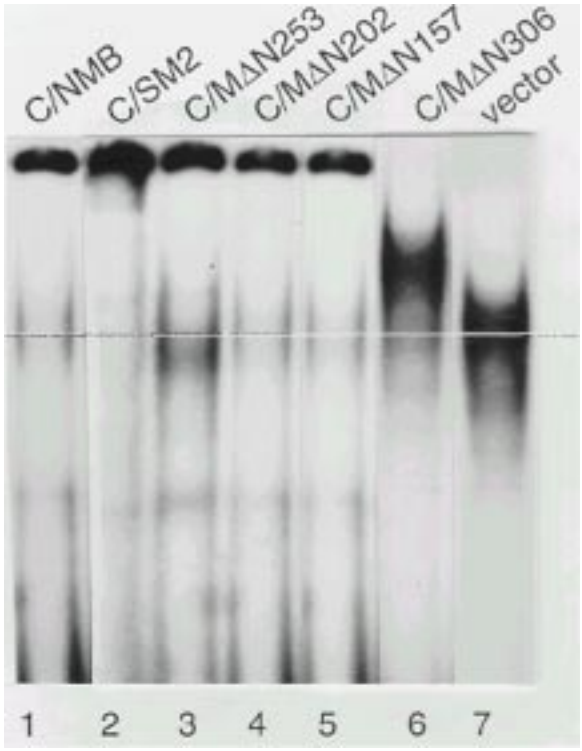


Figure 1. Original Western blot data. The results of an electrophoretic mobility shift assay to show bands reflecting the gene expression of separate mutant proteins. However, the shape of the bands and the pattern of the background in the 1st, 4th, and 5th lanes look alike.

links to find other material.

Reasons for Examination and Some Principles of the Image Analysis Methods

The usual motivation for image analysis is to examine the authenticity of a particular document or to determine whether two purportedly different images really were derived from the same experiment.[2] In fact, image analysis provides information that addresses other issues. For example, features can be detected that reveal the source of the image, whether it is compatible with laboratory records such as autoradiograms or prior blots (see note 2), and whether the questioned image existed on a computer as a file, or on a website as a component of someone else's homepage. Second, the analysis of the latter sources can provide dates of creation, which can be corroborated with laboratory records, etc. Third, image enhancement may reveal evidence for the mechanics of the figure's construction, such as edges of photographic prints and presence of "white-out" and may uncover "hidden" details, such as erasures of labels. Fourth, an analysis of the new facts produced, such as the sources, dates, and incidence of re-use, may establish whether a pattern of misrepresentation existed that rules out honest error. Examples from ORI's cases illustrate these points.

Figure 1 represents a photographic mock-up of Western blot data, consisting of five photographic strips, in which the 2nd to 4th lanes were on one strip. Although purportedly showing differentdeterminations of protein created by separate mutant gene constructs, the 1st, 4th, and 5th lanes look unexpectedly similar, but it is difficult to say for certain that they are the same.

One generic principle in making comparisons to determine the authenticity of data is to look at the features that would otherwise be un-noteworthy, such as fine features hidden in the background.[3] There may be random features that are hidden from our perception. The human eye, which responds to contrast, can distinguish only ~50 shades of gray (7) or less (8), but it can detect 100 shades of color (8).[4] However, the computer's response is not dependant on contrast; it can selectively amplify very slight differences in shade. The ability to detect such differences can be affected by the "depth" used to digitize the image, which in this case is 256 shades of gray.[5] The amplified differences in gray shades can next be shadowed and assigned false-colors to make faint differences even more visible, as shown in Figure 2.

These steps reveal faint artifactual features that were "hidden" in the background which are common to three of the lanes. Thus the respondent's claim, that at least two of the three lanes (1, 4, or 5 in Figure 1) represented evidence for gene expression of different mutant proteins, was a clear falsification of data.

Enhancement of the small difference in shades can also expose minute structural details in the morphology of bands, which otherwise would look smooth and featureless. Figure 3 illustrates a photo-montage from the above case; the bands appear similar in the 1st and 5th lanes.

Contrast enhancement and false-coloring of the above image as shown in Figure 4 demonstrate that the respective bands share similar miniature features. Thus, the image analysis showed that the first and the last lanes were from the same experiment.

In both examples above, the falsification was associated with false labeling of data that had been "re-used" from another experiment. The

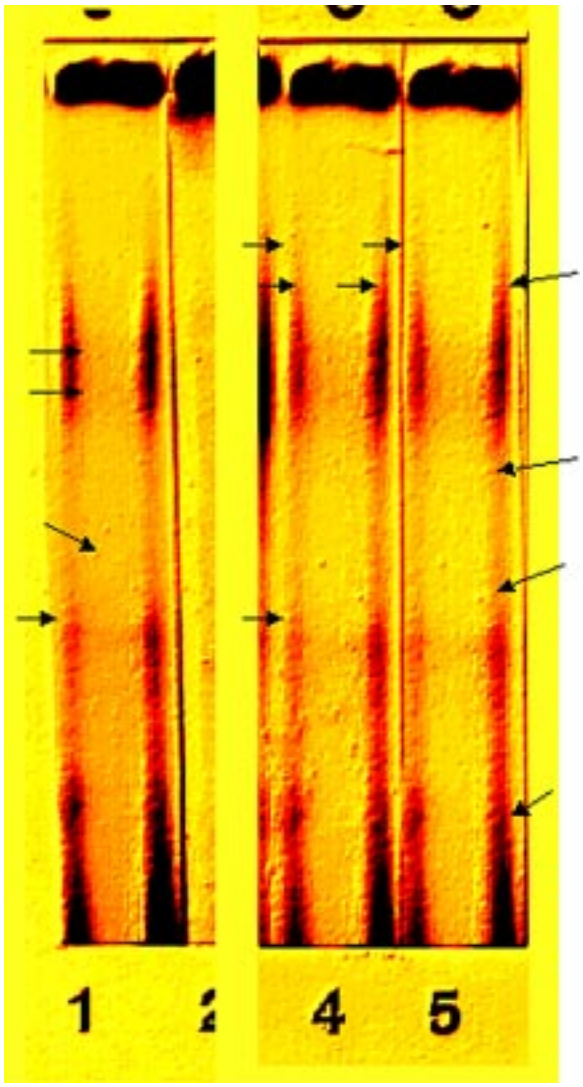


Figure 2 (left). Image enhancement of the questioned Western blot data. This ORI image analysis figure shows only the 1st, 4th, and 5th lanes from Figure 1. Contrast enhancement of the monochromatic gray-scale image, followed by shadowing and false-coloring (using NIH Image), revealed small features in the background artifact that are common to all three lanes (arrows) which the respondent had falsely represented as different. Note that in this case some differences can also be found, such as an horizontal artifact under the top band in the 4th lane, but they are in the background and represent artifacts that were introduced at some later point.

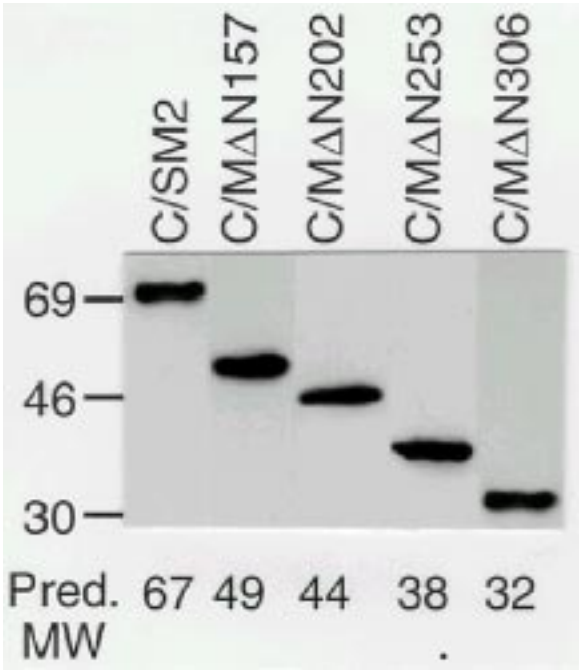


Figure 3. Western blot data. The results purportedly found a good agreement between the observed and the predicted size of five mutant proteins. However, the 1st and the 5th lanes' bands look similar.

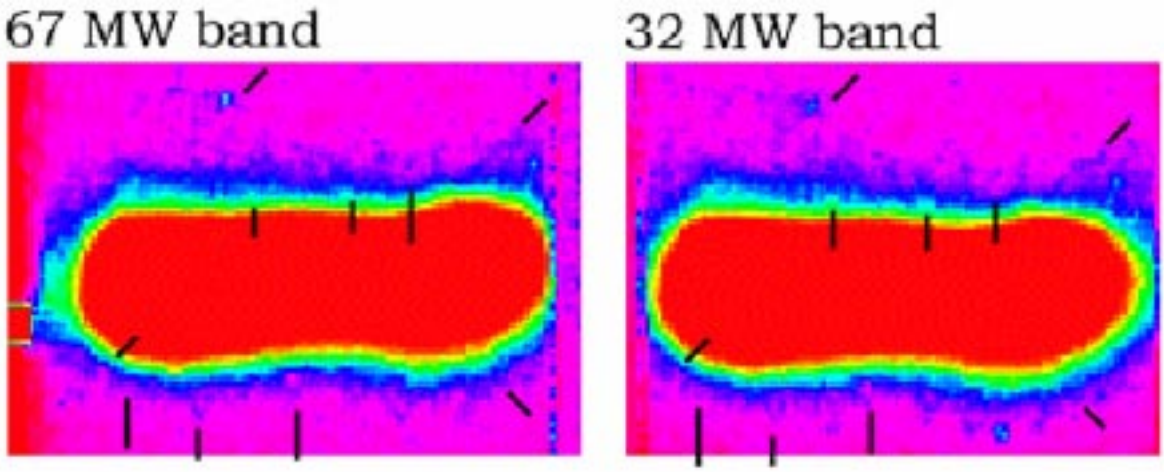


Figure 4. Image enhancement of the 67 kDa MW and 32 kDa MW bands from Figure 3. The bold lines denote miniature features in the bands' morphology that indicate both were actually the same data, which the respondent had falsified in re-use.

second example showed an additional falsification involving a false claim that the molecular weights had been determined. In this case, the intent to falsify the data is *prima facie*, because the molecular weight could not have been measured for the last re-used band. Finally, because the molecular weights were purported to approach the predicted values, the evidence also indicates that the falsifications are significant. These elements strengthen the findings.

Background detail and miniature features cannot be examined by image enhancement in all specimens. Fortunately, numerous other approaches are available in image processing to compare two questioned images. In general a combination of methods is determinative. For example, the morphology, size, and vertical arrangement of the bands and the existence of larger artifacts are the most obvious features to compare. Moreover, the horizontal spacing between the bands should not be overlooked; because substances rarely migrate on gels absolutely parallel, there may be slight differences in lateral disposition that are also significant. Some forms of artifact might re-occur, such as that introduced by a faulty film dryer and/or the edge of a blot on an autoradiographic film. The key question in cases of "replicating" artifacts is whether a relationship to other features should exist.

How to best visually represent the results of an image overlay is always a challenge. A visually effective and efficient method is to overlap color-coded plots of the "contour" map of the intensities in two separate blots, where the areas of overlap generate a third color. If two gray scale images are overlaid, the interpretation of the origin of features in the overlay becomes problematic unless each is first converted to a suitably chosen monochrome color scheme.

Reconstruction of a Missing Document: Analysis of an image can also be used to test the proffered source of a questioned image under circumstances in which the original raw data are missing. Figure 5 represents a composite image, which was created by combining a figure of a questioned Northern blot in the mentor's manuscript with a figure of a different experiment shown in the student's thesis. Unfortunately, the original blot and its PhosphoImager computer file were missing, but the mentor provided laboratory data purporting to be a different representation of the same blot (an ethidium bromide stain) that showed two groups

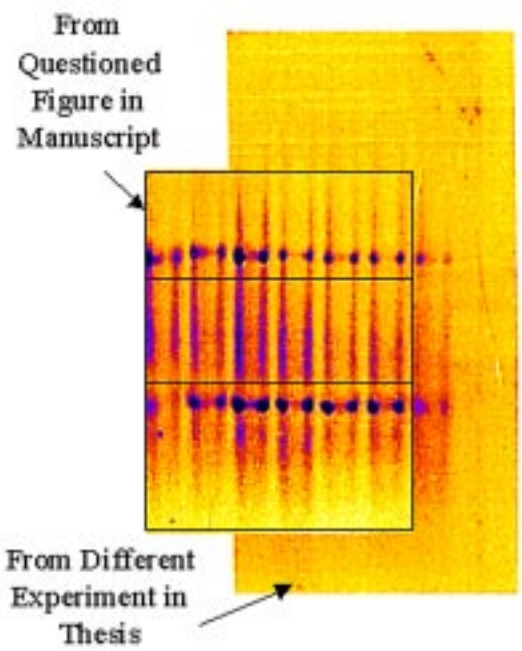


Figure 5. Overlay of the mentor's Northern blot mRNA data (small rectangle) with a figure from a different experiment from the student's thesis (tall rectangle). In this ORI image analysis, the actual fit was determined mathematically and showed the missing blot actually had at least *seven* lanes, indicating the respondent's claim was false.

of six lanes, separate by an empty lane. However, the overlay, shown in Figure 5, which was established as the best mathematical fit between the two sources, demonstrated that the missing original blot had to have had at least seven lanes. Thus, the proffered laboratory records could not be evidence of the mentor's "missing" data.

Analysis of Poor Images: The poor quality of an image is not necessarily a deterrent to the application of the above tools to its examination. The left side of Figure 6 shows a poor quality photocopy of data that was submitted in a mentor's National Institutes of Health (NIH) grant application, which purported to be a Western blot of an immunologic protein, "P-48," using ^{125}I-labeled human lymphocytes. The figure on the right side of Figure 6 represents the enhanced image of an autoradiogram from his student's experiments, which used ^{35}S - methionine labeling of cultured rats cells.

The distracting artifact due to over-photocopying could be minimized by image processing. This step revealed additional bands in the former with more certainty, and it more

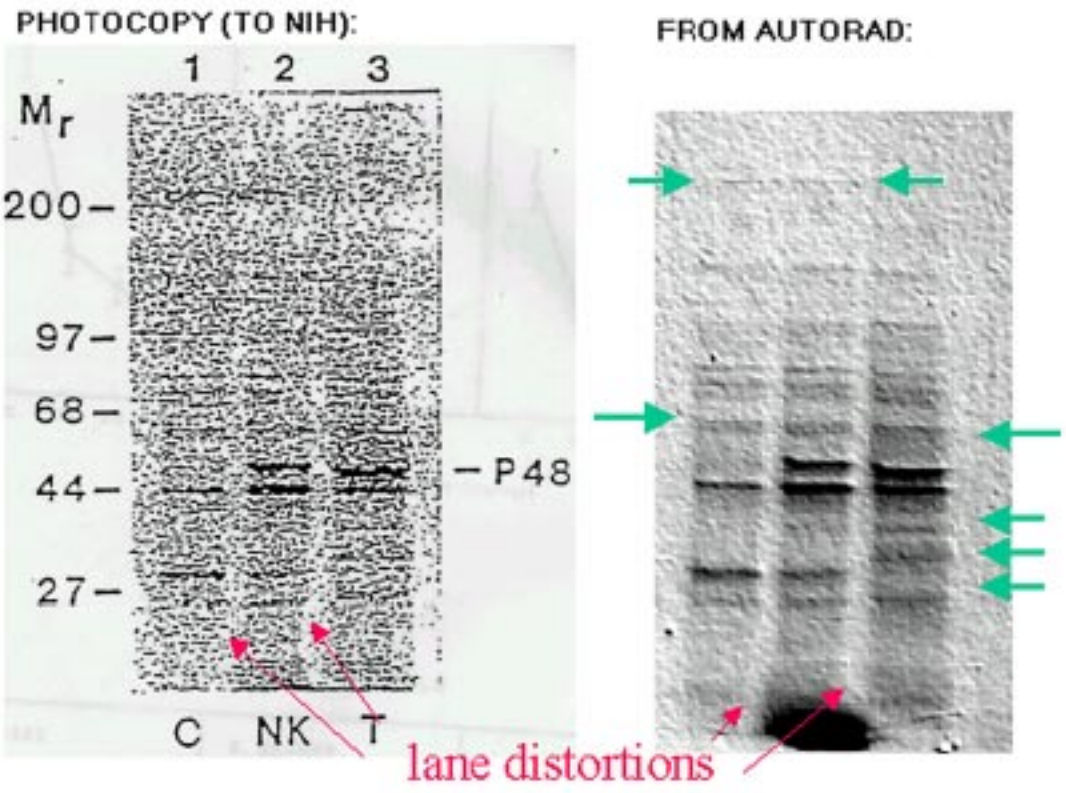


Figure 6. Examination of a poor quality photocopy. The mentor submitted the left hand "^{125}I-labeled" figure in an NIH application. At right is shown the student's ^{35}S-labeled autoradiogram, in which the band pattern was shadow-enhanced (green arrows). An artifactual lane distortion is denoted by the red arrows, which is weakly indicated in the photocopy.

clearly exposed a similar artifactual distortion of the lanes, as shown in Figure 7. The mentor had falsified the preparation, the experimental conditions, and the molecular weights in the photocopy that he had submitted to the NIH.

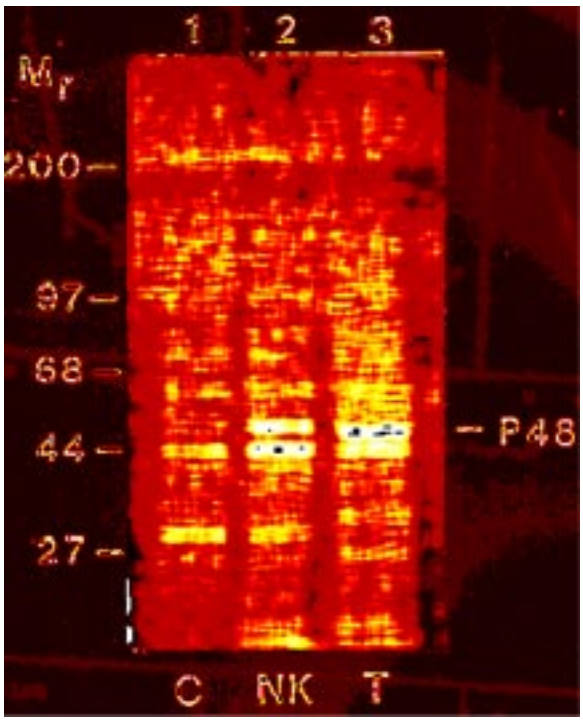


Figure 7. Computer enhancement of the bad photocopy shown in Figure 6. In ORI's image analysis, the distracting artifact in the photocopy can be removed by filtering, while false-coloring further enhanced the bands. The lane distortion artifact, present in the student's autoradiogram (Figure 6) was apparent in the same relation to the bands in the enhanced image, showing the student's results were falsified by the mentor to NIH.

Recovery of Probative Details:
Examinations of images may even reveal new evidence that bears upon other elements that are required for a finding of scientific misconduct. In another case, the allegation involved six instances where different sets of autoradiograms were allegedly falsely labeled and presented as different experiments. The student claimed these were honest errors, due, in part, to her inadvertent use of unlabeled autoradiograms. However, image enhancement by one of the institution's committee found evidence that the original label on one autoradiogram had been

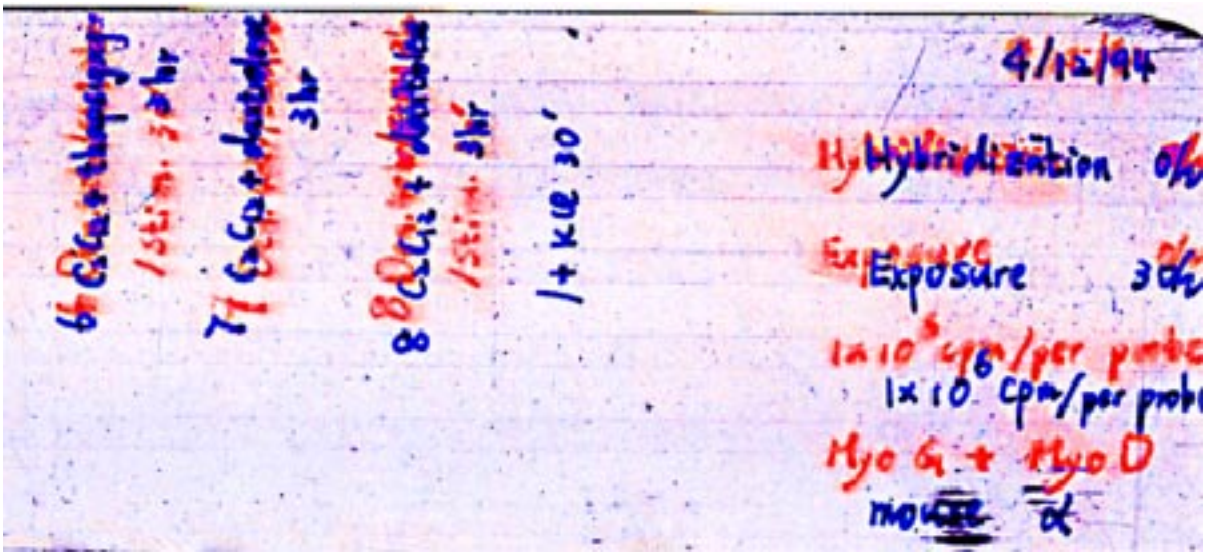


Figure 8. An example from one of six sets in which an autoradiogram had been falsely labeled and re-used. The institution's image analysis found evidence that the label for the prior experiment had been erased on the corner of the autoradiogram. The visible ink is blue, while evidence for the enhanced erasures is shown in red. Originally barely visible only as a faint and diffuse blue smear, the erased label differed from the film's background by only one gray level out of 256. The erasures were visualized here by ORI, after the film had been scanned at 42 bit resolution and the erasures had been selected for enhancement using their hue. The erased "MyoG" and "MyoD" denoted experiments on chickens and not mice. Thus, the re-use of the autoradiogram could not have been an honest error from mixup of unlabeled films, as the respondent originally claimed.

erased, but not fully (Figure 8). Thus, image processing revealed evidence that the falsification was not an honest error. ORI's subsequent analysis of figures in publications found that there was a pattern as to the six instances of re-use that was not consistent with their selection by chance.

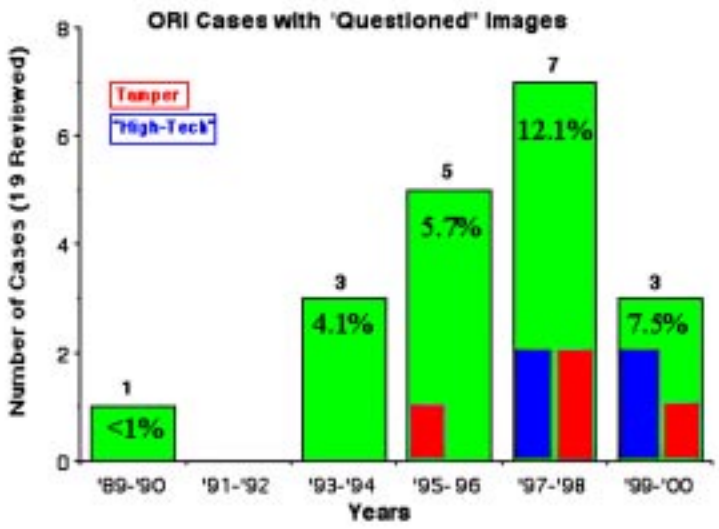


Figure 9. Incidence of 19 ORI cases involving contested scientific images. The data reflect when ORI's case file was opened; this formal step can occur at any phase in a case's history (i.e., at the allegation assessment, inquiry, or investigation stages). Thus the act of misconduct differs slightly in its date of occurrence. The percentages indicate the fraction of all ORI cases opened in those years. "Tamper" refers to allegations where the intensity of bands was selectively altered. "High-Tech" indicates manipulation by a computer to modify the image.

A scientific image is simply a picture constituting evidence that a test was carried out and/or that the test produced a certain outcome. In this context, the image is construed as qualitative "data." It could also be the basis for quantitative measurements, i.e., by measuring the "size" of a substance, or as the raw data for determine the amount of a substance. Thus, one consequence of discovering the falsification in an image is that there may be false claims elsewhere in a paper.

Compilation of Information from 19 ORI image analysis cases

In all of the cases above, the questioned image qualified as

Type of Misconduct Alleged	Number of ORI Cases
Falsely label as a different experiment (re-use)	13
Falsify molecular weight marker positions	>13
Cut, Graft, and Reuse, alter lane positions to fabricate data	5
Tampering: selective manipulation of image content, enhance/subtract bands, move position	4
Plagiarism of images (from Internet or journals), with false claims of element(s) from above	3

Table 1. Falsification of data with images–compilation from review of 19 ORI cases. This compilation indicates the incidence as number of cases, which under- represents the instances of misconduct, i.e., the actual number of figures or publications involved. The impact of the acts in each case was, in some cases, dramatically higher; one case involved 40 separate figures and multiple publications.

data, the intentional fabrication or falsification of which is the key element of research misconduct. On three occasions, a component of the allegation also involved plagiarism. The allegations reviewed by ORI generally involved use of fairly low-tech mechanisms for misrepresenting the data (Figure 9), such as re-use with false labels; in one case there were multiple instances of use of a felt-tip pen to add a band. Use of a computer to alter the content of the image has been seen less frequently.[6]

Table 1 compiles the nature of the misrepresentations involving questioned images in 19 ORI cases. The most common allegation was falsification of data by misrepresenting the results as being a different experiment, which also includes the attendant falsification of molecular weights. Only five examples occurred in which the lane markers were cut, re-grafted, and shifted so as to fabricate a test that was never run. Purposeful tampering with the image to selectively enhance or remove bands has occurred, but it was not very common. The allegations of plagiarism involved falsification of

Image Source	Respondent
Thesis (student)	8 (5 students) (3 mentors)
Others (plagiarized)	3
Prior publication (self)	2
Status:	
Senior Faculty	7
Junior Faculty	4
Fellows	3
Students	5
Allegation Source:	
Student/Mentor/Co-Invest.	9
Reviewers	5
Inquiry Committee	2
Audiovisual Technician	1
Audience	1

Table 2. Characteristic of allegations of falsification of images in 19 ORI cases.

figures copied from published journal figures or by use (and falsification) of images obtained from the Internet homepages of other scientists.

Other aspects of these image-related allegations are described in Table 2. Thesis research appears to provide a relatively frequent source of questioned images, falsified by both students and mentors. In three cases, the images were allegedly obtained from others, and in two other cases they involved falsification of images that had been published earlier by the same laboratory. The source of most of these allegations was co-workers, although in five cases it was a reviewer who recognized the image as being from another source, or saw intrinsic evidence that the image could not be authentic. Most allegations did not arise because the images looked inauthentic, but simply because they were either recognized as false or represented claims that a reviewer frankly disbelieved. The questioned image was often the one concern in a case that could not be easily dismissed.

Discussion

The facts uncovered by the forensic examination of questioned images can often be corroborated by other evidence, such as absence of records/ experiments on the requisite date(s), the existence of dated computer files or other versions, parallel images in publications, etc. In addition to the basic image processing, a clear

follow-up analysis is important.

The most useful analysis of questioned scientific images is done with a clear understanding of the experiment in question. This often requires collaboration between the image analysis and individuals who have a direct familiarity with the conduct of the scientific experiments at issue. (9) To date, only two institutions have reported to ORI using a computer-based image analysis. Only one institution documented those results; in that instance, image processing by a committee member uncovered details that were determinative (see Figure 8). The information from ORI's cases indicates that most allegations involved "reuse" of the image to represent data from a purportedly different experiment. Occasionally, photographs of gels or blots were "cut and pasted" together in different combinations. Manipulations by computer were less common.

An image by itself creates a mantle of authenticity, if only because we give unusual weight to what we see. Yet in those cases where scientific misconduct was found, discovery of one falsified image often led to the discovery of another, and in all the "original' laboratory records were "missing." Thus good laboratory practice may help to deter or to minimize the impact of falsification.

Notes

[1] Any views expressed in this article are my own and do not necessarily reflect those of the Office of Research Integrity. The citation of items in this article does not connote a product endorsement.

[2] The questions are not limited to examining items that look *alike*. For example, immunoblots from the same gel can be stripped and re-probed with entirely new labeled antibody to reveal different protein bands.

[3] The forensic value of the background information is completely analogous to the basis for numerical forensic analyses developed by Dr. James Mosimann in another presentation at this meeting.

[4] A simple "thought" experiment makes the point more elegantly than examining the underlying physiology of visual perception: any two gray levels, so alike that they could be fairly represented as one shade, could still be assigned two separate colors, say red and blue, of the same intensity. (8)

[5] Notice that digitizing at greater bit depth, such as 12 bit, would *in principle* detect fainter differences in shading to $^1/_{4096}$ parts, rather than the $^1/_{256}$ parts shown here.

[6] It is debatable as to whether it would be more or less difficult to detect computer alterations. What can be said is that an allegation rarely arose because an image on its face appeared inauthentic.

Bibliography

1. Blitzer, Herbert, "Digital Imaging," Institute of Forensic Image Processing, Purdue School of Engineering and Technology, IUPUI. PDF document available on the Internet at http://www.engr.iupui.edu/ifi/articles. Published in Law Enforcement Technology, March 2000.
2. Scientific Working Group for Imaging Technologies, "Definitions and guidelines for the use of Imaging Technologies in the Criminal Justice System" (vs. 2.1, June 1999). Department of Justice PDF document available on the Internet at http://www.for-swg.org/swgitin.htm.
3. NIH Image is public domain software that was developed at the National Institutes of Health. A large number of analytical macros for NIH Image are available via an active electronic bulletin board. NIH Image is particularly useful for any quantitative and analytical measurements, false-coloring to better visualize or code details of processed images, but it is limited to 8 bit gray scale images on the Macintosh. NIH Image is available on the Internet at http://rsb.info.nih.gov/nih-image/.
4. Photoshop is a popular commercial program for both computer platforms (Windows and Macintosh). Photoshop records the history of the computer's actions, to document the steps used in the image processing, but its analytical capabilities are limited. Photoshop can be obtained from Adobe Systems, Inc., San Jose, CA 95110-2704.
5. Photoshop-compatible "plugins" are available commercially as the Image Processing Tool Kit® (IPTK), Reindeer Games, Inc., 20 Battery Park Ave., Suite 502, Asheville, NC 28801. IPTK is also compatible with NIH Image; it greatly extends the analytical capabilities of either program for analysis of 8 bit gray or 24 bit color images. Information can be found on the Internet at http://members.aol.com/ImagProcTK.
6. The base in support for analytical macros for ImageJ is not as large as the forerunner, NIH Image, but it is growing. ImageJ and its plugins can be obtained on the Internet at http://rsb.info.nih.gov/ij/plugins.
7. Inoué, Shinya, "Video Microscopy," Plenum Press, New York, N.Y. 1986. pp 80-83.
8. Russ, John C. "The Image Processing Handbook," CRC Press, Inc., Boca Raton, 2nd Edition. 1994. p. 213.
9. Indeed, the examples presented in this paper represent the collaboration and assistance from my ORI colleagues, Drs. John Dahlberg, Kay Fields, and Nancy Davidian.

Terminal Digits and the Examination of Questioned Data

James E. Mosimann, Office of Research Integrity, DHHS, USA

John E. Dahlberg, Office of Research Integrity, DHHS, USA

Nancy M. Davidian, Office of Research Integrity, DHHS, USA

John W. Krueger, Office of Research Integrity, DHHS, USA

Keywords. Digits, Research misconduct, Statistical forensics, Terminal Digits, Uniform distribution

Our objective is to illustrate the use of statistical methods to examine the authenticity of data in the investigation of research misconduct. We present examples of statistical analyses of questioned data from several cases that illustrate the experience of the Office of Research Integrity. We show that the statistical examination of numbers that are normally *unrepeatable* when experiments are repeated, or otherwise are of inconsequential meaning, may reveal substantial clues as to the authenticity of questioned data when compared with numbers in data that are unquestioned. We illustrate the occurrence of the uniform distribution of non-leading (insignificant rightmost) digits in unquestioned numbers, along with examples of deviation from such uniformity for fabricated or falsified numbers. (Most people are unable to choose digits randomly.) We describe several cases in which a variety of anomalies in data sets provided the impetus for the examination of rightmost digits. The anomalous behavior of rightmost digits, when added to testimony and other physical evidence, can greatly enhance or decrease the credibility of witnesses. The cases discussed involve: 1 and 2, Anomalous behavior of terminal digits in published or recorded numbers; 3, Terminal *odd* digits in event times that should have exhibited only *even* digits (and why); and 4, Data that were falsified by calculations from computer spreadsheets (detected by the inclusion of an additional digit of accuracy).

Introduction

Allegations of research misconduct[1] often are of the form that a particular experiment was not done as described, or not done at all. In considering such allegations it is often necessary to examine "questioned" data. Such data can establish that the experiment was performed as described. However, if the allegation is true, then these questioned data are necessarily falsified or fabricated.

A useful way to assess questioned data is to examine inconsequential components of data sets that are not directly related to the scientific conclusions of the purported experiment. Thus if the allegation is true and the data are falsified, the falsifier typically devotes attention to numbers that establish the desired scientific outcome. Properties of the numbers that are not directly related to the desired outcome are less likely to receive consideration by the falsifier.

The same principle of examining details inconsequential to the scientific outcome appears valid

Corresponding author: J.E. Mosimann (present address), ABL Associates Inc., 5 Balmoral Court, Rockville, MD 20850, jemosimann@erols.com.

whether the data are expressed in non-numeric form (images, written descriptions) or as numbers. Here we consider several cases where the data are numeric and lend themselves to immediate statistical description.

In all these cases we stress the importance of comparing "questioned" data with similar unquestioned data from the same laboratory or individuals.

Rightmost digits

Consider counts of radioactivity for a biological preparation; for example, 5071. In a recount of the sample, or in a replication of the assay, it is highly unlikely that the rightmost digits will be the same. Thus with two repetitions of the experimental procedure, instead of 5071, one might obtain respectively, 5109 and 4966. The rightmost, non-leading digits of these three numbers are not the same. Thus _071 differs from _109, and in turn both differ from _966.

Digits are often recorded well beyond the repeatability of the experimental procedure. For such rightmost digits, theoretically[2] there is a tendency to be uniformly distributed as expected in a lottery. For example, a uniform distribution of digits is expected in the Maryland Lottery. Figure 1 shows the frequencies of the digits 0 to 9 found in 5,106 winning "Pick-3" numbers (of 3 digits each) for the past ten years.[3] This distribution is not significantly different from uniform. All digits have occurred with nearly the same frequency, as they should in a lottery.

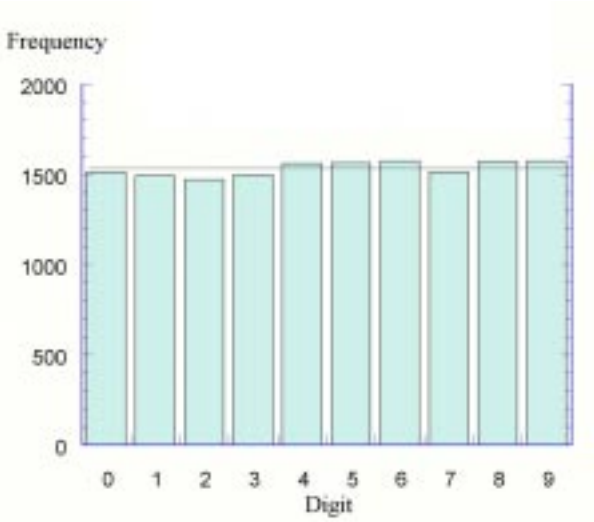


Figure 1. Ten years of Maryland Lottery Pick Three Digits, January 2, 1990 to December 31; 15,318 digits.

Case 1: Uniformly distributed rightmost digits in scintillation counts

In the first case, experimental measurements were known not to have been done because radioactive spots on the experimental sheets had not been excised and hence could not have been counted in the scintillation counter. Yet the respondent's notebook contained (falsified) handwritten counts for that experiment. In this case, faced with the evidence, the respondent admitted to the falsification of the numbers in the notebook.

In addition to the questioned counts, the notebook contained handwritten counts that were supported by counter output, and thus not falsified. Both questioned and unquestioned numbers occur in pairs (a numerator and denominator) and have large numbers of digits (Table 1).

The following procedure was used to find digits. The rightmost digit of a number was designated as occupying "Place 1," then the digit to its left occupied "Place 2," etc. Digits were examined in four places for each number, except that the leftmost digit was never included in the analysis. Thus by way of example, the underlined digits would be included in the analysis: 1<u>078</u>, 25<u>1183</u>, 7<u>35</u>, 6<u>2034</u>. It is clear that a three-digit number contributes two digits for analysis and a four-digit number, three digits. Numbers of five or more digits contribute four digits.

Chi-Square tests for uniformity of digit distributions from 252 falsified counts from notebook pages 141-152 are presented in Table 2. The distributions are not uniform. Three of the four Chi-Square values have probabilities less than .05, and when digits from all four places are grouped together, the total distribution is far from uniform (Chi-Square = 30.94, df = 9, p=.0003).

Chi-Square tests for uniformity of the digit distributions from 222 unquestioned counts also are presented in Table 2. The distributions are not significantly different from uniform. All of the four Chi-Square values have probabilities greater than .05, and when digits from all four places are grouped together, the total distribution is not significantly different from uniform (Chi-Square = 11.09, df = 9, p=.27).

The unquestioned counts have uniform or nearly uniform rightmost digits, whereas the falsified counts do not.[5]

Falsified counts (Notebook page 145)		Unquestioned counts supported by counter printouts (Notebook page 135)	
Numerator	Denominator	Numerator	Denominator
1078	251183	82267	170679
1770	217763	105584	190994
1091	225853	87592	181133
1434	238995	83341	197822
1247	241139	88426	172062
1131	260074	105068	194570
54350	220938	90707	150614

Table 1. Illustrative falsified and unquestioned counts from the respondent's laboratory notebook. Numerator (summation of reaction produced counts) and denominator (residual substrate count) are associated with a given clone, and activity is expressed by the ratio, numerator divided by denominator. Note that the 28 counts illustrated each contain from four to six digits.[4]

Chi-Square Results
For Falsified and Unquestioned Counts

Digits from 252 Falsified Counts

	Place 4	Place 3	Place 2	Place 1	Total
Number	185	250	252	252	939
Chi-Square	34.8	29.3	13.2	27.1	30.94
D. Freedom	9	9	9	9	9
Probability	.00006	.00058	.1521	.0013	.0003

Digits from 222 Unquestioned Counts

	Place 4	Place 3	Place 2	Place 1	Total
Number	195	218	222	222	857
Chi-Square	14.3	9.89	8.72	11.33	11.09
D. Freedom	9	9	9	9	9
Probability	.11	.36	.46	.25	.270

Table 2. Chi-square results for tests of uniformity of digit frequencies for falsified and unquestioned counts. The rightmost place is "Place 1"; the next place to left is "Place 2" etc. (Leftmost digits of numbers were excluded, so there are fewer "Place 4" digits than "Place 3," etc.)

Case 2: Unlikely Patterns in Rightmost Digits.

In this case, we again demonstrate the ability of uniformly distributed digits to distinguish questioned from unquestioned data. However, the digit analyses lead further to the identification of unlikely patterns in numbers that should not be related, given the purported experiment.

Table 3 (next page) reproduces the means and standard deviations from a published Table that was challenged by a coworker. Lipopolysaccharide extracts (LPS) were purified from endotoxin from various bacteria. The five rows in each half represent, respectively, different bacterial sources for the endotoxin. LPS was added at various concentrations to the cell cultures. Thus the five columns of the Table represent different levels of LPS (left to right, respectively: 5000, 500, 50, 5, and .5 ng/ml). The upper half of Table 3 represents cultures to which endotoxin and stimulator cells were added at the same time. The lower half represents cultures to which endotoxin was added 24 hours prior to the addition of stimulator cells. However, while supporting notebook data could be found for the first four columns, no supporting

Column 1		Column 2		Column 3		Column 4		Column 5	
Mean	Std Dev	Mean	Std Dev	Mean	Std Dev	Mean	Std Dev	Mean	Std Dev
17697	1739	17399	1680	15085	1342	18262	2934	27191	1404
20164	3540	16746	1171	19397	1133	17889	3919	26999	7107
23323	3861	24154	722	19094	1340	28763	3373	28611	967
24474	4042	18918	4459	14224	828	24596	6327	29152	1407
29711	1519	21855	8458	23840	1695	29669	3222	28765	7104
24752	1455	22498	4591	21639	1347	32825	3063	70714	2106
32683	8535	26321	2753	20015	2020	34030	3917	68177	7155
43411	4682	41980	1705	34026	3906	47703	1894	66004	3924
26535	2349	41592	5699	31262	2796	54588	5065	74316	2192
33216	3762	37036	2071	27513	5062	32033	8307	71117	6817

Table 3. Published Table (Column 5 has questioned data).

	Column 1	Column 2	Column 3	Column 4	Column 5	Columns 1-4
Number	70	69	69	70	69	278
Chi-Square	8.57	5.93	8.54	9.14	26.22	4.45
D. Freedom	9	9	9	9	9	9
Probability	0.478	0.747	0.481	0.424	0.0019	0.880

Table 4. Tests of Uniformity of Digits for the Columns of the Published Table. Chi-Square tests of rightmost digits for, respectively, Columns 1 to 5 of the Published Table, and for Columns 1-4, together.

notebook data could be found for the questioned numbers in column 5.

Of statistical importance is the fact that means and standard deviations in this Table are reported to several places. Thus numbers are recorded with greater precision than the repeatability that the biological experiment allows, permitting a digit analysis.

The treatment of rightmost digits is the same as that for the previous case. Digits are analyzed in four places with no leftmost digit included in the analysis.

Only the digits of the questioned Column 5 are significantly different from uniform (p = .0019). Columns 1 to 4 separately are not different from uniform (the probability ranges from .424 to .747). In the aggregate, columns 1 to 4 are again not different from uniform (p = .88).

Based on the contrast between the digit distributions for the questioned Column 5 and the unquestioned columns, the complainant's assertion that the experiment for Column 5 was not done is strengthened.

Place 4	Place 3	Place 2	Place 1
1	4	0	4
7	1	0	7
	9	6	7
1	4	0	7
7	1	0	4

Table 5. Vertical Pattern of Digits

Furthermore, examination of the standard deviations in the upper half of Column 5 of Table 3 reveals a remarkable "vertical" pattern. These numbers should be statistically independent from row to row. However moving vertically downward at each digit place reveals a symmetrical recurrence of digits: 1,7, blank, 1, 7; 4, 1, 9, 4,1; then 0, 0, 6, 0, 0; and finally, 4, 7, 7, 7, 4 (Table 5).

The vertical pattern does not appear consistent with the five presumably statistically independent experiments depicted by the separate

Journal 1		Journal 2		Book	
Trauma patients		Cancer patients		Trauma patients	
26428	406	6428	406	116428	3406
7824	376	7824	376	17824	3761
24840	1107	24840	1107	124840	7107
26660	345	6501	355	116660	34511
7791	407	7906	348	17791	407
9276	1498	12016	1476	9276	1498

Table 6.

rows of Table 3. Such a pattern is consistent with the formation of the numbers after the outline of the published Table had been established.

Finally, to check for the possible existence of a pattern, three publications by the respondent (two journal articles and a book chapter) were examined. Examination of these publications reveals patterns of digits that are inconsistent with biological observations. Consider Table 6 (above), which contains numbers from tables in three different publications by the author, all for a similar experimental procedure.

In these three publications, rightmost digits that should not be reproducible are the same in the first and third rows, and they would be the same in the second row except for the arbitrary addition of a "1" after the "376" in the last column. Further, in the fifth row two of the standard deviations are "407" while the corresponding means are "7791" and "17791." Note that the standard deviation 7107 occurs in the book chapter and also in Column 5 of the published table already discussed. The respondent in this case agreed that the data were "flawed" and retracted the relevant articles.

Case 3: Banker's rounding and "odd" terminal digits

For the purposes of a genetic study, electro-physiological measurements of spontaneous "firings" (action potential spikes) of isolated muscle fibers were made. Initially, a firing was determined to occur whenever a peak on the recording of current equaled or exceeded 10 picoAmps. Since the spontaneous "firings" were infrequent, the continuous record of the electrical signal was not retained. Instead, an "event detection" device sent the time and the amplitude of the firing to Excel spreadsheets as a permanent record of the experiment.

To graph the activity of muscles from different genetic crosses, the firings of various amplitudes were accumulated into bins of 5-picoAmp width (10-15, 15-20, 20-25, etc), with accumulation continuing until some bin contained 100 firings.[6] The resulting frequency distribution represented the pattern of firings (for Experiment 1, see Figure 2, below, in which there are just over 100 events in the 20-25 bin).

Prior to publication, the respondent's coworkers thought that firings should only be defined as those peaks 20 picoAmps or greater. Thus they asked the respondent to prepare a new graph like that of Figure 2, but sampling only peaks 20 picoAmps or greater (i.e. resampling the Excel spreadsheet until some bin contained 100 such firings.)

The respondent submitted a new frequency graph that appeared like the first, but truncated at 20 rather than 10. Since one would expect the shape of the new graph (above 20 picoAmps) to differ, the coworkers questioned the result.

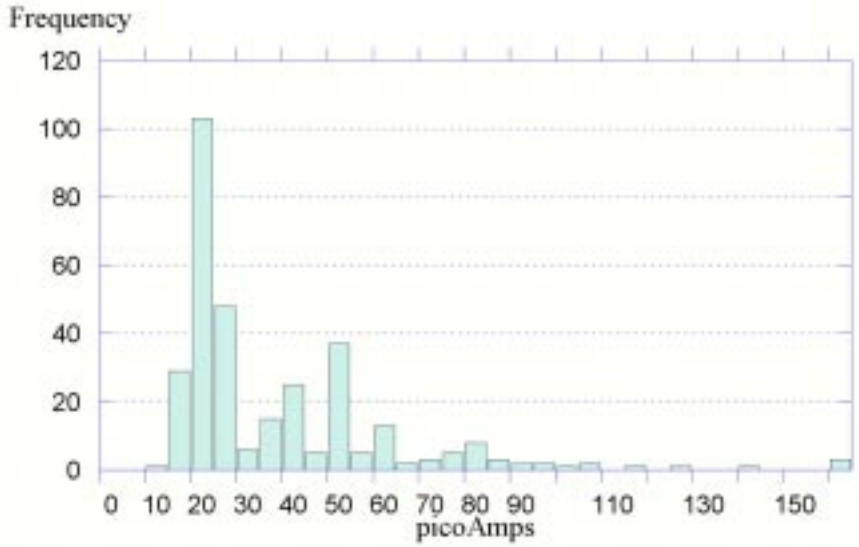


Figure 2. Binning of amplitudes into bins of 5-picoAmp width (initial 321 records of Experiment 1).

The respondent asserted that the new graph was not simply a truncated version of the first, but represented a fresh sampling of peaks greater than 20 picoAmps. He asserted that he had properly sampled the peaks in an Excel spreadsheet by counting beyond the initial 321 records on which the first graph (Figure 2) was based. The respondent furnished an Excel worksheet, "Experiment 1," of 551 records in support of the new graph. This worksheet contained the initial 321 records along with 230 additional records.

In addition to the Excel worksheet for Experiment 1, the respondent also provided a worksheet of unquestioned data "Experiment 2" with 1026 records. For Experiment 1 and the 10 picoAmp peaks, the initial 321 records of Experiment 1 are largely determined since the initial Figure 2 is known. Thus the last 230 records of Experiment 1 are more questionable. Since all 551 records were provided after the allegation, the opportunity existed to falsify or fabricate time points, but if falsifications occur, most would be expected in the last 230 records. Table 7, below, presents the first 12 records of Experiment 1.

It is interesting to note that all of the time values in Table 7 terminate in an even digit. The occurrence of only even time values can be explained by a long-used[7] practice sometimes known as "Banker's Rounding.[8]"

A simple explanation of the even terminal digits for time values is that two successive time-values are used in determining a peak, and the mid-point of the two is recorded. Thus when successive time values are added and divided by 2, the resulting terminal digit is 5 and would be rounded to an even digit, for example: (1000 + 1001)/2 = 1000.5 rounds to 1000, and (108.7 + 108.8)/2 = 108.75 round to 108.8. Therefore if numbers ending in 5 are rounded, only even numbers occur. The rounding of terminal 5's to the nearest even digit is the ANSI/IEEE standard[9] for rounding terminal 5's in computers. Examination of the terminal digits of the 1026 time values of the unquestioned data in Experiment 2 reveals no times ending in an odd digit. (The distribution of the 1026 penultimate digits of the times for Experiment 2 is not different from uniform (Chi-Square = 14.6, df = 9, p = .10).) In contrast, the questioned Experiment 1 contains time values that end in odd digits, reflecting insertions and alterations. In the initial 321 time points, six terminate in an odd digit (Figure 3). (The distribution of the 315 *penultimate* digits from the potentially unaltered *even* times is not different from uniform (Chi-Square = 8.14, df = 9, p = .52).)

Examination of the graph (Figure 4) of the final 230 records of Experiment 1 reveals many more (58) time values with odd terminal digits[10] than Figure 3. (The distribution of the 172 *penultimate* digits from the *even*, potentially unaltered, times is not different from uniform (Chi-Square = 12.3, df = 9, p = .20), whereas the distribution of 58 *penultimate* digits from falsified times ending in an odd digit deviates significantly from uniform (Chi-Square = 33.0, p = .00013).

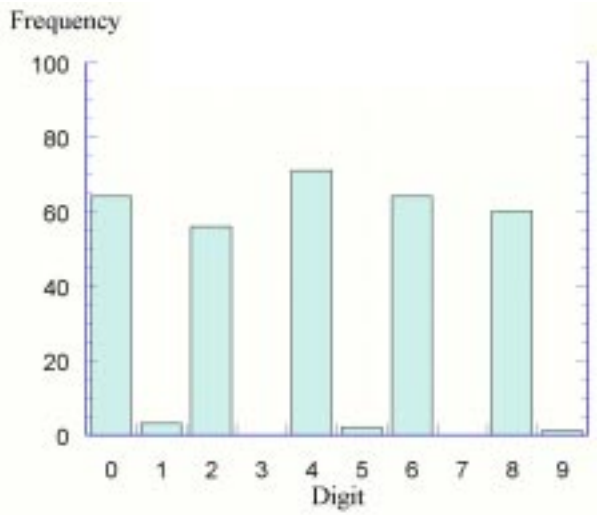


Figure 3. Experiment 1: first 321 time points; 321 terminal digits from 321 numbers. (Note presence of six odd digits.)

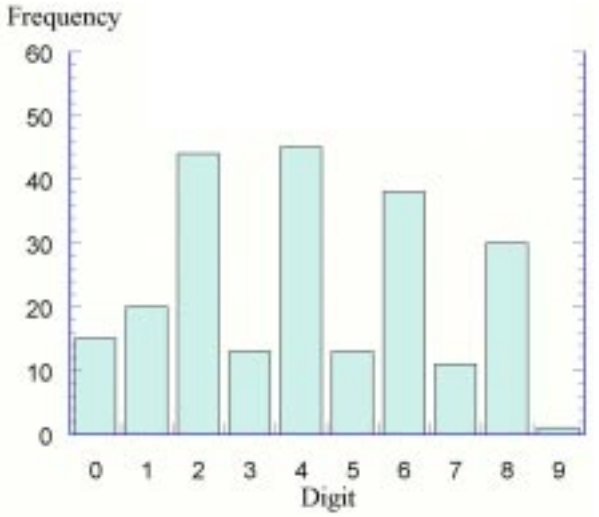


Figure 4. Experiment 1: last 230 time points; 230 terminal digits from 230 numbers. (Note presence of 58 odd digits.)

Experiment 1 - First 12 Records		
Time in Minutes	**Amplitude in picoAmps**	**Terminal Digit of Time**
0.0648	16.1	8
0.4904	22.7	4
0.4952	33.2	2
0.5398	19.8	8
0.9454	36.1	4
1.7182	44.4	2
2.6950	20.5	0
3.3626	19.3	6
3.7294	17.6	4
3.8586	14.9	6
4.3494	12.9	4
4.3712	45.4	2

Table 7. *The first 12 records of Experiment 1. Note that amplitudes include values less than 20, as expected. Also note that the terminal digit of the time is an "even" number for all 12 records.*

Many more time values terminate in odd digits in the final portion of Experiment 1, as expected if falsification occurred. The occurrence of time values ending in odd digits, mostly in the latter part of Experiment 1 (and the lack of uniformity of their penultimate digits) indicates data falsification. The timing of the occurrence of the minutes ending in odd digits is illustrated in Figures 5 and 6.

From Figure 6 it can be seen that not only do most of the odd time values occur in the last part of Experiment 1 (after minute 137.3006); it also appears from the denseness of the plot in the latter that the values immediately after this time point are quite close together. Further statistical tests of the intervals between events confirms the increased density in the latter part of Experiment 1, indicating the insertion of fabricated firing events.

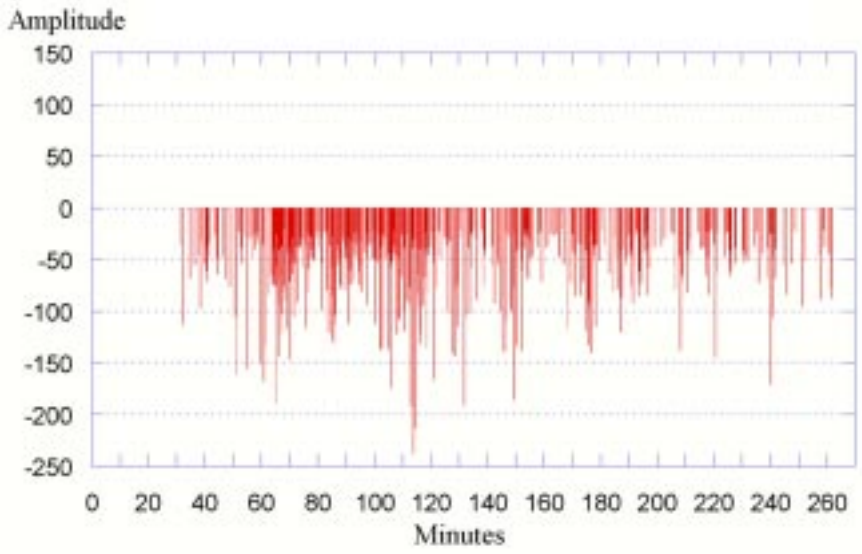


Figure 5 (above). Experiment 2, unquestioned, 699 amplitudes (abs>20). (No amplitude is associated with an odd minutes.)

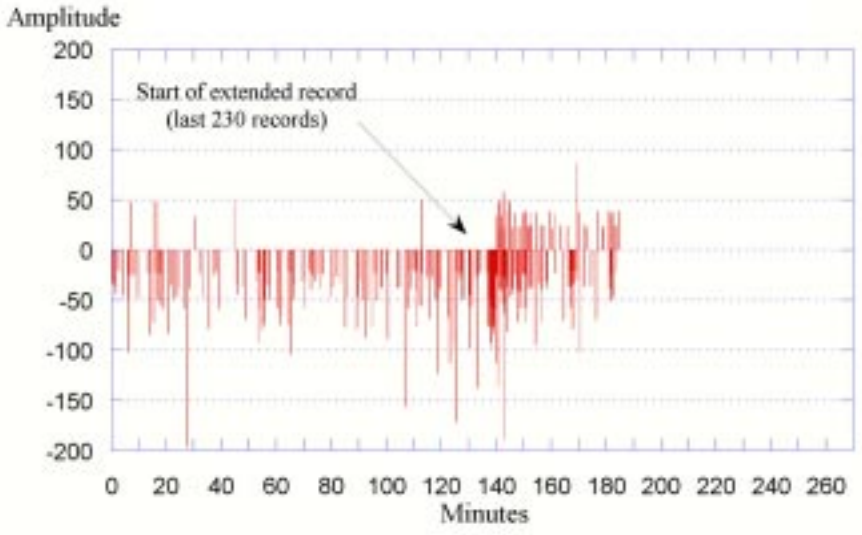


Figure 6 (right). Experiment 1, questioned; 371 amplitudes with abs>20, 52 with odd minutes. (Negative values, even minutes; positive values, odd minutes.)

	Rats					
	Weights-1	**Weights-2**	**Weights-3**	**Weights-4**	**Weights-5**	**Weights-6**
M-1	2.495	3.008	2.7515	4.631	2.250	3.4405
M-2	1.695	2.272	1.9835	3.019	0.702	1.8605
M-3	0.738	1.495	1.1165	1.768	0.843	1.3055
M-4	0.780	0.231	0.5055	0.394	0.085	0.2395
M-5	0.276	0.122	0.199	0.155	0.205	0.180
M-6	4.128	3.413	3.7705	2.261	1.187	1.724
M-7	1.131	1.224	1.1775	2.805	0.726	1.7655

Table 8. Portion of Excel spreadsheet with weights of muscles of rats 1-6. Note that some entries for columns Weights-3 and Weights-6 have four decimal digits and end in 5, whereas other entries have at most three decimal digits.

Case 4: One terminal digit too many

An investigator conducted studies on the effect of rhythmic contractions of skeletal muscle on blood flow using the hind limbs of rats. Blood flow was measured at rest and during nerve stimulation. In addition to measurements of blood flow, weights of individual skeletal muscles were recorded on data sheets. The experimental results for six rats were presented in a laboratory seminar. Sometime later a co-worker discovered that two of six data sheets were blank, and became suspicious that the measurements (blood flow/weights) had not been made for those rats. Suspicions were confirmed when frozen rat carcasses were checked. Although four had the hind limb muscles dissected, two were still intact and un-dissected. When confronted, the investigator (now respondent) admitted to falsifying data for two experimental animals. However, he subsequently withdrew the admission and denied the charges. The respondent stated that there was no evidence to support the claims that the research was falsified,[11] and that the university had not followed timely procedures.

The respondent presented to university officials blood flow and weight data for six rats on an Excel spreadsheet as well as purportedly original data sheets with handwritten entries for the muscle weights for six rats. Weights of 28 muscles and three other body parts for six rats extracted from the Excel printout are presented in Table 8.[12] Further weights as found in handwritten entries on separate data recording sheets for six rats are presented here in Table 9.

In Table 8, columns Weights-1, Weights-2, Weights-3 and Weights-6 correspond, respectively to columns 314-1, 314-2, 315-1 and 316-2 in Table 9. Thus the handwritten "original" data on the four data recording sheets (314-1, 314-2, 315-1 and 316-2) correspond to the columns labeled, respectively, Weights-1, Weights-2, Weights-3, and Weights-6 on the Excel spreadsheet. The columns Weights-4 and Weights-5 do not correspond to two additional data recording sheets labeled 315-2 and 316-1.

When values within a spreadsheet are calculated, rather than transcribed, the numbers may display more digits of accuracy than the original numbers that are the source of the calculated values. Therefore, looking for enhanced precision in spreadsheet numbers can indicate that certain numbers have been calculated or randomly generated by the spreadsheet software.

Since data are presented for six rats,

	Rats					
	314-1	**314-2**	**315-1**	**315-2**	**316-1**	**316-2**
M-1	2.495	3.008	2.725	3.859	3.479	3.440
M-2	1.695	2.272	1.984	2.087	1.881	1.861
M-3	0.738	1.495	1.117	1.464	1.320	1.306
M-4	0.780	0.231	0.506	0.269	0.242	0.240
M-5	0.276	0.122	0.199	0.202	0.182	0.180
M-6	4.128	3.413	3.771	1.933	1.743	1.724
M-7	1.131	1.224	1.178	1.980	1.785	1.766

Table 9. A portion of rat muscles weights from handwritten entries on six data recording sheets. Note that all numbers have a precision of three decimal places.

	Rat-3				Rat-6			
	Mean 1,2	**Weights-3**	**315-1**	**Difference**	**Mean 4,5**	**Weights-6**	**316-2**	**Difference**
M-1	2.7515	2.7515	2.725	0.0265	3.4405	3.4405	3.440	0.0005
M-2	1.9835	1.9835	1.984	-0.0005	1.8605	1.8605	1.861	-0.0005
M-3	1.1165	1.1165	1.117	-0.0005	1.3055	1.3055	1.306	-0.0005
M-4	0.5055	0.5055	0.506	-0.0005	0.2395	0.2395	0.240	-0.0005
M-5	0.199	0.199	0.199	0	0.18	0.18	0.180	0
M-6	3.7705	3.7705	3.771	-0.0005	1.724	1.724	1.724	0
M-7	1.1775	1.1775	1.178	-0.0005	1.7655	1.7655	1.766	-0.0005

Table 10. A Portion of the Weights for Rat 3 and Rat 6. The weights for Rat 3 are precisely the means of the respective weights for Rats 1 and 2. Additionally, the weights for Rat 3 correspond to three decimals to the handwritten weights for Rat 315-1. (The only exception is the weight for M-1 (shaded) where the rounded 2.752 is transcribed as 2.725. Correspondingly, the weights for Rat 6 are precisely the means of the respective weights for Rats 4 and 5. Additionally, the weights for Rat 6 correspond to three decimals to the handwritten weights for Rat 316-2, without exception.

and at most four allegedly were measured, the spreadsheet was evaluated for signs that some of the columns contained calculated values, rather than valid data entered from experimental records. The columns Weights-3 and Weights-6 in the Excel spreadsheet (Table 8) contain a number of entries that are recorded to one more decimal accuracy than the other columns (Weights-1, Weights-2, Weights-4, Weights-5). Additionally, these same entries for Weights-3 and Weights-6 contain one more digit than the purported original handwritten data as recorded on the sheets labeled 315-1 and 316-2 (Table 9). This extra precision could not occur from manual entry of the weights from the raw data sheets.

Instead, the presence of an extra digit indicates the possibility that these two columns represent calculated data. Further, where the extra digit occurs, it is always a "5." This indicates the calculation may have involved division by "2," suggesting that those numbers could be the means of two columns. (When the sum of the two numbers is even, there is no increase of the non-zero digits; however, when the sum is odd, division by 2 produces an additional "5" digit.)

In fact, the column Weights-3 is precisely the mean of columns Weights-1 and Weights-2 (see Table 10, below). Correspondingly, the column Weights-6 is the mean of columns Weights-4 and Weights-5 (Table 10).

Since these two columns are calculated on the spreadsheet, the "original" data on the recording sheets 315-1 and 316-2 are copied, respectively, from the spreadsheet-calculated columns Weights-3 and Weights-6. The only modification is that the "original" copied data are only transcribed to three-decimal accuracy as found on the (presumably) valid sheets labeled 314-1 and 314-2.

Lacking muscle-weight data for two rats, the respondent generated weights by twice forming means of measurements of other rats. The presence of the extra digit in the Excel spreadsheet provided the needed clue. When the respondent was shown that the two rats' weights were clearly produced as means, not measures, he accepted the finding of scientific misconduct.

Notes

1. 65 Federal Register 76260, December 6, 2000.
2. A theoretical discussion is found in J. E. Mosimann and M. V. Ratnaparkhi, "Uniform occurrence of digits for folded and mixture distributions on finite intervals," Communications in Statistics, 1996, **25***(2)*, pp *481-506*. Among other issues, this paper discusses approximations to continuous distributions by histogram-distributions for which the uniformity of terminal digits up to a specified place is known. Such theoretical issues are important, but our emphasis here is on direct comparison of questioned data with unquestioned data.
3. On May 1, 1995, the Maryland Lottery initiated a midday pick-3 drawing for weekdays. This is in addition to the nightly drawing. Thus there are more than 3,650 winning pick-3 numbers over the ten-year period. Maryland Lottery results may be found at the official website, *http://www.mdlottery.com*.
4. In all there are 474 counts: 252 admittedly falsified (notebook pages 141-152) and 222 unquestioned counts that are supported by counter printouts (notebook pages 104-106, 130-131, 134-135). Each count, falsified or unquestioned, contains from three to six digits. Digits were tested in four places, but no digit that was itself the leftmost digit was included in the analysis. Total analyses included 939 digits from 252 falsified numbers and 857 digits from 222 unquestioned numbers.

5. See "Data Fabrication: Can people generate Random Digits?" J. E. Mosimann, C. V. Wiseman and R. E. Edelman, *Accountability in Research*, **4**, *31-55*, 1995. This study shows that many people have difficulty fabricating random digits, even when trying to do so.
6. "Inverse" sampling until a certain number of a particular event occurs has a long history, particularly where rare events are to be studied. (*For example, see* J. E. Mosimann, "On the compound negative multinomial distribution and correlations among inversely sampled pollen counts," 1963, *Biometrika*, **50**, 47-54).
7. "It is conventional to round off to the nearest even digit when the number to be rounded is exactly half way between two successive digits." *pp*. 13-14, Paul S. Dwyer, *Linear Computations*, 1951, John Wiley & Sons Inc., *i-xi* , 1 – 344. (See also the next two footnotes.)
8. "PowerBASIC always rounds towards the closest even number. For example, both 1.5 and 2.5 would be rounded to 2. This is called banker's rounding. …" p. 169, *User's Guide*, 1997, PowerBASIC, Inc. 316 Mid Valley Center, Carmel, California, *i-vi*, 1-318.
9. ANSI/IEEE Std 854-1987, October 5, 1987, "ANSI" denotes the *American National Standards Institute* and "IEEE" denotes the *Institute of Electrical and Electronic Engineers, Inc.* "4.1 Round to Nearest. …if the two nearest representable values are equally near, the one with its least significant digit even shall be delivered." "5.4 Round Floating Point Number to Integral Value. …when rounding to nearest, if the difference between the unrounded operand and the rounded result is exactly one half, the rounded result is even."
10. 46 of these 58 time values that terminate in odd digits occur with amplitudes greater than 20 picoAmps. In the initial 321 records of Experiment 1, 6 of 6 odd time values occur with amplitudes greater than 20 picoAmps.
11. It is only after the respondent denied the charges and findings of the institution that the ORI demonstrated *which* two rats on the spreadsheet represented falsified data, and *the manner of falsification*.
12. The spreadsheet also contains columns of numbers representing blood pressure measurements and radioactive counts, some of which the university committee regarded as falsified. These are not presented here.

9. Publication Practices

Guidelines on Plagiarism and Paraphrasing in Writing Manuals Across Various Disciplines

Miguel Roig, *Department of Psychology, St. John's University, NY, USA*

Jaclyn de Jacquant, *Department of Psychology, St. John's University, NY, USA*

Keywords: Paraphrasing, Plagiarism, Writing guides

Discussions of plagiarism in conventional writing manuals typically focus on acknowledging the source of borrowed ideas and text. Such coverage often includes guidelines for proper attribution and citation practices. A number of manuals also provide specific guidelines for correct paraphrasing. By correct paraphrasing, we mean the extent to which text from an original source should be modified in order for it not to be considered a potential case of plagiarism. Those manuals that cover proper paraphrasing practices (1-3), generally suggest that, in addition to providing a citation, authors should always paraphrase others' work using their own words and expressions and avoid the use of the original author's language. For example, in a widely used guide, the authors state "When you paraphrase or summarize, you should use your own words and sentence structure (4). Imitating syntax, rearranging words and phrases, and borrowing phrases even as brief as two or three words do not change the original sufficiently to avoid plagiarism" (pg. 66).

Aside from the above guideline on paraphrasing, we are not aware of any other major writing manual that provides as close an operational definition for correct paraphrasing as the above example illustrates. However, the examples of proper paraphrasing provided by conventional manuals that offer such coverage suggest that a correct paraphrase must represent a very substantial modification of the original text, otherwise the paraphrase may constitute plagiarism. Moreover, some manuals such as the one quoted above, even suggest that, to avoid plagiarism when paraphrasing, not only should the original words be changed, but also the sentence structure of the newly paraphrased text must be different from that of the original (4-7).

As the reader might suspect, the criteria for correct paraphrasing appear to differ from writer to writer, particularly for inexperienced writers. For example, recent studies by one of the present authors have reported wide differences in plagiarism/paraphrasing criteria among college students (8, 9). Furthermore, similar differences also appear to exist among professionals, including physicians, English professors, and journal editors, and between college professors from a variety of disciplines (10-11). Some authors have even begun to express concern about the writing practices of those who engage in 'light' paraphrasing of others' works and terms, such as 'patchwriting' and 'paraphragiarism', have been offered to describe some of these inappropriate paraphrasing practices (12-14).

Depending on a number of factors, federal agencies, such as the National Science Foundation and the Office of Research Integrity do not classify inappropriate paraphrasing as instances of research

Corresponding author: Miguel Roig, Department of Psychology, Notre Dame Division of St. John's College, St. John's University, 300 Howard Avenue, Staten Island, NY 10301, 718-390-4513 (voice), 718-390-4347 (fax), roigm@stjohns.edu.

misconduct (15). However, based on definitions provided by conventional writing manuals and, depending on the context, others may still judge such inappropriate writing practices as potential instances of plagiarism. Thus, the 'light' paraphrasing of others' text, an innocuous writing practice to some, can have serious consequences and possibly result in disciplinary actions by the individual institutions and/or the academic disciplines involved.

A matter that we believe to be of major concern is evidence that suggests that inappropriate paraphrasing practices on the part of academics may be much more common than most people assume. For example, in a recent series of studies (11), one of the present authors found substantial differences in paraphrasing criteria among college professors from a variety of disciplines, including professors in the sciences. In one of the studies, differences in paraphrasing criteria arose even among members of a single academic discipline: Psychology. These findings led the author to review the official guidelines for avoiding plagiarism published by the American Psychological Association (APA), the parent association of psychologists (16, 17). A close examination of these guidelines revealed a certain degree of ambiguity in how correct paraphrasing and plagiarism are defined in that discipline. That particular finding is noteworthy because one of the sources reviewed is not only used by psychologists, but also by members of other disciplines (e.g., sociology and education) (17).

Given the importance of avoiding plagiarism in scholarly and scientific writing, the above findings raise a number of important questions: How do other disciplines in the sciences and the humanities define plagiarism? What are their guidelines regarding correct paraphrasing? How similar are these definitions across disciplines? In an attempt to address these questions, we surveyed the writing manuals of various disciplines within the sciences and humanities for their coverage of plagiarism. We were interested in the extent to which definitions of plagiarism, specifically guidelines for correct paraphrasing, are covered in these manuals and the degree to which such definitions are consistent across disciplines.

Method

We located the latest edition available to us of writing manuals of various disciplines (Appendix 1). First, we proceeded to determine each manual's extent of coverage of plagiarism by reviewing its index and table of contents for entries for 'plagiarism' and for 'paraphrasing'. If no entries were found for those terms we proceeded to examine sections on citation and documentation procedures.

Results

Most of the manuals were found to provide some discussion of citation and quotation procedures. Indeed, these sections are designed primarily for the purpose of identifying the source of ideas and thus, prevent an interpretation of plagiarism. Surprisingly, only 3 of the writing manuals examined (1, 17-18) listed entries for plagiarism in their index. The extent to which plagiarism was covered in these three sources varied somewhat. All three manuals provided some discussion of plagiarism. But, only two, the Modern Language Association (MLA) manual (1) and the American Medical Association (AMA) manual (18) defined this type of transgression and provided specific examples of instances of plagiarism (e.g., word for word lifting of a passage without attribution; presenting others' ideas without attribution).

The two writing guides that included coverage of paraphrasing (17, 18) defined it as restating text in the author's original words, but only the APA manual (17) provided an example of proper paraphrasing. However, as one of the present authors has pointed out, the definition for paraphrasing provided by the APA (i.e., "Summarizing a passage or rearranging the order of a sentence and changing some of the words is paraphrasing.") appears to be somewhat at odds with the actual example offered (11). That example, which shows the original text to have been substantially modified, is consistent with other conventional manuals' examples paraphrasing.

Discussion

Given the importance of avoiding plagiarism, we are somewhat concerned with the fact that the writing manuals of several academic disciplines, particularly many disciplines in the sciences, do not appear to have explicit sections on these matters. We note that other important resources on writing in the humanities and in the biomedical sciences also appear to lack entries on plagiarism (19-21).

It is possible that at least some of these

manuals do provide some coverage of plagiarism. But, in addition to not listing the term 'plagiarism' in the manuals' index or table of contents, any coverage, if it occurs, is probably very minor at best and takes place in sections other than those we reviewed.

If most of these manuals do not provide coverage of plagiarism the reason may be an assumption on the part of authors and editors of these reference materials that contributors to the professional literature are already knowledgeable about such fundamental matters of scholarship. Indeed, some manuals written for students in disciplines, such as biology, psychology, and sociology provide definitions of plagiarism and paraphrasing that are consistent with those of conventional writing manuals that provide coverage of these issues (22-25). Such detailed coverage at the undergraduate level supports the assumption that, at the professional level, authors already know the rules. Another reason for not including coverage may be that, as an ethical issue, plagiarism is likely to be addressed in other sources of information, such as a discipline's code of ethics. Finally, sections on citation procedures represent, to a great extent, a discipline's way of insuring that authors of original works are properly credited. Therefore, although explicit sections on plagiarism might not be provided in many of the writing guides reviewed, there is an implicit message in these guides that authors must duly credit others whose ideas, text, or data are being borrowed.

In spite of the above considerations, and in view of the fact that plagiarism continues to flourish, we believe that writing manuals across all disciplines should provide explicit sections on plagiarism that include clear definitions and examples of the various forms that plagiarism can take. In addition, given that a significant portion of scholarly writing involves summarizing and paraphrasing others' ideas and text, writing manuals should pay particular attention to this area and offer clear guidelines as to what forms of writing constitute proper summarizing and paraphrasing techniques. Finally, and perhaps most difficult of all, definitions of plagiarism and guidelines for summarizing and paraphrasing text should be standardized across disciplines. We believe that, in the absence of such standardization and given the increasing nature of cross-disciplinary collaborations, there is the potential for an even greater number of plagiarism cases in the future.

Bibliography

1. Gibaldi J. MLA Style Manual and Guide to Scholarly Publishing. 2nd ed. Modern Language Association of America; 1998.
2. Kirszner LG, Mandell SR. The Holt Handbook. 5th ed. Fort Worth: Hartcourt Brace; 1999.
3. Nadell J, McMeniman L, Langan J. The MacMillan Writer: Rhetoric and Reader. 2nd ed. New York: MacMillan; 1994.
4. Campbell WG, Ballou SV, Slade C. Form and style: Theses, reports, term papers. 8th ed. Boston: Houghton Mifflin; 1990.
5. Aaron JE. The Little, Brown Compact Handbook. 3rd ed. New York: Longman; 1998.
6. Hacker D. Writer's Reference. 4th ed. Boston, Bedford; 1999.
7. Troyka LO. Simon & Schuster Handbook for writers. 5th ed. Upper Saddle River, NJ: Prentice Hall; 1999.
8. Roig M. Can undergraduate students determine whether text has been plagiarized? The Psychol Record 1997;47:113-122.
9. Roig M. When college students' attempts at paraphrasing become instances of potential plagiarism. Psychol Rep 1999;84:973-982.
10. Julliard K. Perceptions of plagiarism in the use of other author's language. Fam Med 1994;26:356-360.
11. Roig M. Plagiarism and paraphrasing criteria of college and university professors. Ethics and Behav 2001; 11:307-323.
12. Howard RM. Plagiarisms, authorships, and the academic death penalty. College English 1995;57:788-806.
13. Howard RM. The new abolitionism comes to plagiarism. In: Buranen, L, & Roy LM editors. Perspectives on plagiarism and intellectual property in a postmodern world. NY: State University of New York; 1999. p. 87-95.
14. Levin JR, Marhsall H. Publishing in the Journal of Educational Psychology: Reflections at Midstream (Editorial). J Educational Psychol 1993;85:3-6.
15. Parish D. Scientific misconduct and the plagiarism cases. J College and University Law 1995;21:517-554.
16. American Psychological Association. Ethical principles of psychologists and code of conduct. Am Psychol 1992;47:1597-1611.
17. Publication Manual of the American Psychological Association. 4th ed. Washington, D.C. American Psychological Association; 1994.
18. Iverson C. et al. American Medical Association Manual of Style. A Guide for Authors and Editors. 8th ed. Baltimore: Williams and Wilkins; 1998.
19. The Chicago Manual of Style. 14th ed. Chicago: University of Chicago Press; 1993.
20. Reif-Lehrer L. Grant application writer's handbook. Jones & Bartlet; 1995.
21. Uniform requirements for manuscripts submitted to biomedical Journals. Ann Intern Med 1997;126:36-47.
22. McMillan VE. Writing papers in the biological sciences. 2nd ed. Boston: Bedford/St. Martin's; 1997.

23. Pechnik JA. A short guide to writing about biology. Boston: Little Brown & Company; 2000.
24. Szuchman LT. Writing with style: APA style made easy. Pacific Grove, CA: Brooks/Cole; 1999.
25. Richlin-Klonsky JR, Stresnki E. A guide to writing sociology papers. 4th ed. New York: St. Martin's Press; 1998.

Appendix 1: Writing manuals reviewed for their coverage of plagiarism and paraphrasing.

American Institute of Physics. AIP Style Manual. 4th ed. New York: AIP; 1990.
American Mathematical Society. A Manual for Authors of Mathematical Papers. Revised ed. Providence: AMS; 1990.
Associated Press Staff. Associated Press Stylebook and Libel Manual. Reading, MA: Addison; 1998.
Council of Biology Editors. CBE Style Manual: A Guide for Authors, Editors, and Publishers in the Biological Sciences. 5th ed. Bethesda: CBE; 1994.
Dodd JA. Editor. The ACS Style Guide: A Manual for Authors and Editors. Washington: American Chemical Society; 1998.
Gibaldi J. MLA Style Manual and Guide to scholarly writing. 2nd ed. New York: Modern Language Association of America; 1998.
Holoman DK. Writing about Music: A Style Sheet from the Editors of 19th-Century Music. Bekeley: U of California; 1988.
Iverson C. et al. American Medical Association Manual of Style. A Guide for Authors and Editors. 8th ed. Baltimore: Williams and Wilkins; 1998.
Linguistic Society of America. "LSA Style Sheet". LSA Bulletin 1999 Dec; 166: 64.
Publication Manual of the American Psychological Association. 4th ed. Washington, D.C. American Psychological Association; 1994.
The Bluebook: A Uniform System of Citation. 16th ed. Comp. Editors of Columbia Law Review et al. Cambridge: Harvard Law Review; 1996.
The Chicago Manual of Style. 14th ed. Chicago: University of Chicago Press; 1993.

Instructions to the Author: An Integrity Issue

Mary D. Scheetz, Office of Research Integrity, DHHS, USA

Keywords

Corresponding author:

Erratum Citation and Accuracy in the Publication Record

Marshall Thomsen, Department of Physics and Astronomy. Eastern Michigan University, USA

Christopher Aubin, Department of Physics and Astronomy. Eastern Michigan University, USA

Barbara Hall, Department of Physics and Astronomy. Eastern Michigan University, USA

Matthew Knee, Department of Physics and Astronomy. Eastern Michigan University, USA

Keywords: Erratum, Citation, Publication

Technological advances have greatly influenced the ways in which science is communicated. However, the refereed journal remains an important element of the system, providing a permanent record of information with some quality control over the scientific content. In trying to keep abreast of recent developments in a field or when entering a field of study for the first time, scientists often rely on the refereed journal as their primary information source. Thus accuracy of the written record becomes a significant issue.

While much has been written about the publication process in general, (1) we will focus on a small piece of the process that lends itself to accumulation of basic statistical information and, we hope, provides some insight into other broader aspects of publication. In particular we will look at physics papers that have an erratum associated with them and study how these papers are cited in subsequent literature. There are several issues we will examine. If an erratum is written, how likely is it that those who have read the original paper also will have read the erratum? If a corrected paper is cited, how likely is it that the authors who cited the paper also cited the erratum? Is it misleading to cite the original paper but not the erratum? Do authors typically cite their own errata?

Some of these questions have been addressed before. For instance a 1990 study of retracted medical papers showed that retractions tended to reduce, but not eliminate, citation rates. (2) A 1995 study of errata in physics journals showed that when corrected papers are cited, most often the corresponding erratum is not cited. (3) The authors of the study commented at the time that part of this citation problem was associated with the logistical issue of locating an erratum. It is much easier to search the publication record backward in time by studying citations. Moving forward in time to locate errata requires scanning journal contents or using an index (such as the *Science Citation Index*). The authors speculated that as more journals were provided in an electronic format, locating errata would be easier since the original paper presumably would be linked electronically to the erratum.

The American Physical Society now has a large collection of its journals available online via subscription. All of their recent online papers that have an associated erratum have a link to that erratum. We thus undertook a new study to determine if this electronic linking has improved the

Corresponding author: Marshall Thomsen, Department of Physics and Astronomy, Eastern Michigan University, Ypsilanti, MI, 48197, 734-487-8794 (voice), 734-487-0989 (fax), phy_thomsen@online.emich.edu.

citation rate of errata. Curiously, we find that, if anything, the citation rate for errata seems to have decreased since the introduction of the electronic format.

Study Design

Our study consisted of an examination of the citations of 14 papers from *Physical Review Letters* in 1995 and 1996 that had errata making nontrivial corrections. We included papers with calculational errors requiring replotting graphs or reproducing tables, papers in which derivations needed modifications, papers in which data needed to be reanalyzed due to misinterpretation, etc. We excluded papers in which simple typographical errors were corrected or acknowledgments of related work were added. The goal was to focus on papers in which there was a scientific error of substance being corrected by the erratum. At the same time, none of the errata reported on here amount to a complete retraction of a paper. For clarity in discussions below, we refer to these 14 papers as the *primary* papers.

We selected our primary papers from amongst the first papers to become available in the *Physical Review Letters* online collection. Hence the primary papers all have electronic links to their errata, and they have all been available in the literature for several years (thus increasing their chances of being cited).

Physical Review Letters is one of the most selective physics journals, containing papers describing some of the most recent and significant advances across all physics disciplines. We focussed on these papers since they are higher profile and hence likely to produce a greater set of citation data. In contrast, the 1995 study of errata in physics journals included papers from both *Physical Review Letters* and *Physical Review B*, the latter being a more specialized journal. That study showed that papers in *Physical Review Letters* tend to be cited two to three times as often as papers in *Physical Review B.* The 1995 study also showed the citation rate for errata in *Physical Review Letters* was substantially higher than that for *Physical Review B.* Thus our present study focuses on a journal with a *relatively* high erratum citation rate.

We attempted to identify all papers that had cited the primary papers and/or their associated erratum, using the *Science Citation Index* as our main tool. We located 507 papers citing the primary papers and/or their errata. We refer to this collection of papers as *secondary* papers. It is interesting to note that a small portion of these secondary papers cited only the erratum and not the corresponding primary paper. As a spot check on the accuracy of *Science Citation Index*, we used the citation list provided by the online version of *Physical Review Letters*. It should be noted that the journals used in this citation index are much more limited in scope than those used to assemble the *Science Citation Index*, listing citations by only American Physical Society journals. We selected three primary papers from our list that, according to the *Science Citation Index*, had no citations to their erratum. We verified this finding with all available listings on the more limited *Physical Review Letters* citation data base and also confirmed that all 21 secondary papers appearing on this database also appeared on the *Science Citation Index* data base. That is, we discovered no evidence that *Science Citation Index* was omitting papers appropriate for our secondary category.

Results and Discussion

The collection of secondary papers was divided into two categories. The first category contained those papers in which there was an overlap between the authors of the secondary paper and those of the cited primary paper. The second category consisted of those secondary papers in which there was *not* any overlap of authorship with the cited primary paper. The purpose of this division was to address separately the questions of how often authors cite their own errata and how often independent authors cite errata. The cases with overlapping authors will be considered first.

Table 1 shows data for authors citing their own errata. We exclude from consideration in the secondary paper data set those papers published prior to the appearance in print of the erratum. We are left with 59 secondary papers that could have cited an erratum. Of these, 25 (42%) actually did cite the erratum. The reason for the remaining 58% of the secondary papers not including the erratum citation is not clear. One possibility is that the author of the primary paper and erratum chose not to cite the erratum. Another possibility is that the person or persons of the secondary paper who took the most responsibility for writing that paper were not among the authors of the primary paper. In this case, it would be possible for the writer of the

Paper Identification Number	Potential Erratum Citations	Actual Erratum Citations
1	1	1
2	5	0
3	2	0
4	4	2
5	4	3
6	3	2
7	0	0
8	5	2
9	3	1
10	1	0
11	5	0
12	3	0
13	23	14
14	0	0
Total	59	25

Table I: Analysis of citations by one or more authors of the original (corrected) paper. Potential erratum citations represent the total number of papers citing the original paper, its erratum, or both. Actual citations represent the number of times the erratum was cited. Only papers appearing after the publication date of the erratum were considered in columns 2 and 3.

secondary paper to be unaware of the existence of the erratum. However, assuming the erratum author read through the secondary paper prior to publication, then either that author chose not to add the erratum citation to the list or overlooked the absence of the erratum in the references. We will return to this issue later.

Table 2 shows data for secondary papers sharing no authors in common with the cited primary paper. We exclude from the secondary paper data set those papers that did not appear in print at least one year after the publication date of the erratum. This is to ensure that the authors of the secondary paper had the chance to see the erratum at the time they were writing their own paper. After reducing the data set as described, 355 secondary papers remained. Of these, just 59 (17%) cited the erratum. The 1995 study of 9 primary papers in *Physical Review Letters* had a citation rate of 39% (51 of 131) when a similar approach to data analysis was used. While there are obviously statistical fluctuations associated with this sampling, it is worth noting that only 4 of the 14 primary papers in the present study had an erratum citation rate exceeding the 39% average from the previous study. It is thus safe to conclude that the advent of electronic journals has not had the desired impact on erratum citation.

We now return to the issue of the extent to which it is a problem that errata are not generally being cited. There are three fundamental questions. First, does the reader of the secondary paper need to be aware of the erratum? Second, will a reader discover an erratum based on information provided by the authors of a secondary paper? Third, whose responsibility is it to locate the erratum?

We will examine the first question in the context of the errata discussed here: those providing substantive corrections. The 1995 study of erratum citations showed that a little more than half of the primary papers examined were cited “in passing” in the secondary reference. In these cases, the secondary authors were primarily acknowledging the work of others in the field rather than laying down specific ground work for their own paper. These citations typically occur in the introductory section. The remaining citations to the primary papers

Paper Identification Number	Potential Erratum Citations	Actual Erratum Citations
1	6	0
2	4	0
3	22	0
4	13	8
5	8	2
6	15	7
7	3	1
8	6	6
9	2	0
10	2	2
11	8	0
12	17	1
13	248	32
14	1	0
Total	355	59

Table II: Analysis of citations not involving authors of the original (corrected) paper. Potential erratum citations represent the total number of papers citing the original paper, its erratum, or both. Actual citations represent the number of times the erratum was cited. Only papers appearing one year or more after the publication date of the erratum were considered for columns 2 and 3.

indicated that the authors of the secondary paper were using one or more results or ideas from the primary paper to support their own work. This latter group of citations raises the erratum citation question in a direct way. Even if the erratum did not have any direct bearing on the portion of the primary paper that was drawn upon, citing the erratum is still significant in that it indicates that the secondary authors are aware of its existence and took it into account (if necessary) in preparing their paper. Furthermore, a reader of the secondary paper who is inclined to investigate the topic more thoroughly can be misled if unaware of the existence of the erratum.

Returning to the citations "in passing," there are typically two motivations for providing such a citation. First, one may wish to pay tribute to predecessors in a particular field. Second, one may wish to direct the reader to papers with relevant background information. Papers cited for the second reason also should have their corresponding errata cited as a service to the reader.

We now consider the second question: Will a reader discover an erratum based on information provided by the authors of a secondary paper? Obviously, if the authors have cited the erratum, the answer is yes. If the authors have not cited the erratum, then there are a number of ways in which the reader may discover the erratum. For instance, the reader may look up the primary paper electronically and discover a link to the erratum. This is constrained by the fact that not everyone has access to journals in electronic form and not all journals are available in this format. When using journals in printed format, the reader must rely on techniques such as searching the journal index for errata or using a more comprehensive index such as the *Science Citation Index*. Otherwise, the erratum might only be discovered by chance while browsing through an issue.

Perhaps authors of the secondary papers assume that the interested reader will be able locate errata on their own. Short of taking a survey, we can only speculate as to whether this is the rationale for authors not citing errata. However, given the fact that this citation problem predates the electronic journal format, it is unlikely that most authors are consciously electing not to cite an erratum on these grounds. It is possible, however, that this rationale may explain the *drop* in the erratum citation rate between the 1995 study and the present study.

This brings us to our final question: Who is responsible for locating the erratum? It is reasonable to view a reference to a paper as a recommendation of a source to consult for further information. In making that recommendation, an author thus has some responsibility to ensure that it is a sound recommendation. However, a reader of a secondary source who is making an in depth study that requires consulting cited references also bears some responsibility for seeking out relevant errata. While it is difficult to say who has the greater responsibility, neither side can be removed from the equation.

It is worth noting that the secondary author is somewhat more likely to be aware of the erratum than the reader of the secondary paper, because often one cites papers written by people with whom one has had some direct or indirect association or by people whose work one has followed closely. This correlation of course is particularly true in the case of a secondary author also being a primary author. This observation coupled with the fact that erratum citation is not routine even when there is an overlap between primary and secondary authors leads us to speculate that secondary authors are not always citing errata even when they are aware of their existence. Why is this the case? One possible argument is that some perceive there is a stigma associated with publishing an erratum and hence they prefer not to call attention to it. Arguably, however, publishing an erratum is a sign of both integrity and attention to detail. It is likely most physicists who have done any significant amount of research have encountered papers that should have had errata but the authors chose not to write one. Clearly there is more damage to the field by uncorrected papers than by those properly corrected. The irony is that if one takes the time to do the right thing—to write the erratum—it is not clear how many people are going to read it.

Conclusions

We conclude as the previous study did with the hope that eventually the conversion of printed journals into electronic databases will resolve the erratum citation problem. In particular, if we reach a point where all journals are in a dynamic electronic database that is updated with appropriate links as errata are written and electronic access is as pervasive as printed access, then it becomes unnecessary to cite errata. While many physics journals are headed in this direction, it is not clear if and when all

will get there. Particularly problematic is the task of going through older journals and converting them to electronic format. In the meantime, citing errata will continue to be an important part of the service provided by authors in their reference sections.

Even if the erratum citation problem is resolved, the fact that it has existed raises more general questions concerning the integrity of the publication record. Specifically, is the accepted norm that authors *do* have a responsibility to cite errata or is the expectation that the reader is responsible for locating them? More generally, is this problem a sign of pervasive sloppy practices in publication or is it merely a situation of ill-defined responsibility? The answers to these questions will become clearer only after more discussion within the scientific community.

Acknowledgments

It is a pleasure to acknowledge support from the Graduate School, the College of Arts and Sciences, and the Department of Physics and Astronomy, all at Eastern Michigan University.

Bibliography

1. See, for instance, LaFollette MC. Stealing into print: fraud, plagiarism, and misconduct in scientific publishing. Berkeley (CA): University of California Press; 1992.
2. Pfeifer MP, Snodgrass GL. The continued use of retracted, invalid scientific literature. Journal of the American Medical Association 1990 March 9; 263(10): 1420-1423.
3. Thomsen M, Resnik D. The effectiveness of the erratum in avoiding error propagation in physics. Science and Engineering Ethics 1995; 1(3):231-240.

Paper Number	Paper Reference	Erratum Reference
1	**74**:694	**75**:355
2	**74**:1839	**76**:4097
3	**74**:4101	**76**:4293
4	**75**:1447	**75**:3781
5	**75**:394	**75**:1874
6	**75**:3549	**77**:2345
7	**75**:4413	**76**:3242
8	**76**:014	**76**:2826
9	**76**:1031	**77**:4278
10	**76**:2848	**77**:5148
11	**76**:3955	**78**:3227
12	**77**:127	**78**:3587
13	**77**:3865	**78**:1396
14	**77**:4066	**78**:162

The table above provides references to the papers used in this study. All are from Physical Review Letters, *published by the American Physical Society. The format is volume:beginning page.*

10. Theory and Models from other Disciplines

An Epistemic Model for Moral Hazards in Scientific Enterprises

Jean Maria Arrigo, Irvine, CA, USA

Maj-Britt Juhl Poulsen, Department of Medical Philosophy and Clinical Theory, University of Copenhagen, Denmark

KEY WORDS ***Adversarial epistemology, Biomedical research, Cooperative epistemology, Dynamics of competition and cooperation, Epistemic model, Ethics of science, Moral hazards in scientific enterprises, Political and military intelligence***

The immediate connection between successful science and ethical science is weak, so any urgency for successes may invite ethical lapses. We present a model of the dynamics between methods and morals in scientific enterprises. The developmental course of scientific enterprises generates characteristic moral hazards and opportunities, as we exhibit in our case study of a collaboration between two biomedical research teams. Lastly, we argue that our model offers conceptual gains in unifying "ethics of research" and "ethics of application" (1, p. 503) and offers practical gains in guiding codes of science ethics.

Interviews with biomedical researchers (2) and with military intelligence professionals, together with archived oral histories of weapons researchers, underlie our model (3). Reviews by military intelligence interviewees improved it iteratively.

A Model of Methods and Moral Hazards in Scientific Enterprises

In our model, the 17th Century Enlightenment vision of science constitutes the prototype of a *cooperative epistemology* (theory of knowledge). Epistemic partners freely share targets of inquiry, observations, and analyses. Empirical inquiry generates layers of knowledge through: (a) observation of a phenomenon, (b) analysis of observations, (c) meta-analysis of analyses, and so on.

Political and military intelligence, in contrast, constitutes the prototype of an *adversarial epistemology*. Epistemic adversaries may conceal targets of inquiry, observations, and analyses, and may spy on or sabotage the Adversary's inquiries. Empirical inquiry by agent and adversary generate interleaved layers of knowledge through: (a_1) Agent's investigation of a phenomenon (e. g., Manhattan Project study of nuclear fission), (a_2) Adversary's possible investigation of the phenomenon (e. g., possible Japanese study of nuclear fission in World War II), (b_1) Agent's investigation of Adversary's possible investigation through espionage, (b_2) Adversary's investigation of Agent's possible espionage, and so on. Each investigation by Agent or Adversary includes all the processes of observation and analysis in cooperative investigation above—and is often performed by an epistemic subcontractor, such as a scientist or historian, with cooperative methods. The

Corresponding author: Maj-Britt Juhl Poulsen, Department of Medical Philosophy and Clinical Theory, Build. 22.3, University of Copenhagen, Panum Institute, Blegdamsvej 3, DK-2200 Copenhagen N, Denmark, +45 3532 7596 (voice), +45 3532 7938 (fax), m.poulsen@pubhealth.ku.dk; for reprints: Jean Maria Arrigo,110 Oxford St., Irvine, CA 92612, jmarrigo@pacbell.net .

	Adversarial Epistemology		Cooperative Epistemology
I.	Partisanship: the goal of inquiry is conscious, strategic advantage over an Adversary.	I'.	Impartiality: the goal of inquiry is knowledge *per se* or its nonpartisan utility.
II.	Deceptiveness of phenomena: all observations are vulnerable to deliberate deception by the Adversary, whether by omission or commission.	II'.	Accessibility of phenomena: the natural world is not inherently deceptive (René Descartes' premise).
III.	Urgency: the Adversary is dangerous and implacable so decision is urgent in the short run.	III'.	Deliberation: method leads to superior results in the long run (Charles Peirce's "self-corrective" hypothesis).
IV.	Subordination: researchers' clients govern the broad topics, opportunities, and constraints of inquiry.	IV'.	Independence: researchers themselves govern the topics and methods of inquiry.

Table 1. Poles of the epistemic continuum

adversarial epistemology is thus far more complex than the *cooperative epistemology*. The complexity encourages intuitive or artful approaches in attributing intentions and meanings to the adversary's behavior. Our formulation contrasts the rational basis of military and political intelligence with the rational basis of science.

A recent headline story of scientific misconduct illustrates the interleaving layers of adversarial investigation. In 1981 the archeologist Shinichi Fujimura (Agent) unearthed the oldest artifacts in Japan, 40,000 years old. Some critics (Adversary) became skeptical of his celebrated discoveries, which by 1993 had pushed the origins of Japanese civilization back 700,000 years. Mindful of their suspicions, Fujimura surreptitiously buried artifacts that he later "discovered" in the presence of witnesses. Journalists documented with hidden video cameras his burials of artifacts. Anticipating Fujimura's defenders, the journalists filmed Fujimura's fraud at a second site before exposing him. Aware of the limitations of the journalists' investigations, Fujimura denied planting artifacts at sites he had excavated previously. Japan's Cultural Affairs Agency, now doubting Fujimura, plans reviews of his earlier excavations. Critics speculate that Fujimura's subterfuge may have set back Japanese archeology a decade (4).

The Epistemic Continuum

The adversarial and cooperative prototypes stand as opposite poles on a continuum of epistemic commitments. Cosmology and plant taxonomy lie towards the cooperative pole. Biological warfare research and forensic psychiatry lie towards the adversarial pole. Biometrics, clinical trials, educational testing, and research in the social constructionist paradigm occupy intermediate positions.

Four principles separate the most extreme positions of the *adversarial epistemology* from the corresponding principles of the most extreme positions of the *cooperative epistemology*, as stated in Table 1 (5).

In the *adversarial epistemology*, deception by the adversary leads to secrecy, compartmentalization of knowledge, reward of researchers on the basis of loyalty as well as ability, and organizational structures that limit the scope of inquiry. Repeated use of any research technique or conceptual schema offers the adversary an opportunity for sabotage, which raises the value of innovation over perfection. Urgency creates trade-offs between accuracy and utility. Fear of surprises from the adversary promotes crude study of broad fields in preference to fine study of narrow fields. Researchers' subordination to decision makers creates a distinction between the complex pattern "knowledge" held by researchers and the simplistic linear "information" provided to clients for decision making.

Consideration of typical epistemic adversaries in science-related enterprises suggests the pervasiveness of adversarial epistemic methods, as indicated in Table 2.

A Case Study of Competition and Cooperation among Biomedical Teams

Biomedical research can be described as collective research, for cooperation among individuals is necessary to reach the goals of research. At the same time, biomedical researchers need credit for their work and discoveries to make a career and to bring their own research programs to fruition. In this way the interplay of cooperation and competition is an

Domains of Inquiry	Common Epistemic Adversaries of Researchers	Historical Prototypes
Basic sciences	Colleagues, rivals, proponents of conflicting paradigms, ethics committees, institutional authorities, peer reviewers, funding agencies	Watson & Crick, *The Double Helix*
Medical sciences	Institutional Review Boards, regulatory agencies, Health Maintenance Plans, alternative healthcare practitioners, patients, animal rights advocates, hospital administrations, malpractice attorneys, news media	Tuskegee syphilis study
Social sciences	Cultural or identity groups, privileged economic and social classes, legislators, courts, admissions and ethics committees, hate groups	*Brown v. Board of Education* (school desegregation)
Industrial research	Industrial competitors, whistleblowers, labor unions, customers, consumer advocates, regulatory agencies, environmentalists	Tobacco industry cancer research
Weapons research	Enemy states, allied states, spies, terrorists, news media, social activists	Manhattan Project

Table 2. Common epistemic adversaries

essential element of daily practice. These internal dynamics may swing a project between the cooperative and adversarial modes. Therefore, during the course of scientific enterprises researchers may face new or unexpected moral challenges.

Identifiable conditions make a biomedical project swing either in the competitive direction or the cooperative direction. Adversarial conditions, for example, include colleagues' overlapping goals and/or methods; proximity to project completion and therefore to allocation of credits; limited resources; and, at a personal level, researchers' hidden agendas and breach of confidence. Cooperative conditions include (also) colleagues' overlapping goals and/or methods of project; complementary skills or resources; the need for face-to-face meetings; and, at a personal level, friendship and trust.

Our case study of competition and cooperation among three biomedical research teams illustrates the natural fluctuation between the adversarial and cooperative poles. A virology research team sent a young researcher to study abroad for a year with a cellular biology team, because of its expertise with a certain protein. His hosts encouraged the visiting virologist to present his research results at a conference. They also urged him to publish the "hot" results quickly, but he delayed. The biology team subsequently confirmed the results experimentally and proposed to take a subordinate role in a joint publication with the virologists. However, the virologists wished to publish alone first. Meanwhile, a second cellular biology group contacted the first and implied that they (also) had experimentally confirmed the results announced by the visiting virologist at the conference, and this second group of biologists intended to publish independently right away. The first biology group communicated this turn of events to the virology group, which began writing immediately.

In this narrow case we identify conditions that support cooperative practices, such as complementary skills, and conditions that support adversarial practices, such as allocation of credit at publication. The cooperation between the first cellular biology team and the virology team was established due to complementary expertise. The swing towards the cooperative pole was enhanced by a longer stay of one virologist with the cellular biology team. The desire to obtain credit made the project swing towards the adversarial pole and was enlarged by a hidden agenda on the part of the virology team, who wished to publish alone first. Competition from another cellular biology team made the project swing back towards the cooperative pole. Indeed, the methods and norms of science drive research projects along a typical trajectory of moral hazards and opportunities.

As this and other studies show (e.g., 6) biomedical projects can be seen as intermediary between the adversarial and cooperative poles. In particular situations, it can be very difficult to distinguish adversarial from cooperative epistemic orientations. The model provides a point of reference by stating that the key difference between adversarial and cooperative epistemologies is deliberate deception.

Utility of the Model

The epistemic model offers both conceptual and practical gains to science ethics. Conceptually, the model serves as a unifying schema for issues in science ethics. Two classes of scientific misconduct are commonly distinguished. The "ethics of research" is largely concerned with the means of competition among researchers, such as strategic secrecy. The "ethics of application" is concerned with the means used to attain scientific and technological ends, such as creation of toxic wastes (1, p. 503). These two classes are distinguished by the types of harm produced. The epistemic continuum accommodates the ethics of research and the ethics of application in a single schema. The harms change, but the adversarial epistemic principles that lead to the harms remain the same! Deception of colleagues in recording data and deception of research subjects in promising medical cures both follow the same adversarial epistemic principle of deception of the adversary, although the corruption of science and the injury to persons are ontologically different harms. The epistemic model identifies misconduct in science according to the principles of adversarial inquiry employed in the misconduct rather than the nature of the harm.

Further, the model guides study of the interaction between cooperative and adversarial epistemic methods. Cooperative epistemic methods lead to specialization, perfection of methods, and accountability in applications. Adversarial epistemic methods lead to expansion of domains, innovation in methods, and speed of application. To what extent are adversarial methods actually separable from cooperative methods in scientific projects? What are the costs and benefits of eliminating adversarial methods? How can beneficial and destructive competition be characterized?

As a practical contribution to science ethics codes, the model translates ethical problems in science—which philosophy of science cannot directly address—into products of a competing epistemology—which philosophy of science is better equipped to address. For typical research projects, epistemic adversaries and collaborators can be specified across the stages of the project, and typical moral risks and opportunities can be assessed.

The model highlights what we call *the tracking problem*: the original moral rationale for a project may cease to apply as the project evolves. For an example from ethics of application, the Manhattan Project authorized a metabolic plutonium experiment on unwitting, terminal patients, to gauge effects of plutonium exposure on bomb production workers. In 1944 many people would have agreed that the national security interest morally superseded the rights of patients, who were expected to die before the plutonium affected them adversely. But some of the patients survived for decades and suffered severe damages from plutonium injections, which invalidated the original moral rationale. For an example of the tracking problem from ethics of research, in our case study of three biomedical research teams, the rationale for the project appeared to change during the course of the project. At first the advancement of knowledge was the ultimate objective, which includes the obligation to publish results as soon as possible. This objective was superseded in later stages by the objective to obtain credit for discovery. A key ethical requirement for a scientific project would be to show how the original moral rationales, if needed, track along with the anticipated course of the project.

The fluctuation between cooperative and adversarial modes addresses the limitations of front-end solutions to moral problems in science, such as voluntary informed consent of subjects and authorship agreements. As a further contribution to science ethics codes, the epistemic model invites consideration of the most effective points of intervention for ethical codes. The model also suggests addressing potentially adversarial roles with support for the weaker party instead of only admonitions to the stronger. For example, to moderate the potentially adversarial roles of researcher and graduate student assistant, ethical codes might standardize support for the student in the form of a mentor at another institution.

Philosopher Henry Sidgwick, who laid the foundations for 20th Century ethics, considered whether society would be more improved by correction of character flaws, so as to gain the capacity to follow our moral convictions, or by moral understanding, so as to gain insight into the consequences of our actions. Sidgwick (7) advocated education of moral understanding on the grounds that strong character coupled with conviction leads to the most serious moral offenses. Historically, this has been the danger

for science. The epistemic model for scientific misconduct follows Sidgwick in offering moral understanding for science ethics education.

Acknowledgments

Arrigo acknowledges Harold William Rood for consultations on military intelligence and Kurt Smith for pointing out the symbiosis between cooperative and adversarial epistemologies. Poulsen was supported by the Danish Social Science Research Council which she gratefully acknowledges. Furthermore Poulsen extends her thanks to the scientists who gave their time and attention to her interviews. Jean Maria Arrigo is largely responsible for the epistemic model. Maj-Britt Juhl Poulsen is solely responsible for the case study.

Bibliography

1. Forge J. Merton revisited.... . Metascience 1999; 8(3):503-504.
2. Poulsen MJ. Competition and cooperation: what roles in scientific dynamics? International Journal of Technology Management. In press. To appear in Vol.22, No.8, 2001.
3. Arrigo JM. Sins and salvations in clandestine scientific research: A social psychological and epistemological inquiry. Claremont Graduate University; 1999. Dissertation.
4. Magnier M. Japan Amazed as Archeologist's Magic Exposed as Sleight of Hand. Los Angeles Times 2000 November 9; Sect. A i 1, 13, 14.
5. Arrigo JM. The ethics of weapons research: A framework for moral discourse between insiders and outsiders. Journal of Power and Ethics [On-line journal] 1999; *1*(4). Available WEB: http://spaef.com/JPE_PUB/v1n4.html
6. Atkinson P, Batchelor C, Parsons E. Trajectories of collaboration and competition in a medical discovery. Science, Technology, & Human Values 1998; 23:259-84.
7. Sidgwick H. The morality of strife. In: Practical Ethics: A collection of addresses and essays. London: Swan Sonnenschein; 1909. p. 83-112.

Scientific Misconduct as Organizational Deviance

Robert Dingwall, Institute for the Study of Genetics, Biorisks and Society, University of Nottingham, UK

Keywords: Human subjects; Organizational deviance

Although, as Steneck points out in his background report for this meeting, scientific misconduct is usually understood to involve "fabrication, falsification and plagiarism in proposing, conducting or reporting the results of research", human subjects protection cannot be excluded from this agenda. There are two reasons for this. First, it may be argued that research misconduct is in itself a form of human subjects abuse, since people have taken part in procedures that break the contract between researcher and participants by not making a valid contribution to scientific knowledge. Second, as Steneck also notes, integrity is a "measure of the degree to which researchers adhere to the rules or laws, regulations, guidelines and commonly accepted professional codes and norms of their respective research areas." To the extent that human subjects protection is the objective of much of this regulatory framework, we may argue both that researchers who compromise on the truthfulness of their reporting may be more likely to commit other abuses and that the success or failure of strategies for human subjects protection may offer relevant lessons for strategies to limit misconduct.

The death of Jesse Gelsinger in the course of a gene therapy trial at the University of Pennsylvania Institute for Human Gene Therapy (IHGT) in September 1999 has cast a long shadow over the adequacy of the regulatory framework in this area of medical science. It has led to significant restructuring of IHGT, has been used to justify changes in Federal regulatory structures and has provoked a bout of intense internal and external scrutiny of practice in clinical trials throughout the international community. While the narrative of events at IHGT is now reasonably well-established, there is still much to be understood about the reasons for the regulatory breaches brought to light by the subsequent investigations, particularly given the lack of evidence for any causal relationship between these and Gelsinger's death. How significant are the breaches identified? If they are relatively insignificant, have the correct regulatory conclusions been drawn? Will the changes proposed or introduced through the spring and summer of 2000 actually make trials safer, as opposed to satisfying public and political demands that "something be done?"

Traditionally, failures of the kind represented by the Gelsinger case have led to a search for blameworthy individuals, whose errors or omissions produced the negative consequences that have given rise to public scandal. The conventional response has been to call for individual sanctions and a strengthening of regulations or their enforcement. However, social scientists have become increasingly critical of this approach, arguing that organizational failures or misconduct are nowadays rarely the result of individual negligence or deceit. More typically, these failures arise as the unintended consequences of personnel carrying out their routine work under conditions of

---

Corresponding author: Robert Dingwall, Institute for the Study of Genetics, Biorisks and Society, Law and Social Science Building, University of Nottingham, University Park, Nottingham NG7 2RD, UK, +44-0115 951-5418 (voice), +44-0115 846-6349 (fax), robert.dingwall@nottingham.ac.uk.

organizational or environmental complexity that fail to give them appropriate feedback on the implications or results. Policy responses that increase complexity may actually further obstruct feedback, or introduce new opportunities for unpredictable system interactions to occur, rather than eliminating those that proved troublesome in the past. This argument, originating with the work of Charles Perrow (1) in the US and Barry Turner (2, 3) in the UK, has been developed over recent years by Diane Vaughan (4, 5) in her studies of the 1977 Ohio Revco Medicaid fraud and the Challenger space shuttle disaster. In the latter, for example, Vaughan shows how the social structure of NASA and its contractors, and the dispersion of information about problems with the O ring seal, allowed correct engineering reasoning to produce advice to launch that had devastating consequences. For present purposes, however, the Revco study may be a more useful model with its deliberate attempt to merge the understandings of social scientists who have studied organizations, regulatory bodies, and white collar crime. How do "respectable folks" end up in situations where they breach regulations intended to keep them honest? Why do organizations fail to prevent this?

This paper falls into three parts. The first briefly reconstructs the Gelsinger case from published sources available over the Internet. (It is not claimed that this is an exhaustive account, given the time and resources available.) Some of the main ideas put forward by Vaughan are then introduced, as a way of thinking about the kind of issues represented by this incident. Finally, these ideas are used to look at the Gelsinger narrative, with some reference to a brief period of participant observation in a British university's genetic science laboratories during summer 2000.

Gene Therapy at the IHGT

According to an official Food and Drug Administration (FDA) version (6), although gene therapy is an attractive idea, it has been slow to fulfil its theoretical promise. It has proved difficult to package correctly-functioning versions of disease-related genes in a way that allows them both to be delivered into the appropriate cells of a patient and to switch on. US researchers have generally looked to modified adenoviruses as the delivery vehicles, although UK researchers have been more attracted by lipids. The general principles have been known since the 1970's, giving rise to public concern about the possible implications of the release of genetically engineered material. In the US, National Institutes of Health (NIH) established the Recombinant Advisory Committee (RAC) to oversee development. However, RAC's formal powers were limited, and unlicensed experimentation took place as long ago as 1980, although the clinician involved was heavily censured. The first FDA approved trial began in September 1990, to treat an inherited immune disorder, and more than 400 trials are known to have taken place, worldwide, during that decade. However, clinical benefit has been hard to demonstrate. In 1995, Harold Varmus, then Director of NIH, created an ad hoc committee to review NIH investment in a field that seemed to have so much potential and to be realizing so little of it. This committee reviewed more than 100 approved protocols but its report to the RAC meeting in December 1995 underlined the lack of progress and the fundamental scientific problems that remained unsolved.

Coincidentally, the IHGT trial was approved at the same RAC meeting. The trail was intended to investigate possible treatment for a condition known as ornithine transcarboxylase deficiency (OTCD). This condition arises when a baby inherits a broken gene that is needed for the liver to produce an enzyme that breaks down ammonia. The IHGT researchers wanted to package this gene with a modified adenovirus and inject it into the hepatic artery to get the most direct delivery to the liver. Although there were some anxieties expressed about this delivery route, both RAC and FDA eventually agreed to approve the trial. In 1999, Jesse Gelsinger was the eighteenth and final patient to be recruited. Gelsinger was eighteen years old and in good health at the time but could not be described as a healthy teenager. He had a long history of OTCD problems, which had finally been brought under some control by a combination of medications and a highly restricted diet. He received the experimental treatment in September 1999 and died four days later, apparently from an overwhelming immune response to the carrier virus.

The subsequent FDA investigation found a series of regulatory breaches committed by the IHGT (7). Gelsinger had been entered into the trial as a substitute for another volunteer, although his high ammonia levels at the time of treatment should have led to his exclusion. IHGT

had failed to report serious side effects experienced by two previous patients in the trial, and the deaths of two monkeys given similar treatment had not been mentioned to Gelsinger or his father at the time informed consent was obtained. FDA shut down the OTCD trial immediately. FDA Form 483 issued to Dr. James Wilson, IHGT Director, on January 19, 2000, listed a number of concerns, which were summarized in a letter from FDA dated January 21, 2000, as failing to ensure the following:

> conduct of the study in accordance with the clinical protocols that are contained in the IND; obtaining adequate informed consent from subjects prior to participation in a study of an investigational agent or performance of invasive procedures; compliance with reporting protocol changes and adverse events to the responsible IRB; filing of safety reports as outlined in 21 CFR 312.64; and maintenance of complete and accurate records (8).

This letter suspended authorization for all IHGT clinical trials. A nationwide review of other approved trials revealed a high level of under-reporting of serious adverse events and possibly associated deaths. General shortcomings included: eroded adherence to requirements or standards of informed consent; lack of investigator adherence to good clinical practices and current Federal requirements; lack of adequate quality control and quality assurance programs for the gene therapy products used in trials; weak IRB processes; financial conflicts of interest; lack of public access to safety and efficacy data; limited regulatory enforcement options for Federal authorities; inadequate resources for enforcement; scope for improved co-ordination between FDA, NIH and OPRR; and poor understanding by investigators of FDA and NIH roles in gene therapy oversight. Several other trials were suspended for regulatory breaches or because of technical similarities to the OTCD trial. Other funders also suspended trials for review (9).

In March 2000, FDA and NIH launched a Gene Therapy Trial Monitoring Plan, increasing reporting requirements and requiring researchers to communicate more with each other about safety issues. In May 2000, President Clinton announced plans for legislation to allow FDA to impose civil penalties on researchers and institutions for regulatory violations. In June 2000, the NIH Office for Protection from Research Risks (OPRR), established in 1972, was reconstituted as the Office for Human Research Protections (OHRP), as advised by an NIH review submitted in 1999 before the Gelsinger incident. At the same time, the newly constituted OHRP was given expanded authority and relocated in the Office of the Assistant Secretary for Health in the Department of Health and Human Services (DHHS), placing it closer to the line of direct political authority. The overall response was summarized in evidence to a US Senate Subcommittee on May 25, 2000, under five headings: education and training; informed consent; improved monitoring; conflicts of interest; and civil money penalties. All clinical investigators receiving NIH funds would have to show that they had received appropriate training in research bioethics and human subjects protection, as would Institutional Review Board (IRB) members in their institutions. Audits of informed consent records would be performed and IRBs would be required to monitor informed consent elicitation more closely. Informed consent would have to be re-confirmed after any significant trial event. A wider range of Clinical Trial Monitoring Plans would have to be reviewed by both NIH and local IRBs. Conflict of interest rules for investigators would be reviewed to ensure that research subjects and findings were not manipulated for commercial gain. Finally, as mentioned earlier, legislation would be proposed to allow FDA to levy civil fines for regulatory breaches (9,10).

Meanwhile, IHGT and the University of Pennsylvania had initiated their own actions. IHGT filed a response to FDA Form 483 on February 14, 2000. In contrast to the FDA version, IHGT noted that it had promptly informed FDA, RAC, and the relevant IRB of Jesse Gelsinger's condition and that, in contrast to the FDA version above, IHGT had taken the initiative in suspending the trial. Moreover, IHGT could demonstrate that every trial participant had given informed consent and their eligibility for participation was fully documented. There had been delays of 3-4 months in submitting toxicity information on some early participants, which should have been discussed with FDA before proceeding with the next cohort. Nevertheless, FDA had these reports in its possession for more than six months prior to August 1999 when it approved the trial's continuation for the cohort that included Jesse Gelsinger. IHGT had Standard Operating

Procedures that met the regulatory requirements in force. The study in which two primates had died was unrelated, using different genetic material to treat a different disease. One primate had shown a mild reaction to a viral vector from the same generation but at a much higher dose–seventeen times higher–than in the OTCD trial. Available evidence did not establish any causal link between Gelsinger's plasma ammonia level prior to the infusion and his death (11). FDA reacted critically to the IHGT response. In a Warning Letter on March 3, 2000, there was a parallel exchange over the non-clinical laboratories at IHGT (12).

The University President set up an independent external panel to review IHGT. The panel reported on April 27, 2000 (13). The panel noted the discrepancies between the FDA Form 483 and the IHGT response but disclaimed sufficient regulatory expertise to comment. The panel focused on the operations of IHGT, noting its commitment to good practice and any necessary revision of operating procedures. IHGT had already contracted out the monitoring of its trials to an independent organization. However, the panel noted the growing costs of compliance and the need for the university to invest more resources in this area. The panel made the following recommendations. The university needed better internal monitoring and lower workloads for each of its IRBs. Bioethicists should cease to be involved in operational decision-making but act as consultants to investigators who would be responsible for their own actions. Conflict of interest policies should be reviewed. There should be closer scrutiny of informed consent procedures to ensure compliance with the letter as well as the spirit of FDA regulations. The panel also questioned the lack of continuing review for university institutes, the wisdom of concentrating all gene therapy work in one organization, the training of young clinical investigators in the special issues of investigational drugs, and the desirability of the university itself being simultaneously involved in the production of vectors, research, and the monitoring of standards. The President's response was delivered on May 24, 2000 (14). She announced a new assessment of all clinical trials by the University's Office of Regulatory Affairs (ORA). Where regulatory affairs professionals were not already involved, as in trials sponsored by pharmaceutical companies, the ORA would monitor the trials themselves or recruit external consultants to do so. The IHGT vision of a combined unit for basic, pre-clinical, and clinical work in gene therapy would be abandoned. The Center for Bioethics would become a free-standing department. IRB procedures would be strengthened and given extra resources. Ultimately, principal investigators and research coordinators would require certification before being allowed even to submit protocols to the IRB. The University already restricted investigators from having financial stakes in companies sponsoring trials but would review and strengthen this restriction.

At the time of writing (October 2000), a number of loose ends remained, particularly the final determination of FDA's response to IHGT and University of Pennsylvania's actions and the nature of any new legislation. However, there is no doubt that the Gelsinger case has come to be seen as iconic of problems in the regulation of scientific research and of public and political mistrust of this process, not just in the US but also in the UK and other countries with advanced levels of science. The regulatory and institutional responses will be widely studied. How much faith should we place in them?

Understanding Organizational Misconduct

Over the last thirty years, researchers in the fields of law and society and of organizational studies have become increasingly sceptical about the effectiveness of regulatory interventions as incentives for corporate bodies to act in a lawful fashion. Vaughan has summed up the alternative as a view that organizational misconduct is produced by social structure:

> By social structure, I mean (1) the stable characteristics in American society that form the environment in which organizations conduct their business activities: sets of social relations, laws, norms, groups, institutions; and (2) the stable characteristics of organizations themselves: internal structure, processes, and the nature of transactions. (4, p. 54)

Vaughan elaborates on a model first suggested by Merton (15) that locates the incentives for deviant action in the tension between cultural goals of economic success and social structures that limit access to legitimate means for achieving these goals. Merton set out a range of possible responses, but the one that interests Vaughan is "innovation". This is the attempt to achieve the valued goals by expedient but

prohibited means, justified on the basis that the unequal access to legitimate means compromises the norms that distinguish legitimacy from illegitimacy. If this distinction is perceived to be arbitrary or discriminatory, then it may fail to command moral respect. In the context of science, for example, Barber and colleagues (16) showed that those most likely to cheat on the norms of the professional community were those who felt unjustly treated in their careers. Vaughan notes that Merton focused mainly on the impact of the tension between culturally valued goals and social structures for individuals in lower social classes. However, Vaughan argues that this approach is at least as well suited to the analysis of organizations, which may be more strongly driven than individuals by the requirements of profit-maximization but where competition undercuts the force of norms. The processes of change that are the dynamic of a market economy continually challenge the normative order of that economy. The formalization of norms into law has limited effectiveness. Legal responses to "innovation" occur after the event and are skewed by the extent to which both rules and their enforcement rest on negotiations between regulatory agencies and the firms they regulate (17).

As Vaughan points out, unlawful behavior cannot be explained solely in terms of these social structural tensions. Opportunities must arise that offer the possibility of unlawful acts and the regulatory environment must be such that there is a reasonable chance of escaping sanctions. Vaughan points to the processes, structures, and transactions of modern complex organizations as the sources of opportunity. As the literature on white-collar crime shows, these create the conditions for individuals to act illegitimately: her claim is that they also make organizational misconduct possible. Organizational processes create a moral and intellectual world for members, encouraging them to identify with the organization and its goals. The survival of one becomes linked to the survival of the other. Those most exposed to temptation are those in the subunits most relevant to the resource or profit-seeking goals, with information linking subunit performance to the achievement of those goals and some responsibility for that achievement. Their choices reflect their awareness of the organization's relative rewards for achievement and its sanctions for illegality and of the structural visibility of their actions. Complex organizations multiply opportunities for misconduct through their structural differentiation and task segregation.

The result is what Vaughan terms "authority leakage", the loss of capacity for internal control. The actions of subunits may become effectively invisible, particularly where they involve specialized knowledge that is not shared elsewhere in the organization. A rational process of internal censorship designed to match upward information flows to the processing capacity of senior managers, obscures misconduct, and diffuses personal responsibility. Finally, the nature of transactions both provides legitimate opportunities for illegitimate behavior, and further minimizes the risk of detection and sanctioning. Transactions between complex organizations have four distinguishing characteristics: formalization; complex processing and recording methods; reliance on trust; and general rather than specific monitoring procedures. Because of the difficulty of monitoring each individual transaction, organizations tend to rely on signals that can be manipulated to present an appearance of legitimacy to outside observers, whether transaction partners or regulators.

Vaughan discusses the particular example of Medicaid fraud where determinations of eligibility for participation tend to rest on data submitted by would-be service providers. The complexity of the government paperwork and the lack of resources for verification create conditions where willful misrepresentation can occur. This also indicates a problem of system interface, where the culture and structure of two organizations, in this case government bureaucracies and relatively small for-profit enterprises, conflict. If these cannot be brought into alignment, one or both organizations may choose unlawful actions as a means of achieving their goals. Vaughan notes how Revco executives felt justified in false billing the Ohio Welfare Department for an amount equal to the claims for payment that had been denied on what Revco felt to be excessively bureaucratic grounds. The Welfare Department wanted Revco to internalize a government agency culture that Revco found incompatible with a private, for-profit enterprise.

Regulating Science

Vaughan's analysis of the Revco case focuses on

the potential sources of misconduct in profit-seeking organizations, although she makes some suggestions about its possible relevance to other sorts of enterprise. Scientific research organizations have some peculiar features, and may vary somewhat according to whether they are in universities, not-for-profit corporations, or commercial companies. However, it is arguable that, whether or not scientists are overtly engaged in profit-seeking, the incentives that they face are functionally equivalent. Profit, as Vaughan notes, is merely the most obvious indicator of an organization's success in locating and securing resources for its operations and survival. Scientific work depends upon flows of grant and contract income which, in turn, depend upon the production of results which lead to further income flows. These may derive from patentable innovations or from peer esteem, which leads to publication in high-quality journals, professional networking opportunities and so on. For the individual scientist, personal rewards may be symbolic rather than material, but these virtual profits are converted into economic resources for the research organization (18). Science is reward-driven in the same way as other enterprises and, as elsewhere, a failure to win rewards leads to bankruptcy, whether personal or corporate. In the British university department that I studied, for example, laboratories began almost literally as shells, which faculty were expected to equip for both the capital and consumable needs of their research through their income-generating activities. A run of unsuccessful grant applications could lead to a downward spiral where the investigator simply ran out of resources. The department claimed to be unusual in having an internal taxation system that could provide some support for a member in this position, at least for a period, in the hope that their luck would turn. This was said to be unpopular with funders who would have preferred to see a purer market system with no socialization of resources.

If this leads us to accept that Vaughan's analysis could be broadly applicable, we also need to acknowledge that there may be some differences between scientific research organizations and other kinds of enterprise. The most important may be the way in which the problems of the reactive nature of regulation are accentuated by the defining characteristic of science, namely its engagement with uncertainty. Regulation is an institutionalized means of managing risk. It can work reasonably effectively in mature environments where risks are well-understood. In many engineering situations, for example, there is a recognizable cycle of risk and regulation. A new technology generates a number of accidents that lead to a definition of hazards and a regulatory response that produces a safe environment until the next significant change in technology comes along. Although there are also routines in scientific research, science is ultimately about pushing into the unknown and taking unknowable risks. A regulatory regime that prevented all risk would prevent all scientific innovation. However, to the extent that contemporary societies have a low tolerance for risk, there is an inherent tension for regulators between the demand that risk be averted and the functioning of the regulated enterprise at all. A level of regulation that stifles enterprise is not in the regulators' interest any more than a failure to regulate sufficiently that leads to legitimacy problems with the public or the political system. In any clinical trial, participants assume some measure of risk: regulators may do their best to manage this, but it cannot be eliminated because of the variability of human response and possible interactions with other idiosyncratic features of the participant's biology or environment. The question is whether participants are adequately informed about this risk and compensated for adverse outcomes. If the risks were eliminated, so would be the possibility of discovery. Regulators must always trail behind and the letter of regulation can never be more than a partial solution to the management of risk.

If the effectiveness of regulation is necessarily limited, we may need to look more closely at the social norms of research organizations and the structures in which they are embedded (19). The university department that I studied was a relatively compact physical group, where the principal investigators had offices in the corner of the laboratories in which their postdocs, research assistants, technicians, and graduate students worked. Laboratory work was highly visible to colleagues. There was also an active tradition of seminars, journal clubs, gathering for coffee and lunch breaks, and departmentally-based socializing. This facilitated the development of a departmental culture, although it did not prevent perceptible differences emerging in the climate of different faculty member's laboratories. Clinical trials,

however, as the Gelsinger documents clearly show, tend to have a much longer chain of command, which makes important parts of the process substantially invisible to principal investigators.

The scale and complexity of the clinical trial process has generated an increasingly intricate division of labor. At the top are the principal investigators (PIs), whose strategic vision and social networks are crucial to generating the flow of resources that keep the enterprise going. In the middle are the trial managers and coordinators who keep the process on track. Patients, however, actually have direct contact with much lower level people who obtain informed consent, administer the interventions, and collect the test data on the results. The "hired hand" problem has long been recognized by those social sciences that make extensive use of survey techniques (20). How do you guarantee that low-level workers doing rather mundane jobs do not simply make up data or ignore the code book when entering it? Computerized interview techniques have reduced the opportunities for misconduct, but it has historically been a considerable challenge to the management processes of survey organizations. It represents the same problem of authority leakage and internal censorship that Vaughan describes. Structural differentiation and task segregation make operational performance invisible to senior managers. Whatever performance or quality standards are set, managers are unable to follow them through. At the same time, information from lower-level personnel is censored as it rises to match the capacity of supervisors and managers to handle it.

Various solutions have been tried, two of which are worth further discussion here. One is more detailed organizational rule-making to try to govern lower-level personnel by command and control methods. The result of this is usually to reduce further commitment to organizational goals and to sacrifice the potential gains from a degree of flexibility at the point of operational activity. If we take the specific example of informed consent, this has become the subject of increasingly elaborated procedural rules. Consent may now be deemed to be informed only if it is in accordance with these rules, something that may account for the discrepancy in view between FDA and IHGT. FDA finds that the paperwork is not in order, while IHGT claims that, although not recognized by the FDA, adequate documentation for consent does exist. However, the elicitation of consent is also a difficult interactional task. How do you ask someone voluntarily to assume a risk that can be broadly described but is ultimately unknowable until after the event. Lower-level personnel charged with the execution of the task tend to deal with this by a measure of improvisation. They seek to comply with the spirit of the regulation rather than the letter.

The result is a degree of variance that is hard to reconcile with the command and control approach. Both the University of Pennsylvania and FDA seem to have responded by trying to toughen the regime. Indeed there are even proposals that IRB members should monitor the consent process by direct observation. The problem would seem to be that you could reduce the process to a script, force the consent-takers to read the script aloud to the patient by recording or observing them, as in call centers, and then discover either that hardly anyone is willing to volunteer, because the process has been made regulator-friendly rather than user friendly, or that consent is formal rather than substantive and that patients who experience adverse outcomes can still reasonably claim to have been deceived or not to have understood the nature, purpose, and risk/benefit ratio of the trial.

In effect, this reproduces the Revco problems of the organizational interface between a Federal regulatory bureaucracy and, in this case, the professional traditions of university science. Traditionally, universities have been federations, or even confederations, of professionals, with a high degree of internal autonomy and limited collective responsibility. Although this model has come under some pressures from demands for greater social accountability in recent years, these have been opposed by the encouragement of entrepreneurial science. The difficulties of raising student fee income to a level where salaries competitive with the general commercialization of professions (21-23) can be paid have been met by a shift in culture that allows those who can to top up their incomes with consultancy earnings and stakes in spin-off companies. Although academics may be able to raise their market price by trading on their university's reputation, they are simultaneously less constrained by the university's employment discipline, since their salary may be a relatively small proportion of their income. This poses a

considerable management problem for universities, since bureaucratization may cost them faculty whose presence is crucial to their general competitive position. The University of Pennsylvania, for example, proposes to introduce certification for PIs: if this is perceived as burdensome, the result may be that the university loses star talent to less intrusive competitors.

The result, as is evident from the FDA response to the Gelsinger events, is often a division of rules into those taken seriously and those on the book but disregarded unless something goes wrong and a source of sanctioning is required. There is a hierarchy of rules, some of which "really" matter and some of which are there for use only if needed. The IHGT/FDA clashes seem to suggest that something similar has happened. Having complied with what IHGT seems to have been led to understand were the "important" rules, it clearly feels aggrieved that the FDA inspection has produced an exhaustive list of breaches, arguably to cover the agency's own collusion in the procedures at the Institute. One might note particularly the counter-charge that FDA had been in possession of toxicity reports on earlier trial participants for six months without comment before approving the recruitment for the final cohort that included Gelsinger.

When bureaucratic command-and-control fails to defend the organization from regulatory pressures or liability suits, one response can be its replacement by a network of outsourced sub-contractors, as the University of Pennsylvania seems to envisage. PIs or research organizations lay off the risk by sub-contracting the work through contracts that specify performance and quality but locate the responsibility outside the core business. The difficulty with this model is that exhaustive performance contracts are essentially impossible to write and that further incentives for misconduct tend to be created. If a sub-contractor is required to deliver a certain number of patients and associated paperwork for a fixed price, they clearly have reason to see where corners can be cut. The PI sacrifices control over data quality and, to some extent, ethics in favor of protection from the professional or legal implications of failing to control either personally, provided that there are adequate risk-shifting clauses in the original contract. It is, however, probably naive to assume that such risk-shifting will be an effective defense, particularly given the tendency of US courts to look behind the letter of such contracts to the responsibility of those issuing them to audit the performance of contractors. The growing liability of hospitals for the acts of physicians afforded admitting privileges is an obvious parallel. The result is likely to be an organizational internalization of law, as the alternative to bureaucratization, with PIs required to attend to the compliance of the documentation of their work with the forms of private rather than public law (24). It is simply a different kind of interface problem.

Ultimately, there is probably no substitute for the more active engagement of PIs with their projects and methods of countering authority leakage and internal censorship. The paradox is that the enhanced systems of scrutiny, whether bureaucratic or legal, will tend to make this more difficult by enhancing the competing calls on this pool of senior investigators to participate in peer oversight of others. To the extent that their time is drawn into this system, by the sorts of measures that FDA envisions in terms of more frequent sharing of trial experiences or the expansion of IRB membership to spread workload and allow more intensive scrutiny of proposals, then the problem that internal censorship solves will grow worse. Internal censorship, remember, is the solution to the limited time and attention that senior organizational actors can give to any particular problem. If time becomes more restricted, then censorship will increase. The FDA's measures may mean that PIs become much better informed about other people's problems and less well informed about their own. Which is most likely to contribute to safer research for human subjects?

This is obviously a brief account of a complex story that is still some way from completion. It is also heavily reliant on the public record and would obviously benefit from interview data of the kind that Vaughan had access to in her work. However, it may serve to exemplify an approach to the study of scientific misconduct and, in particular, to illustrate some of the very real difficulties of imposing a strong external regulatory regime on practice. The issues of compliance that arose in the human subjects protection of Jesse Gelsinger are immediately parallel to those that arise in controlling falsification, fabrication, and plagiarism, which also are compromised by the structural and cultural problems that lead to

authority leakage and internal censorship. It is only by recognizing and engaging with these underlying problems that effective interventions can be designed.

Acknowledgements

Material on the Gelsinger case was assembled by Mark Allan. I am grateful to Diane Vaughan and Liz Sockett for their comments on an earlier draft.

Bibliography

1. Perrow C. Normal Accidents: Living with high risk technologies. New York: Basic Books; 1984.
2. Turner BA. Man-made disasters. London: Wykeham Publications; 1978.
3. Turner BA, Pidgeon NF. Man-made disasters. Second ed. Boston: Butterworth-Heinemann; 1997.
4. Vaughan D. Controlling unlawful organizational behavior: social structure and corporate misconduct. Chicago: The University of Chicago; 1983.
5. Vaughan D. The challenger launch decision: risky technology, culture and deviance at NASA. London: The University of Chicago Press; 1996.
6. Human Gene Therapy > Harsh Lessons, High Hopes. FDA Consumer Magazine 2000 Sep-Oct.
7. Statement of Jay P Siegel, MD, Director, Office of Therapeutics Research and Review, Center for Biologics Evaluation and Research, Food and Drug Administration, Department of Health and Human Services before the Subcommittee on Public Health, Committee on Health, Education, Labor and Pensions, United States Senate, February 2, 2000.
8. Letter from the FDA to the IHGT. University of Pennsylvania *Almanac* 2000 Jan 21; 46 (18).
9. Statement by William F Raub PhD, Deputy Assistant Secretary for Science Policy, US Department of Health and Human Services before the Subcommittee on Public Health, Committee on Health, Education, Labor and Pensions, United States Senate, May 25, 2000.
10. Statement of Greg Koski, PhD, MD, Director, Office for Human Research Protections, Office of the Secretary, Department of Health and Human Services for the Hearing on Human Subjects Protections in VA Medical Research before the Subcommittee on Oversight and Investigations, Committee on Veterans Affairs, US House of Representatives, September 28, 2000.
11. Summary of IHGT response to the FDA. University of Pennsylvania *Almanac* 2000 Feb 22; 46 (22).
12. Warning Letter from Steven A Masiello to James M Wilson, March 3, 2000. Warning Letter from Steven A Masiello to James M Wilson, July 3, 2000. [I have been unable to locate the university's response to the first of these or details of the Form 483 issued on or about 1 March 2000 and the university's response on May 4, 2000, which precede the second letter.]
13. Report of the Independent Panel Reviewing the University of Pennsylvania's Institute for Human Gene Therapy. University of Pennsylvania *Almanac* 2000 May 30; 46 (34).
14. Action by the University of Pennsylvania in Response to "Report of the Independent Panel Reviewing the Institute for Human Gene Therapy". University of Pennsylvania *Almanac* 2000 May 30; 46 (34).
15. Merton RK. Social Structure and Anomie. In: Merton RK, editor. Social Theory and Social Structure. New York: Free Press; 1968. p.185-214.
16. Barber B, Lally J, Makarushka JL, Sullivan D. Research on human subjects: problems of social control in medical experimentation. New Brunswick, NJ: Transaction Books; 1979.
17. Albiston C. The rule of law and the litigation process: the paradox of losing by winning. Law and Society Review 1999; 33 (4): 869-910.
18. Latour B, Woolgar S. Laboratory life: the social construction of scientific facts. London: Sage; 1979.
19. Etzioni A. Social norms: internalization, persuasion and history. Law and Society Review 2000; 34 (1): 157-178.
20. Roth JA. Hired hand research. American Sociologist 1966; 1: 190-196.
21. Hanlon G. Professionalism as enterprise: service class politics and the redefinition of professionalism. Sociology 1998; 32 (1): 43-63.
22. Hanlon G. The changing nature of professionalism and the fracturing of the service class. International Review of Sociology 1999; 9 (1): 87-100.
23. Hanlon G. The commercialisation of accountancy. Basingstoke: Macmillan, 1994.
24. Edelman LB, Suchman MC. When the 'haves' hold court: speculations on the organizational internalization of law. Law and Society Review 1999; 33 (4): 941-991.

A Market Approach to Research Integrity

Aditi Gowri, University of Texas, Austin, USA

Keywords: Buyer's market, Collective agency, Ethics, Market demand , Research integrity, Research funding

Introduction: Supply Side Ethics

The standard approach to research integrity can be characterized as supply side ethics because of its emphasis on individual persons who supply research products. These suppliers—researchers—are screened, educated, exhorted, given incentives, and as a last resort threatened in the interest of getting them to be honest, to keep careful, accurate records, to test their hypotheses fairly, to use reporting techniques fairly, and to train their subordinates and encourage their colleagues in turn to do likewise. This paper discusses the effects of the *demand for research* on the collective integrity of research fields.

Briefly, the argument is as follows: where research takes place within a market for research products, effective demand in this market will affect the distribution of knowledge produced. Therefore, the body of scientific knowledge will be skewed by demand. The analysis here suggests that this skew can result in a form of malignant bias resulting from demand. Paradoxically, this form of bias occurs in the absence of corrupt researchers, research designs or grantors. It is a form of bias nonetheless, since it can lead to misleading research-based knowledge and less than optimal policy decisions. Thus, it should be of concern to researchers in research ethics.

There is a Market for Research Products

In the market for research, scientists are the suppliers and those who offer to pay research expenses through grants or contractual funding are the "consumers" of research products. Any offer to fund research activities constitutes a demand for the kind of research that is expected to result from those activities. Principal Investigators can be seen as entrepreneurs who compete with their peers for contracts. Those who are successful in obtaining contracts become suppliers and then use the resources they have obtained to hire labour and buy raw materials—essential components in the production of research.

Objection: funding cannot produce science

It is normal to accept that there is a competitive market for research funding but to separate this conceptually from the idea of a market for research products. Knowledge, as usually understood, is not something that can be sold to order like a car or a bunch of carrots. On this standard model, scientists choose to pursue particular research questions because of their intrinsic value and expected fertility. Funding comes to those with skill who choose a fruitful line of inquiry, as a reward for past

Corresponding author: Aditi Gowri, LBJ School of Public Affairs, University of Texas, Austin, TX 78713-8925, 512-471-8245 (voice), 512-471-1835 (fax), gowri@mail.utexas.edu.

successes and to support the *promise* of future productivity. But there is no simple buyer-seller transaction. Any appraisal of research products themselves must be based on examination of the autonomous technical pursuit of the research craft: experimental design, data collection, record keeping and the interpretation of results. There is of course an acknowledgement that getting research funding is a competitive pursuit, but the funding transaction is seen as completely external to the generation of research products. In other words, there may be a market for funding but there is no market for knowledge. This conventional separation is unsatisfactory because funding affects the actual content of research products in at least three ways.

How does demand affect research knowledge?

First, some researchers will modify their research questions, design and methodology to receive funding. It is not difficult to think of colleagues who have changed their research questions or design slightly to obtain the interest of a funding agency. Indeed, at least in the social sciences, often the tail wags the dog, with research proposals and even research programs developed in response to offers to fund. Generally this is not considered to be dishonorable, provided the proposed studies are intrinsically legitimate and carried out fairly.

Second, researchers who really want to pursue research interests or designs that do not fit some effective demand for knowledge are like sellers in a market with no buyers. They can still conduct research, but only to the extent that personal funds or their general institutional budgets are adequate to support its costs. In general, smaller budgets will limit the scale and type of work they can do. Since there is always competition for scarce research dollars, grantors have the prerogative of declining proposals from researchers who do not offer what is demanded. In other words, the market for research is a buyer's market.

Third, if we look at a researcher's life cycle, the effect of demand is most strongly felt at the early stages of a career. Doctoral and post-doctoral researchers usually must serve as apprentices to a more senior researcher to begin to earn a living in the research trades. Because it is an apprenticeship phase, junior researchers are expected to develop a package of skills and competences that will then affect their approach to doing research over the course of a professional life. On the other hand, a senior researcher with better funding is likely to attract more and brighter young scholars than his less generously funded colleagues. Thus the demand for knowledge, operating through the demand for junior collaborators and research assistants, plays a part in developing the competences and commitments of each new generation of researchers. Demand not only has an immediate market effect but also a life cycle effect on the researcher's capacity for—and commitment to—future research projects. Again, it is hardly blameworthy for a junior researcher to consider the size of available fellowships before choosing to work in a particular sub-discipline or laboratory.

Demand calls forth its supply

In all of these ways, the economic demand for research will affect the supply of research products developed. To accept this conclusion we do not have to believe that the demand for research can produce its own supply (although this is the way an economist might put it), nor that research in the absence of funding is impossible. We must only accept that some researchers will respond to the incentives offered by granting agencies and that those who do so will be better situated to generate research than the rest. In other words, research flourishes in the presence of money, and generating research products without money is very difficult and rare. Grantor sovereignty certainly is not absolute; it is no more than a form of consumer sovereignty, resulting from the prerogative of buyers in any market to demand the products that give them most satisfaction.

Contrast the demand for corrupt research

It is worth emphasizing that the effect of demand on knowledge does not entail any individually discreditable conduct on the part of either buyer or supplier. A demand for corrupt research products probably exists. For instance, a grantor with a preferred ideology may put pressure on a researcher to design not quite fair tests of hypotheses, to address data selectively, or to misreport or over-generalize findings. Perhaps more subtly but no less deceitfully, a pharmaceutical company might commission more than one study of a drug, publicize only those favorable to its product, and bury the rest. Each of these is an example of corrupt(ing)

demand, but neither is our concern here. While the demand for corrupt research is certainly worthy of study, a discussion of its extent and effects does not lie within the scope of this paper. Throughout this discussion our concern is rather with the demand for legitimate, honest research products to be supplied by researchers whose integrity in conducting each separate research project is not under question. The problem raised here does not result from any individual wrongdoing but rather centers on a robustly collective effect of individually blameless acts (1, 2).

What is the Problem?

Those who accept the analysis so far will concede that the market for research funding affects the distribution of research products; however they may still deny that this is an ethical problem. For instance, if one subscribes to the "marketplace of ideas" model of truth (cf. 3, 4), then a free market for ideas, for their sponsorship and dissemination—such as has been described—is the most efficient system for allowing the truest views to emerge. As long as each seller and buyer of ideas is free to make her own choices for her own reasons, the invisible hand of the market will guarantee that the best (i.e., the most sought after) ideas flourish. If an area of research truly has merit, surely some clever grantor will see that there are returns to be obtained and enter the market. This model presupposes that within a free market for research funding, the best quality science will receive the best funding simply by virtue of its quality.

Two different rebuttals to a marketplace of ideas model are offered here; each based on an accepted standard for assessing the inherent quality of research products, independent of market demand. The first argument is democratic, while the second is elitist.

The democratic argument: knowledge is a public good

Although effective demand for research is exerted by grantors, research products do not serve only grantors. Knowledge is a *public good* in at least three different senses.

First, knowledge is public in the technical economist's sense: knowledge products are often non-excludable or offer positive externalities to people other than the purchaser. Research products are not only there to be used by a purchaser, they also become part of the common stock of knowledge. Research produced for one purpose will often have unexpected "external" benefits and uses. (Proprietary approaches to knowledge present only an apparent challenge to this argument, because they do not change the underlying quality of knowledge as public, they only change the way our legal systems sanction its use.)

Second, knowledge is public in a proprietary sense. That is, the public owns it by virtue of having paid for its production through taxes. Not only do public grant funds pay for much research directly, there are also many implicit forms of subsidy that enable scientific education and practice—the public school and university system being only one large example.

Third, knowledge is public in a normative sense. We pursue research as a calling—as something we do for our fellow humans—as much as for our own livelihood and reputation. The cobbler usually does not take up this trade so that the feet of the world may be shod, but researchers often are motivated by a desire to contribute to the progress of humankind's knowledge. Most of us believe that knowledge exists to serve society or humanity, not only for the "consumers" who pay for the production of research.

A free market of interactions between purchasers and suppliers of research (or any) products might perhaps optimize the satisfaction of direct parties to these transactions. However, the interests of the public are not directly represented in reaching this theoretical market equilibrium. A bias away from the public interest will result, to exactly the extent that research-demanding grantors and the broader research-using public have systematically different interests.

The elitist argument: good science is an autonomous pursuit

A body of scientific knowledge is not simply a collection of individual researchers' products. It is produced by a community of scientists. Individual researchers may have unconscious biases (5) and may certainly commit honest errors. These flaws can only be corrected from another's perspective. Thus the quality of scientific *knowledge* emerges from interaction among knowledge producers, not only from the quality of any one producer's activity. This self-correcting feature of scientific knowledge is

historically traced to the work of Herschel, Merton and Popper (6, 7, 8), but the motif of a self-correcting, autonomous body of science-producing experts is also implicit in Kuhn's classic account of progress through revolution and in post-Kuhnians such as Laudan (9, 10). If one subscribes to any such elitist model, the proper advance of scientific knowledge results from the intellectual judgements made by a community of qualified researchers, not from the economic demand for research. If aggregate demand for research does not correspond to the range of projects that researchers would choose to pursue on solely intellectual grounds, then to this extent, the body of knowledge being produced will exhibit a form of bias.

Why does collective bias matter?

Ultimately, the main reason we care about integrity of research at the individual level is that the intellectual adequacy of a body of research is vitiated by research corruption. Corrupt practices produce dubious, misleading results. From either a democratic or an elitist perspective, we should care about collective bias for exactly the same reason—because a body of research formed by demand may mislead researchers, students, the public at large, and policymakers. In any field based on multi-causal or probabilistic systems, the problem of collective bias resulting from the demand for research should be of particular concern.

Case: Causes of Disease

Sylvia Tesh remarked in 1988 that studies based on a contagion model of disease were best funded, most prestigious and generally dominant in American medical research (11). Today (in 2001) contagion has been joined or perhaps displaced by genetics as the dominant cause of disease to be researched. A third model underlying research studies is lifestyle theory, the idea that modifiable personal behaviors result in illness. All three of these causal models fall under an overarching individualistic framework, where disease is located within the person, whether in her genes, in a viral or bacterial agent she has taken in, or in her choice of (un)healthy behaviours. By contrast, environmental, economic and psycho-social causes of disease receive far less attention (and far less funding). Evidence from other First World countries suggests that these would be highly fruitful areas of inquiry. To take only one instance, the Whitehall studies in Great Britain showed that age-adjusted mortality from nearly all causes varied inversely and quite significantly with *civil service grade* even when controlled for individual health variables such as smoking. In other words, the higher the civil service grade, the less likely these civil servants were to get ill or to die, all other things being equal. Similar relationships between social status and biochemical health indicators have been found in experimental monkeys (12).

The nearly exclusive emphasis on one or two modes of causation is problematic because the others might equally and perhaps more cheaply lead to better public health. If prevention is intrinsically better than cure, then controlling large scale correlates of disease is better than using genetic or pharmaceutical technology to treat disease. To make this concrete: a breast cancer gene may be significantly correlated with breast cancer, but possibly not more so than poverty, radiation, or other environmental and economic factors. If the public and policymakers become aware of the first relationship but few researchers are pursuing the rest, a misplaced emphasis will be put on genetic therapy and too little effort on other possible methods for addressing this disease.

As long as there is a predominant demand for genetic research, we will continue to get genetic results. What is more, a disproportionate number of apprentice researchers will continue to be trained in the area of genetic medical research (not environmental or social medicine) and to develop a commitment to being *geneticists* rather than some other kind of health researcher. They in turn will have incentives to conduct and to support future medical research on a genetic model. Thus demand is not only affecting research in the present, it is also influencing the shape of the future research producing community.

Why is it an Integrity Problem?

If the analysis of the paper is accepted, then the demand for research poses some kind of social problem. Yet as an *ethical* problem it is paradoxical because we cannot find the wrongdoer. For this form of research corruption to arise, there need not be any demand for corrupt research nor any suppliers of research who are willing to be corrupted. No personal misconduct or violation of individual research autonomy needs to take place. There must only

be a situation where funding organizations freely select the type of research they will fund from among various projects and models being proposed. In other words, corruption of research due to the demand for research is a robustly collective problem; it is not a problem that can be resolved by making individual people behave more honestly or fairly. The reader may wonder, therefore, whether this is actually a problem of research integrity, or just some kind of market imperfection or political problem. The reply to this last objection lies in the professional status of researchers.

Research is a profession

Professionals are characterized by most ethicists as the bearers of many social privileges including a monopoly on legitimate practice within their domain, control of entry into that domain, and evaluation of one another's competence (13). Following this definition, scientific researchers are professionals. In exchange for their privileges, the members of a profession are collectively responsible for the character of their practice as a whole: they must ensure that it benefits a society as much as possible, and at least that it does no harm. If researchers are professionals then they are not only responsible for doing research honestly, they are also custodians of their realm of research. Collective responsibility of this kind has been accepted by traditional professions including medicine and law, and by many newer ones such as nursing, accounting and insurance (14). Of course researchers in a field may not be the *only* persons responsible for the collective integrity of that field.

What can be done?

In this paper I have called attention to a type of failure of research integrity that has not yet been addressed in research on research integrity. I do not pretend that it will be easy to address the problem of collective integrity in knowledge production: indeed, intrinsically collective problems tend to be philosophically and practically difficult (cf. 2). However, just because a problem is not easy to fix, this does not mean we should ignore it.

The existence argument for market effects on the integrity of research must be supplemented with research on the magnitude of these effects. Such empirical studies could document the effect of demand on research programs through historical and international comparisons, qualitative social studies of market effects on mentoring and career choice, or quasi-experimental studies of factors involved in research problem choice, for example. Finally, I do not expect it will be easy to fund research about collective market effects on research integrity, since funding agencies can hardly be expected to have an interest in demanding this kind of knowledge that would, after all, challenge their own role in directing the course of knowledge production. Such research would, however, offer valuable insight to the research professions and to the public.

Bibliography

1. Niebuhr R. Moral man and immoral society. New York: Scribner's Sons; 1932. Reprinted; 1960.
2. Gowri A. Towards a moral ecology: the relationship between collective and human agents. Social Epistemology 1997;1(1):73-95.
3. Wright JS. On courts and democracy. Miller A, ed. Greenwood Press; 1984.
4. Gowri A. Speech and spending: corporate political speech rights under the first amendment. J Bus Eth 1998;16:1835-1860.
5. Barnes B. Interests and the growth of knowledge. Routledge & Kegan Paul; 1977.
6. Herschel J. A preliminary discourse on the study of natural philosophy. 1830. Reprinted: University of Chicago Press; 1987.
7. Merton R. The normative structure of science. 1942. Reprinted in: Merton R. The sociology of science. Storer A, ed. University of Chicago Press; 1973.
8. Popper K. The logic of scientific discovery. Harper & Row; 1959.
9. Kuhn T. The structure of scientific revolutions. University of Chicago Press; 1962.
10. Laudan L. Progress and its problems. University of California Press; 1977.
11. Tesh, S. Hidden arguments: political ideology and disease prevention policy. Rutgers University Press;1988.
12. Wilkinson R. Health inequalities: relative or absolute material standards? Brit Med J 1997;314: 591-5.
13. Airaksinen T. Professional ethics. Encyclopedia of Applied Ethics 1998;3: 671-682.
14. Horn R. On professions, professionals and professional ethics. American Institute of Property and Liability Underwriters.

Methods for Research on Research Integrity: Doing Research on Sensitive Topics

Margot L. Iverson, Department of History of Science and Technology, University of Minnesota, Twin Cities, USA; formerly with the Program on Scientific Freedom, Responsibility and Law, American Association for the Advancement of Science, USA

Mark S. Frankel, Program on Scientific Freedom, Responsibility and Law, American Association for the Advancement of Science, USA

Keywords: Academic environments, Deviant behavior, Evaluation research, Methodology, Science careers, Scientific misconduct

Promoting research integrity requires a greater understanding than we now have of the factors that influence the full range of research conduct. There is a dearth of empirical research addressing issues related to research integrity and misconduct in science. It is critical, therefore, that more research on these issues be supported, not only to provide useful guidance to researchers and to the formulation of appropriately measured policy, but also to stimulate a critical mass of scholars to develop research on research integrity as a legitimate field of scientific inquiry. Such research must employ rigorous research designs and methods of evaluation.

The "Session on Methods for Research on Research Integrity," co-organized by Mark S. Frankel and Felice Levine, considered the methodological challenges faced by researchers studying research integrity and discussed research approaches best-suited to this topic. Four speakers presented different models and strategies for conducting research on research integrity and suggested promising areas for future research. The session concluded with discussion of a possible research agenda for research on research integrity. This account is a summary of the session.

Contextual Effects in the Study of Academic Misconduct

Melissa Anderson, Associate Professor of Higher Education at the University of Minnesota, presented conceptual models of scientific misconduct that could be used to guide research on the role of the academic environment on research misconduct. Studying different aspects of the research context in which incidents occur can move researchers away from focusing on prevalence, which is difficult to determine and of limited utility, to examining other research questions useful to institutions trying to promote research integrity. Researchers face several methodological challenges in investigating research misconduct. Misconduct is a sensitive topic that individuals wish to keep hidden from researchers (and others), making it hard to observe, and incidents are relatively rare, making them difficult to find and compare. The academic context in which misconduct is to be studied also can create methodological difficulties. Research areas in which perpetrators of scientific misconduct

Corresponding author: Mark S. Frankel, Program on Scientific Freedom, Responsibility and Law, American Association for the Advancement of Science, 1200 New York Avenue, NW, Washington, DC 20005, 202-326-6793 (voice), 202-289-4950 (fax), mfrankel@aaas.org.

function can be very technical, requiring investigators to possess some mastery of the specialized subject matter (or to collaborate with someone who does). Another problem can be the autonomous nature of academic researchers, which makes their behavior difficult to observe or to confirm independently. Additionally, research integrity research is not always welcome by institutions or departments, out of fear of media or legal attention, and individuals and organizations may not cooperate with researchers.

Rather than artificially disassociating misconduct from research, conceptualizing it as linked to unavoidable research error is one way in which misconduct can be understood in the context of the research process. Error and misconduct both involve issues of intention and acceptability, with misconduct being both intentional and unacceptable, and inadvertent error being the reverse—acceptable and unintentional. Anderson identified two other categories as well, avoidable error, which is unintentional but also unacceptable, and "minor hypocrisies," which are intentional but acceptable. Studying these categories of avoidable error and minor hypocrisies, which presumably are much more common than misconduct, may provide information on the contextual influences on misconduct that is difficult to obtain by other means. And since intent is hard to determine, some instances of avoidable error may be incidents of misconduct that have never been so identified. Other topics for further research that grow out of this linkage between misconduct and error are how scientists decide what separates misconduct from these categories, and if and how they deal with error as well as misconduct.

Another way to examine context is to consider not just the actual incident of misconduct, but rather to understand cases as having four distinct stages: the context (institutional, disciplinary, and immediate lab) in which the incident occurs, the misconduct event itself, the exposure of the misconduct, and the consequences for the perpetrator and others. This framework provides a way of considering and comparing different aspects of misconduct so that interactions between each stage can be explored. For example, what impact does the context of funding sources and mechanisms have on incidents of misconduct? Longitudinal research of patterns of interactive effects over time presents many possible research projects. The contextual influences of the broader research environment on these four stages, from such sources as disciplinary societies, journals, industry, government policies, and elsewhere, also suggest many useful research topics.

Scientific Misconduct as a Form of Deviant Behavior

Researchers who engage in scientific misconduct are behaving in a presumably deviant way that violates both legal and social norms. Conducting empirical research on research integrity and misconduct therefore requires that researchers consider the implications of studying deviant behavior in designing and conducting their research. In her presentation, Eleanor Singer, from the Institute for Social Research at the University of Michigan, discussed some methodological considerations arising from this understanding of research misconduct as a form of deviant behavior. In addition, she also presented some applications of more universal research principles to research on research integrity.

Deviant behavior is difficult to study because there are strong incentives for both perpetrators and the institutions at which it takes place to keep it hidden. This makes it difficult to observe directly, and so researchers must resort to asking subjects to report incidents. Two of the most common methods used are self-administered surveys and interviews. These are more likely to produce honest answers if the confidentiality of those participating can be guaranteed. Surveys that are self-administered, further ensuring privacy, also can improve rates of subjects' veracity. Another useful research method for some research questions is to present subjects with vignettes of ethical quandaries in research and to ask them how the researcher in the vignette would behave. Such vignettes are most useful when the type of research and the status of the researcher in the vignette parallels those of the subject, as this increases the chance that the answer will reflect their own behavior. Vignettes also can be used to study what behaviors actually are regarded as violations of standards of conduct by members of a particular field.

Like other forms of deviant behavior, opportunities to engage in scientific misconduct as well as opportunities for observing it can vary depending on factors such as the discipline of

research and the size of a department. Also, motivations for deviant behavior may vary, based on incentives and reward systems present.

Singer also presented several other principles of empirical research that are critical to producing rigorous empirical research on research integrity. Researchers must establish the questions they wish to answer with their research. To obtain consistent answers and meaningful results, terms used also must be defined. For example, since norms and definitions of research integrity and misconduct vary, these terms must be clarified so that all researchers and subjects are using a standard definition. If not, ambiguity may be introduced into the data. (Research that explores differences in norms and definitions of misconduct could be very useful in helping to interpret current data on prevalence.) The populations to be studied also must be selected so that comparisons can be made. When choosing research methods, the match between method and research question should be carefully considered. Direct observation, deliberate experimentation, questioning subjects, and analysis of official records are all possible methods, and each has advantages and disadvantages. The choice of method also involves a selection of the indicators the study will use. Official records of complaints of research misconduct, for example, will yield different information about incidence than data collected through surveys of bystanders or perpetrators. Since descriptive statistics are much more meaningful in a comparative context, it is important to consider how different parts of a study can be made sufficiently equivalent so that data can be analyzed comparatively. The research conducted by Judith Swazey, Melissa Anderson, and Karen Seashore Louis on integrity issues in graduate education is a good example of the effective application of these research principles to research on research integrity (1-2).

Influences on Research Integrity at Different Stages of Academic Science Careers

Another research model that can be applied to research on research integrity is the effect of the academic environment on researchers at different stages of their careers. Although many scientists take a class on research ethics early in their training, the major influence on how they learn to conduct ethical research is usually the environment in which they work. Rachel Rosenfeld, a Professor of Sociology at the University of North Carolina, presented a sociological framework to consider how scientists learn about ethical research practices at different career stages, what it is they learn, from whom, and why sometimes they learn the wrong lesson (i.e., unethical behavior). At each stage of a scientist's career, several nested contexts influence research integrity. In the immediate research environment, researchers are exposed to peers, mentors, teachers, collaborators, and students. Surrounding and overlapping this immediate environment are the context of department and institution and the broader context of journals, professional societies, and federal policies.

Rosenfeld discussed some potential research projects at each stage of a scientist's career, from undergraduate through senior scientist. Currently available research on undergraduates has focused on the conduct of science students in the classroom and has indicated distressingly high rates of plagiarism and fudging data. Are advanced students engaged in independent research projects more or less likely to fudge or plagiarize data in the research environment? This would be an especially interesting research topic since those undergraduates who do participate in research are more likely to continue on to graduate school than other students. For graduate students, research has suggested that the interaction between them and their mentors is critical to their subsequent ethical behavior. More research is needed on how aspects of this interaction affect the information on research integrity transmitted. The role of other graduate students, and the effect of isolation from peers on ethical behavior are other potential topics. To what extent are graduate students who interact frequently with their peers learning ethical (or unethical) behaviors from them? As researchers move from being graduate students to post-doctoral trainees to junior scientists, the broader research community context becomes more important. Journals and scientific societies may become more influential in shaping junior scientists' behavior. Do varying standards of evidence adopted by different journals influence researchers' research practices? For example, if a journal requires that all underlying data be accessible, does that have an effect on the accuracy of the researcher's analysis of the data? How does the pressure to publish affect what and how researchers conduct research? Regarding

scientific societies, how does leadership on research integrity from societies impact the behavior of members? Do society ethics codes and ethics prizes influence members? And for senior scientists, who are likely to become part of the leadership of societies and departments, how do these roles influence their own research conduct?

Contextual questions exist for each stage of a scientist's career, and studying these questions can identify the conditions under which interactions in a particular context lead to the learning of ethical or less ethical research practices. That researchers might receive mixed messages from the different contextual environments was noted by an audience member, and Rosenfeld concurred, noting that some contextual messages may promote unethical behavior and that it is important to assess how competing messages are dealt with by scientists. Another factor to consider in research is how these nested contexts affect individual researchers in different ways. A researcher's gender, race, country of origin, or sexual orientation can all impact the individual's interactions with the surrounding environment.

Utilizing Evaluation Research to Assess Research Integrity Programs

Joyce Iutcovich, President of Keystone University Research Corporation in Erie, Pennsylvania, presented an overview of the contributions that evaluation research can make to research on research integrity. Along with basic research, which addresses questions about causality and contributes to theory development, evaluation research provides the link between theory and practice. When research institutions and scientific societies develop research integrity programs based in part on theory, evaluation research plays an important role in assessing the effectiveness of these programs. Further, it offers a system for transferring knowledge gained through research to program improvement efforts over time.

Evaluation research is conducted within the context of social action programming. It focuses on an assessment of the implementation process as well as the outcomes for targeted groups. Process evaluation determines whether a program has been implemented as planned; outcome evaluation determines the short- and long-term impact of a program on the target group(s). To conduct a process and outcome evaluation, the following programmatic and research design elements need to be in place.

First, program goals must be clearly defined for a specified target audience (e.g., graduate students will be made aware of the ethical standards for research and the strategies for adhering to these standards). Second, activities to achieve these goals must be designed and implemented (e.g., an educational program consisting of a one-credit course is established as a graduation requirement; it is taught every fall semester). Next, a plan for the evaluation of the program's implementation process and outcomes needs to be delineated, including measurements and instrumentation (e.g., measures of knowledge using a paper/pencil test or measures of decision making using case scenarios), timing of data collection (at the end of each course), methods of analysis (quantitative), and format for reporting the results and implications for an organization's activities, since it is essential to incorporate a system for linking knowledge gained through research to organizational planning and action.

Evaluation research assesses the overall effectiveness of an organizational program and is used to improve programming so that goals are met and resources are used efficiently. It is based on an open system's model of organizations ("open" because the organization is open to political, social, and economic influences from the external environment). As conceptualized using this model, evaluation research provides evidence, which becomes part of the continuous feedback loop that constantly works to improve programmatic efforts. Ideally, programmatic efforts that address issues and concerns related to research integrity are based on theoretical models that provide an understanding of research integrity and how to ensure it within a population of researchers and scientists. Once implemented, evaluation data on these programs are collected, analyzed, and used for program improvement. Evaluation research also provides another critical assessment of the theoretical model, which establishes the framework for the program. This further enhances theoretical development by providing evidence about what works and what doesn't work as predicted by a theoretical model.

Session Conclusion—Developing a Research Agenda

Felice Levine, Executive Officer of the American Sociological Association, addressed the scope of

research integrity and misconduct concerns, the challenges for undertaking study of such issues, and the need to attract researchers with broad expertise. Also, synthesizing many of the topics raised in the presentations, she concluded by suggesting steps needed to establish a research agenda for studying research integrity.

Prior to designing an agenda, the scope of the research and related topics on research integrity and misconduct must first be determined. Along with fabrication, falsification, and plagiarism, issues of conflicts of interest, human research participants, confidentiality, authorship determination, data access/sharing, data design, and accurate representations and interpretations of data all may fall within this subject area. The complexities involved in conducting research on research integrity also must be considered. Since deviant behavior is often hidden from outside view or occurs among powerful elites, there are many challenges to obtaining empirical data on research integrity. Political concerns within and between organizations also may inhibit research. Also, since this research could benefit from research methodologies and frameworks from a variety of disciplines, attracting researchers from a broad range of disciplines is crucial. Across disciplines, important areas of expertise for such research include history and sociology of science; work, occupations, and professions; research ethics; deviance and white collar crime; decisionmaking; and organizational behavior.

Levine then presented initial steps to be taken to establish an agenda. The stakeholders in research integrity—including the individual investigators, research teams, scientific societies, potential funders, subjects to be studied, policymakers, and the public—must be identified. Data sources already available from federal agencies and other organizations as well as resources needed but not available should be assessed. Funding sources and mechanisms should be identified, and structures—including conferences, working groups, panels, and large-scale collaborations—should be set in place to provide frequent opportunities for scholars to communicate. Finally, to develop a community of researchers working in this area, a substantial investment is needed to provide educational opportunities for researchers from different disciplines and at different career stages. These opportunities could include internships for students, postdoctoral and mid-career incentives or awards, and specialized training programs.

The session ended with some questions and comments from the audience. Among the final comments was the observation that many of the presentations focused more on context than on individual behavior and that this seemed to reflect a shift from individual character to research context in understanding research misconduct. The need to include “organizational misconduct” in this field of research also was voiced.

Studying research misconduct presents several kinds of methodological challenges, including difficulties in observing deviant behavior and in conducting research in an academic environment. Researchers, institutional review boards, and funders must be sensitive to these matters and give due diligence to research design and methods. Nothing could set the field back more, even before it takes shape, than sloppy, inappropriate, or poorly designed or applied research methods.

Bibliography

1. Swazey JM, Anderson MS, Louis KS. Ethical problems in academic research. Academic Scientist 1993; 81 (November/December): 542-555.
2. Anderson MS. Uncovering the covert: research on academic misconduct. In: Braxton JM, editor. Perspectives on Scholarly Misconduct in the Sciences. Columbus, Ohio: Ohio State University Press; 1999. p. 283-314.

Research Misconduct: A Multiperspectival Approach

Robert J. Silverman, Fielding Graduate Institute, Santa Barbara, CA, USA

Keywords: Scholarship norms, Social science misconduct

This paper engages the topic of research integrity with a foundation that differs from that which appears in more traditional treatments of misconduct (1, 2). These other treatments are grounded in the notion of "role;" that is, scholars are in role and guided or controlled by certain norms whose abrogation are role sins.

But what if one, at base, defines the scholarly life in different terms, not on the basis of distinct research practices, mandated by professional role, but where there is no separation between how one lives one's life and how one produces as a scholar? Or, what if one uses one's scholarly work as a way of living one's life in the broader society and culture such that one's impact (3) will reshape the environments in more acceptable terms to oneself and others? Current misconduct considerations need to pre-suppose social and psychological patterns under girding professional life that are, in fact, more varied than often assumed.

The traditional norms, such as communality, universality, organized skepticism, or even their opposite as counter norms–which continue the salience of the dominant terms–and which have historical grounding (4) do not capture these definitional differences, as the wars among the epistemologies make rather clear. How might we engage this thicket? We need to do so less as watchdogs, which is the current preferred pattern. Charging and defending individuals and social institutions is but one set of approaches. Holding individuals to norms that are irrelevant to their definitions of scholarly work and life will not motivate us to consider how we might address the heart of the knowledge creating enterprise in its foundational diversity. This is a time for rethinking what norms and standards should obtain, and we need to do so with energy and with expectations of their significant complexity.

Unfortunately, this complexity and inventiveness is not apparent in the Report of the Commission on Research Integrity, entitled *Integrity and Misconduct in Science* (5), though it invites definitions from various single and multidisciplinary fields. It argues that:

> Research misconduct is significant misbehavior that improperly appropriates the intellectual property or contributions of others, that intentionally impedes the progress of research, or that risks corrupting the scientific record or compromising the integrity of scientific practice....(5)

It develops new terms, in place of "fabrication, falsification and plagiarism," namely, "misappropriation, interference, and misrepresentation" (5). The first includes plagiarism and also notes the improper use of what is essentially confidential or privileged information; the second covers the damaging of others' research property; and the last deals with attempts to deceive, from omission or commission. There are other forms of "professional misconduct" added to that of research misconduct, primarily dealing with attempts to obstruct investigations of misconduct. In addition, there are calls for both academic institutions and professional fields or disciplines to develop codes of

Corresponding author: Robert Silverman, Fielding Graduate Institute, 2122 Santa Barbara St., Santa Barbara, CA, 805-687-1099 or 360-566-8188 (voice), 360-566-1101 (fax), rsilverman7@home.com.

conduct and to provide a variety of educational experiences for scholars regarding ways of behaving professionally to avoid research misconduct. While this committee's initiative in rethinking misconduct and its invitation for contingent organizations to focus on "local" concerns appears positive, there appears to be no larger vision in the recommendations, apart from the Committee's observation that a basic principle was "that scientists be truthful and fair" (5). That is, there is no sense that inquiry, the actions of scholars, or the codes of disciplinary/ multidisciplinary communities vary considerably. However, these are not just local factors that are at issue, but different world views, different logics-in-use. Unless our codes take them into account, scholars will be applying the wrong rules and norms for their own and others' research and ignoring the need to create norms that have relevance.

Communities of Scholarship

The focus of this paper is to begin generating different sets of values or norms that have relevance for alternative ways of knowing, not reflected by singular methods or in specific fields, but in light of co-existing epistemological contexts. And, the location is the social sciences and not the natural sciences—which are assumed in nearly all conversations regarding research integrity. We do need to refresh the conversation with creative possibilities that are in keeping with *how* and *with whom* scholars do their work. As these alternatives become more evident there is a responsibility to consider their implications for normative as well as other issues. Alternative warranties for different ways of doing scholarship and being academics are needed. Three years or 36,000 miles is just one solution.

Academic Contexts

This paper differentiates the notion of "academic context" in terms of the generativity or embeddedness of the knowledge and the openness or stability of the communities within which scholars work. It appears that the concepts of "constitutive" and "regulative" can be appropriately applied as significant adjectives to both knowledge and community such that we identify working cohorts of scholars who are engaging with knowledge differently and whose work will have differential value to other stakeholders who identify with each other.

Many scholars evoke images suggested by the distinction between newly conceived and established knowledge, including Lorraine Code, who in *What Can She Know*? (6), extends constitutive and regulative patterns to the community, as well as to the knowledge that transpires among its members. It is the interaction of various approaches to knowledge *and* to community that shapes the academic contexts in which scholars live their lives. Let us use Code (6), though one could focus on many, many others to make the same distinctions under the banners of their foundational goals (e.g., 7,8,9,10).

Academic Communities

First, let us differentiate among "academic communities."

1. In a *constitutive community*: "...every cognitive act takes place at a point of intersection of innumerable relations, events, circumstances, and histories that make the knower and known what they are, at that time... (It focuses on) the complex network of relations within which an organism realizes, or fails to realize its potential... (6)." The community allows for interrogation, renegotiation ; it evidences trust which "involves making oneself vulnerable... (6)."
2. In a *regulative community*: One sees authoritarian knowers who "...claim credibility on the basis of privilege alone or of ideological orthodoxy... (6)." Code suggests that there is an obsession with autonomy and an overemphasis of the self (6).

Regulative communities have historical dimensions with regard to the participation of various actors. These dimensions have significance for an understanding of current professional directions, as exemplified by Hull's work on the consequences of members' contributions to the field of taxonomic biology. Regulative fields are more integrated and planned with professional divisions than are constitutive communities which are more organic, emergent, and fragmented, with collections of individuals coming together for cause, that is, the attraction of a problem or issue rather than a continuing research or theoretical focus.

Much of academic life is geared to the celebration of success, and the study of action, in

a regulative model of community, is reflected by research prizes, citations studies of a contribution's value or impact, and the life of research schools (12). Constitutive communities often are inappropriately placed in the regulative frame as when one discusses a field or area as being "pre-paradigmatic," as if it were "pre-pregnant." It occurs inappropriately when a comparison of citations patterns is made between physicists and educationists, as if research papers are not used by the latter for policy guidance, a value that goes unrecorded, and assume the only value of a paper is in crafting new research. Now let us differentiate among kinds of knowledge:

1. For *constitutive knowledge*: one takes account of testimony and cognitive interdependence (6),..." letting 'objects' of study speak for themselves..., ...understand(ing) difference and accord(ing) it respect (6). It grows by accretion without a preexisting frame.
2. For *regulative knowledge*: there are more standard forms; it is more hierarchical, is informed by such principles as objectivity and value-neutrality at the same time it is also more adversarial and territorial (p. 6).

Constitutive knowledge is developed from many sources; it is constructed piecemeal, whether the source materials are concepts, ideas, or data bases, or some combination derived from resources made proximate by the scholar who creates such a bundle to address a "problematic." A marvelous example is reflected in Shapin's work (13) where he writes in his introduction: "(This book on social history) is concerned with questions ...which have traditionally been the preserve of philosophers; it uses evidence and techniques customarily owned by historians; and the conclusions it arrives at are broadly sociological in form and substance." He disavows an interdisciplinary orientation and hopes to be identified as an historian. Regulative knowledge is more standardized, cumulative, specialized, and stable.

There is an interaction between "knowledge" and "community," that is, alternative kinds of knowledge are crafted in communities of either type. It is important to not necessarily equate the resulting four types of possibilities with disciplines or fields of study. Some fields or disciplines may have contexts of only one type; there are many in the social sciences that embrace those reflecting multiple logics. For example, on a common-sense basis, it is clear that the field of psychology embraces humanistic and behavioral alternatives, as well as areas that are laboratory and clinically-based. To place a field such as this conceptually in one area is to deny its multiplicity. It contains cohorts of scholars who could be differentially placed in the four-fold scheme, as suggested above. But, such locating is not necessarily a "cold-blooded" act of placement; it occurs as well among scholars who react with some emotion to each other's work. This paper acknowledges the legitimacy of all that claim to be knowledge communities and asks the reader to surrender her or his current categories for the suggested set of alternatives. Unless one is so willing, it is possible that researchers may be locked into more narrowly defined debates than relevant. In commonplace language, this is seeing the trees and not the forest.

As suggested, then, there are four knowledge contexts which are the relation of constitutive and regulative possibilities for both the community and the knowledge its' members produce.

1. Scholars in *regulative communities* can develop *regulative knowledge*. This is the traditional context in which fields develop incrementally within well-established paradigms or theoretically/empirically informed schools. While these areas grow in terms of the "agencies" of various human and machinic components, the goal is the stabilization of knowledge (14) by persons who, with a complementary set of scholarly interests, seek answers to research questions of acknowledged importance. This is the context that Kuhn and most other commentators assume, and which they unfortunately assume to be universal.

2. Scholars in *constitutive communities* can also develop *regulative knowledge*. This occurs when cohorts of scholars in a variety of fields or areas center around the work of theorists or schools of thought that pull them into similar logics. For example, persons in a variety of "disciplines" or areas study Piaget, or Kohlberg, or Kuhn, or critical perspectives as developed by Foucault. Concepts or theorists are treated in canonical fashion as they are "applied" to various "problematics" by persons in various locations.

3. Scholars in *regulative communities* can develop *constitutive knowledge*. This occurs when researchers in an area, such as higher education or the sociology of science, study an

issue by bringing together unique knowledge resources, that is, concepts, theories, and methodologies from a number of sources to frame and study the concern at hand. The focus is on the question, with the inventiveness of the scholar addressing others who examine similar questions from very different resources and choices.

4. Scholars in *constitutive communities* can develop *constitutive knowledge*. This occurs when individuals from a variety of locations come together to establish evolving understandings in a particular way. As an example, one might have feminist scholars develop an understanding that has both independent and interdependent sections on ways of doing professional work, such as teaching. The community of interest and understanding are organic, evolving, and this is reflected in various patterns of intersection.

In earlier work, this researcher has attempted to appreciate the meaning of these differences in terms of their implications for electronic publishing and to examine how they shape the nature of argumentation within and between scholars who live in these academic contexts (15, 16). There is work that establishes these distinctions through philosophic attention (e.g., 17, 9, 10). It can be argued, from a broad multidisciplinary base, that these distinctions are primary ones and that the particular labeling of the contingencies in this paper reflects alternatives that have universal meaning, though other scholars have used different language and examples in their areas to denote these alternatives. The question for what follows is, "What are the implications of alternative production in different kinds of academic communities for the meaning of misconduct?" If the community is attempting to appreciate what comprises misconduct one needs to understand the nature of the production within their communities for such an understanding to have value.

Scientific Misconduct Pluralised

Regulative Community/Regulative Knowledge. This context accounts for the observations made by those who are part of the historical and contemporary conversations with which we are familiar.

It is a voluminous set of materials which focuses on such domains as the traditional norms of science and adherence to them in practice; both current and historic examples of misconduct by figures who made foundational contributions (e.g., Newton, Pasteur), contributed significant work (e.g., Burt), and did normal science, all of which are treated in historical studies and journal, news and list-serv accounts; admonitions and suggestions regarding how the mechanisms of the scientific community (e.g., editors and peer reviewers) can act to be aware of misconduct reflected in submissions; how various stakeholders (e.g., lawyers, scientists) focus on different dimensions of misconduct and the actions of institutions (e.g., universities) in such instances; the debates regarding what misconduct includes in practice and the delimiting of the practices that are so situated, to include the on-going conversations sponsored by such organizations as the American Association for the Advancement of Science that provide for public reflection on the issues. There are many voices, from the philosophers, historians, and sociologists, to the aggrieved parties and those who are situated in different parts in the various dramas; to those who police the science community.

As one can note, the concern is primarily for the misguidance of the scientist, and there is limited attention to the community of science as playing a part in the perpetuation of misconduct. While the Public Health Service Report reflects a concern for the role of universities and professional societies and focuses attention on the "whistleblower" as she or he is treated by professional peers, the attention is on the producer of knowledge as an individual and not elements of the community (5).

For example, if a journal editor sends a manuscript out for external review to peers who do not share the methodological bias of the author or the reviewer comments on the paper in ways that suggests that he or she challenges the logic of the approach to inquiry used by the author, this would not be considered misconduct by either the editor or reviewer. Rather, it would be considered as poor judgment or bias, and the actors would have nothing to answer to in any special forum. A journal that fails to publish its policies and fails to send manuscript reviews to authors would not be cited for "killing" the scholar by denying opportunities for access. Trust, not justice, is the focus of the operant norms in this fully regulative sector, and the regulative community has little to answer to under this additional value. Interestingly, there is

a vital literature on how academics "cool out" their colleagues, either because they produce on the "margins" of accepted knowledge or because of gender (18). Practices of disciplinary bodies or forums, and scholars associated with them, are not considered as having relevance for attributions of misconduct. The actions of a majority or those who operate from power are hardly ever placed in such arenas.

Unfortunately, as well, the work on misconduct that has relevance for this context is assumed to be of value more generically. This allows us to ignore what is not normative in the other approaches to knowledge construction.

Constitutive Community/Regulative Knowledge. As noted above, in this context scholars from various disciplines, forming a colleague group, typically focus on the work of particular theorists, or the implications of particular macro concepts or worldviews. For example, this would include scholars from such fields as psychology, sociology, political science, education, social work, and others, engaging with the work of, say, Foucault and referencing colleagues who are similarly involved, rather than colleagues who share a disciplinary-derived designation, such as deviance in sociology, or analytical or mathematical geography.

It is of interest that the constituents are committed to certain ideas that often are in defensive contention with alternatives held by another cohort of scholars. Peers engage in discourse around which one finds a consensus, with argumentation around the fringes as implications and new application of well understood and frequently articulated ideas are evoked and borders are defended. What does it mean to plagiarize when proselytizing and self-affirmation are the orders of the day? The truth of work rests on a self-evident body of original conceptions and supporting material, such as that surrounding the Jungian foundation of the Myers-Briggs personality instrument. Advancing a set of concepts or worldview is the challenge, and the use of additional supporting material only aids the cause. This is not to say that it is appropriate to use someone else's work as one own. But, unlike the scholarship emerging from the previous location, this work is not so much the careful extension or articulation of a body of work growing on the edge of a well-prepared community, as it is the "answer" for viewing a concern. That is, since individual scholars evoke in their language certain commonly-noted truths, or notions, or meanings, one is repeating material, replicating others' language in which the constancy of the foundational work is repeated. If one uses the term "needs" or "paradigm", then the source is known, not only in terms of the original authors, but the wall of support behind the term. Those in the various fields who are using canonical concepts or theories would find little value in using the work of scholars in other fields who use the same literature, say by Piaget, since, it would have no applicability to their work in question. There is a "catechism," then, which like any canonical text comes from well recognized sources and which needs repeating, which migrates, as the parameters it asserts are firmly supported by peers. Growth occurs as additional applications and connections among key spokespersons are articulated and as peers defend and repair boundaries that support the difference "it" makes.

In the heuristic spirit of this paper, misconduct can take a number of forms among academic "true believers." First is the failure of the advocates to examine first principles. It is suggested here that it is inappropriate for a scholar to support and defend certain points of view without giving serious consideration to the origins and consequences of the point of view or scheme and to appreciate it in relation to alternatives. Presenting a persuasive or defensive case in reference to one's advocacy is a foundation that can be avoided by simply acknowledging the value of a work and identifying oneself through such association. It is done frequently, often by graduate students and then by those who find the pull of the network, or invisible college as less a call for excellence than a sinecure for the privilege of self-evident truth. Blind advocacy here is no less significant than being a true believer in any other context, and it has less a place in the academy than in other social institutions. It should be expected that one's agency and voice be earned, not by mimicking a rhetoric that one finds attractive, but by being able to articulate with others, especially those who reflect alternative perspectives, a logic that one has reasons to represent.

The second concern is about the nature of the argumentation that ensues. One may try to besmirch the other, to insult and shame the other, rather than challenge her or him on appropriate grounds. My study examining patterns of advocacy and defense in this context applied the labels "contention and fortification" to describe

the interactions across schools of thought around the same issues (16). At times, when reading the various protagonists, one could call forth an analogy from the larger society, namely, the no-holds-barred punching and kicking strong-person contests that seem to have captured some interests among fans in the larger community. Since one's identity is a function of "a point of view" or the viability of a particular body of work, one finds scholars needing to defend it by any means available. Some of the strategies include attributing motives to the other that she or he has not articulated; finding fault with the other because he or she did, or might have, read certain authors which the writer holds in disrepute; using hostile words in combination such that the rhetorical impact transcends their logical significance; and claiming others have alternative perspectives even when they gave no voice to such. The writer will allow the other view to survive, but in rather tender shape, because it is easier to support oneself by arguing against another view than doing so without an enemy. Simply put, members of constitutive communities working with and articulating regulative knowledge should make judgments on evidence and need not unfairly represent their positions or challenge alternatives or prevent their articulation.

It can be argued that the content of the regulative knowledge at issue is of some importance. There are academic "belief" systems that create inequities for others. Since theory often outstrips its empirical base, or creates the possibility of work that can only support the contentions of those involved, the impact of theoretical systems on life worlds needs attention—ideas which are removed from serious possibilities of destruction and upon which whole careers are based. Is it a matter of misconduct if one's scholarship creates obstacles to others' human rights—either in terms of their status possibilities or relative success? An example of such an obstacle would be the use of a Caucasion-normed personality instrument and associated theory to classify the development of African American students, especially when an instrument based on African American student growth and change is available. We are responsible for the worlds we create.

Regulative Community/Constitutive Knowledge. In this context, scholars from regulative communities, such as from within the sociology of science, develop an understanding of a topic by bringing theories, findings, concepts, and methodologies together that will enable her or him to understand an issue in a way that is deemed appropriate for the stakeholders involved. The titles of two articles, the first in *Configurations: A Journal of Literature, Science, and Technology*—(a journal title that itself makes the point)— "The Pathology of Painting: Tuberculosis as Metaphor in the Art Theory of Kazimir Malevich" and in *Social Studies of Science*—"Literary Genres and the Construction of Knowledge in Biology: Semantic Shifts and Scientific Change" are exemplars here. It is not unusual for a person to claim that he or she is a multi or interdisciplinarian, or some variant thereof. Typically, the questions being investigated are not of state-of-the-art significance to a subset of the academic community, but instead better reflect the unique interests of the involved scholars and present their special and idiosyncratic ways of dealing with approaches to questions that might be shared with peers.

Certainly authors can engage in falsification, fabrication, and plagiarism, and possibly with a lesser chance of being found out than one might were one in the fully regulative context. There is less concentration of similarly educated and concerned peers and a greater variety of folks who have different backgrounds and who roam the literatures and methodologies in seeking "fits" with their disciplinary and value-constituted dispositions. What in a more specific sense does this suggest regarding misconduct?

It is suggested that not recognizing and dealing with one's own constructions as constructions is "misconduct." That is, it is anormative to consider one's own construction as beyond reflexivity. Unlike work in the fully regulative environment, here there is no historical and progressive justification for a line of reasoning. Treating a solution to a problem as self evident, such that it is not justifying itself in relation to alternative treatments to the same or a similar problem, does not allow others to appreciate the added value that may accompany a way of understanding. So, while one can gauge the meaning of a work in the fully regulative context by its specific and its particular use of references, in this third context, such is not possible. There needs to be the willingness to justify the connections among the elements of a work and to engage in academic conversations, which interestingly enough, are regular features

of *Social Studies of Science*. Of course, others need to be willing to engage, and this suggests the need for relevant regulative communities to have a meta-language available so that persons from different vantage points can engage each other.

One might also suggest that it is cause for concern, that is, misconduct, when the scholar uses the work of others in ways that fundamentally change the elements such that the sources of origin would likely object to the implications of the uses of the work should the source have potential voice here, or when such objections are ignored. This use could reflect a number of possibilities, from the location of a work, that is, where it is used, and how the sources are modified in a new treatment. For example, today there is a major concern in the management literature on the concept of "resilience." Some relevant questions might be, "How can leaders have more of this?" "Do successful CEO's have this attribute?" In approaching these types of questions, a scholar uses narrative accounts of those who demonstrated resilience for their survival: returning POW's and those liberated from Nazi concentration camps. It can be argued that one dishonors the events and the lives of those who perniciously suffered to use their accounts as conversation pieces for a cocktail party, and to allow the CEO to think that he or she walks in the POW's shoes. Such use reflects back on the original stories. They are re-storied, and arguably, in ways that lessens their deeper meaning and the meaning of those who found themselves as unwilling participants.

It is also a concern when scholars do not allow their solutions to a problem to be engaged and modified through additional empirical and theoretical treatments. Certainly, the concatenation or assemblage of new material will allow each of the contributing pieces to develop alternative textures and tones, if not be challenged in new ways. A scholar who attempts to prevent such consequences, say, as a peer reviewer who also is the author of an earlier text that is being revised and does not allow another author to continue the development of or challenge to the work, could be considered to be acting in a way I would label "misconduct."

Constitutive Community/Constitutive Knowledge. In the fully constitutive context, we have a high degree of organicism as emergent forms and emergent knowledge continually develop in conjunction. There is a "collective integrity," to use Mary Ann Caws's term, with the consequence being greater illumination, which leads both to interactants' ensembling and external stakeholder mutual appreciation. Since, misconduct was originally defined within a fully regulative context, one might wonder if it has any applicability at all in this fully constitutive one (19).

Individuals in this context reveal themselves through conversation, narrative, anecdotes, personally situated histories, and engaged professional settings. They do not use rhetorical strategies to persuade, but attempt to present themselves as evolving, with the risks that such confession might have. They reveal how they have grown or changed through various encounters with persons with whom they engage in a scholarly way or with situations that provide environments for learning and action. It would seem that it would be a misconduct situation if one used the information so revealed for private gain, either as the author or the audience. One is expressing or bringing forth an emotive/affective connection to the rational material; values are clearly articulated, and a kind of privileged relationship is being established among the parties. While not referring to confidentiality in terms of content, here one should maintain confidentiality based on one's respect for the person who has revealed something that is "personal." Thus, to ridicule the narrative of another or to suggest one's superiority in terms of intelligence, motives, or values is antithetical to the orientation of this knowledge community. It freezes the logic of interaction and unnaturally shapes the content of the exchange.

It is also anormative in this context not to listen. Interestingly, hearing is the locus of interaction in this context, not seeing (20). So, not listening is misconduct, as would be those practices that chill the aural environment, such as intruding on others' exposition, translating a person's words into alternative words, attributing an exposition to a rationale or condition that has analytic rather than personalistic origins. For example, saying that this person speaks a certain way because she is of a certain psychological type is to reduce the individual's being to a set of variables and should be considered unethical.

Conclusion

Defining scientific misconduct and discussing examples and exemplars has become, if not

popular, then a more broadly based consideration than was evident even a few years ago. However, in spite of a great deal of commentary, we have failed to extend our considerations to issues regarding misconduct within various types of academic communities, especially those that are reflected in the social sciences. It is not enough to extend concepts having value in one domain of scholarship and then apply them conveniently to others. The social sciences and the humanities reflect alternatives that have meaning for misconduct, and not only the conduct of work. While academic communities have legitimate interests regarding fabrication, falsification, and plagiarism, there has been no previous attempt to go the basics and to consider what might be anormative for different ways of knowing in different settings, not in a methodological, but in an epistemological way.

This paper has attempted to explore "misconduct," with the explicit understanding that the ideas and possibilities discussed here are not presented as answers or solutions but as heuristic tools to carry the initial discussion. It is likely, even hoped, that what has been noted here will be revised in the continuing dialogue that transcends these notions and goes to deeper and more critical hearts of the matter. And for that conversation to be based on research would allow for such engagement.

Bibliography

1. Broad W, Wade N. Betrayers of the truth: betrayers in the halls of science. New York: Simon & Schuster; 1982.
2. Journal of Higher Education. Special issue on research misconduct. 1994; 65, #3.
3. Alcoff L, Potter E, editors. Feminist epistemologies. New York: Routledge; 1993.
4. Mitroff I, Kilmann R. Methodological approaches to social science. San Francisco: Jossey-Bass; 1978.
5. Commission on Research Integrity. Integrity and misconduct in science. Washington: U.S. Dept of Health and Human Services: 1995.
6. Code L. What can she know?: Feminist theory and the construction of knowledge. Ithaca, N.Y.: Cornell University Press; 1991.
7. Fuchs S. The professional quest for truth: a social theory of science and knowledge. Albany, N.Y.: State University of New York Press; 1992.
8. Sayre K M. Belief and knowledge: mapping the cognitive landscape. Lanham, MD.: Rowman & Littlefield Publishers; 1997.
9. McKeon R. On knowing: the natural sciences. Chicago: University of Chicago Press; 1994.
10. Pepper S. World hypotheses. Berkeley, CA: University of California Press; 1942.
11. Hull D. Science as process: an evolutionary account of the social and conceptual development of science. Chicago: University of Chicago Press; 1988.
12. Geison, G L, Holmes, F L, editors. Research schools: historical reappraisals. Osiris, 9. Chicago: University of Chicago Press; 1993
13. Shapin S. A social history of truth: civility and science in seventeenth century England. Chicago: University of Chicago Press; 1994.
14. Pickering A. The mangle of practice: time, agency, & science. Chicago: University of Chicago Press; 1995.
15. Silverman R. The impact of electronic publishing on the academic community. Peek, R P, Newby G B, editors. Scholarly publishing: the electronic frontier. Cambridge: MA: MIT Press; 1995.
16. Silverman R. Comments and replies: academic conversations. Science Communication 1994; 2: 132-151.
17. Detienne M. The masters of truth in archaic Greece. New York: Zone; 1996.
18. Silverman R. Publishing in an immature field: the case of 'higher education.' Higher Education Research and Development 1986; 5,2: 123-134.
19. Caws, M A. The art of interference: stressed readings in verbal and visual texts. Princeton: Princeton University Press; 1989.
20. Levin D M. The listening self: personal growth, social change, and the closure of metaphysics. London: Routledge; 1989.

Keyword Index

U

W