Internal and External Validity and Method of Control

December 13, 2011 , 2-3:30pm Drew MSRTP, Los Angeles, CA

Ron D. Hays, Ph.D. (drhays@ucla.edu)

- UCLA Department of Medicine: Division of General Internal Medicine and Health Services Research
- UCLA School of Public Health: Department of Health Services
- RAND, Santa Monica

http://gim.med.ucla.edu/FacultyPages/Hays/

Listed below are a few statements about your relationships with others. How much is each statement TRUE or FALSE for you?

- I am always courteous even to people who are disagreeable.
- There have been occasions when I took advantage of someone.
- I sometimes try to get even rather than forgive and forget.
- I sometimes feel resentful when I don't get my way.
- No matter who I'm talking to, I'm always a good listener.

Give yourself 1 point for each item that you answered as shown below
(Possible score range is 0-5)

- I am always courteous even to people who are disagreeable. **DEFINITELY TRUE**
- There have been occasions when I took advantage of someone. **DEFINITELY FALSE**
- I sometimes try to get even rather than forgive and forget. **DEFINITELY FALSE**
- I sometimes feel resentful when I don't get my
- way. **DEFINITELY FALSE**
- No matter who I'm talking to, I'm always a good listener. **DEFINITELY TRUE**

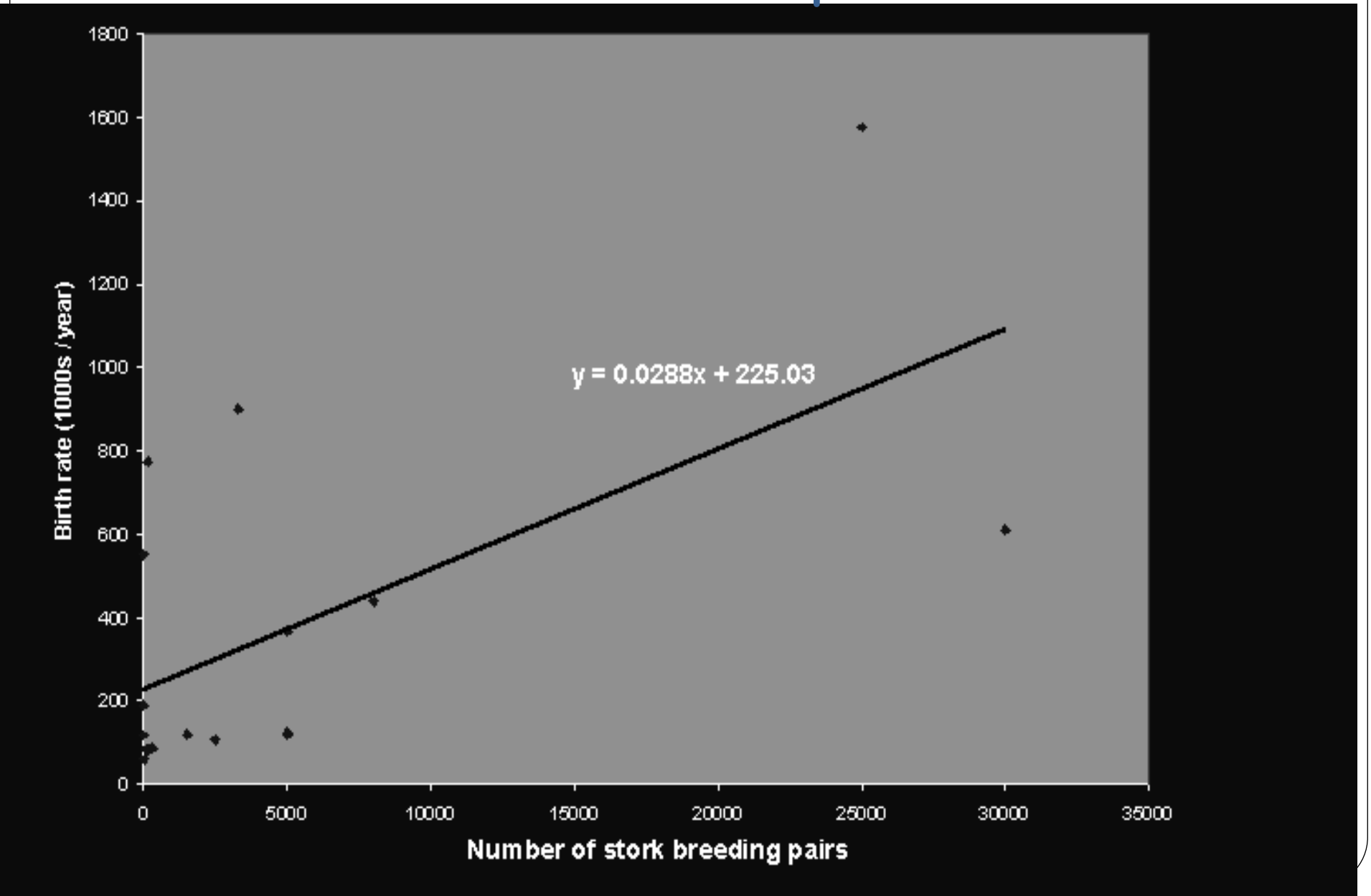
How many people scored?

- 0 points? _____
- 1 point? _____
- 2 points? _____
- 3 points? _____
- 4 points? _____
- 5 points? _____
- Mean was 1-2 points in MOS

Basic Concepts

- Internal validity
  - Data support conclusions about the hypothesis in the specific instance studied
- External validity
  - Findings of the study can be generalized to other populations and settings
- Extraneous variable
  - Capable of explaining the study findings without invoking the hypothesis (alternative explanation for the results)

Correlation = 0.62 between number of breeding pairs of storks and births in 17 European Countries



Country	Area (km^2)	Storks (pairs)	Humans (10^6)	Birth rate (10^3/yr)
Albania	28,750	100	3.2	83
Austria	83,860	300	7.6	87
Belgium	30,520	1	9.9	118
Bulgaria	111,000	5000	9.0	117
Denmark	43,100	9	5.1	59
France	544,000	140	56	774
Germany	357,000	3300	78	901
Greece	132,000	2500	10	106
Holland	41,900	4	15	188
Hungary	93,000	5000	11	124
Italy	301,280	5	57	551
Poland	312,680	30,000	38	610
Portugal	92,390	1500	10	120
Romania	237,500	5000	23	367
Spain	504,750	8000	39	439
Switzerland	41,290	150	6.7	82
Turkey	779,450	25,000	56	1576

SAS - [Output - (Untitled)]

File Edit View Tools Solutions Window Help

The CORR Procedure

4 Variables: area storks humans birth

Simple Statistics

Variable	N	Mean	Std Dev	Sum	Minimum	Maximum
area	17	219675	219561	3734470	28750	779450
storks	17	5059	8833	86009	1.00000	30000
humans	17	25.55882	23.52871	434.50000	3.20000	78.00000
birth	17	370.70588	410.04630	6302	59.00000	1576

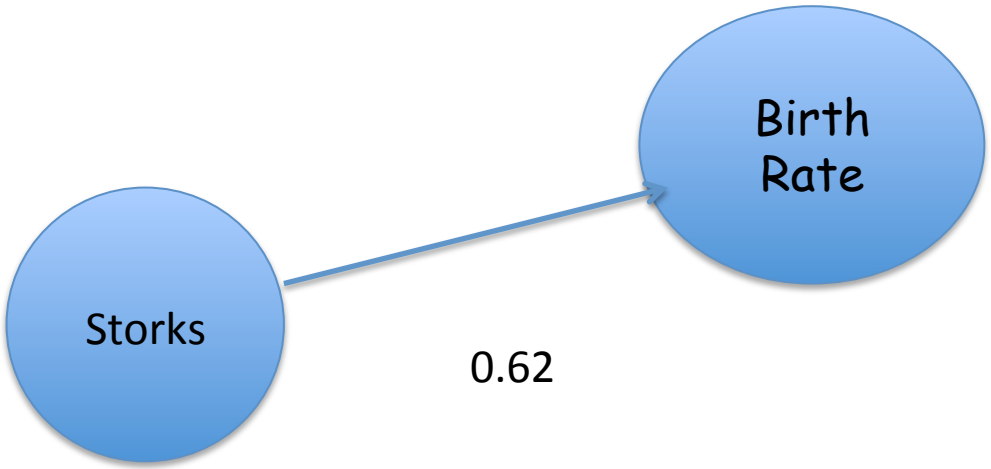
Pearson Correlation Coefficients, N = 17
Prob > |r| under H0: Rho=0

	area	storks	humans	birth
area	1.00000	0.57934 0.0148	0.81223 <.0001	0.92254 <.0001
storks	0.57934 0.0148	1.00000	0.35424 0.1630	0.62027 0.0079
humans	0.81223 <.0001	0.35424 0.1630	1.00000	0.85121 <.0001
birth	0.92254 <.0001	0.62027 0.0079	0.85121 <.0001	1.00000

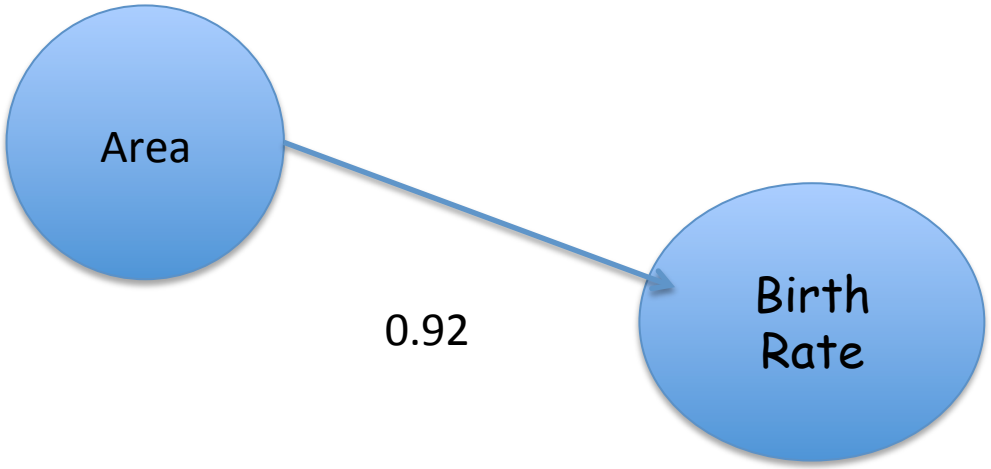
Output - (Untitled) Log - (Untitled) drew-stork

C:\Users\Dr. Ron D. Hays

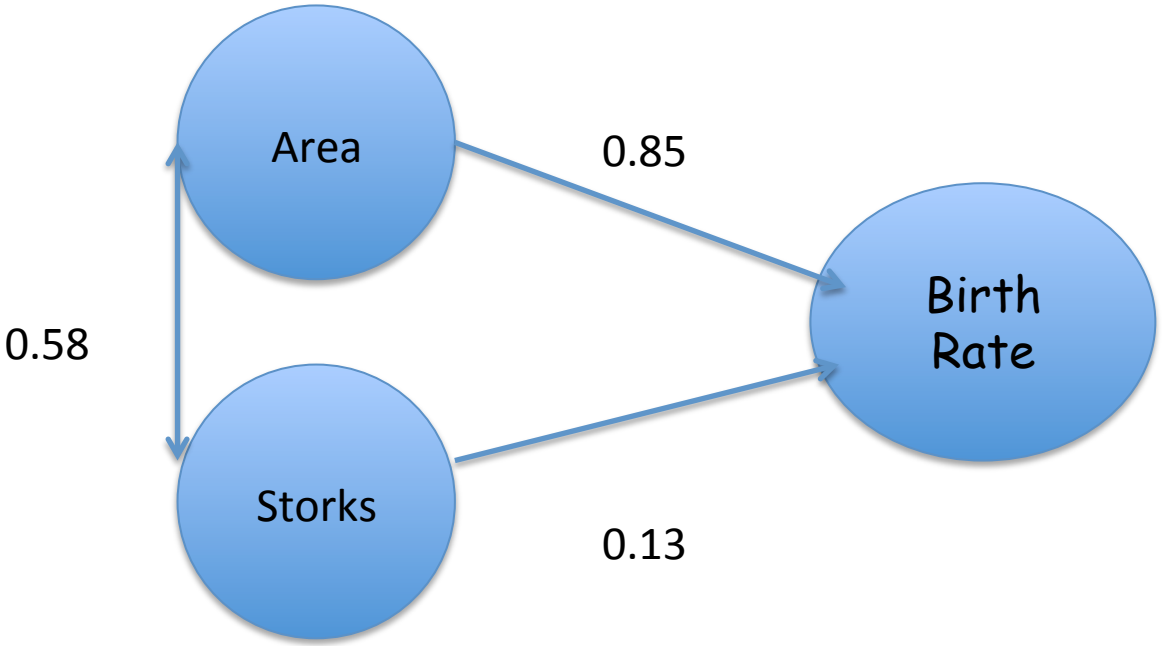
Do storks cause (increase the number of) births in Europe?



An extraneous variable, area



Standardized regression coefficients show that storks do not cause births



Factors that Compromise Internal Validity

- "On Stage" Effects
  - Social desirability
  - Evaluation apprehension
  - Demand characteristics
  - Looking bad
- Hawthorne effects
- Placebo effects
- (Researcher/teacher) Expectancy effects
- Personal relationship effects
- Recall bias
- Biased sample

"On Stage" Effects

- Study participants may begin to "act" when they know they are being observed.
- Social desirability
  - Participant tells the observer what they think they "should" say.
- Evaluation apprehension
  - Participant tries to do what mentally health people are supposed to do
- Demand characteristics
  - Participant picks up by subtle clues about what the researcher wants the study to show and behaves so as to please the researcher.

"On Stage" Effects

- Looking bad
  - Participant tries to look bad to sabotage research or because it might lead to personal gain (e.g., student who wants to be able to take exam late claims to be sick)

Hawthorne Effects

- Every time something was done to change the work routine, productivity increased initially but then went back to baseline.
- Participants are aware they are being studied and given special treatment so they work harder.

Placebo effects

- Just expecting a treatment to work can lead to improvement
- Power of suggestion by quacks and charlatans

Expectancy effects

- If the researcher expects people to behave in a certain way, it may come to pass by the way she behaves toward them.
- Pygmalion effect
  - Teachers were told that some students were "late bloomers" (randomly) and those students had greater improvement in IQ scores than students not so labeled.

Personal relationship effects

- The extent to which the researchers becomes known personally by the study participant may affect their behavior

Selective or distorted memory/ recall bias

- Study participant's memory may be distorted to fit her opinion
- Loss of memory for distant events
- Telescoping

Biased sample

Coverage Error

- Does each person in population have an equal chance of selection? (sample frame)

Sampling Error

- Are only some members of the population sampled?

Nonresponse Error

- Do people in the sample who respond differ from those who do not?

(Measurement Error--Are inaccurate answers given to survey questions?)

Methods of Control

- Unobtrusive measures
- Extended observation
- Cross-checking
- Deception
- Masked ("blind") measurement
- Placebo and demand characteristics control groups
- Controls for social desirability

Unobtrusive measures

- Measure wear and tear on carpets in museum to determine popularity of different exhibits
- Go through trash cans looking for discarded medicine bottles

Extended Observation and Cross-Checking

- Extended Observation
  - Effects of being observed diminish over time
- Cross-Checking
  - Multiple observers
  - Multiple time points

Deception

- Purpose of study secret kept secret
- In extreme, those observed are not told you are the researcher
- Misinform participants deliberately in order to get more honest answers
  - http://en.wikipedia.org/wiki/Milgram_experiment
  - http://en.wikipedia.org/wiki/Stanford_prison_experiment
  - http://en.wikipedia.org/wiki/Tuskegee_syphilis_experiment

Masked Measurement

- Researcher doesn't know which group the participant is randomized to be in
- Participant doesn't know either

Control Groups

- Demand characteristics control
  - Experimenter's opinion about what is presented is shared
- Placebo control

Controls for Social Desirability

- Use of well-written survey questions
- Forced choice between equally socially desirable options
- Socially desirable response scale

Listed below are a few statements about your relationships with others. How much is each statement TRUE or FALSE for you?

- I am always courteous even to people who are disagreeable.
- There have been occasions when I took advantage of someone.
- I sometimes try to get even rather than forgive and forget.
- I sometimes feel resentful when I don't get my way.
- No matter who I'm talking to, I'm always a good listener.

Study Designs

- Randomized experimental designs
- Observational Study
  - Quasi-experimental designs
    - Matching
    - Statistical control (e.g., propensity scores)

TABLE 1. Characteristics (%) the Sample Included and the One Dropped From the Analysis Due to Nonmatching

Variable	All Hispanics	Matched Sample Used	English Speakers Not Matched	Spanish Speakers Not Matched
Sample size	10,078	703	7793	879
Age (%)			*	*
18-44 y	1	0	1	1
45-64 y	8	0	8	11
65-69 y	25	29	24	21
70-74 y	30	34	29	30
75-79 y	20	24	20	20
80+ y	16	14	16	16
Missing	1	0	1	1
Sex (%)			*	
Male	45	43	45	44
Female	54	57	54	55
Missing	1	0	1	0
Education (%)			*	*
Eight grade or less	36	60	28	61
Some high school	19	14	21	13
High school graduate	23	17	25	11
Some college	12	6	14	7
College graduate	4	2	4	4
More than 4y of college	3	1	4	1
Missing	3	0	4	4
Self-rated health (%)			*	*
Excellent	8	7	8	15
Very good	17	8	20	10
Good	33	35	34	27
Fair	32	45	28	38
Poor	9	6	9	9
Missing	1	0	1	1

As the matched sample was based on an identical one-to-one match of these characteristics, the Spanish and English matched sample have the same characteristics. Test statistics conducted are for comparison of the nonmatched Spanish and English speakers to the matched sample.

Thank you