**Excluding Those Who Report Having "Syndomitis" or "Chekalism" Improves the Reliability of PROMIS-29+2 v2.1 Scales**

# Ron D. Hays, Ph.D.

October 23, 2023 (KC 305)

11:15 am – 12:15 pm session



**9th Annual PROMIS International Conference, Banff, Alberta, Canada**

**https://labs.dgsom.ucla.edu/hays/pages/presentations**
**https://www.jmir.org/2023/1/e46421**
.

# Acknowledgements

*I have no conflicts of interest*

https://www.jmir.org/2023/1/e46421

# Online Research "Panels"

- ➢ Quick data collection
- ➢ Large samples
- ➢ Cost-effective

*Probability/Opt-in Panels*

Growth of panels for data collection is: **"one of the most compelling stories of the last decade"** Baker et al. (2013, p. 715).

Summary report of the AAPOR Task Force on Non-probability Sampling. Journal of Survey Statistics and Methodology, 1, 90–143.

# Amazon Mechanical Turk (MTurk) Data Collection (2021)

- Opt-in (crowdsourcing) platform hosted by Amazon.
- Jobs or tasks are referred to as human intelligence tasks and include:
  - Completing surveys
  - Coding
  - Identifying content in images or videos
  - Writing product descriptions

- MTurk workers compared to U.S. general population
  - Younger age
  - More educated
  - Less likely to be have household income of $100k or higher
  - Worse self-reported health

# MTurk Questionnaire (~190 items)

- ==61 PROMIS® items (including PROMIS-29)==
- ==9 demographic items, 24 health conditions==
- Back pain-targeted measures
  - 7 chronicity items, 13 pain management
  - PEG (Pain intensity, interference with Enjoyment of life, interference with General activity),
  - Oswestry Disability Index (ODI)
  - Roland-Morris Disability Questionnaire (RMDQ)
  - Orebro Musculoskeletal Pain Questionnaire (OMPQ)
  - Subgroups for Targeted Treatment (STarT) Back Tool
  - Graded Chronic Pain Scale (GCPS) disability score.

# Data Quality Steps

- Quality workers - $\geq$95% approval rating; 500+ HITs
- Deployed in small batches (9 surveys per hour) – reduce selection bias
- Screened for back pain without revealing this was our target to minimize reporting it just to get paid
- Eliminated those <1 second per item
- Checked MTurker forums (e.g., www.mturkcrowd.com)

*mturk forum*

# Data Quality Steps

- Quality workers - $\geq$95% approval rating; 500+ HITs
- Deployed in small batches (9 surveys per hour) – reduce selection bias
- Screened for back pain without revealing this was our target to minimize reporting it just to get paid
- Eliminated those <1 second per item
- Checked MTurker forums for chatter on survey
- Pilot study showed that 20% of respondents endorsed all health conditions
  - So we added fake conditions (Syndomitis, Chekalism)

# Have you EVER been told by a doctor or other health professional that you had

1) hypertension

2) high cholesterol

3) heart disease

4) angina

5) heart attack

6) stroke

7) asthma

8) cancer

9) diabetes

10) chronic obstructive pulmonary disease

11) arthritis

12) anxiety disorder

13) depression

14) Syndomitis

# Do you currently have

1) allergies or sinus trouble

2) back pain

3) sciatica

4) neck pain

5) trouble seeing

6) dermatitis

7) stomach trouble

8) trouble hearing

9) trouble sleeping

10) Chekalism

**Characteristics of Those Endorsing/Not Endorsing a Fake Health Condition at Baseline**

| Variable | Did not Endorse Fake Health Condition (n = 5836) | Endorsed Fake Health Condition (n = 996)* |
|---|---|---|
| **Gender** | | |
| Female | 46% | 32% |
| Male | 53% | 67% |
| | | |
| **Non-White** | 18% | 28% |
| **Age** | 40 years old | 38 years old |
| **Number of conditions** | 4 | 15 |
| | | |

**\* 15% reported having 1 or both fake conditions, and were more likely to be male, non-White, younger, and report more health conditions.**

**Internal Consistency Reliability of PROMIS Scales at Baseline**

| Scale | Did not Endorse Fake Health Condition (n = 5836) | Endorsed Fake Health Condition (n = 996) |
|---|---|---|
| Physical function | 0.89 | 0.69 |
| Pain interference | 0.94 | 0.80 |
| Fatigue | 0.92 | 0.80 |
| Depression | 0.92 | 0.78 |
| Anxiety | 0.90 | 0.78 |
| Sleep disturbance | 0.84 | -.27* |
| Ability social roles/ activities | 0.92 | 0.77 |
| Cognitive function | 0.77 | 0.65 |

*Sleep109 & Sleep116, and Sleep20 & Sleep44 positively correlated.

## PROMIS Scale Means at Baseline

| Scale | Did not Endorse Fake Health Condition (n = 5836) | Endorsed Fake Health Condition (n = 996) | Overall Sample (n = 6832) |
|---|---|---|---|
| Physical function | 49 | 41 | 48 |
| Pain interference* | 51 | 63 | 53 |
| Pain intensity* | 52 | 64 | 54 |
| Fatigue* | 50 | 58 | 51 |
| Depression* | 53 | 63 | 54 |
| Anxiety* | 54 | 63 | 56 |
| Sleep disturbance* | 50 | 51 | 50 |
| Ability social roles/ activities | 53 | 43 | 52 |
| Cognitive function | 50 | 47 | 49 |
| P-29 Physical Health Summary | 49 | 40 | 48 |
| P-29 Mental Health Summary | 50 | 39 | 48 |
| PROPr (SD 0.25) | 0.45 | 0.20 | 0.41 |

*Higher scores represent worse health.

T-scores: U.S. mean = 50, SD = 10 for all measures except

PROPr: U.S. mean = 0.52, SD = 0.24, possible range: -0.022 to 1.00

**Internal Consistency Reliability of PROMIS Scales at 3 Months**

| Scale | Did not Endorse Fake Health Condition (n = 972) | Endorsed Fake Health Condition (n = 59) |
|---|---|---|
| | | |
| Physical function | 0.92 | 0.53 |
| Pain interference | 0.95 | 0.76 |
| Fatigue | 0.94 | 0.77 |
| Depression | 0.93 | 0.81 |
| Anxiety | 0.92 | 0.80 |
| Sleep disturbance | 0.88 | -.21 |
| Ability to participate in social roles and activities | 0.94 | 0.78 |
| Cognitive function | 0.70 | 0.44 |

6% of the 3-month survey respondents endorsed a fake condition.

**PROMIS Scale Means at 3-Months Follow-up**

| Scale | Did not Endorse Fake Health Condition (n = 972) | Endorsed Fake Health Condition (n = 59) | Overall Sample (n = 1031) |
|---|---|---|---|
| Physical function | 46 | 41 | 46 |
| Pain interference* | 54 | 62 | 55 |
| Pain intensity* | 56 | 62 | 56 |
| Fatigue* | 54 | 57 | 54 |
| Depression* | 55 | 62 | 55 |
| Anxiety* | 56 | 63 | 56 |
| Sleep disturbance* | 53 | 51 | 54 |
| Ability social roles/ activities | 51 | 44 | 51 |
| Cognitive function | 50 | 46 | 50 |
| P-29 Physical Health Summary | 47 | 40 | 46 |
| P-29 Mental Health  Summary | 46 | 41 | 46 |
| PROPr | 0.37 | 0.22 | 0.37 |

*Higher scores represent worse health.

# Summary

- Based on the 15% faker rate at baseline and 6% at 3-months, we estimate a 25%* faker rate in the MTurk sample.

- Excluding those who endorsed a fake condition improved
  - ➤ reliability of measurement
  - ➤ estimated mean health
    - ➤ PROMIS-29+2 v2.1 T-scores by 1-2 points
    - ➤ PROPr preference-based score by 0.04 (~0.16 effect size).



* f = % of fakers; p = probability of getting caught using the fake conditions
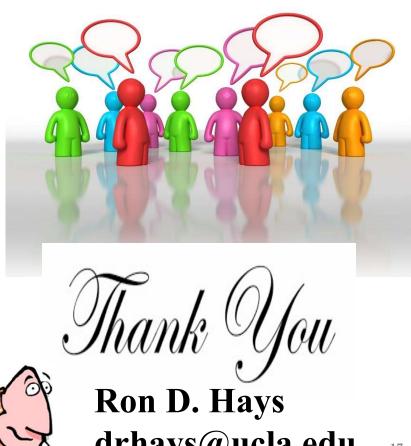p(f)=0.146 and p(f)(1-p) = 0.061… then 0.146(1-p) = 0.061 -> p =0.58, f=0.061/(p(1-p))

# Implications

- Asking about fake health conditions can help screen out respondents who misrepresent themselves.

- Its usefulness could fade over time if information about it spreads among survey respondents.

  - e.g., urban dictionary warns readers not to select "Bindro" on surveys of drug use because doing so "voids the whole test."

  - https://www.urbandictionary.com/define.php?term=Bindro

# Implications

- Asking about fake health conditions can help screen out respondents who misrepresent themselves.

- Its usefulness could fade over time if information about it spreads among survey respondents.
    - e.g., urban dictionary warns readers not to select "Bindro" on surveys of drug use because doing so "voids the whole test."
    - https://www.urbandictionary.com/define.php?term=Bindro

Thank You

**Ron D. Hays**
**drhays@ucla.edu**