Mass Pre-registered Replications of Classic JDM Findings in Two HKU Courses: Second Phase Findings

(December 20, 2018)

Fili (Gilad Feldman)

http://giladfeldman.org
gfeldman@hku.hk
Twitter: @giladfeldman
Today: Why we're here

• Present the mass pre-registered replication project phase 2.
• Thank the students and TAs for their hard work.
• Share what we did and insights, if you wish to do similar things.
• Invite you to join us.
Principles I set out for courses

• Student-led. Students seek out answers and solutions.
• Students do real science, hands-on, publishable work.
• Students do projects that have real impact.
• Students take part in the scientific process (peer review, pre-reg, etc.).
• Students use latest tools/trends in psychological science.

• No more books with problematic findings. Focus on recent articles.
• No instructor "truth". Discussion based. I learn together with them.
Open-science principles

Complete transparency
• I share everything I have.
• Everything we do is shared with everyone.
• Everything will be made available to the world.

Rigor & Collaboration
• Focus on getting things right.
• Collaborative effort
• No hierarchy.

Rethinking science process
• Rigor/accuracy > novelty.
• Simplicity > complexity.
Let's start from last semester
On Twitter:

https://twitter.com/giladfeldman/status/1033267774351523842

Gilad Feldman
@giladfeldman

For a presentation, I summarized all the JDM preregistered replications I ran since 2016, when I decided to prioritize these to examine the empirical foundations of my JDM research domain

See mgto.org/pre-registered ...

So far: 14 successful, 3 semi-successful, 4 need reexamination

UI replicants

Effect (Kahnejian & Tversky, 1982): Replicated several times (> 8).

Effect (Zoehnder et al., 2002): Replicated Experiment 1 several times (Inks, Oppenau, Mirnil, & Baron, 2011).

Replicated two scenarios from E

nately effect (Kahnejian & Miller, 1986): Replicated two experiments. (Kahnejian & Miller, 1986)

nately effect (Seta et al., 2001): Replicated 3 times.

rreplication (Nutini, 1987): Replicated the main experiment.

rexpansion & transaction demand (Mandel, 2002). Replicated Experiment [9].

Impressions

2,081

Total engagements

164

Will present JDM mass-replication effort findings in 2 HKU courses on MTurk & Hong Kong samples. You’re invited.

Initial results:

10 successful

1 found unexpected effect

2 likely “failed” 1 unsure

1 flawed (fail)

Fairly consistent MTurk-HK effects

Summary: ~11/15 is great news

http://twitter.com/giladfeldman/status/993056624007823360

Gilad Feldman
@giladfeldman

Mass-Replication Project of Classic JDM Findings in two HKU Courses

3:30 p.m. - 4:30 p.m., May 10, 2018 (Thursday)

Run 13, A/F, The Jockey Club Tower | Centennial Campus | The University of Hong Kong

Dr. Gilad Feldman
Assistant Professor
Department of Psychology
University of Hong Kong

Abstract

Reproducibility and replicability are at heart of science, yet increasing evidence from recent years suggests that many of the findings in psychological science are reproducible and non-
replicable in what some termed as a “replication crisis” and a new movement calling for significant changes in the way we do science. How can we do better? How can we inform colleagues and students about these issues and train students to rigorously reproducible science? In this talk I will discuss a mass-replication effort I headed in HKU courses PSYC208 and PSYC502 to conduct 1531 pre-registered replications of classic findings in judgment and decision-making literature. With the help of four wonderful TAs students analyzed articles and tried to reproduce methods and materials to conduct effect-size calculations and power analyses, design Quirics experiments, and adopt latent tools and templates and pre-register the replications on the Open Science Framework. We then ran the experiments on (1) a limited sample of HKU students (N = up to 99) and (2) high-power Amazon Mechanical Turk American online samples (power ~ 0.95-0.99, N = 300-800).

PSYC208 course was overhauled to discuss the replication crisis in-depth and involve students in thinking of its implications and improving. I will briefly review the process, the overarching course designs, the students’ mass pre-registered replications findings, as well as my main take-aways from the process. I conclude the experience as an invaluable learning experience, not only for the students, but also for myself and the team, with insights and contributions to the literature and the academic community.
Previous semester... (link)

14 successful
3 semi-successful
4 unsuccessful

17/21 = 80%

Successful replications
1. Action effect (Kahneman & Tversky, 1982): Replicated several times (> 8).
2. Inaction effect (Zeelenberg et al., 2002): Replicated Experiment 1 several times (> 4).
3. Omission bias (Spranca, Minsk, & Baron, 1991): Replicated two scenarios from Experiment 1.
4. Exceptionality effect (Kahneman & Miller, 1986): Replicated two experiments (hitchhiker and car accident scenarios).
5. Exceptionality effect (Seta et al., 2001): Replicated 3 times.
10. Bias blind spot / Actor observer bias (Pronin et al., 2002): Replicated Experiments 1b and 2. Found an effect for shortcomings when none was expected.
11. Actor-observer bias in free will attributions (Pronin et al., 2010): Replicated twice in US/HK.
13. Preference for indirect harm (Rozzman & Baron, 2002): Replicated Experiment 2 twice (US/HK) and Experiment 3 once (HK).

Semi-successful replications
1. Exceptionality effect (Miller & McFarland, 1986): Replicated twice using a regret DV, but not using original compensation DV.
2. Doing/allowing morality asymmetry (Dushman et al., 2008): Replicated Experiment 1 in US but not in a small HK sample.
3. Folk intentionality (Malle & Knobe, 1997): Twice (US/HK) found an effect when none was expected (actor-observer asymmetry).

Unsuccessful replications, needs to revisit further
1. Endowment effect and goal relevance (Irmai, Waskul, & Trope, 2013): Failed to replicate the second experiment in paper.
2. Force-intention in moral judgment (Greene et al., 2002): Failed to replicate Experiment 1b in both HK and US samples.
3. Actor-observer bias (Pronin et al., 2007): Failed to replicate Experiment 1 twice (US/HK).
All materials from courses were made available on OSF


Course materials for concluded courses:

- **HKU PSYC2020 – Spring 2017-8 – Fundamentals of social psychology – Gilad Feldman**: Social psychology and judgment decision-making with lots of funky in-class experiments. (Cite as DOI 10.17605/OSF.IO/E4PXZ)
  - A list of all PSYC2020 in-class experiments conducted.

- **HKU PSYC3052D – Spring 2017-8 – Advanced social psychology – Gilad Feldman**: Discussing the science crisis and challenging students to reflect on the crisis and implications. (Cite as DOI 10.17605/OSF.IO/BFETX)
I recently ran this...

**JAMOVI / JASP workshop**

**HKU**

**December 6, 2018**

Workshop Google Drive materials

https://tinyurl.com/hku2018jamovi

Mailing list:

https://tinyurl.com/giladmailinglist

JAMOVI/JASP/R collaborative manual:

https://tinyurl.com/hkujamoviguide

Pre-registered replication guide collaborative manual:

https://tinyurl.com/replicationguide

Other resources on **my website** (giladfeldman.org)
And asked this...

Poll is full and no longer accepting responses

- Yes, very: 18%
- Yes, heard of it: 50%
- No, what's that?: 33%
And asked this...
And asked this...

Are you familiar with the "new statistics" (AKA, effect sizes and confidence intervals)?

- Yes, apply them regularly: 38%
- Yes, but don't use them regularly: 48%
- No, what's an effect size? What's confidence intervals?: 15%

Poll is full and no longer accepting responses.
And asked this...

What stats (programming) software do you use for your science?

- Excel: 6%
- SPSS: 79%
- SAS
- STATA
- Matlab
- R / RStudio: 6%
- JAMOVI: 9%
- Python

When poll is active, respond at PollEv.com/giladfeldman999
Text GILADFELDMAN999 to 6452 2795
Brian Nosek
@BrianNosek

Many Labs 2: 28 findings, 60+ samples, ~7000 participants each study, 186 authors, 36 nations.

Successfully replicated 14 of 28 psyarxiv.com/9654g

ML2 may be more important than Reproducibility Project: Psychology.

2016

Michael Inzlicht
@minzlicht

Big news: RRR of ego depletion reveals no effect. Nada. Zip. Nothing. @ME_McCullough called it first #spsp2016

March 2018

Summary of Vohs RRR findings

- Effect d = 0.08 (very weak ~r=.03)
- Comparable/weaker than first RRR
- Need N = 4908 (power = 80%; alpha = 0.05) to study these effects
Brian Nosek's Summary

Across 6 large-scale replication projects, replication rate is 90 of 190 (47%).

ML1: econtent.hogrefe.com/doi/full/10.10...
ML2: psyarxiv.com/9654g
ML3: sciedirect.com/science/article ...
SSRP: nature.com/articles/s4156 ...
EERP: science.sciencemag.org/content/351/62 ...

RPP:

Evaluating the replicability of social science exper...
Camerer et al. carried out replications of 21 Science and Nature social science experiments, successfully replicating 13 out of 21 (62%). Effect sizes of replication...
nature.com

Also, we replicated evidence that surveys & markets can anticipate replication success. There are now 4 studies showing evidence for this.
Scholars worldwide: Do we think we have a “crisis”?

90%: Yes
No, science’s reproducibility problem is not limited to psychology

In Medicine, the Science Has Stopped Working

By PASCAL-EMMANUEL GOBRY | November 15, 2017 4:25 PM

Can Reproducibility in Chemical Research be Fixed?

The replication crisis has engulfed economics

Cancer Research Is Broken

There’s a replication crisis in biomedicine—and no one even knows how deep it runs.
I am convinced we're in a crisis.
It's okay if you're not.

At the very least ...

make sure you are knowledgeable of what's happening and have an informed opinion.
Course 2nd run: Advanced Social Psychology (PSYC3052)

Focusing on the replication / reproducibility crisis / credibility revolution and open-science.

Each week:
• 2 mandatory + 2/3 optional readings
• Facilitator group summarizes readings, and issues report of class conclusions
New course: Judgement and Decision Making (PSYC2071)

- Weeks 3-10: 2/3 readings each week.
- Pairs assigned to each article.
- Pairs first run in-class experiments on classmates.
- Students conduct replications independently, with peer-review.
- Students design individual extensions on the replications.

<table>
<thead>
<tr>
<th>C1</th>
<th>Date</th>
<th>Topic</th>
<th>TA</th>
<th>Tasks due end of week (Sunday 11:59pm)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>06/09</td>
<td>Judgment &amp; Decision making: Introduction lecture #1</td>
<td></td>
<td>Register for HKU Qualtrics account hku.psych.qualtrics.com</td>
</tr>
<tr>
<td>2</td>
<td>13/09</td>
<td>Judgment &amp; Decision making: Introduction lecture #2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>20/09</td>
<td>Asymmetries (Time / self-other / outcome bias)</td>
<td>T1 Qualtrics survey</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>27/09</td>
<td>Choices</td>
<td>T2 Article analysis (Wednesday class)</td>
<td>Deadline 30/09 11:59pm: Qualtrics survey</td>
</tr>
<tr>
<td>5</td>
<td>04/10</td>
<td>Intuitive statistics</td>
<td>T2 Article analysis (Tuesday class)</td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>11/10</td>
<td>Instructor session on open-science pre-reg/extension</td>
<td>T3 Pre-registration</td>
<td>Deadline 14/10 11:59pm: Article analysis</td>
</tr>
<tr>
<td>7</td>
<td>25/10</td>
<td>Hindsight/money</td>
<td></td>
<td>Deadline 28/10 11:59pm: Peer review of pre-registration (includes article analysis and Qualtrics survey)</td>
</tr>
<tr>
<td>8</td>
<td>01/11</td>
<td>Escalation / effort</td>
<td></td>
<td>Deadline 06/11 2:00pm: Revised pre-registration (includes article analysis and Qualtrics survey)</td>
</tr>
<tr>
<td>9</td>
<td>08/11</td>
<td>Emotions &amp; Counterfactuals</td>
<td>T4 Data analysis</td>
<td>(Students receive data collection results)</td>
</tr>
<tr>
<td>10</td>
<td>15/11</td>
<td>Course JDM summary + Academic writings and journal submissions</td>
<td>T5 Writing APA style replication reports</td>
<td>Deadline 20/11 11:59pm: Data analysis</td>
</tr>
<tr>
<td>11</td>
<td>22/11</td>
<td>Presentations</td>
<td></td>
<td>Deadline 02/12 11:59pm: Submission of final report</td>
</tr>
<tr>
<td>12</td>
<td>29/11</td>
<td>Presentations</td>
<td></td>
<td></td>
</tr>
<tr>
<td>13</td>
<td>06/12</td>
<td>No class</td>
<td></td>
<td>Deadline 09/12 11:59pm: Submission of peer review on final report</td>
</tr>
<tr>
<td>14</td>
<td>13/12</td>
<td>No class</td>
<td></td>
<td>Deadline 16/12 11:59pm: Submission of revised final report</td>
</tr>
</tbody>
</table>
Process

Qualtrics survey
- TA/Instructor feedback

Guide
- Blue: students
- Green: peers/classmates
- Red: TA/Instructor

Article Analysis
- TA/Instructor feedback

Pre-registration Report

Revised pre-registration Report

Peer review, class pre-test

Instructor: Pre-registration & data collection

Data Analysis
- TA/Instructor feedback

Final report

Peer review

Revised Final report

Project joint class presentations
Overall view of projects

• 51 reports, 18 projects pre-registered and data collected
• ~5900 participants recruited on MTurk (turkprime.com)
• 56,498.5 HKD (7215USD$) spent. ~1/2 of my seed funds.
What changed? We got better

• 2 undergrad advanced courses with individual projects

• 18 replication projects

• Higher complexity / risk
  – Articles are more complex / complicated.
    • Advanced statistics - higher expectations.
What changed? We got better

• Process / structure

• Collaborative manuals
  – Pre-registered replication guide
  – JAMOVI / JASP/ R guides
  – Extension guide

• Workshops / materials
  – Much more comprehensive materials. Full examples.
  – JAMOVI / JASP/ R workshop

• Twitter prediction markets
More structured: Design and Extensions

A detailed guide on how to analyze design and add extensions.

Three types of extensions on top of original design:

• Adding an individual differences predictor
• Adding a DV
• (adding an IV)

<table>
<thead>
<tr>
<th>IV1: Outcome manipulation</th>
<th>IV1: Experimental 1 condition</th>
<th>IV1: Experimental 2 condition</th>
<th>IV1: Control condition</th>
</tr>
</thead>
<tbody>
<tr>
<td>IV2: Self-other manipulation</td>
<td>Title example: Negative outcome condition</td>
<td>Title example: Positive outcome condition</td>
<td>Title example: No outcome</td>
</tr>
<tr>
<td>Manipulation example: Investment described as resulting in a loss of money</td>
<td>Manipulation example: Investor scenario described as resulting in a gain/win of money</td>
<td>Manipulation example: Investor scenario described with no indication of outcome</td>
<td></td>
</tr>
</tbody>
</table>

2 IVs - IV1 (1/2/control) x IV2 (A/B):

<table>
<thead>
<tr>
<th>IV2: Experiment A condition</th>
<th>Dependent variable</th>
<th>IV2: Experiment A condition</th>
<th>Dependent variable</th>
</tr>
</thead>
<tbody>
<tr>
<td>Title example: self condition</td>
<td>Title example: Evaluations of investment decision</td>
<td>Title example: other condition</td>
<td>Title example: Evaluations of investment decision</td>
</tr>
<tr>
<td>Manipulation example: Investment described as made by self - you're the investor</td>
<td>Specific DV item: Please evaluate your investment decision on a scale of 0-6 (very bad to very good)</td>
<td>Manipulation example: Investment described as made by other - someone else is the investor</td>
<td>Specific DV item: Please evaluate the person's investment decision on a scale of 0-6 (very bad to very good)</td>
</tr>
</tbody>
</table>
Inclass demonstrations
Of the experiments
+
Discussion
Compassion fade and psychic numbing


Two versions

Given a middle-class Hong Kong student status, how much money would you be willing to donate to save Rokia?

Given a middle-class Hong Kong student status, how much money would you be willing to donate to save Rokia and Moussa?
MENTIMETER

Go to www.menti.com and use the code 127069

Given a middle-class Hong Kong student status, how much money would you be willing to donate to save
Inclass voting results

- Version A: Rokia - 332.9
- Version B: Rokia and Moussa - 466.9
- Nothing (0HKD) - 196.8
- Version A: 2,000 birds - 378
- Version B: 20,000 birds - 378
Prediction markets: Setup

Dec 20 I will summarize findings from 18 replications we conducted this semester. We wanted prediction markets, but were overwhelmed by work just running the replications.

Instead, a Twitter vote, you're welcome to rate replicability. Below are studies we attempted to replicate.
Prediction markets: Attention

Gilad Feldman @giladfeldman · Dec 11
Dec 20 I will summarize findings from 18 replications we conducted this semester. We wanted prediction markets, but were overwhelmed by work just running the replications.

Instead, a Twitter vote, you're welcome to rate replicability. Below are studies we attempted to replicate.

<table>
<thead>
<tr>
<th>Impressions</th>
<th>26,337</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total engagements</td>
<td>468</td>
</tr>
<tr>
<td>Detail expands</td>
<td>221</td>
</tr>
<tr>
<td>Profile clicks</td>
<td>101</td>
</tr>
<tr>
<td>Likes</td>
<td>80</td>
</tr>
<tr>
<td>Retweets</td>
<td>61</td>
</tr>
</tbody>
</table>
## Prediction markets: Results

<table>
<thead>
<tr>
<th>St</th>
<th>Authors</th>
<th>Studies</th>
<th>Bias</th>
<th>Votes</th>
<th>Success</th>
<th>Failure</th>
<th>Direction</th>
<th>binom pval</th>
<th>Success sig&lt;&gt;0.5</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Baron, &amp; Hershey, 1988</td>
<td>1</td>
<td>Outcome bias</td>
<td>93</td>
<td>76%</td>
<td>24%</td>
<td>Success</td>
<td>.000</td>
<td>TRUE</td>
</tr>
<tr>
<td>2</td>
<td>Epley &amp; Gilovich 2006</td>
<td>1b</td>
<td>Anchoring-and-adjustment heuristic</td>
<td>114</td>
<td>88%</td>
<td>12%</td>
<td>Success</td>
<td>.000</td>
<td>TRUE</td>
</tr>
<tr>
<td>3</td>
<td>Epstein, Lipson, Holstein, &amp; Huh 1992</td>
<td>1 &amp; 2</td>
<td>Irrational reactions to negative outcomes</td>
<td>64</td>
<td>55%</td>
<td>45%</td>
<td>Success</td>
<td>.532</td>
<td>FALSE</td>
</tr>
<tr>
<td>4</td>
<td>Fischhoff, 1975</td>
<td>2</td>
<td>Hindsight bias</td>
<td>97</td>
<td>98%</td>
<td>2%</td>
<td>Success</td>
<td>.000</td>
<td>TRUE</td>
</tr>
<tr>
<td>5</td>
<td>Hamill, Wilson, &amp; Nisbett, 1980</td>
<td>1</td>
<td>Insensitivity to sample bias</td>
<td>92</td>
<td>68%</td>
<td>32%</td>
<td>Success</td>
<td>.001</td>
<td>TRUE</td>
</tr>
<tr>
<td>6</td>
<td>Hsee &amp; Weber, 1997</td>
<td>1</td>
<td>Fundamental predictor error</td>
<td>77</td>
<td>70%</td>
<td>30%</td>
<td>Success</td>
<td>.001</td>
<td>TRUE</td>
</tr>
<tr>
<td>7</td>
<td>Hsee, 1998</td>
<td>1, 2, 4</td>
<td>Less is better</td>
<td>88</td>
<td>53%</td>
<td>47%</td>
<td>Success</td>
<td>.749</td>
<td>FALSE</td>
</tr>
<tr>
<td>8</td>
<td>Kruger etal, 2004</td>
<td>1 &amp; 2</td>
<td>Effort heuristic</td>
<td>73</td>
<td>74%</td>
<td>26%</td>
<td>Success</td>
<td>.000</td>
<td>TRUE</td>
</tr>
<tr>
<td>9</td>
<td>Kruger, Wirtz &amp; Miller 2005</td>
<td>2</td>
<td>First instinct fallacy</td>
<td>60</td>
<td>60%</td>
<td>40%</td>
<td>Success</td>
<td>.155</td>
<td>FALSE</td>
</tr>
<tr>
<td>10</td>
<td>Mellers, Hertwig, &amp; Kahneman, 2001</td>
<td>1</td>
<td>Conjunction effect</td>
<td>85</td>
<td>85%</td>
<td>15%</td>
<td>Success</td>
<td>.000</td>
<td>TRUE</td>
</tr>
<tr>
<td>11</td>
<td>Miller, &amp; McFarland, 1987</td>
<td>1</td>
<td>Pluralistic ignorance</td>
<td>109</td>
<td>68%</td>
<td>32%</td>
<td>Success</td>
<td>.000</td>
<td>TRUE</td>
</tr>
<tr>
<td>12</td>
<td>Schwarz, Strack, Hilton, &amp; Naderer, 1991</td>
<td>1</td>
<td>Relevance of Irrelevant information</td>
<td>91</td>
<td>48%</td>
<td>52%</td>
<td>Failure</td>
<td>.675</td>
<td>FALSE</td>
</tr>
<tr>
<td>13</td>
<td>Shafir, 1993</td>
<td>1 to 8</td>
<td>Choosing versus rejecting</td>
<td>82</td>
<td>48%</td>
<td>52%</td>
<td>Failure</td>
<td>.741</td>
<td>FALSE</td>
</tr>
<tr>
<td>14</td>
<td>Shafir, Diamond, &amp; Tversky, 1997</td>
<td>1 to 4</td>
<td>Money illusion</td>
<td>70</td>
<td>59%</td>
<td>41%</td>
<td>Success</td>
<td>.188</td>
<td>FALSE</td>
</tr>
<tr>
<td>15</td>
<td>Slovic &amp; Fischhoff, 1977</td>
<td>1</td>
<td>Hindsight bias</td>
<td>111</td>
<td>96%</td>
<td>4%</td>
<td>Success</td>
<td>.000</td>
<td>TRUE</td>
</tr>
<tr>
<td>16</td>
<td>Staw, 1976</td>
<td>1</td>
<td>Escalation of commitment</td>
<td>78</td>
<td>86%</td>
<td>14%</td>
<td>Success</td>
<td>.000</td>
<td>TRUE</td>
</tr>
<tr>
<td>17</td>
<td>Tversky &amp; Shafir, 1992</td>
<td>1</td>
<td>Disjunction effect</td>
<td>83</td>
<td>76%</td>
<td>24%</td>
<td>Success</td>
<td>.000</td>
<td>TRUE</td>
</tr>
<tr>
<td>18</td>
<td>Zeelenberg etal 1996</td>
<td>1</td>
<td>Regret aversion</td>
<td>75</td>
<td>73%</td>
<td>27%</td>
<td>Success</td>
<td>.000</td>
<td>TRUE</td>
</tr>
</tbody>
</table>
Prediction markets: Reactions

Michał Białek @mbialek82 · 4m
Replies to @giladfeldman
More scientists believe in hindsight bias than in global warming....
Before results: How do we evaluate replications?


Original Study

<table>
<thead>
<tr>
<th>Replication #1</th>
<th>Replication #2</th>
<th>Replication #3</th>
<th>Replication #4</th>
<th>Replication #5</th>
<th>Replication #6</th>
<th>Replication #7</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Replication outcome

- signal - consistent
- signal - inconsistent, larger
- signal - inconsistent, smaller
- signal - inconsistent, opposite
- no signal - consistent
- no signal - inconsistent
- no signal - inconsistent (less precise)
Time for some data.

Ready for the replication results?

WARNING: Preliminary student findings, need to be rechecked and verified
First, we were pretty well powered.
Results (with extensions)

<table>
<thead>
<tr>
<th>St</th>
<th>Authors</th>
<th>Studies</th>
<th>Bias</th>
<th>Votes</th>
<th>Success</th>
<th>Failure</th>
<th>Direction</th>
<th>sig &lt; 0.5</th>
<th>N</th>
<th>xN</th>
<th>CS</th>
<th>CNS</th>
<th>IS</th>
<th>INS</th>
<th>Conclusions</th>
<th>Extension1</th>
<th>Extension2</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Baron, &amp; Hershey, 1988</td>
<td>1</td>
<td>Outcome bias</td>
<td>93</td>
<td>76%</td>
<td>24%</td>
<td>Success</td>
<td>TRUE</td>
<td>20</td>
<td>692</td>
<td>34.6</td>
<td>1</td>
<td></td>
<td></td>
<td>Successful</td>
<td>DV</td>
<td>DV</td>
</tr>
<tr>
<td>2</td>
<td>Epley &amp; Gilovich 2006</td>
<td>1b</td>
<td>Anchoring-and-adjustment heuristic</td>
<td>114</td>
<td>68%</td>
<td>40%</td>
<td>Success</td>
<td>TRUE</td>
<td>102</td>
<td>401</td>
<td>3.9</td>
<td>Too messy</td>
<td></td>
<td></td>
<td>Estimated low replicability</td>
<td>N/A</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>Epstein, Lipson, Holstein, &amp; Huh 1992</td>
<td>1 &amp; 2</td>
<td>Inattentional reactions to negative outcome</td>
<td>64</td>
<td>55%</td>
<td>40%</td>
<td>Success</td>
<td>FALSE</td>
<td>83</td>
<td>1038</td>
<td>12.5</td>
<td>2</td>
<td>1</td>
<td></td>
<td>Mostly successful</td>
<td>N/A</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>Fischhoff, 1975</td>
<td>2</td>
<td>Hintbias</td>
<td>97</td>
<td>98%</td>
<td>2%</td>
<td>Success</td>
<td>TRUE</td>
<td>260</td>
<td>894</td>
<td>3.4</td>
<td>11</td>
<td>4</td>
<td></td>
<td>1 Successful</td>
<td>DV</td>
<td>Surprise Mixed</td>
</tr>
<tr>
<td>5</td>
<td>Hamill, Wilson, &amp; Nisbett, 1980</td>
<td>1</td>
<td>Sensitivity to sample bias</td>
<td>94</td>
<td>68%</td>
<td>32%</td>
<td>Success</td>
<td>TRUE</td>
<td>124</td>
<td>850</td>
<td>7.2</td>
<td>2</td>
<td></td>
<td>2</td>
<td>Inconclusive, found typical</td>
<td>DV</td>
<td>Conscious Failed</td>
</tr>
<tr>
<td>6</td>
<td>Hsee &amp; Weber, 1997</td>
<td>1</td>
<td>Fundamental predictor error</td>
<td>77</td>
<td>70%</td>
<td>30%</td>
<td>Success</td>
<td>TRUE</td>
<td>99</td>
<td>401</td>
<td>4.1</td>
<td>2</td>
<td>1</td>
<td></td>
<td>Mostly successful</td>
<td>DV</td>
<td>Regret Success</td>
</tr>
<tr>
<td>7</td>
<td>Hsee, 1998</td>
<td>1, 2, 4</td>
<td>Less is better</td>
<td>88</td>
<td>53%</td>
<td>47%</td>
<td>Success</td>
<td>FALSE</td>
<td>256</td>
<td>401</td>
<td>1.6</td>
<td>2</td>
<td>1</td>
<td></td>
<td>Mostly successful</td>
<td>DV</td>
<td>Attractive Success</td>
</tr>
<tr>
<td>8</td>
<td>Kruger, et al, 2004</td>
<td>1 &amp; 2</td>
<td>Effort heuristic</td>
<td>73</td>
<td>74%</td>
<td>26%</td>
<td>Success</td>
<td>TRUE</td>
<td>200</td>
<td>709</td>
<td>3.5</td>
<td>1</td>
<td>1</td>
<td></td>
<td>1 Semi-successful, 3 failed</td>
<td>Combination</td>
<td>Mixed</td>
</tr>
<tr>
<td>9</td>
<td>Kruger, Wirtz &amp; Miller 2005</td>
<td>2</td>
<td>First instinct falacy</td>
<td>60</td>
<td>60%</td>
<td>40%</td>
<td>Success</td>
<td>FALSE</td>
<td>23</td>
<td>401</td>
<td>17.4</td>
<td>3</td>
<td></td>
<td>3</td>
<td>Successful, stronger effects</td>
<td>DV</td>
<td>Strategy/S Success</td>
</tr>
<tr>
<td>10</td>
<td>Melchers, Hertwig, &amp; Krahman, 2001</td>
<td>1</td>
<td>Conjunction effect</td>
<td>85</td>
<td>85%</td>
<td>15%</td>
<td>Success</td>
<td>TRUE</td>
<td>432</td>
<td>2032</td>
<td>2.4</td>
<td>2</td>
<td></td>
<td>2</td>
<td>Successful</td>
<td>DV</td>
<td>Populatio Success</td>
</tr>
<tr>
<td>11</td>
<td>Miller, &amp; McFarland, 1987</td>
<td>1</td>
<td>Pluralistic ignorance</td>
<td>109</td>
<td>68%</td>
<td>32%</td>
<td>Success</td>
<td>TRUE</td>
<td>116</td>
<td>400</td>
<td>3.4</td>
<td>2</td>
<td>1</td>
<td></td>
<td>Mixed findings, main hypothesis</td>
<td>IV</td>
<td>Friend Mixed</td>
</tr>
<tr>
<td>12</td>
<td>Schwartz, Struck, Hilton, &amp; Naderer, 1991</td>
<td>1</td>
<td>Relevance of irrelevant information</td>
<td>91</td>
<td>48%</td>
<td>52%</td>
<td>Failure</td>
<td>FALSE</td>
<td>44</td>
<td>604</td>
<td>13.7</td>
<td>1</td>
<td></td>
<td>1</td>
<td>Failure to replicate</td>
<td>DV</td>
<td>Bayesian est Success</td>
</tr>
<tr>
<td>13</td>
<td>Shafir, 1993</td>
<td>1 to 8</td>
<td>Choosing versus rejecting</td>
<td>82</td>
<td>48%</td>
<td>52%</td>
<td>Failure</td>
<td>FALSE</td>
<td>277</td>
<td>1026</td>
<td>3.7</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>Failure to replicate</td>
<td>DV</td>
<td>Attractive N/A</td>
</tr>
<tr>
<td>14</td>
<td>Shafir, Diamond, &amp; Tversky, 1997</td>
<td>1 to 4</td>
<td>Money illusion</td>
<td>70</td>
<td>59%</td>
<td>41%</td>
<td>Success</td>
<td>FALSE</td>
<td>323</td>
<td>604</td>
<td>1.9</td>
<td>4</td>
<td></td>
<td></td>
<td>Successful</td>
<td>DV</td>
<td>Confident Success</td>
</tr>
<tr>
<td>15</td>
<td>Slovic &amp; Fischhoff, 1977</td>
<td>1</td>
<td>Hindsight bias</td>
<td>111</td>
<td>96%</td>
<td>4%</td>
<td>Success</td>
<td>TRUE</td>
<td>184</td>
<td>604</td>
<td>3.3</td>
<td>5</td>
<td>3</td>
<td></td>
<td>Successful</td>
<td>DV</td>
<td>Confident Failed</td>
</tr>
<tr>
<td>16</td>
<td>Staw, 1976</td>
<td>1</td>
<td>Escalation of commitment</td>
<td>78</td>
<td>86%</td>
<td>14%</td>
<td>Success</td>
<td>TRUE</td>
<td>240</td>
<td>403</td>
<td>1.7</td>
<td>1</td>
<td></td>
<td>1</td>
<td>Inconclusive -&gt; Failure</td>
<td>N/A</td>
<td></td>
</tr>
<tr>
<td>17</td>
<td>Tversky &amp; Shafir, 1992</td>
<td>1</td>
<td>Disjunction effect</td>
<td>83</td>
<td>76%</td>
<td>24%</td>
<td>Success</td>
<td>TRUE</td>
<td>298</td>
<td>894</td>
<td>3.0</td>
<td>1</td>
<td>1</td>
<td></td>
<td>1 Successful, 1 Failed</td>
<td>DV</td>
<td>Predictor intolerance/Not analyzed</td>
</tr>
<tr>
<td>18</td>
<td>Zeelenberg et al 1996</td>
<td>1</td>
<td>Regret aversion</td>
<td>75</td>
<td>73%</td>
<td>27%</td>
<td>Success</td>
<td>TRUE</td>
<td>78</td>
<td>452</td>
<td>5.8</td>
<td>1</td>
<td></td>
<td></td>
<td>1 Inconclusive -&gt; Failure</td>
<td>Predictor</td>
<td>Choice req Failed</td>
</tr>
</tbody>
</table>

Warning: Student calculations, needs to be rechecked, and checked. CS = Consistent signal CNS = Consistent no signal IS = Inconsistent signal INS = Inconsistent no signal

**WARNING:** Preliminary student findings, need to be rechecked and verified
**Results: Zoom in**

<table>
<thead>
<tr>
<th>St</th>
<th>Authors</th>
<th>Studies</th>
<th>Bias</th>
<th>Prediction market</th>
<th>Original Replication</th>
<th>Conclusion</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Votes</td>
<td>Success</td>
<td>Failure</td>
</tr>
<tr>
<td>1</td>
<td>Baron, &amp; Hershey, 1988</td>
<td>1</td>
<td>Outcome bias</td>
<td>93</td>
<td>76%</td>
<td>24%</td>
</tr>
<tr>
<td>2</td>
<td>Epley &amp; Gilovich 2006</td>
<td>1b</td>
<td>Anchoring-and-adjustment heuristic</td>
<td>114</td>
<td>88%</td>
<td>12%</td>
</tr>
<tr>
<td>3</td>
<td>Epstein, Lipson, Holstein, &amp; Huh 1992</td>
<td>1 &amp; 2</td>
<td>Irrational reactions to negative outcome</td>
<td>64</td>
<td>55%</td>
<td>45%</td>
</tr>
<tr>
<td>4</td>
<td>Fischhoff, 1975</td>
<td>1</td>
<td>Hindsight bias</td>
<td>97</td>
<td>98%</td>
<td>2%</td>
</tr>
<tr>
<td>5</td>
<td>Hamill, Wilson, &amp; Nisbett, 1980</td>
<td>1</td>
<td>Insensitivity to sample bias</td>
<td>92</td>
<td>68%</td>
<td>32%</td>
</tr>
<tr>
<td>6</td>
<td>Hsee &amp; Weber, 1997</td>
<td>1</td>
<td>Fundamental predictor error</td>
<td>77</td>
<td>70%</td>
<td>30%</td>
</tr>
<tr>
<td>7</td>
<td>Hsee, 1998</td>
<td>1, 2, 4</td>
<td>Less is better</td>
<td>88</td>
<td>53%</td>
<td>47%</td>
</tr>
<tr>
<td>8</td>
<td>Kruger etal, 2004</td>
<td>1 &amp; 2</td>
<td>Effort heuristic</td>
<td>73</td>
<td>74%</td>
<td>26%</td>
</tr>
<tr>
<td>9</td>
<td>Kruger, Wirtz &amp; Miller 2005</td>
<td>2</td>
<td>First instinct fallacy</td>
<td>60</td>
<td>60%</td>
<td>40%</td>
</tr>
<tr>
<td>10</td>
<td>Mellers, Hertwig, &amp; Kahneman, 2001</td>
<td>1</td>
<td>Conjunction effect</td>
<td>85</td>
<td>85%</td>
<td>15%</td>
</tr>
<tr>
<td>11</td>
<td>Miller, &amp; McFarland, 1987</td>
<td>1</td>
<td>Pluralistic ignorance</td>
<td>109</td>
<td>68%</td>
<td>32%</td>
</tr>
<tr>
<td>12</td>
<td>Schwarz, Strack, Hilton, &amp; Naderer, 1997</td>
<td>1</td>
<td>Relevance of irrelevant information</td>
<td>91</td>
<td>48%</td>
<td>52%</td>
</tr>
<tr>
<td>13</td>
<td>Shafir, 1993</td>
<td>1 to 8</td>
<td>Choosing versus rejecting</td>
<td>82</td>
<td>48%</td>
<td>52%</td>
</tr>
<tr>
<td>14</td>
<td>Shafir, Diamond, &amp; Tversky, 1997</td>
<td>1 to 4</td>
<td>Money illusion</td>
<td>70</td>
<td>59%</td>
<td>41%</td>
</tr>
<tr>
<td>15</td>
<td>Slovic &amp; Fischhoff, 1977</td>
<td>1</td>
<td>Hindsight bias</td>
<td>111</td>
<td>96%</td>
<td>4%</td>
</tr>
<tr>
<td>16</td>
<td>Staw, 1976</td>
<td>1</td>
<td>Escalation of commitment</td>
<td>78</td>
<td>86%</td>
<td>14%</td>
</tr>
<tr>
<td>17</td>
<td>Tversky &amp; Shafir, 1992</td>
<td>1</td>
<td>Disjunction effect</td>
<td>83</td>
<td>76%</td>
<td>24%</td>
</tr>
<tr>
<td>18</td>
<td>Zeelenberg etal 1996</td>
<td>1</td>
<td>Regret aversion</td>
<td>75</td>
<td>73%</td>
<td>27%</td>
</tr>
</tbody>
</table>

**Warning:** Student calculations, needs to be reverified, and checked.

**CS = Consistent signal | CNS = Consistent no signal | IS = Inconsistent signal | INS = Inconsistent no signal**

**WARNING: Preliminary student findings, need to be rechecked and verified**
Zoom in: Compare predictions to conclusions

<table>
<thead>
<tr>
<th>St</th>
<th>Authors</th>
<th>Studies</th>
<th>Bias</th>
<th>Predictor</th>
<th>Replication</th>
<th>Conclusion</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Baron, &amp; Hershey, 1988</td>
<td>1</td>
<td>Outcome bias</td>
<td>Success</td>
<td>1</td>
<td>Successful</td>
</tr>
<tr>
<td>2</td>
<td>Epley &amp; Gilovich 2006</td>
<td>1b</td>
<td>Anchoring-and-adjustment heuristic</td>
<td>Success</td>
<td>Too messy</td>
<td>Estimated low replicability</td>
</tr>
<tr>
<td>3</td>
<td>Epstein, Lipson, Holstein, &amp; Huh 1992</td>
<td>1 &amp; 2</td>
<td>Irrational reactions to negative outcomes</td>
<td>Success</td>
<td>2</td>
<td>Mostly successful</td>
</tr>
<tr>
<td>4</td>
<td>Fischhoff, 1975</td>
<td>2</td>
<td>Hindsight bias</td>
<td>Success</td>
<td>11 4 1</td>
<td>Successful</td>
</tr>
<tr>
<td>5</td>
<td>Hamill, Wilson, &amp; Nisbett, 1980</td>
<td>1</td>
<td>Insensitivity to sample bias</td>
<td>Success</td>
<td>2 1</td>
<td>Inconclusive, found typical</td>
</tr>
<tr>
<td>6</td>
<td>Hsee &amp; Weber, 1997</td>
<td>1</td>
<td>Fundamental predictor error</td>
<td>Success</td>
<td>2 1</td>
<td>Mostly successful</td>
</tr>
<tr>
<td>7</td>
<td>Hsee, 1998</td>
<td>1, 2, 4</td>
<td>Less is better</td>
<td>Success</td>
<td>2 1</td>
<td>Mostly successful</td>
</tr>
<tr>
<td>8</td>
<td>Kruger etal, 2004</td>
<td>1 &amp; 2</td>
<td>Effort heuristic</td>
<td>Success</td>
<td>1 1</td>
<td>1 semi-successful, 1 failed</td>
</tr>
<tr>
<td>9</td>
<td>Kruger, Wirtz &amp; Miller 2005</td>
<td>2</td>
<td>First instinct fallacy</td>
<td>Success</td>
<td>3</td>
<td>Successful, stronger effects</td>
</tr>
<tr>
<td>10</td>
<td>Mellers, Hertwig, &amp; Kahneman, 2001</td>
<td>1</td>
<td>Conjunction effect</td>
<td>Success</td>
<td>2</td>
<td>Successful</td>
</tr>
<tr>
<td>11</td>
<td>Miller, &amp; McFarland, 1987</td>
<td>1</td>
<td>Pluralistic ignorance</td>
<td>Success</td>
<td>2 1</td>
<td>Mixed findings, main hypothesis</td>
</tr>
<tr>
<td>12</td>
<td>Schwarz, Strack, Hilton, &amp; Naderer, 1991</td>
<td>1</td>
<td>Relevance of irrelevant information</td>
<td>Failure</td>
<td>1 1</td>
<td>Failure to replicate</td>
</tr>
<tr>
<td>13</td>
<td>Shafir, 1993</td>
<td>1 to 8</td>
<td>Choosing versus rejecting</td>
<td>Failure</td>
<td>2 2 2 2</td>
<td>Failure to replicate</td>
</tr>
<tr>
<td>14</td>
<td>Shafir, Diamond, &amp; Tversky, 1997</td>
<td>1 to 4</td>
<td>Money illusion</td>
<td>Success</td>
<td>4</td>
<td>Successful</td>
</tr>
<tr>
<td>15</td>
<td>Slovic &amp; Fischhoff, 1977</td>
<td>1</td>
<td>Hindsight bias</td>
<td>Success</td>
<td>5 3</td>
<td>Successful</td>
</tr>
<tr>
<td>16</td>
<td>Staw, 1976</td>
<td>1</td>
<td>Escalation of commitment</td>
<td>Success</td>
<td>1 1</td>
<td>Inconclusive -&gt; Failure</td>
</tr>
<tr>
<td>17</td>
<td>Tversky &amp; Shafir, 1992</td>
<td>1</td>
<td>Disjunction effect</td>
<td>Success</td>
<td>1 1</td>
<td>1 successful, 1 failed</td>
</tr>
<tr>
<td>18</td>
<td>Zeelenberg etal 1996</td>
<td>1</td>
<td>Regret aversion</td>
<td>Success</td>
<td>1 1</td>
<td>Inconclusive -&gt; Failure</td>
</tr>
</tbody>
</table>

CS = Consistent signal | CNS = Consistent no signal | IS = Inconsistent signal | INS = Inconsistent no signal

WARNING: Preliminary student findings, need to be rechecked and verified
Summary (out of 18)

Successful (9)
• Strong replication: 6
• Mostly successful: 3

Inconclusive (5)
• Inconclusive: 4
• Messy, hard to tell, likely failure: 1

Failure (4)
• Most likely failure: 2
• Inconclusive but likely failure: 2

WARNING: Preliminary student findings, need to be rechecked and verified
How well did the students do?
Hindsight (Fischhoff 1975)

<table>
<thead>
<tr>
<th>Event A Outcome 1</th>
<th>Event A Outcome 2</th>
<th>Event A Outcome 3</th>
<th>Event A Outcome 4</th>
</tr>
</thead>
</table>

<table>
<thead>
<tr>
<th>Event B Outcome 1</th>
<th>Event B Outcome 2</th>
<th>Event B Outcome 3</th>
<th>Event B Outcome 4</th>
</tr>
</thead>
</table>

<table>
<thead>
<tr>
<th>Event B Outcome 1</th>
<th>Event B Outcome 2</th>
<th>Event B Outcome 3</th>
<th>Event B Outcome 4</th>
</tr>
</thead>
</table>

<table>
<thead>
<tr>
<th>Event B Outcome 1</th>
<th>Event B Outcome 2</th>
<th>Event B Outcome 3</th>
<th>Event B Outcome 4</th>
</tr>
</thead>
</table>

<table>
<thead>
<tr>
<th>Result</th>
<th>Cohen's d &amp; 95% Confidence Intervals</th>
<th>Some Direction?</th>
<th>Replication CI include self?</th>
<th>Effect in target CI?</th>
<th>Effect magnitude (Cohen's d)</th>
<th>Note</th>
</tr>
</thead>
<tbody>
<tr>
<td>Present Study</td>
<td>E2 Δ Outcome 1</td>
<td>1.02 [0.339, 1.605]</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Similar</td>
</tr>
<tr>
<td></td>
<td>E2 Δ Outcome 2</td>
<td>0.109 [-0.235, 0.453]</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
<td>Different</td>
</tr>
<tr>
<td></td>
<td>E2 Δ Outcome 3</td>
<td>0.609 [0.022, 0.887]</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
<td>Different</td>
</tr>
<tr>
<td></td>
<td>E2 Δ Outcome 4</td>
<td>0.599 [0.001, 0.504]</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Different</td>
</tr>
<tr>
<td></td>
<td>E2 B Outcome 1</td>
<td>1.022 [0.333, 1.677]</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Similar</td>
</tr>
<tr>
<td></td>
<td>E2 B Outcome 2</td>
<td>0.907 [0.447, 1.354]</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Similar</td>
</tr>
<tr>
<td></td>
<td>E2 B Outcome 3</td>
<td>0.750 [0.293, 1.161]</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Different</td>
</tr>
<tr>
<td></td>
<td>E2 B Outcome 4</td>
<td>0.420 [0.029, 0.829]</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Different</td>
</tr>
<tr>
<td></td>
<td>E2 C Outcome 1</td>
<td>1.231 [0.755, 1.718]</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Similar</td>
</tr>
<tr>
<td></td>
<td>E2 C Outcome 2</td>
<td>0.825 [0.365, 1.277]</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Similar</td>
</tr>
<tr>
<td></td>
<td>E2 C Outcome 3</td>
<td>0.578 [0.009, 1.064]</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
<td>Different</td>
</tr>
<tr>
<td></td>
<td>E2 C Outcome 4</td>
<td>0.184 [-0.448, 0.379]</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
<td>Different</td>
</tr>
<tr>
<td></td>
<td>E2 D Outcome 1</td>
<td>0.401 [-0.023, 0.826]</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
<td>Different</td>
</tr>
<tr>
<td></td>
<td>E2 D Outcome 2</td>
<td>0.510 [0.174, 0.805]</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Different</td>
</tr>
<tr>
<td></td>
<td>E2 D Outcome 3</td>
<td>0.510 [0.065, 0.965]</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Different</td>
</tr>
<tr>
<td></td>
<td>E2 D Outcome 4</td>
<td>0.522 [0.083, 0.962]</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Different</td>
</tr>
</tbody>
</table>

*p<.05; **p<.01
## Hindsight (Fischhoff 1975)

<table>
<thead>
<tr>
<th>Event Type</th>
<th>Cohen's d &amp; 95% Confidence Intervals</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fischhoff 1975</td>
<td>1.130 [0.440, 1.820]</td>
</tr>
<tr>
<td>Event A Outcome 1</td>
<td>1.002 [0.530, 1.465]</td>
</tr>
<tr>
<td>Event A Outcome 2</td>
<td>0.199 [-0.230, -0.625]</td>
</tr>
<tr>
<td>Event A Outcome 3</td>
<td>0.409 [-0.022, 0.837]</td>
</tr>
<tr>
<td>Event A Outcome 4</td>
<td>0.539 [0.098, 0.974]</td>
</tr>
<tr>
<td>Event B Outcome 1</td>
<td>1.022 [0.558, 1.477]</td>
</tr>
<tr>
<td>Event B Outcome 2</td>
<td>0.907 [0.447, 1.358]</td>
</tr>
<tr>
<td>Event B Outcome 3</td>
<td>0.705 [0.262, 1.141]</td>
</tr>
<tr>
<td>Event B Outcome 4</td>
<td>0.450 [0.026, 0.869]</td>
</tr>
<tr>
<td>Event C Outcome 1</td>
<td>1.251 [0.755, 1.736]</td>
</tr>
<tr>
<td>Event C Outcome 2</td>
<td>0.825 [0.368, 1.273]</td>
</tr>
<tr>
<td>Event C Outcome 3</td>
<td>0.358 [-0.069, 0.781]</td>
</tr>
<tr>
<td>Event C Outcome 4</td>
<td>-0.034 [-0.448, 0.379]</td>
</tr>
<tr>
<td>Event D Outcome 1</td>
<td>0.401 [-0.023, 0.820]</td>
</tr>
<tr>
<td>Event D Outcome 2</td>
<td>0.503 [0.074, 0.926]</td>
</tr>
<tr>
<td>Event D Outcome 3</td>
<td>0.530 [0.095, 0.960]</td>
</tr>
<tr>
<td>Event D Outcome 4</td>
<td>0.525 [0.093, 0.952]</td>
</tr>
</tbody>
</table>
Hypothesis 1: Enriched option will be chosen and rejected more often than the impoverished option > Enriched options share will exceed 100%

Table 1: Comparison of the enriched options observed share with the expected share of 100% with z-value, p-value and effect size with the 95% confidence interval.

<table>
<thead>
<tr>
<th>Problem</th>
<th>Enriched options share</th>
<th>z-value</th>
<th>p-value</th>
<th>effect size (Cohen’s d)</th>
<th>confidence interval (CI = 95%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Problem 1:</td>
<td>1.018</td>
<td>0.58</td>
<td>-.56</td>
<td>0.0362 (no effect)</td>
<td>[-0.0862, 0.1586]</td>
</tr>
<tr>
<td>Sole-custody case</td>
<td>0.958</td>
<td>-1.35</td>
<td>.18</td>
<td>-0.0843 (no effect)</td>
<td>[-0.2067, 0.0382]</td>
</tr>
<tr>
<td>Problem 3:</td>
<td>0.951</td>
<td>-1.57</td>
<td>.12</td>
<td>-0.0981 (no effect)</td>
<td>[-0.2205, 0.0244]</td>
</tr>
<tr>
<td>Coarse selection</td>
<td>1.130</td>
<td>4.16</td>
<td>&lt;.001</td>
<td>0.2619 (small effect)</td>
<td>[0.1389, 0.3847]</td>
</tr>
<tr>
<td>Problem 5:</td>
<td>1.186</td>
<td>5.96</td>
<td>&lt;.001</td>
<td>0.3787 (small effect)</td>
<td>[0.2551, 0.5021]</td>
</tr>
<tr>
<td>Lottery A:</td>
<td>1.033</td>
<td>1.06</td>
<td>.29</td>
<td>0.0662 (no effect)</td>
<td>[-0.0562, 0.1886]</td>
</tr>
<tr>
<td>Problem 7:</td>
<td>0.802</td>
<td>-6.34</td>
<td>&lt;.001</td>
<td>-0.4038 (small effect)</td>
<td>[-0.5273, -0.2801]</td>
</tr>
<tr>
<td>Election</td>
<td>0.899</td>
<td>-3.24</td>
<td>&lt;.001</td>
<td>-0.2033 (small effect)</td>
<td>[-0.3259, -0.0806]</td>
</tr>
<tr>
<td>Problem 8:</td>
<td>1.018</td>
<td>0.58</td>
<td>-.56</td>
<td>0.0362 (no effect)</td>
<td>[-0.0862, 0.1586]</td>
</tr>
<tr>
<td>Lottery C (non-binary)</td>
<td>0.958</td>
<td>-1.35</td>
<td>.18</td>
<td>-0.0843 (no effect)</td>
<td>[-0.2067, 0.0382]</td>
</tr>
</tbody>
</table>
Shafir 1993: We addressed most issues in reply to Manylabs2

Hypothesis 2: Higher chance of the enriched option in the choose-condition

<table>
<thead>
<tr>
<th>Problem</th>
<th>Chance of the enriched option in the choose- vs. reject condition</th>
<th>z-value</th>
<th>p-value</th>
<th>effect size (Cohen’s d)</th>
<th>confidence interval (CI = 95%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Problem 1:</td>
<td>Same-custody case</td>
<td>0.543</td>
<td>0.58</td>
<td>0.0511 (no effect)</td>
<td>[-0.1219, 0.2246]</td>
</tr>
<tr>
<td>Problem 2:</td>
<td>Vacation spot</td>
<td>0.444</td>
<td>.135</td>
<td>-0.1193 (no effect)</td>
<td>[-0.2923, 0.0538]</td>
</tr>
<tr>
<td>Problem 3:</td>
<td>Course selection</td>
<td>0.545</td>
<td>0.35</td>
<td>-0.1397 (no effect)</td>
<td>[-0.3127, 0.0335]</td>
</tr>
<tr>
<td>Problem 4:</td>
<td>Lottery A</td>
<td>0.817</td>
<td>0.79</td>
<td>0.4323 (small effect)</td>
<td>[0.2572, 0.6070]</td>
</tr>
<tr>
<td>Problem 5:</td>
<td>Lottery B</td>
<td>0.852</td>
<td>0.06</td>
<td>0.6461 (medium effect)</td>
<td>[0.4684, 0.8232]</td>
</tr>
<tr>
<td>Problem 6:</td>
<td>Ice-cream flavor</td>
<td>0.564</td>
<td>1.06</td>
<td>0.0936 (no effect)</td>
<td>[-0.0794, 0.2666]</td>
</tr>
<tr>
<td>Problem 7:</td>
<td>Election</td>
<td>0.093</td>
<td>0.05</td>
<td>-0.7596 (medium effect)</td>
<td>[-0.9383, -0.5802]</td>
</tr>
<tr>
<td>Problem 8:</td>
<td>Lottery C (non-binary)</td>
<td>0.681</td>
<td>-3.65</td>
<td>-0.1262 (small effect)</td>
<td>[-0.5001, -0.1520]</td>
</tr>
</tbody>
</table>

Table 3: Chance of the enriched option in the choose-condition vs. reject-condition with z-value, p-value and effect size with 95% confidence interval.
Comparison: Shafir (1993) vs. Replication

Hypothesis 1:
Enriched options share exceeds 100%
➢ No (consistent) replication

Hypothesis 2:
Higher chance of the enriched option in the choose-condition
➢ No (consistent) replication
I showed some reports to collaborators

The general response?

THAT WAS AMAZING.

That is the most amazing thing I have seen.
I showed some reports to original authors

Thanks for sharing your student's work with us. I'm truly amazed by her quality of work! I don't think our year 1 ph D students in Business school are able to do anything like that. U must have given them lots of guidance.

we should seriously consider asking our ph d students to do the same (conduct replication study) in the method class.
This is how I summarized things for the students....
Think of all you accomplished in this course

You experienced things first hand!

- You experienced experiments as researchers, as participants, as instructors, as audience.
- You took part, you were active and engaging.
- You led an inclass experiment (JDM) / class discussions/presentations (ASP)!
- You reflected on high-level academic articles, contemplated real life implications, and designed extensions in every class!

We (I) had fun

You learned by thinking and doing
Think of all you accomplished in this course

For your projects:
• You analyzed a classic article in depth
• Effect-size calculations
• Confidence interval calculations
• Power-analyses
• Pre-registration templates
• You designed a replication experiment
• (Some of you designed extensions)
• You mastered Qualtrics
• You did stats, you mastered JAMOVI/JASP/R
• You joined the academic community on OSF & RG
• You did a peer-review, you revised based on peer review
• You communicated your findings in a presentation
• You wrote a very high-standard scientific report + Replication Recipe
Final remarks

• I do know this was
  – A lot of work
  – Confusing at times
  – Different from what you're used to
    [It was that for ALL of us (me, Boley, Bill)]

BUT, keep this in mind...

• You have done **REALLY WELL** on this course
• You have learned a LOT
• You should feel REALLY PROUD of your achievements
• I think, no, I KNOW, this was worth it
Now what?
Interesting, but... Replications!? are these publishable?!

https://blogs.royalsociety.org/publishing/reproducibility-meets-accountability/#.W8UMUvTYRsk.twitter

Reproducibility meets accountability: introducing the replications initiative at Royal Society Open Science

15 October 2018 by Chris Chambers

Today marks the launch of a new initiative in which the Psychology and Cognitive Neuroscience section of Royal Society Open Science guarantees to publish any close replication of any article published in our journal, and from most other major journals too.

- Royal Society Open Science guarantees to publish any close replication of any study previously published in its Psychology and Cognitive Neuroscience section. This commitment extends to replication studies themselves, with no limit on the number of acceptable repeats.

- One concern with results-blind review (where results are known to the authors but not the reviewers) is that reviewers may assume that the results are negative or confusing, leading to biased reasoning when assessing the paper. Therefore, reviewers will initially be blinded to whether the article has been submitted via the Results-Blind or Fully Preregistered track. Submissions in both categories will be written in past tense.
That's just one journal we don't know, anything else?

- Social Psychological and Personality Science: Simine Vazire
- Cognition & Emotion: Sander Koole & Daniel Lakens
- Journal of Experimental Social Psychology: Roger Giner-Sorolla
- Psychological Science: Stephen Lindsay (Chris Chambers running to replace)

New journals enthusiastic of replications:

- Advances in Methods and Practices in Psychological Science
- Collabra: Psychology (commits to publish rejected in others due to lack of interest)
- Meta-psychology

And there are many others making the change...
Students were a bit worried if anyone would care about our project/findings

But then this happened a week after we posted the project on Research Gate:

Lucas Kutscher

added a project goal

The aim of this project is to conduct a pre-registered replication and a meta-analysis on how past behavior norms affect perceived regret over a negative outcome.

Sander L. Koole

Dear Lucas and Gilad, at Cognition and Emotion (where Klaus Rothermund and I are incoming editors), we are seeking to publish more pre-registered research. Please consider C&E as an outlet for this project.

Best, Sander
Mar 28, 2017
Yeah, replications & extensions can be published...

In Cognition & Emotion:

The impact of past behaviour normality on regret: replication and extension of three experiments of the exceptionality effect

Lucas Kutscher\textsuperscript{a} and Gilad Feldman\textsuperscript{a,b}

\textsuperscript{a}Department of Work and Social Psychology, Maastricht University, Maastricht, the Netherlands; \textsuperscript{b}Department of Psychology, University of Hong Kong, Hong Kong SAR, China

In Social Psychological and Personality Science:

Laypersons’ Beliefs and Intuitions About Free Will and Determinism: New Insights Linking the Social Psychology and Experimental Philosophy Paradigms

Gilad Feldman\textsuperscript{1} and Subramanya Prasad Chandrashekar\textsuperscript{2}
With 3 of them...

Sander L. Koole asked a question in History of Ideas

At Cognition and Emotion, we are seeking to publish more pre-registered research. Would you consider C&E as an outlet for this project?

Question 2 Answers  Asked a year ago

Dear Tijen and Gilad, at Cognition and Emotion (where Klaus Rothermund and I are incoming editors), we are seeking to publish more pre-registered research. Would you consider C&E as an outlet for this project?

Best, Sander

Yajing Gao  Mar 3, 2017
added a project goal

The aim of this project is to conduct a pre-registered replication and a meta-analysis on mere ownership effect.

Comment  Recommend  Share

Sander L. Koole  Mar 28, 2017

Dear Yajing and Gilad, at Cognition and Emotion (where Klaus Rothermund and I are incoming editors), we are seeking to publish more pre-registered research. Would you consider C&E as an outlet for this project?

Best, Sander
Join us in publishing those (ECRs)

My students and I invite you to help us in finalizing their very high quality submissions:

For each replication, we have: 1-4 (!) APA ready submissions, complete with pre-registrations (power analyses->high power), full open-science data/code, replication recipe...

An opportunity for you to have a lead author pre-registered replication submission and learn "new" science

Already working with 5 collaborators world-wide (HK, Canada, Norway, France). Have 20-30 more projects that need collaborators.
What's the "catch"?

• Need to put in the work, and learn with/from me.
• Open science: Need to change many old habits.
• Social Psychology journals (JESP/SPPS/CogEmo/RSOS etc.).

Why do these?

• Build an early career researcher network that cares about open-science. It's up to us to change things and set an example.
• I want you to do well in the job market.

You need to lead quick doable projects, with low risk/uncertainty and show commitment to science and getting things right.
Join us next academic year

Things you can do:

• Work with us on collaborative manuals
• Integrate/Implement these in your courses
• Run our surveys on your courses' students as additional samples
• Take part in our workshops/Hackathons
• Spread the word

• Implement open-science in your labs.
• Tell me what I can do to help you do better.
Thank you

- About me and my research: [http://giladfeldman.org](http://giladfeldman.org)
- Contact: gfeldman@hku.hk
- Twitter: @giladfeldman