Dr. Strangelove or How I Learned to Stop Worrying and Love the Replication Crisis

Jeffrey Inman

University of Pittsburgh

February 16, 2018

Under Siege

- Replication crisis
- p-hacking
- p-curve
- Researcher df
- Data police

Poster used in McCarthy's successful 1946 Senate campaign against his Democratic rival, Howard McMurray. *Marquette University Archives*

Proffered Solutions

- Increase sample size
- Decrease critical p-value (α)
- Preregistration
- Increase power

Additional Approaches

- Tackle a consumer-relevant problem
- Build the case through a robustness philosophy

Tackle a Consumer-Relevant Problem

#Influencers

What do these companies have in common?

What does CX stand for?

- A) Consumer Excess
- B) Canadian Express
- C) Customer Experience

What does CX stand for?

- A) Consumer Excess
- B) Canadian Express
- **C) Customer Experience**

WIKIPEDIA
The Free Encyclopedia

- **Customer experience** (**CX**) is the product of an interaction between an organization and a customer over the duration of their relationship.[1] This interaction is made up of three parts: the customer journey, the brand touchpoints the customer interacts with, and the environments the customer experiences (including digital environment) during their experience.[*citation needed*]

Additional Approaches

- ✓Tackle a consumer-relevant problem

- o **Build the case through a robustness philosophy**

Robustness

- James Woodward (2006)
  - Inferential robustness: Degree to which inference from a dataset depends on underlying assumptions
  - Measurement robustness: Triangulation of a finding by different means of measurement
  - Causal robustness: Stability in the inherently noisy real world
    - Field experiments, Quasi experiments, etc.

Inferential Robustness

- Typically what we mean by sensitivity analysis
  - For example, how results vary as a function of choice of background variables
  - Leamer (1983) investigated the effect of capital punishment on the murder rate.
    - Each execution deters twenty-two murders
    - But using a different set of (but no less plausible) control variables, each execution results in twelve additional murders
    - Results depended on which control variables were entered into the analysis – **low inferential robustness**

Aldrich's Robustness Theorem (1989)

- Assume we are interested whether a hypothesis is true with probability greater than or equal to some value (Type 1 error <.05),
- If one of the potential background assumptions is true (completeness), and
- If the hypothesis is supported under all of the assumptions Ai (robustness)
- Then the probability of the hypothesis being true is high.

An Example

- Our hypothesis: Using a home assistant device like the Amazon Echo to shop leads to more hedonic purchases
- There is equivocal theoretical/empirical support for the effect of weekend shopping (completeness)
- If our hypothesis is supported under both the inclusion and the omission of weekend in the analysis (robustness), then confidence in our hypothesis is increased.
- **Should include a robustness analysis discussion and report results**

Other Empirical Aspects Could Benefit from a Robustness Philosophy

- Operationalization of the DV
  - Raw $ or Log $
  - Donation amount or incidence
- Operationalization of the IV
  - Scale points (1-7)
  - Mid-point of each range [10 if 1= < $20K/year]
- Treatment of outliers
  - Rather than defending dropping observations, could show the cutoff point for their effect to manifest (i.e., “sample reduction floodlight analysis”)

Measurement Robustness

- Persistence of a result to variations in measurement
- While a given procedure is subject to various sources of error, studies using different procedures are unlikely to be subject to the same kinds of error
- Measurement robustness requires testing the hypothesis with
  - Different procedures
    - Not a stylized procedure that just tweak and rerun
  - Different contexts
    - Lab and field
  - Different DVs
    - Preferably including consequential DVs
  - Different IV manipulations

What's the difference?

- In inferential robustness, various aspects of the analysis are varied for the same data.
  - **Within-Study Robustness**
- In measurement robustness, different measurement procedures are used and these produce different bodies of data (multiple studies)
  - **Cross-Study Robustness**

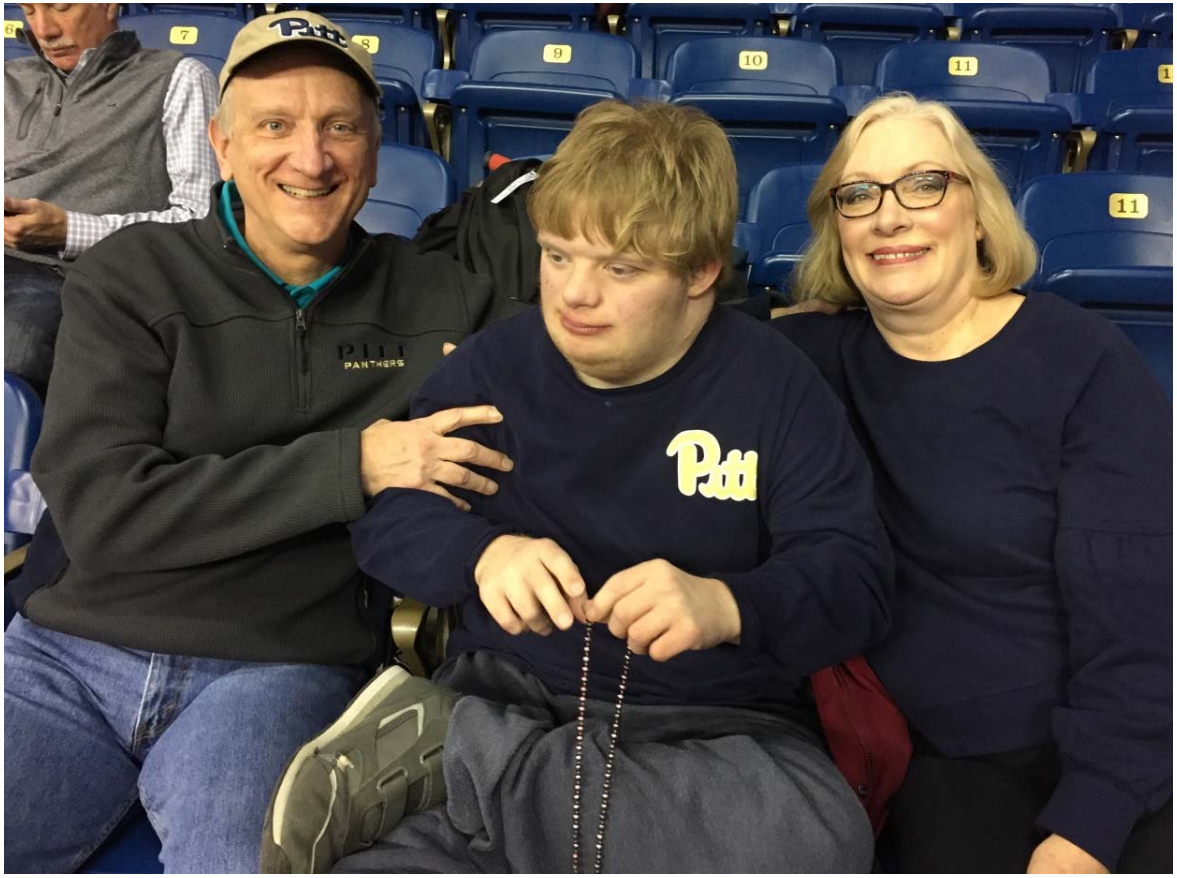
Building the Case via Robustness Philosophy

- Empirical Studies - "The Body of Evidence"
  - Package that tests the predictions, triangulates on the focal phenomenon, and reinforces shortcomings (*Ed stds*)
  - Balance validities (internal, construct, nomological, ecological, external, etc.)
  - Within-study and cross-study robustness

Approaches to Greater Impact

- ✓Tackle a consumer-relevant problem
- ✓Build the case via a robustness philosophy
- ✓**Meaningful consumer-relevant insights**

Closing Thoughts




SCP
SOCIETY FOR
CONSUMER
PSYCHOLOGY