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



Department of Psychology The University of Hong Kong

范大学心理学系

Departmental Seminar

Mass-replication Project of Classic JDM Findings

in two HKU Courses

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

Reproducibility and replicability are at heart of science, yet increasing evidence from recent

years suggests that many of the findings in psychological science are irreproducible 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 for rigorous replicable

reproducible science? In this talk I will discuss a mass-replication effort I headed in HKU courses PSYC2020 and PSYC3052 to conduct 13/11 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 effectsize calculations and power analyses, design Qualtrics experiments, and adopt latest tools and templates and preregister the replications on the Open Science Framework. We then ran

the experiments on (1) a limited sample of HKU students (N = up to 49) and (2) high-power Amazon Mechanical Turk American online samples (power = 0.95-0.99; N = 300-800). PSYC3052 course was overhauled to discuss the replication crisis in-depth and involve students in thinking of its implications and improving. I will briefly present 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 TA team, with insights and

Dr. Gilad Feldman

Department of Psychology University of Hong Kong

contributions to the literature and the academic community.

Assistant Professor

Rm 813, 8/F, The Jockey Club Tower | Centennial Campus | The University of Hong Kong

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...

ffect (Kahneman & Tversky, 1982): Replicated several times (> 8).

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

ul replications

scenarios).

er & McFarland, 1986): Replic

olications

effect (Zeelenberg et al., 2002): Replicated Experiment 1 several times (mmetry (Cushman et al, 20

n bias (Spranca, Minsk, & Baron, 1991): Replicated two scenarios from E nality effect (Kahneman & Miller, 1986): Replicated two experiments (hits Knobe, 1997): Twice (US/HI

asymmetry).

cations, needs to revi

al relevance (Irmak, Wakslak, &

ent effect & transaction demand (Mandel, 2002): Replicated Experimenties

2.081 **Impressions**

Total engagements

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

tter effect (Nuttin, 1987): Replicated the main experiment.

164

d spot (Pronin & Kugler, 2007): Replicated twice in US/HK, in atleast 2/3 ce for indirect harm (Royzman & Baron, 2002): Replicated Experiment 2 eriment 3 once (HK).

inertia (Tykocinski et al., 1995): Replicated Experiment 1 twice in US/HK

4:20 PM - 25 Aug 2018

5 Retweets 23 Likes 🔐 🚱 🔞 🔞 😭 😭











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



Will present JDM mass-replication effort findings in 2 HKU courses on MTurk & HongKong 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

		Original findings	Replication MTurk 3052	Replication HKU 2020	Conclusio	
Departmental Seminar	st Fallacy	54%-40%	17%-83%	50%-50% 0.54 [08, 3.16]*	Successful Likely over effect Successful	
Mass-replication Project of Classic JDM Findings	st Fallacy	0.32 (0.01,0.64)	0.42 (0.28,0.56)			
in two HKU Courses		Extensions Intent to proceed	1.08 [0.93, 1.23]	Not conducted	In-line with	
	flowing	Doing: 0.84 [0.80, 1.07]	0.89 [0.64, 3.13]	0.05ns	Successful	
Impressions Total engagements				289	Successful Mixed ha Successful Failed	
bstract					much wea	
reproducibility and replicability are at heart of science, yet increasing evidence from secut years suggests that many of the findings in psychological science are irreproducible	d Spot	558: 1.24 [1.02, 1.45] PHE: 0.48 [0.27, 0.69] FAE: 0.86 [0.62, 1.09]	0.68 (0.45, 0.91) 0.15 (-0.07, 0.37) 0.40 (0.18, 0.63)	0.52ns 0.40ns 0.45ns	(MTurk 2/	
nd non-replicable in what some termed as a "replication crisis" and a new movement alling for significant changes in the way we do science. How can we do better? How can	ice for indirect	0.70 [0.14, 1.24]	Sc1: .26 [.04, .49] Sc2: .32 [.10, .34]	0.43*	Successful	
re inform colleagues and students about these issues and train students for rigorous		0.63	Not conducted	0.39*	Successful	
eplicable reproducible science? In this talk I will discuss a mass-replication effort I headed I HKU courses PSYC2020 and PSYC2032 to conduct 13/11 pre-registered replications of lassic findings in judgment and decision-making literature. With the help of four conductal Tax endeate analyzed articles and tried to reproduce methods and materials to	nertia	Ski: 2.43 [2.05, 2.80] Car: 1.07 [0.85, 1.30] Flyer: 2.17 [1.84, 2.50] Flyers: 2.29 [1.95, 2.63]	0.78 [0.59, 0.85] 0.67 [0.54, 0.79] 0.26 [0.15, 0.38] 0.78 [0.86, 0.92]	Medium Medium Medium	Successful, effects	

5:15 PM - 6 May 2018

7 Retweets 31 Likes





















Previous semester... (link)

14 successful3 semi-successful4 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.
- Exceptionality effect (Kahneman & Miller, 1986): Replicated two experiments (hitchhiker and car accident scenarios).
- 5. Exceptionality effect (Seta et al., 2001): Replicated 3 times.
- 6. Name letter effect (Nuttin, 1987): Replicated the main experiment.
- 7. Endowment effect & transaction demand (Mandel, 2002): Replicated Experiment 1.
- 8. Status quo bias (Samuelson & Zeckhauser, 1988): Replicated 4 scenarios from Experiment 1.
- Escalation of commitment (Arkes & Blumer, 1985): Replicated Experiments 1 and 4 twice (US/HK).
- 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.
- 12. Bias blind spot (Pronin & Kugler, 2007): Replicated twice in US/HK, in atleast 2/3 categories.
- 13. Preference for indirect harm (Royzman & Baron, 2002): Replicated Experiment 2 twice (US/HK) and Experiment 3 once (HK).
- 14. Inaction inertia (Tykocinski et al., 1995): Replicated Experiment 1 twice in US/HK samples.

Semi-successful replications

- Exceptionality effect (Miller & McFarland, 1986): Replicated twice using a regret DV, but not using original compensation DV.
- 2. Doing/allowing morality asymmetry (Cushman 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

- Endowment effect and goal relevance (Irmak, Wakslak, & 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)
- 4. Anchoring effect by framing (Wong & Kwong, 2000): Failed to replicate twice in US and HK. Very likely culture/language bad translation issues.

All materials from courses were made available on OSF

See http://mgto.org/teaching-courses/ for updates.

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:

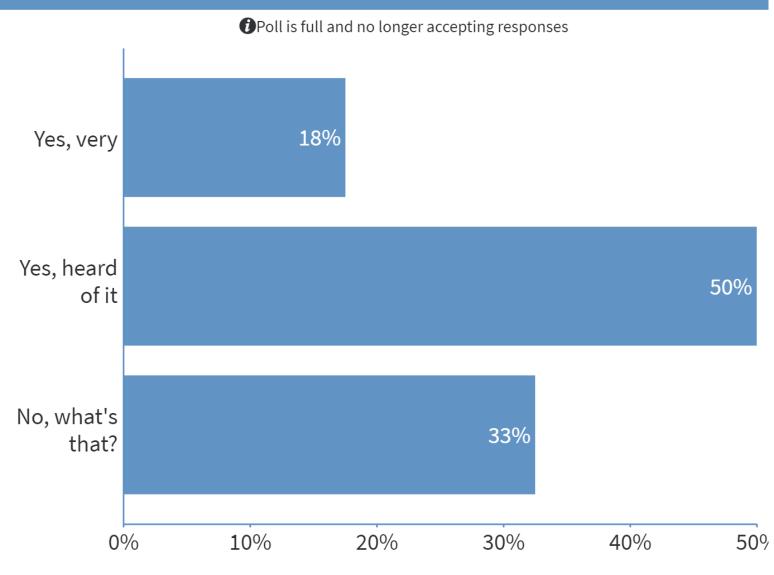
https://tinyurl.com/hkujamoviguide

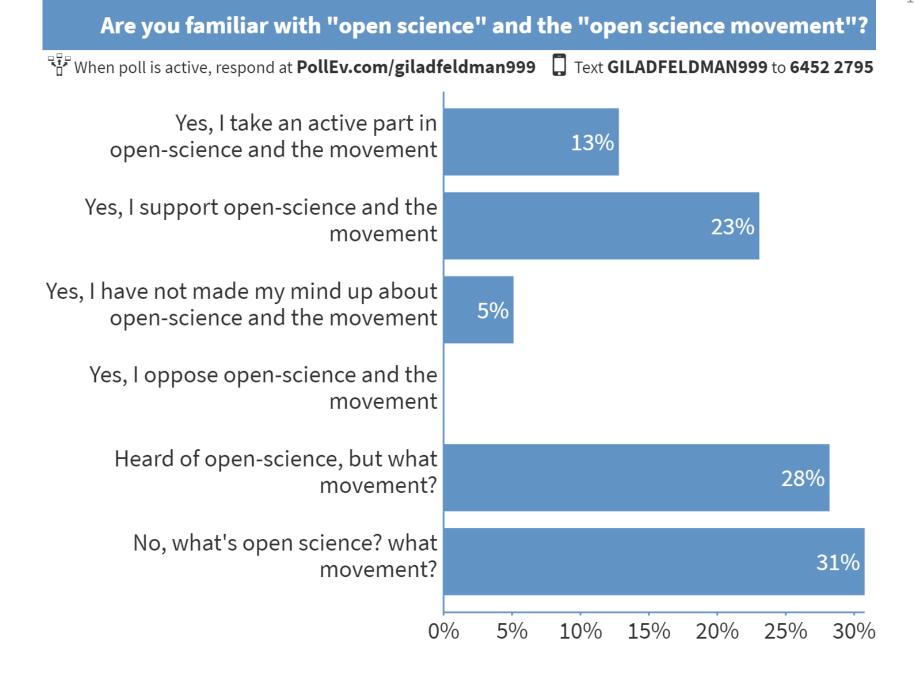
Pre-registered replication guide collaborative manual:

https://tinyurl.com/replicationguide

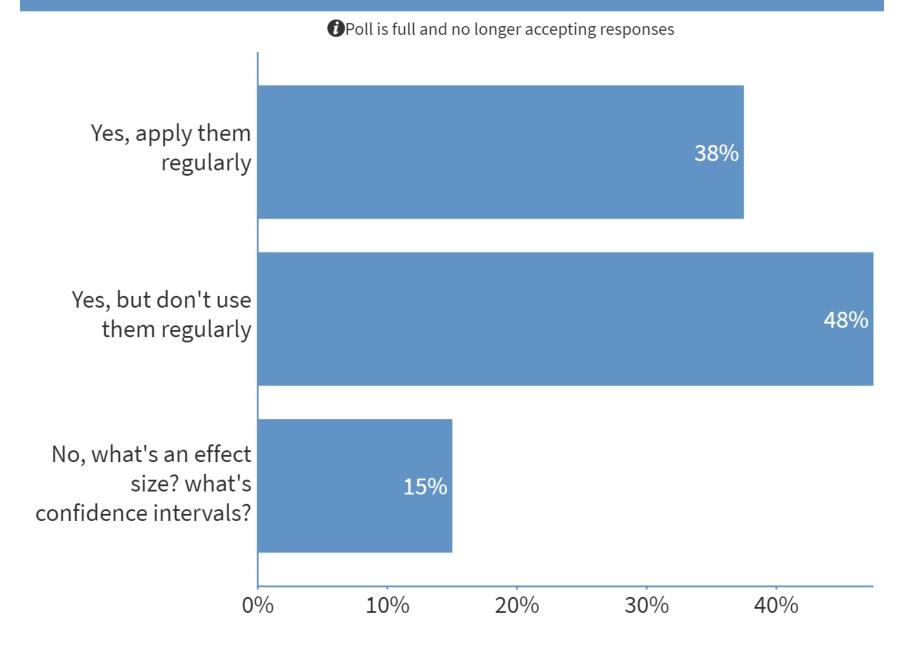
Other resources on my website (giladfeldman.org)

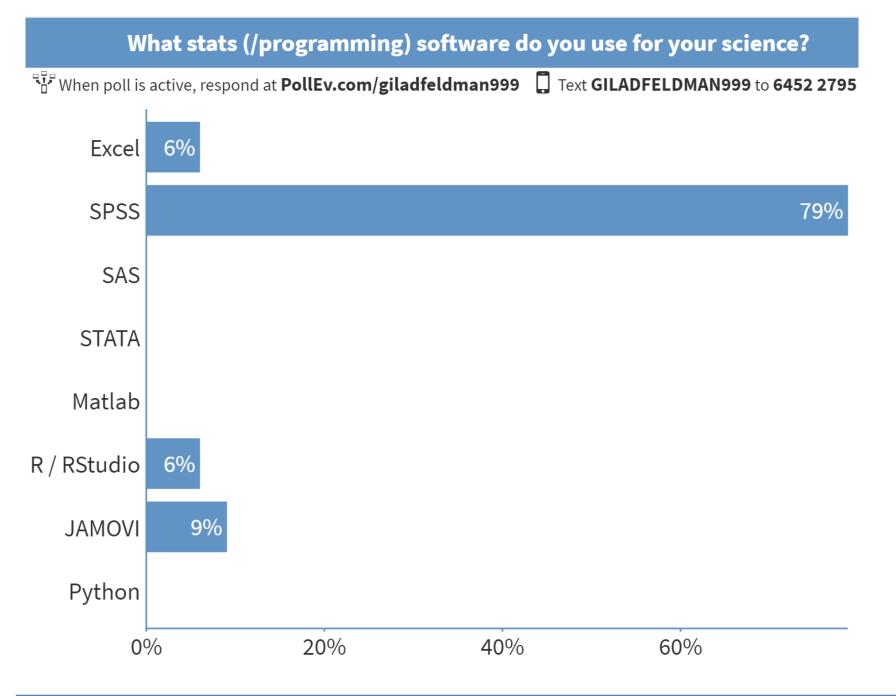
Are you familiar with the replication/reproducibility crisis (or, "credibility revolution")?





Are you familiar with the "new statistics" (AKA, effect sizes and confidence inter





Nov 2018





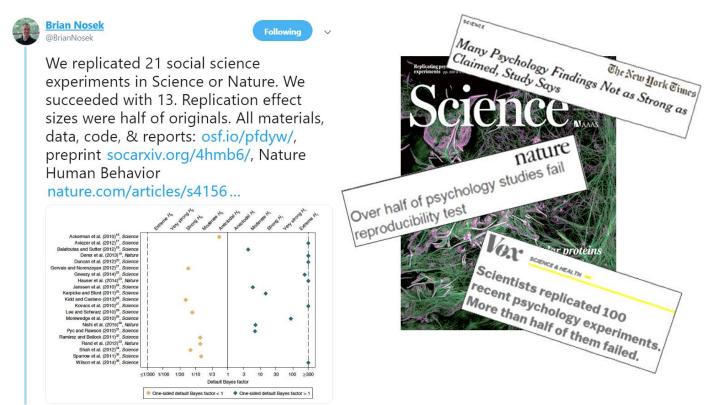
Many Labs 2: 28 findings, 60+ samples, ~7000 participants each study, 186 authors, 36 nations.

Successfully replicated 14 of 28 psyarxiv.com/9654g

ML2 may be more important than Reproducibility Project: Psychology.

Aug 2018

2015



2016





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

RETWEETS 80

74



March 2018

Summary of Vohs RRR findings

- Effect d = 0.08 (very weak $\sim r = .03$)
- Comparable/weaker than first RRR
- Need N = 4908 (power = 80%; alpha = 0.05) to study these effects

Brian Nosek's Summary



Brian Nosek @BrianNosek · Nov 20

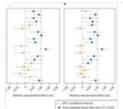
Also, we replicated evidence that surveys & markets can anticipate replication success. There are now 4 studies showing evidence for this.

ML2: psyarxiv.com/9654g

RPP: pnas.org/content/112/50...

EERP: science.sciencemag.org/content/351/62...

SSRP:



Evaluating the replicability of social science experi...

Camerer et al. carried out replications of 21 Science and Nature social science experiments, successfully replicating 13 out of 21 (62%). Effect sizes of replicatio...

nature.com





♡ 51





Following

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: sciencedirect.com/science/articl ...

SSRP: nature.com/articles/s4156...

FFRP:

science.sciencemag.org/content/351/62

RPP:



Estimating the reproducibility of psychological sci...

One of the central goals in any scientific endeavor is to understand causality. Experiments that seek to demonstrate a cause/effect relation most often manipul...

science.sciencemag.org

12:04 AM - 20 Nov 2018

53 Retweets 94 Likes







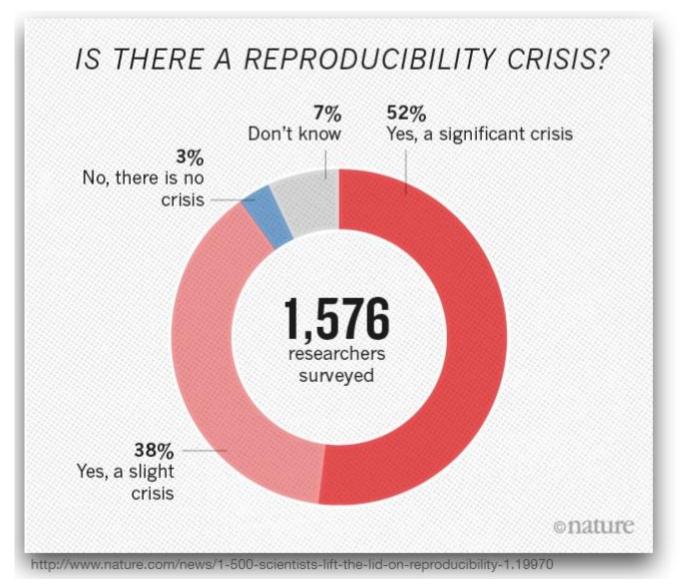








Scholars worldwide: Do we think we have a "crisis"?



90%: Yes

Speaking of Science

No, science's reproducibility problem is not limited to psychology

POLITICS & POLICY

In Medicine, the Science Has Stopped Working

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

Can Reproducibility in Chemical Research be Fixed?

I listen to Post

© Sep 25, 2017 Lenago Academy

The replication crisis has engulfed economics

FROM SLATE, NEW AMERICA, AND ASU

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

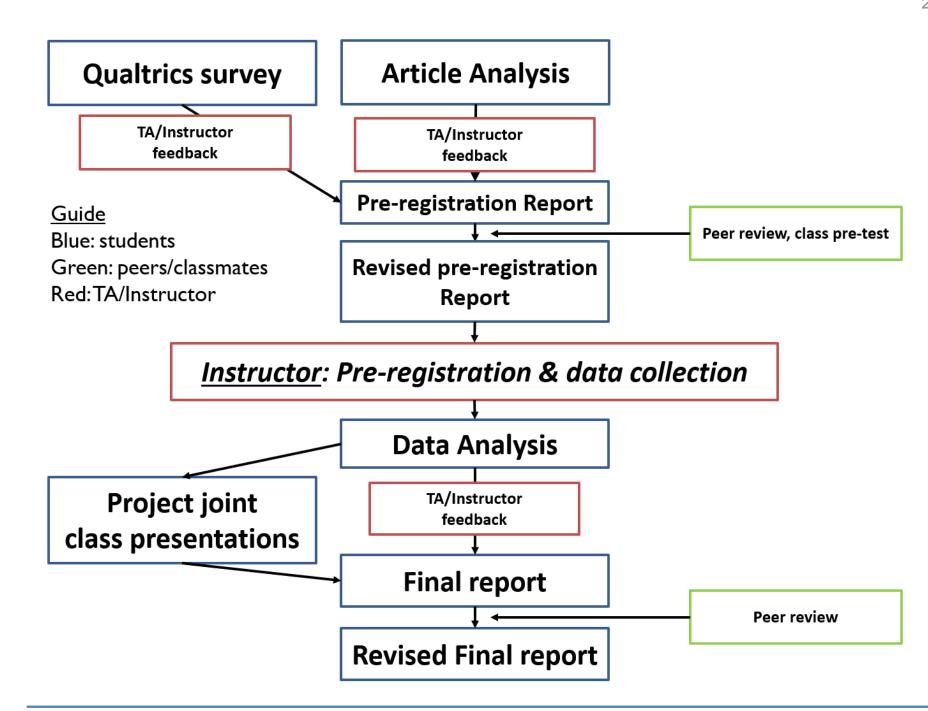
<u>~</u>	ъ.	m ·	T.A.	20
Cl	Date	Topic	TA	Tasks due end of week (Sunday
				11:59pm)
1	05/09	Introduction lecture #1		
		Science under crisis -		
		understanding the crisis		
2	12/09	Introduction lecture #2		Register for HKU Qualtrics
		Addressing the crisis, intro to		account hkupsych.qualtrics.com
		pre-registered replications		Complete quiz on the syllabus
3	19/09	Understanding the ongoing	T1 Qualtrics	
		science crisis	survey	
4	26/09	Why is this happening?	T2 Article analysis	Deadline 30/09 11:59pm:
			(Wednesday class)	Qualtrics survey
5	03/10	Replications and	T2 Article analysis	
		pre-registrations	(Tuesday class)	
6	10/10	Pre-registered replications: Case	T3 Pre-registration	Deadline 14/10 11:59pm:
		studies #1		Article analysis
	17/10	Reading week		Deadline 21/10 11:59pm:
		No class		Pre-registration
7	24/10	Pre-registered replications: Case		Deadline 28/10 11:59pm:
		studies #2		Peer review of pre-registration
				(includes article analysis and
				Qualtrics survey)
				37
8	31/10	"New statistics": Effects and		Deadline 06/11 2:00pm:
		power		Revised pre-registration (includes
		F-11-2		article analysis and Qualtrics
				survey)
9	07/11	Open science & future of	T4 Data analysis	(Students receive data collection
		science		results)
10	14/11	Recommendations to improve	T5 Writing APA	Deadline 20/11 11:59pm:
			style replication	Data analysis
			reports	2 414 41141, 515
11	21/11	Presentations	reports	
12	28/11	Presentations		Deadline 02/12 11:59pm:
				Submission of final report
	05/12	No class		Deadline 09/12 11:59pm:
				Submission of peer review on
				final report
	12/12	No class		Deadline 16/12 11:59pm:
	12/12	110 01100		Submission of revised final report
	L	I	l	Scottission of revised intal report

New course: Judgement and Decision Making (PSYC2071)

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

				21
Cl	Date	Topic	TA	Tasks due end of week (Sunday 11:59pm)
1	06/09	Judgment & Decision making: Introduction lecture #1		
2	13/09	Judgment & Decision making: Introduction lecture #2		Register for HKU Qualtrics account hkupsych.qualtrics.com Complete quiz on the syllabus
3	20/09	Asymmetries (Time / self-other / outcome bias)	T1 Qualtrics survey	
4	27/09	Choices	T2 Article analysis (Wednesday class)	Deadline 30/09 11:59pm: Qualtrics survey
5	04/10	Intuitive statistics	T2 Article analysis (Tuesday class)	
6	11/10	Instructor session on open-science/pre-reg/extensions	T3 Pre-registration	Deadline 14/10 11:59pm: Article analysis
	18/10	Reading week No class		Deadline 21/10 11:59pm: Pre-registration
7	25/10	Hindsight/money		Deadline 28/10 11:59pm: Peer review of pre-registration (includes article analysis and Qualtrics survey)
8	01/11	Escalation/effort		Deadline 06/11 2:00pm: Revised pre-registration (includes article analysis and Qualtrics survey)
9	08/11	Emotions & Counterfactuals	T4 Data analysis	(Students receive data collection results)
10	15/11	Course JDM summary + Academic writings and journal submissions	T5 Writing APA style replication reports	Deadline 20/11 11:59pm: Data analysis
11	22/11	Presentations		
12	29/11	Presentations		Deadline 02/12 11:59pm: Submission of final report
	06/12	No class		Deadline 09/12 11:59pm: Submission of peer review on final report
	13/12	No class		Deadline 16/12 11:59pm: Submission of revised final report

Process



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)

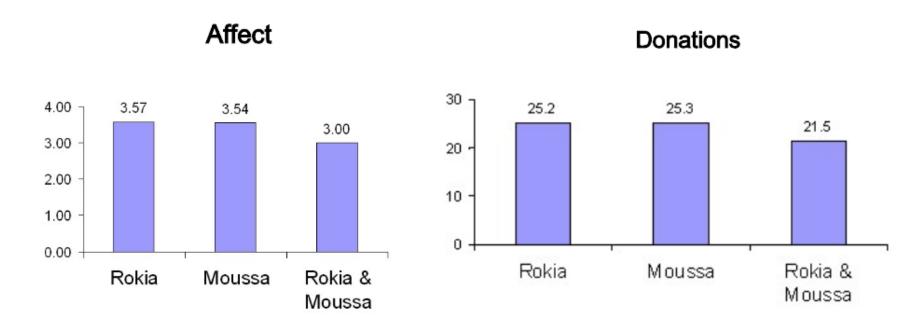
2 IVs - IV1 (1/2/control) x IV2 (A/B):

IV1: Outcome manipulation	IV1: Experimental 1	IV1: Experimental 2	IV1: Control condition			
IV2: Self-other manipulation	condition	condition	Title example: No			
	Title example: Negative	Title example: Positive	outcome			
	outcome condition	Manipulation				
	Manipulation example:	Manipulation example:	example: Investor			
	Investor scenario	scenario described				
	described as resulting in a	described as resulting in a described as resulting in a				
	loss of money	outcome				
IV2: Experiment A condition	Dependent variable					
Title example: self condition	Title example: Evaluations of investment decision					
Manipulation example:	Specific DV item: Please evaluate your investment decision on a scale of 0-6					
Investment described as made	(very bad to very good)					
by self - you're the investor						
IV2: Experiment A condition	Dependent variable					
Title example: other condition	Title example: Evaluations of investment decision					
Manipulation example:	Specific DV item: Please evaluate the person's investment decision on a scale					
Investment described as made	of 0-6 (very bad to very good)					
by other - someone else is the						
investor						

Inclass demonstrations Of the experiments

Discussion

Compassion fade and psychic numbing



Västfjäll D, Slovic P, Mayorga M, Peters E (2014) Compassion Fade: Affect and Charity Are Greatest for a Single Child in Need. PLoS ONE 9(6): e100115. doi:10.1371/journal.pone.0100115

Slovic, P. (2007), If I Look at Mass I Will Never Act: Psychic Numbing and Genocide. In *Judgment and Decision Making*, Volume 2, no. 2, pp. 79-95.

"how do you feel about donating to rokia [moussa][rokia & Moussa]?" -1 to 5

Two versions

Given a middle-class Hong Kong student status, how much money

would you be willing to donate to save Rokia?



Rokia

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







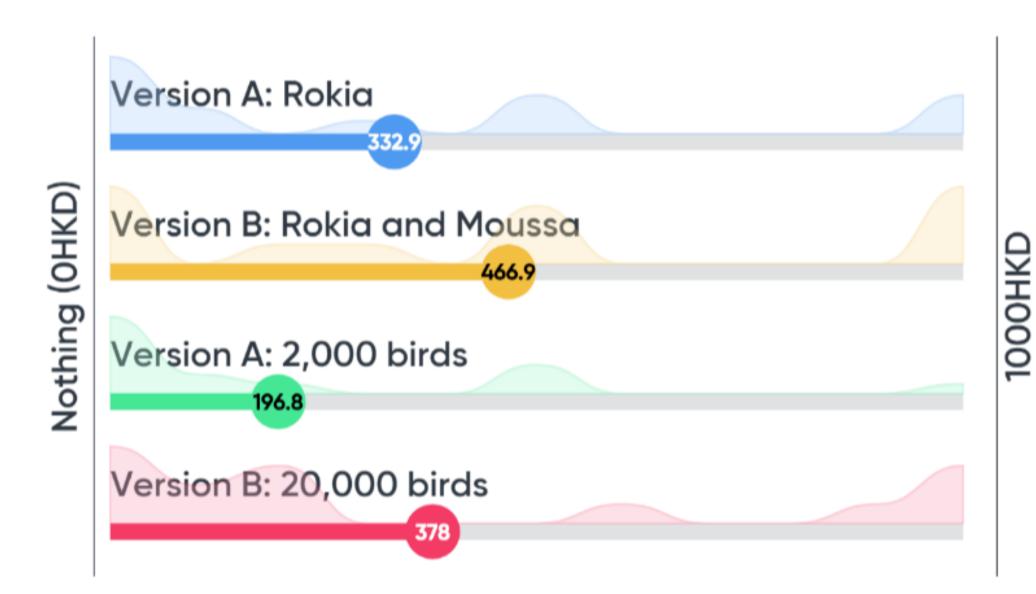
Moussa

MENTIMETER

Go to www.menti.com and use the code 12 70 69

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

Inclass voting results



Prediction markets: Setup



Gilad Feldman

@giladfeldman

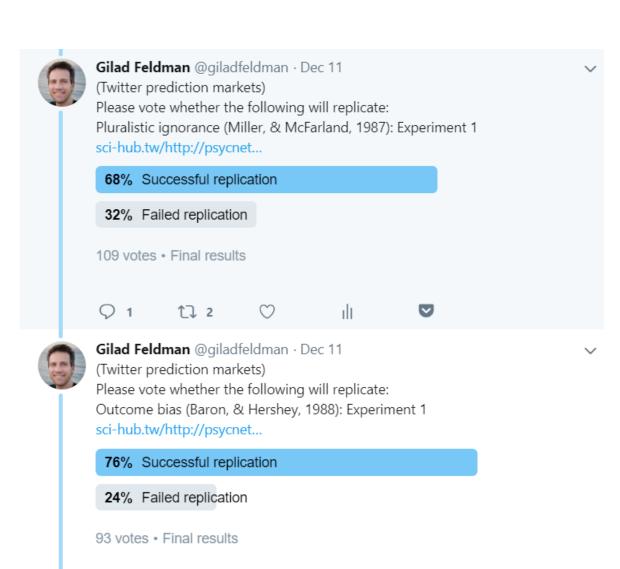
Dec 20 I will summarize findings from 18 replications we conducted this semester. We wanted prediction markets, but were overwhelmed by work just running the replications.

Instead, a Twitter vote, you're welcome to rate replicability.

Below are studies we attempted to replicate.

10:51 PM - 11 Dec 2018





Prediction markets: Attention

Tweets Top Tweets Tweets and replies Promoted Impressions Engagements

Gilad Feldman @giladfeldman · Dec 11 26,307 466



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.

Impressions 26,3	37
Total engagements 4	68
Detail expands	221
Profile clicks	101
Likes	80
Retweets	61

Prediction markets: Results

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

Prediction markets: Reactions



Michał Białek @mbialek82 · 4m

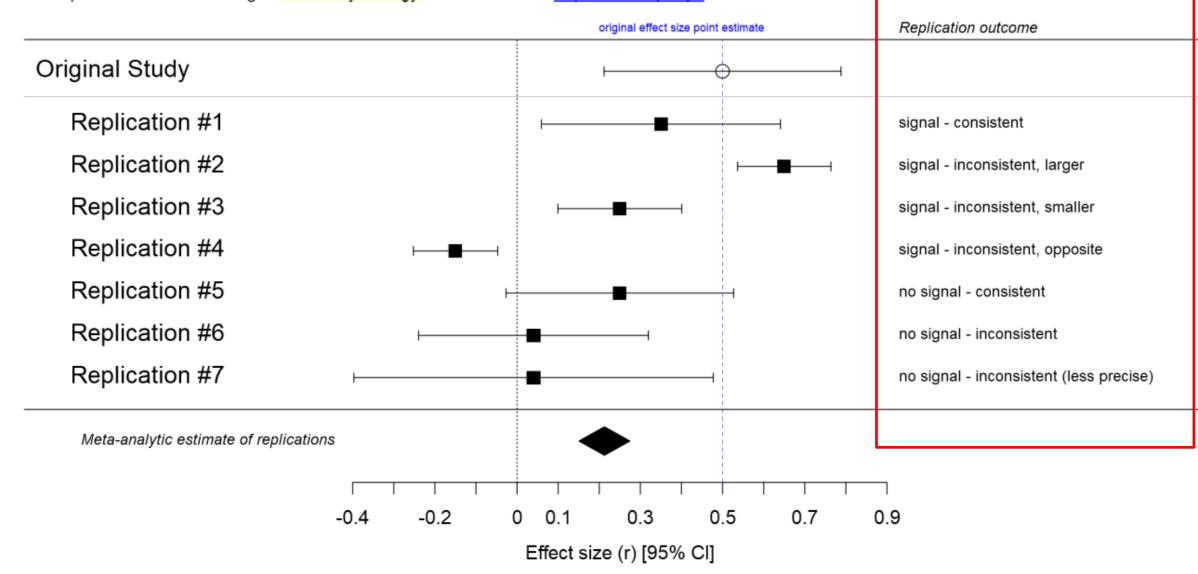
Replying to @giladfeldman

More scientists believe in hindsight bias than in global warming....



Before results: How do we evaluate replications?

LeBel, E. P., Vanpaemel, W., Cheung, I., & Campbell, L. (in press). A Brief Guide to Evaluate Replications. Forthcoming at Meta-Psychology. Retrieved from https://osf.io/paxyn



Time for some data. Ready for the replication results?

First, we were pretty well powered

Original	Replic	ation
N	N	χN
20	692	34.6
102	401	3.9
83	1034	12.5
260	894	3.4
124	890	7.2
99	401	4.1
256	403	1.6
200	705	3.5
23	401	17.4
432	1032	2.4
116	400	3.4
44	604	13.7
277	1026	3.7
323	604	1.9
184	604	3.3
240	403	1.7
298	894	3.0
78	452	5.8

Results (with extensions)

				Prediction	market				Original	Replic	ation					Conclusion	Extension	1		Extension	2	
St	Authors	Studies	Bias	Votes	Success	Failure	Direction	sig<>0.5	N	N	xΝ	CS	CNS	IS	INS	Conclusions	Factor	Construct	Result	Factor	Construct	Result
1	Baron, & Hershey, 1988	1	Outcome bias	93	76%	249	6 Success	TRUE	20	692	34.6	1				Successful	DV	Responsib	Success	DV	Action	Success
2	Epley & Gilovich 2006	1b	Anchoring-and-adjustment heuristic	114	88%	129	6 Success	TRUE	102	401	3.9	Too r	nessy			Estimated low replicability	N/A					
3	Epstein, Lipson, Holstein, & Huh 1992	1 & 2	Irrational reactions to negative outcome	64	55%	459	6 Success	FALSE	83	1034	12.5	2	1	L		Mostly successful	N/A					
4	Fischhoff, 1975	2	Hindsight bias	97	98%	29	6 Success	TRUE	260	894	3.4	11		Ļ	1	Successful	DV	Surprise	Mixed			
5	Hamill, Wilson, & Nisbett, 1980	1	Insensitivity to sample bias	92	68%	329	6 Success	TRUE	124	890	7.2	2			1	Inconclusive, found typica	I <mark>-</mark> DV	Conscious	Failed			
6	Hsee & Weber, 1997	1	Fundamental predictor error	77	70%	309	6 Success	TRUE	99	401	4.1	2	1	L		Mostly successful	DV	Regret	Success	DV	Willingnes	N/A
7	Hsee, 1998	1, 2, 4	Less is better	88	53%	479	6 Success	FALSE	256	403	1.6	2		1		Mostly successful	DV	Attractive	Success	DV	Perceived	Success
8	Kruger etal, 2004	1 & 2	Effort heuristic	73	74%	269	6 Success	TRUE	200	705	3.5	1			1	1 semi-successful, 1 failed	Combinat	ion	Mixed			
9	Kruger, Wirtz & Miller 2005	2	First instinct fallacy	60	60%	409	6 Success	FALSE	23	401	17.4	3				Successful, stronger effects	DV	Strategy/N	Success	IV	Self/other	Success
10	Mellers, Hertwig, & Kahneman, 2001	1	Conjunction effect	85	85%	159	6 Success	TRUE	432	1032	2.4	2				Successful	DV	Populatio	Success			
11	Miller, & McFarland, 1987	1	Pluralistic ignorance	109	68%	329	6 Success	TRUE	116	400	3.4	2		1		Mixed findings, main hypo	IV IV	Friend	Mixed			
12	Schwarz, Strack, Hilton, & Naderer, 1991	1	Relevance of irrelevant information	91	48%	529	6 Failure	FALSE	44	604	13.7		1	L	1	Failure to replicate	DV	Bayes esti	Success			
13	Shafir, 1993	1 to 8	Choosing versus rejecting	82	48%	529	6 Failure	FALSE	277	1026	3.7	2	2	2 2	2	Failure to replicate	DV	attractive	N/A	IV	Problem to	Mixed
14	Shafir, Diamond, & Tversky, 1997	1 to 4	Money illusion	70	59%	419	6 Success	FALSE	323	604	1.9	4				Successful	DV	Confidence	Success	Predictor	Knowledge	N/A
15	Slovic & Fischhoff, 1977	1	Hindsight bias	111	96%	49	6 Success	TRUE	184	604	3.3	5	3	3		Successful	DV	Confidence	Failed	DV	Surprise/c	Mixed
16	Staw, 1976	1	Escalation of commitment	78	86%	149	6 Success	TRUE	240	403	1.7	1		1		Inconclusive -> Failure	N/A					
17	Tversky & Shafir, 1992	1	Disjunction effect	83	76%	249	6 Success	TRUE	298	894	3.0	1			1	1 successful, 1 failed	Predictor	intolerand	Not analy:	zed		
18	Zeelenberg etal 1996	1	Regret aversion	75	73%	279	6 Success	TRUE	78	452	5.8		1		1	Inconclusive -> Failure	Predictor	Choice reg	Failed	IV	Both	
	Warning: Student calculations, needs to	be rever	ified, and checked.	CS = Cons	istent signa	l	CNS = Cor	nsistent no	signal		IS = Inco	nsist	ent sig	nal	INS =	Inconsistent no signal						

Results: Zoom in

				Prediction	market				Original	Replic	ation					Conclusion
St	Authors	Studies				Failure	Direction			_		CS	CNS	IS	INS	Conclusions
1	Baron, & Hershey, 1988	1	Outcome bias	93	76%	24%	Success	TRUE	20	692	34.6	1				Successful
2	Epley & Gilovich 2006	1b	Anchoring-and-adjustment heuristic	114	88%	12%	Success	TRUE	102	401	3.9	Too n	nessy			Estimated low replicability.
3	Epstein, Lipson, Holstein, & Huh 1992	1 & 2	Irrational reactions to negative outcome	64	55%	45%	Success	FALSE	83	1034	12.5	2	1			Mostly successful
4	Fischhoff, 1975	2	Hindsight bias	97	98%	2%	Success	TRUE	260	894	3.4	11	4		1	Successful
5	Hamill, Wilson, & Nisbett, 1980	1	Insensitivity to sample bias	92	68%	32%	Success	TRUE	124	890	7.2	2			1	Inconclusive, found typical-
6	Hsee & Weber, 1997	1	Fundamental predictor error	77	70%	30%	Success	TRUE	99	401	4.1	2	1			Mostly successful
7	Hsee, 1998	1, 2, 4	Less is better	88	53%	47%	Success	FALSE	256	403	1.6	2		1		Mostly successful
8	Kruger etal, 2004	1 & 2	Effort heuristic	73	74%	26%	Success	TRUE	200	705	3.5	1			1	1 semi-successful, 1 failed
9	Kruger, Wirtz & Miller 2005	2	First instinct fallacy	60	60%	40%	Success	FALSE	23	401	17.4	3				Successful, stronger effects
10	Mellers, Hertwig, & Kahneman, 2001	1	Conjunction effect	85	85%	15%	Success	TRUE	432	1032	2.4	2				Successful
11	Miller, & McFarland, 1987	1	Pluralistic ignorance	109	68%	32%	Success	TRUE	116	400	3.4	2		1		Mixed findings, main hypot
12	Schwarz, Strack, Hilton, & Naderer, 1991	1	Relevance of irrelevant information	91	48%	52%	Failure	FALSE	44	604	13.7		1		1	Failure to replicate
13	Shafir, 1993	1 to 8	Choosing versus rejecting	82	48%	52%	Failure	FALSE	277	1026	3.7	2	2	2	2	Failure to replicate
14	Shafir, Diamond, & Tversky, 1997	1 to 4	Money illusion	70	59%	41%	Success	FALSE	323	604	1.9	4				Successful
15	Slovic & Fischhoff, 1977	1	Hindsight bias	111	96%	4%	Success	TRUE	184	604	3.3	5	3			Successful
16	Staw, 1976	1	Escalation of commitment	78	86%	14%	Success	TRUE	240	403	1.7	1		1		Inconclusive -> Failure
17	Tversky & Shafir, 1992	1	Disjunction effect	83	76%	24%	Success	TRUE	298	894	3.0	1			1	1 successful, 1 failed
18	Zeelenberg etal 1996	1	Regret aversion	7 5	73%	27%	Success	TRUE	78	452	5.8		1		1	Inconclusive -> Failure
	Warning: Student calculations, needs to	be rever	ified, and checked.	CS = Consi	stent signa	al	CNS = Con	sistent n	o signal		IS = Inco	onsiste	ent sign	nal	INS =	Inconsistent no signal

Zoom in: Compare predictions to conclusions

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

CS = Consistent signal

CNS = Consistent no signal

IS = Inconsistent signal INS = Inconsistent no signal

Summary (out of 18)

Successful (9)

- Strong replication: 6
- Mostly successful: 3

9/18

50% replication rate

Inconclusive (5)

- Inconclusive: 4
- Messy, hard to tell, likely failure: I

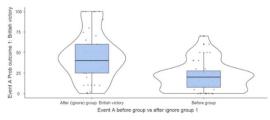
Failure (4)

- Most likely failure: 2
- Inconclusive but likely failure: 2

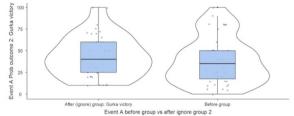
How well did the students do?

Hindsight (Fischhoff 1975)

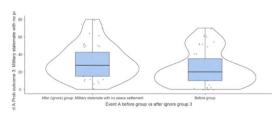
Event A Outcome 1



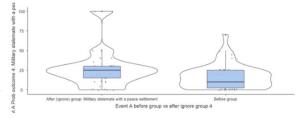
Event A Outcome 2



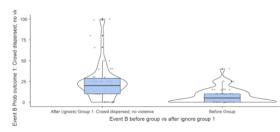
Event A Outcome 3



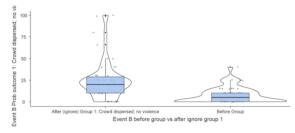
Event A Outcome 4



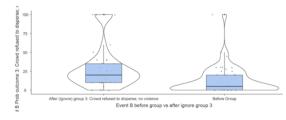
Event B Outcome 1



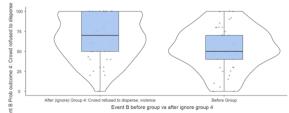
Event B Outcome 2



Event B Outcome 3



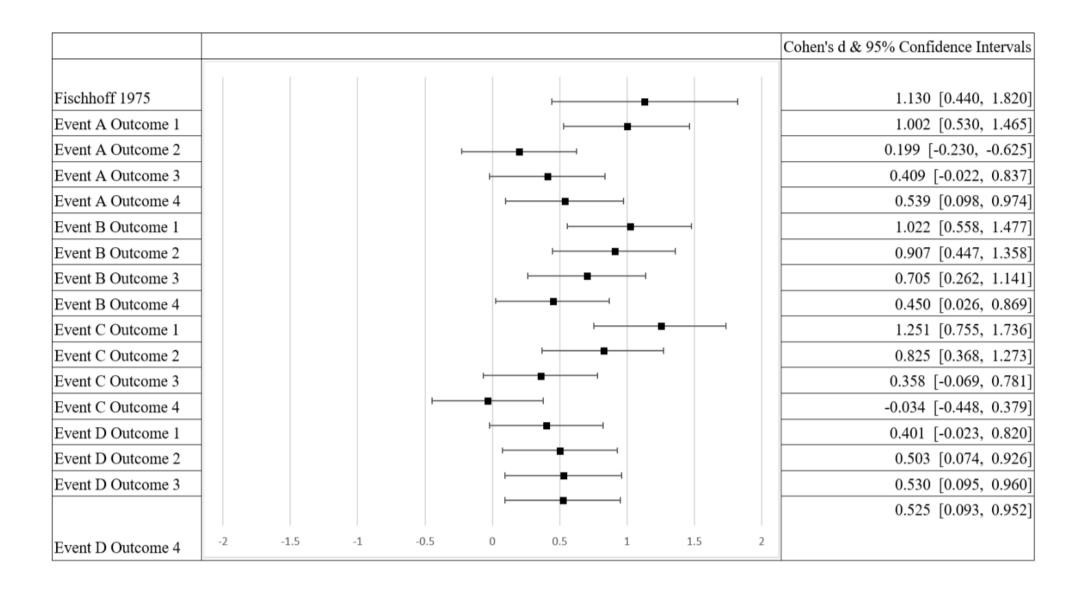
Event B Outcome 4



				Same	Replication	Effect in	Effect	Note
		Cohen's d & 95% Confidence		Direction?	CI include	target CIs?	magnitude	
	Result	Intervals	p-value		null?		(Similar: diff<0.3)	
Fischhoff				/	/	/	/	
1975	Original Article	1.130 [0.440, 1.820]	<.001**					
Present	Event A			Yes	Yes	Yes	Similar	Consistent Signal
Study	Outcome 1	1.002 [0.530, 1.465]	<.001**					
	Event A			Yes	No	No	Different	No signal – consistent
	Outcome 2	0.199 [-0.230, -0.625]	0.279					direction
	Event A			Yes	No	No	Different	No signal – consistent
	Outcome 3	0.409 [-0.022, 0.837]	0.032*					direction
	Event A			Yes	Yes	Yes	Different	Consistent Signal
	Outcome 4	0.539 [0.098, 0.974]	0.009**					
	Event B			Yes	Yes	Yes	Similar	Consistent Signal
	Outcome 1	1.022 [0.558, 1.477]	<.001**					
	Event B			Yes	Yes	Yes	Similar	Consistent Signal
	Outcome 2	0.907 [0.447, 1.358]	<.001**					
	Event B			Yes	Yes	Yes	Different	Consistent Signal
	Outcome 3	0.705 [0.262, 1.141]	<.001**					
	Event B			Yes	Yes	Yes	Different	Consistent Signal
	Outcome 4	0.450 [0.026, 0.869]	0.041*					
	Event C			Yes	Yes	Yes	Similar	Consistent Signal
	Outcome 1	1.251 [0.755, 1.736]	<.001**					
	Event C			Yes	Yes	Yes	Similar	Consistent Signal
	Outcome 2	0.825 [0.368, 1.273]	<.001**					
	Event C			Yes	No	No	Different	No signal – consistent
	Outcome 3	0.358 [-0.069, 0.781]	0.039*					direction
	Event C			No	No	No	Different –	No signal – inconsistent
	Outcome 4	-0.034 [-0.448, 0.379]	0.569				Opposite	direction
	Event D			Yes	No	No	Different	No signal – consistent
	Outcome 1	0.401 [-0.023, 0.820]	0.016*					direction
	Event D			Yes	Yes	Yes	Different	Consistent Signal
	Outcome 2	0.503 [0.074, 0.926]	0.004**					
	Event D			Yes	Yes	Yes	Different	Consistent Signal
	Outcome 3	0.530 [0.095, 0.960]	0.022*					
	Event D			Yes	Yes	Yes	Different	Consistent Signal
	Outcome 4	0.525 [0.093, 0.952]	0.003**					

^{*}p<.05; **p<.01

Hindsight (Fischhoff 1975)



Shafir 1993: We addressed most issues in reply to Manylabs2

Hypothesis 1: Enriched option will be chosen and rejected more often than the impoverished option > Enriched options share will exceed 100%

,	Enriched options share	z-value	p-value	effect size (Cohen's d)	confidence interval (CI = 95%)
Problem 1: Sole-custody case	1.018	0.58	= .56	0.0362 (no effect)	[-0.0862, 0.1586]
Problem 2: Vacation spot	0.958	-1.35	= .18	-0.0843 (no effect)	[-0.2067, 0.0382]
Problem 3: Course selection	0.951	-1.57	= .12	-0.0981 (no effect)	[-0.2205, 0.0244]
Problem 4: Lottery A	1.130	4.16	<.001	0.2619 (small effect)	[0.1389, 0.3847]
Problem 5: Lottery B	1.186	5.96	<.001	0.3787 (small effect)	[0.2551, 0.5021]
Problem 6: Ice-cream flavor	1.033	1.06	= .29	0.0662 (no effect)	[-0.0562, 0.1886]
Problem 7: Election	0.802	-6.34	<.001	-0.4038 (small effect)	[-0.5273, -0.2801]
Problem 8: Lottery C (non-binary)	0.899	-3.24	= .001	-0.2033 (small effect)	[-0.3259, -0.0806]

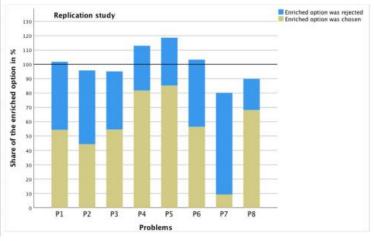
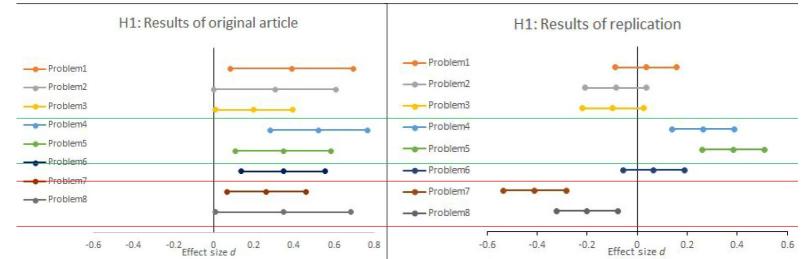


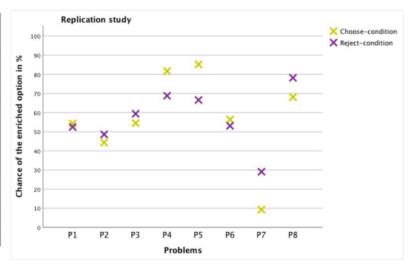
Table 2: Comparison of the enriched options observed share with the expected share of 100% with z-value, p-value and effect size with the 95% confidence interval.

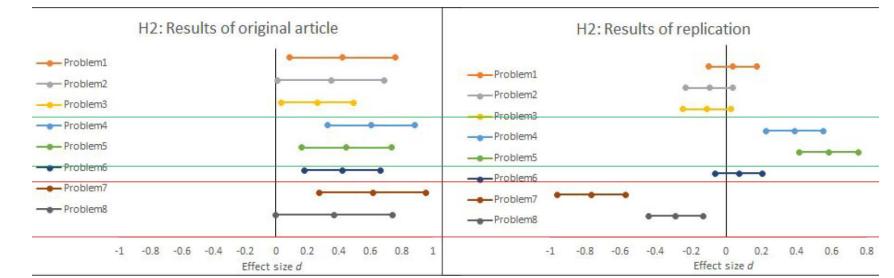


Shafir 1993: We addressed most issues in reply to Manylabs 2

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

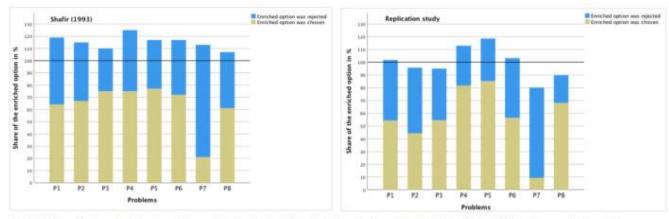
	Chance of the enriched	z-value	p-value	effect size	confidence interval
	option in the choose-		-	(Cohen's d)	(CI = 95%)
	vs. reject condition				
Problem 1:	0.543 vs. 0.525	0.58	= .28	0.0511	[-0.1219, 0.2240]
Sole-custody case				(no effect)	
Problem 2:	0.444 vs 0.486	-1.35	= .09	-0.1193	[-0.2923, 0.0538]
Vacation spot				(no effect)	
Problem 3:	0.545 vs. 0.594	-1.58	= .06	-0.1397	[-0.3127, 0.0335]
Course selection				(no effect)	
Problem 4:	0.817 vs. 0.688	4.79	< .001	0.4323	[0.2572, 0.6070]
Lottery A				(small effect)	
Problem 5:	0.852 vs. 0.666	6.97	< .001	0.6461	[0.4684, 0.8232]
Lottery B				(medium effect)	
Problem 6:	0.564 vs. 0.531	1.06	< .14	0.0936	[-0.0794, 0.2666]
Ice-cream flavor				(no effect)	
Problem 7:	0.093 vs. 0.291	-8.05	< .001	-0.7596	[-0.9383, -0.5802]
Election				(medium effect)	
Problem 8:	0.681 vs. 0.781	-3.65	< .001	-0.3262	[-0.5001, -0.1520]
Lottery C (non-binary)				(small effect)	





Shafir 1993

Comparison: Shafir (1993) vs. Replication



Hypothesis 1: Enriched options share exceeds 100%

No (consistent) replication

Fig. 1. Findings for hypothesis 1; share of the enriched option in % for all problems in the original article by Shafir (1993) and our replication study

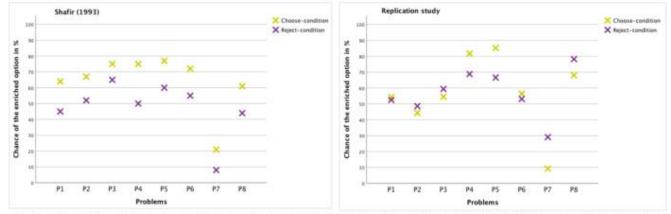


Fig. 2. Findings for hypothesis 2; chance of the enriched option in % for all problems in the original article by Shafir (1993) and our replication study

Hypothesis 2:

Higher chance of the enriched option in the choose-condition

No (consistent) replication

I showed some reports to collaborators

The general response?







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 <u>researchers</u>, as <u>participants</u>, 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?!

ROYAL Society

Publishing blog

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

Journal	2017 Impact factor	Rank	5 year impact factor	Cited half- life	Immediacy index
Royal Society Open Science	2.504	17th out of 64 in 'Multidisciplinary Sciences'	2.642	1.9	0.406

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.

2017 Impact factor

2.504

- 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:



Feb 11, 2017

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.

Comment

Recommend

Share

1 Comment · 15 Reads



Sander L. Koole

레 38.79 · Vrije Universiteit Amsterdam

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 · Recommend · Share ·

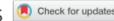
Yeah, replications & extensions can be published...

In Cognition & Emotion:

COGNITION AND EMOTION https://doi.org/10.1080/02699931.2018.1504747







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

Lucas Kutscher^a and Gilad Feldman ^{o a,b}

^aDepartment of Work and Social Psychology, Maastricht University, Maastricht, the Netherlands; ^bDepartment 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**

Social Psychological and Personality Science 2018, Vol. 9(5) 539-549 © The Author(s) 2017



Article reuse guidelines: sagepub.com/journals-permissions DOI: 10.1177/1948550617713254 journals.sagepub.com/home/spp

(\$)SAGE

Gilad Feldman¹ and Subramanya Prasad Chandrashekar²

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

Answer



Yajing Gao 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

1 Comment · 33 Reads

Mar 3, 2017



Sander L. Koole

레 38.79 · Vrije Universiteit Amsterdam

Share

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

Mar 28, 2017 · Recommend · Share ·

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: I-4 (!) APA ready submissions, complete with preregistrations (power analyses->high power), full open-science data/code, replication recipe...

An opportunity for you to have a lead author preregistered 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
- Contact: gfeldman@hku.hk
- Twitter: @giladfeldman

