

Preregistrace a registrované reporty zvyšují kvalitu (nejen) sociálně-vědního výzkumu

Hynek Cígler | Institut pro psychologický výzkum
Fakulta sociálních studií | Masarykova univerzita

Open Science Week | 21.–23. 10. 2024 | online

Obsah přednášky

Kontext: Počátky a podoba replikační krize

Příčiny replikační krize: QRP, publikační zkreslení.

Dopady: Zkreslení existence a síly efektů.

Způsoby řešení.

Preregistrace a registrované reporty.

Kazuistika: Open Science Framework

US

University of Sussex

Why I don't Believe Anything in Psychology

Professor Andy Field

Počátky *replikační krize* v psychologii



Daryl Bem (2011): Feeling the Future



John Bargh (2010): priming stáří

Estimating the reproducibility of psychological science

„We conducted a large-scale, collaborative effort to obtain an initial estimate of the reproducibility of psychological science.“

100 studií a výsledky jejich replikace

- Psychological Science
- Journal of Personality and Social Psychology
- Journal of Experimental Psychology: Learning, Memory, and Cognition

Open Science Collaboration. (2015). Estimating the reproducibility of psychological science. *Science*, 349(6251), aac4716.

<https://doi.org/10.1126/science.aac4716>

Alexander A. Aarts, Joanna E. Anderson, Christopher J. Anderson, Peter R. Attridge, Angela Attwood, Jordan Axt, Molly Babel, Štěpán Bahnik, Erica Baranski, Michael Barnett-Cowan, Elizabeth Bartmess, Jennifer Beer, Raul Bell, Heather Bentley, Leah Beyan, Grace Binion, Denny Borsboom, Annick Bosch, Frank A. Bosco, Sara D. Bowman, Mark J. Brandt, Erin Braswell, Hilmar Brohmer, Benjamin T. Brown, Kristina Brown, Jovita Brüning, Ann Calhoun-Sauls, Shannon P. Callahan, Elizabeth Chagnon, Jesse Chandler, Christopher R. Chartier, Felix Cheung, Cody D. Christopherson, Linda Cillessen, Russ Clay, Hayley Cleary, Mark D. Cloud, Michael Cohn, Johanna Cohoon, Simon Columbus, Andreas Cordes, Giulio Costantini, Leslie D. Cramblet Alvarez, Ed Cremata, Jan Crusius, Jamie DeCoster, Michelle A. DeGaetano, Nicolás Della Penna, Bobby den Bezemer, Marie K. Deserno, Olivia Devitt, Laura Dewitte, David G. Dobolyi, Geneva T. Dodson, M. Brent Donnellan, Ryan Donohue, Rebecca A. Dore, Angela Dorrough, Anna Dreber, Michelle Dugas, Elizabeth W. Dunn, Kayleigh Easey, Sylvia Eboigbe, Casey Eggleston, Jo Embley, Sacha Epskamp, Timothy M. Errington, Vivien Estel, Frank J. Farach, Jenelle Feather, Anna Fedor, Belén Fernández-Castilla, Susann Fiedler, James G. Field, Stanka A. Fitneva, Taru Flagan, Amanda L. Forest, Eskil Forsell, Joshua D. Foster, Michael C. Frank, Rebecca S. Frazier, Heather Fuchs, Philip Gable, Jeff Galak, Elisa Maria Galliani, Anup Gampa, Sara Garcia, Douglas Gazarian, Elizabeth Gilbert, Roger Giner-Sorolla, Andreas Glöckner, Lars Goellner, Jin X. Goh, Rebecca Goldberg, Patrick T. Goodbourn, Shauna Gordon-McKeon, Bryan Gorges, Jessie Gorges, Justin Goss, Jesse Graham, James A. Grange, Jeremy Gray, Chris Hartgerink, Joshua Hartshorne, Fred Hasselman, Timothy Hayes, Emma Heikensten, Felix Henninger, John Hodsoll, Taylor Holubar, Gea Hoogendoorn, Denise J. Humphries, Cathy O.-Y. Hung, Nathali Immelman, Vanessa C. Irsik, Georg Jahn, Frank Jäkel, Marc Jekel, Magnus Johannesson, Larissa G. Johnson, David J. Johnson, Kate M. Johnson, William J. Johnston, Kai Jonas, Jennifer A. Joy-Gaba, Heather Barry Kappes, Kim Kelso, Mallory C. Kidwell, Seung Kyung Kim, Matthew Kirkhart, Bennett Kleinberg, Goran Knežević, Franziska Maria Kolorz, Jolanda J. Kossakowski, Robert Wilhelm Krause, Job Krijnen, Tim Kuhlmann, Yoram K. Kunkels, Megan M. Kyc, Calvin K. Lai, Aamir Laique, Daniël Lakens, Kristin A. Lane, Bethany Lassetter, Ljiljana B. Lazarević, Etienne P. LeBel, Key Jung Lee, Minha Lee, Kristi Lemm, Carmel A. Levitan, Melissa Lewis, Lin Lin, Stephanie Lin, Matthias Lippold, Darren Loureiro, Ilse Luteijn, Sean Mackinnon, Heather N. Mainard, Denise C. Marigold, Daniel P. Martin, Tylar Martinez, E.J. Masicampo, Josh Matacotta, Maya Mathur, Michael May, Nicole Mechin, Pranjal Mehta, Johannes Meixner, Alissa Melinger, Jeremy K. Miller, Mallorie Miller, Katherine Moore, Marcus Möschl, Matt Motyl, Stephanie M. Müller, Marcus Munafo, Koen I. Neijenhuijs, Taylor Nervi, Gandalf Nicolas, Gustav Nilsson, Brian A. Nosek, Michèle B. Nuijten, Catherine Olsson, Colleen Osborne, Lutz Ostkamp, Misha Pavel, Ian S. Penton-Voak, Olivia Perna, Cyril Pernet, Marco Perugini, R. Nathan Pipitone, Michael Pitts, Franziska Plessow, Jason M. Prenoveau, Rima-Maria Rahal, Kate A. Ratliff, David Reinhard, Frank Renkewitz, Ashley A. Ricker, Anastasia Rigney, Andrew M. Rivers, Mark Roebke, Abraham M. Rutchick, Robert S. Ryan, Onur Sahin, Anondah Saide, Gillian M. Sandstrom, David Santos, Rebecca Saxe, René Schlegelmilch, Kathleen Schmidt, Sabine Scholz, Larissa Seibel, Dylan Faulkner Selterman, Samuel Shaki, William B. Simpson, H. Colleen Sinclair, Jeanine L. M. Skorinko, Agnieszka Slowik, Joel S. Snyder, Courtney Soderberg, Carina Sonnleitner, Nick Spencer, Jeffrey R. Spies, Sara Steegen, Stefan Stieger, Nina Strohminger, Gavin B. Sullivan, Thomas Talhelm, Megan Tapia, Annie te Dorsthorst, Manuela Thomae, Sarah L. Thomas, Pia Tio, Frits Traets, Steve Tsang, Francis Tuerlinckx, Paul Turchan, Milan Valášek, Anna E. van 't Veer, Robbie Van Aert, Marcel van Assen, Riet van Bork, Mathijs van de Ven, Don van den Bergh, Marije van der Hulst, Roel van Dooren, Johnny van Doorn, Daan R. van Renswoude, Hedderik van Rijn, Wolf Vanpaemel, Alejandro Vásquez Echeverría, Melissa Vazquez, Natalia Velez, Marieke Vermue, Mark Verschoor, Michelangelo Vianello, Martin Voracek, Gina Vuu, Eric-Jan Wagenmakers, Joanneke Weerdmeester, Ashlee Welsh, Erin C. Westgate, Joeri Wissink, Michael Wood, Andy Woods, Emily Wright, Sining Wu, Marcel Zeelenberg, Kellylynn Zuni

Estimating the reproducibility of psychological science

Původní velikost efektů:

- Průměrná velikost účinku
 $M_r = 0,403$; $SD = 0,188$
- Statistická signifikance: 97 % studií $p < 0,05$

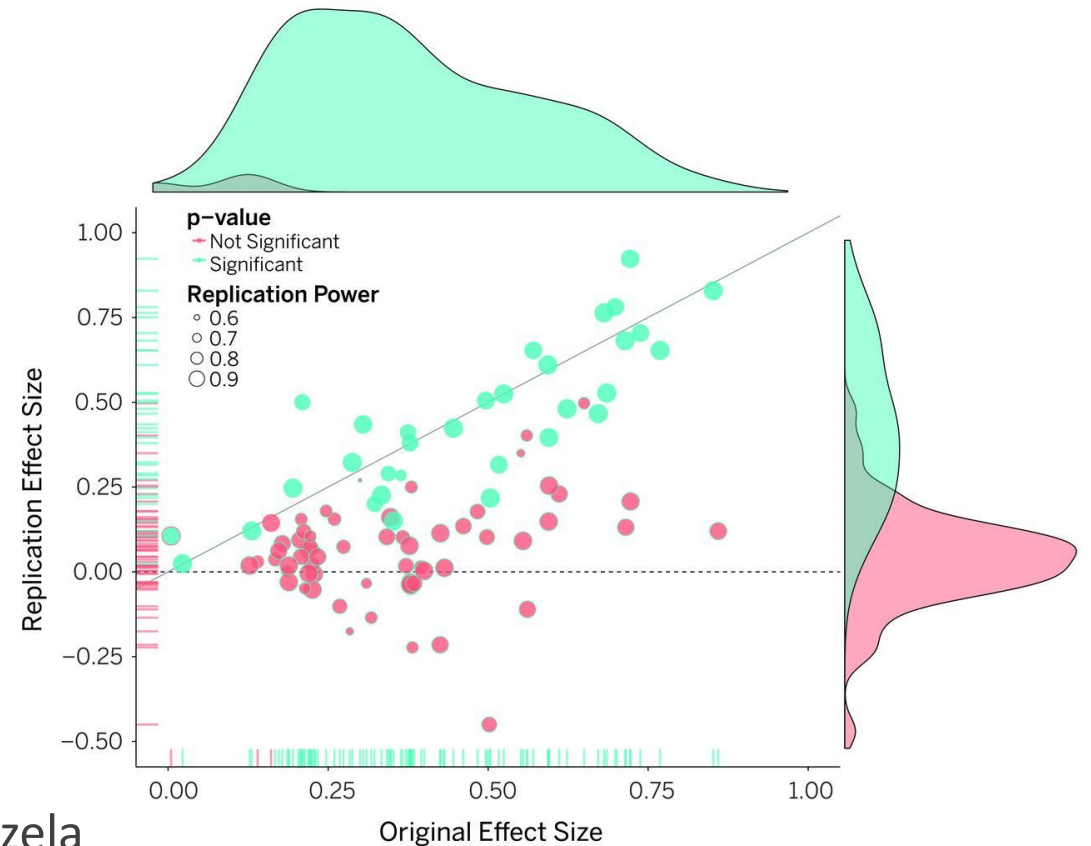
Design replikací: průměrná síla testu $1-\beta = 0,92$.

- → 89 % replikací by mělo být signifikantní.
- Ale: průměrná síla testu originálních studií: **39 %**.

Replikovaná velikost efektů:

- Průměrná velikost účinku
 $M_r = 0,197$; $SD = 0,257$
- Statistická signifikance: **36 % studií $p < 0,05$**

Hodnota velikostí účinku z původních studií se nacházela v 95% intervalu spolehlivosti při replikaci v 47 % případů.



Příklady nereplikovatelných

Priming (social priming).

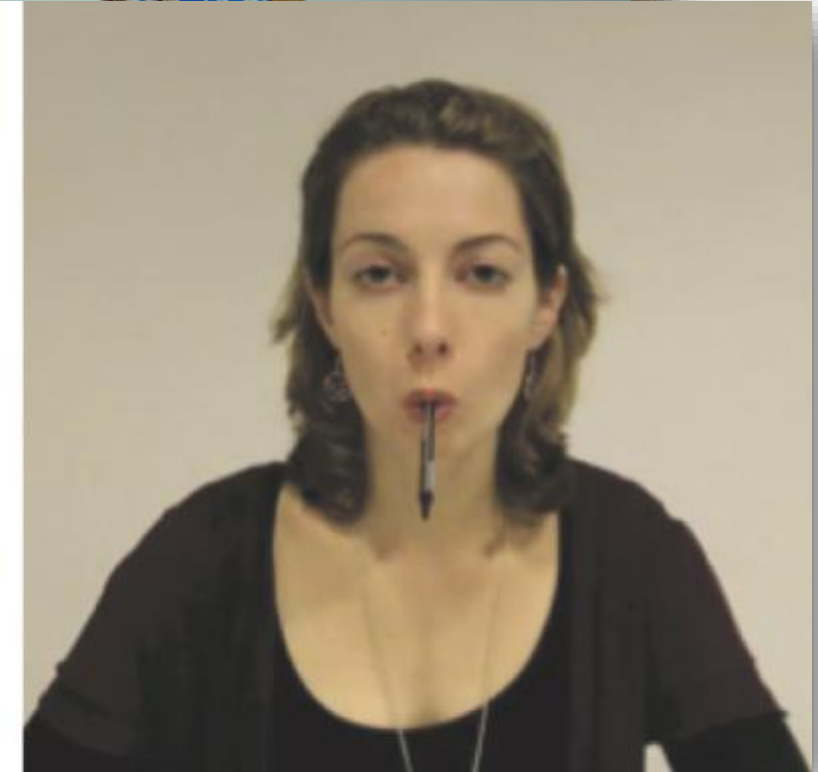
- elderly priming, MacB

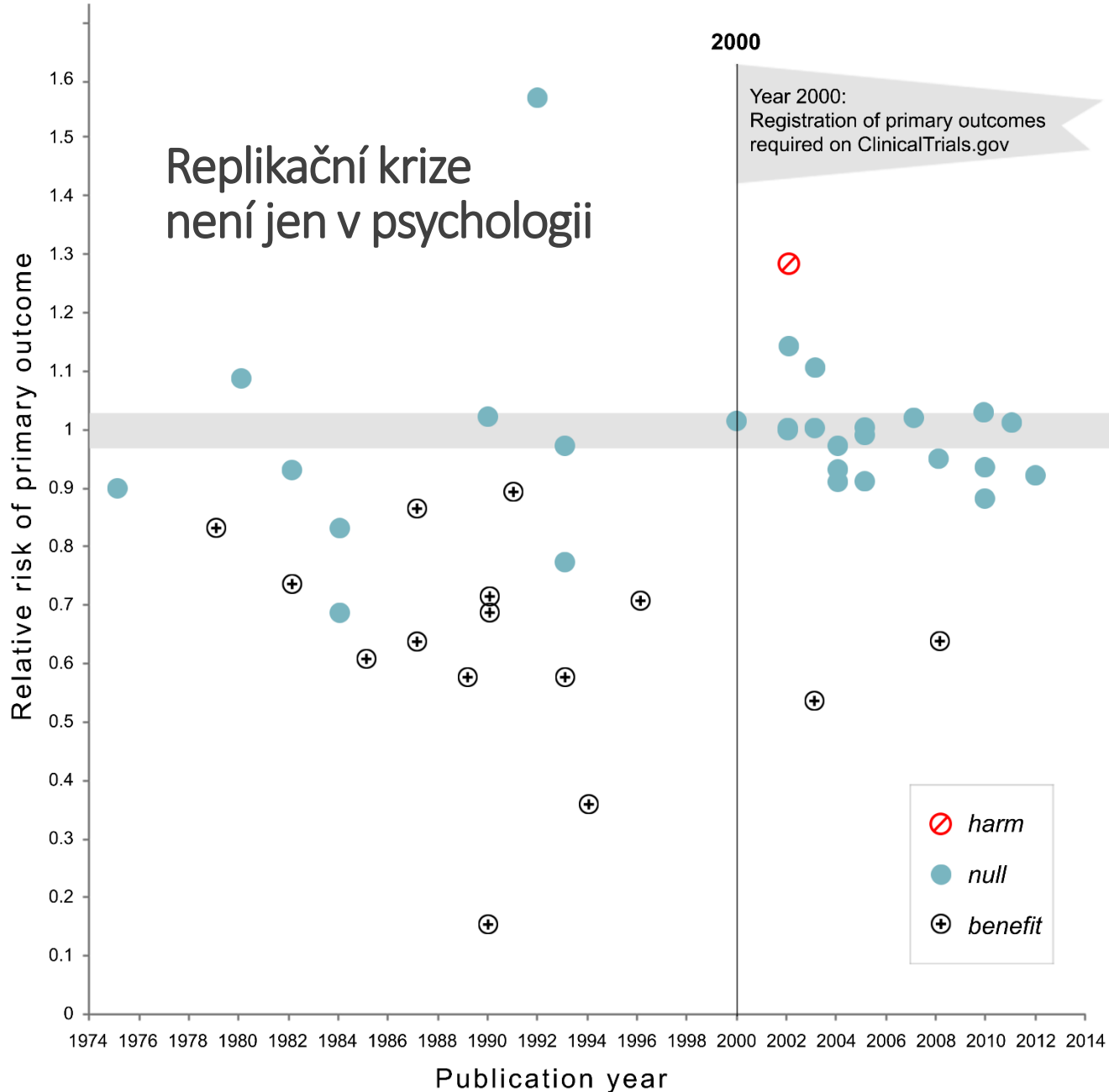
Ego depletion (vyčer

Power posing.

Vybrané aspekty faci

- „Smiling will make you





„We identified all large NHLBI supported RCTs between 1970 and 2012 evaluating **drugs or dietary supplements for the treatment or prevention of cardiovascular disease**. Trials were **included if direct costs >\$500,000/year**, participants were adult humans, and the **primary outcome was cardiovascular risk, disease or death**. [...] The number NHLBI trials reporting positive results declined after the year 2000. Prospective declaration of outcomes in RCTs, and the adoption of transparent reporting standards, as required by *clinicaltrials.gov*, may have contributed to the trend toward null findings.“

Kaplan, R.M., Irvin, V.L. (2015). Likelihood of Null Effects of Large NHLBI Clinical Trials Has Increased over Time. *PLoS ONE* 10(8), e0132382.

<https://doi.org/10.1371/journal.pone.0132382>

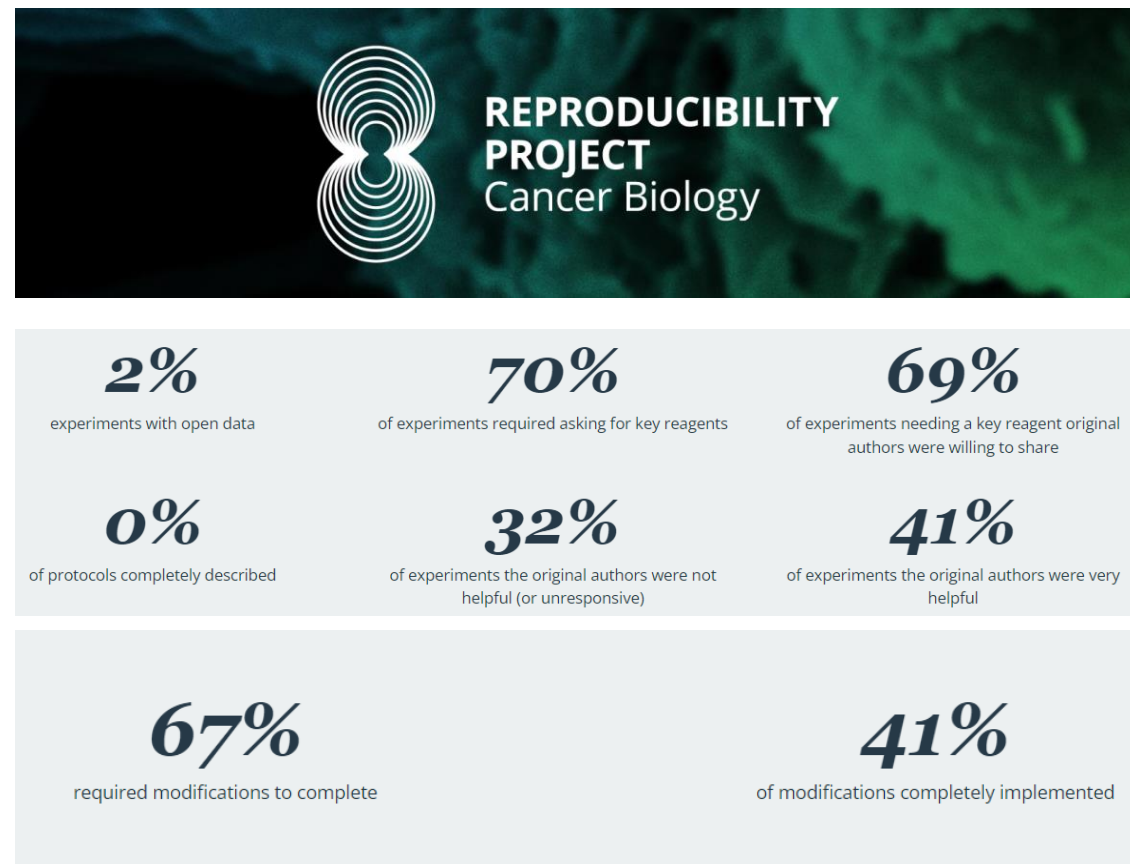
Replikační krize není jen v psychologii

Reproducibility project: Cancer Biology (2021)

- <https://www.cos.io/rpcb>
- 193 navržených replikací celkem
53 preklinických studií z let 2010–2012.

Výsledky:

- Realizace 50 replikací 23 článků
(nedostatek informací, nespolupráce).
- Jen 46 % efektů bylo replikovaných.
- Velikost efektu o 85 % nižší.



Replikační krize napříč obory

Nejčastěji zmiňované obory

- Psychologie (zejména sociální)
- Behaviorální ekonomie
- Experimentální politologie
- Medicína a biomedicínský výzkum
- Neurovědy
- Genetika (zejm. populační)

Psychologie: historicky důraz na kvalitní metodologii a statistické metody.

- Proto se replikační krize projevila zde a je zde nejvíce zkoumaná.
- Pokud jinde není tolik vidět, třeba se zatím jen „schovává“.

Potenciální příčiny replikační krize

Zkoumané efekty jsou místně, časově, kulturně závislé.

Vědecký podvod.

Replikační studie nejsou kvalitní („neschopní replikátoři“).

Původní efekty skutečně neexistují, ale původní vědci „nepodváděli“.

Nedostatky vědecké metody: Statistika, metodologie, publikační proces.

- Publikační zkreslení.
- Pochybné praktiky ve výzkumu (QRP).

Many Labs: Výzkum replikovatelnosti

Replikovatelnost efektu je důsledkem efektu, nikoli intervenujících proměnných.

- **Many Labs 1 (2014):** Jazyk či konkrétní laboratoř nemá vliv.
- **Many Labs 2 (2018):** Charakteristiky laboratoře nemají vliv, heterogenita efektů se však různí.
- **Many Labs 3 (2016):** Výsledky na studentských populacích vycházejí stejně v průběhu roku (např. semestr vs. zkouškové).
- **Many Labs 4 (2022):** Účast původního autora nemá vliv.
- **Many Labs 5 (2020):** Úpravy výzkumného protokolu nemají vliv.

„After 10 Years, ‘Many Labs’ Comes to an End – But Its Success Is Replicable“

- <https://news.virginia.edu/content/after-10-years-many-labs-comes-end-its-success-replicable>

Publikační zkreslení

Každý vědec přirozeně chce na „něco přijít“, snaha „nalézt“ výsledek.

Vědci jsou posuzováni podle citačního „ohlasu“.

- Publikační tlak, „publish or perish“.

Studie se signifikantními výsledky citovány 1,6krát častěji ([Duyx et al., 2017](#)).

- Pokud autoři explicitně napíší, že našli podporu pro své hypotézy, tak dokonce 2,7krát.

Kvalita časopisu je posuzována podle citovanosti jeho studií.

Editoři proto preferují články, u kterých je vyšší pravděpodobnost citování.

To vše dohromady → **publikační zkreslení** (file-drawer effect; Rosenthal, 1972).

Důsledky publikačního zkreslení

Nejde jen o to, že nejsou publikované studie dokládající „neexistenci“ efektu.

Statistické testování hypotéz:

- Výběrová chyba: pozorované efekty se náhodně pohybují okolo „skutečného“ efektu.
- Pravděpodobnost falešně pozitivních výsledků: typicky $\alpha = 0,05$.

Nadhodnocení evidence pro existenci efektu.

- Včetně neexistujících efektů.

Nadhodnocení velikosti účinku.

- A falešně negativní souvislost velikosti vzorku a velikosti účinku.

Pochybné praktiky ve výzkumu

<u>Questionable Research Practices</u>	<u>OK</u>
1. Not reporting “failed” studies.	83%
2. Not reporting DVs if not significant	92%
3. Not reporting Conditions that “did not work”	89%
4. Excluding data based on effect on p-value.	81%
5. Stopping data collection when significant.	89%
6. Reporting unexpected results “as predicted”	75%

Table 1. Likelihood of Obtaining a False-Positive Result

Researcher degrees of freedom	Significance level		
	$p < .1$	$p < .05$	$p < .01$
Situation A: two dependent variables ($r = .50$)	17.8%	9.5%	2.2%
Situation B: addition of 10 more observations per cell	14.5%	7.7%	1.6%
Situation C: controlling for gender or interaction of gender with treatment	21.6%	11.7%	2.7%
Situation D: dropping (or not dropping) one of three conditions	23.2%	12.6%	2.8%
Combine Situations A and B	26.0%	14.4%	3.3%
Combine Situations A, B, and C	50.9%	30.9%	8.4%
Combine Situations A, B, C, and D	81.5%	60.7%	21.5%

(John, Loewenstein, & Prelec, [2012](#))

(Simmons, Nelson, & Simonsohn, [2011](#))

Pochybné praktiky ve výzkumu

P-hacking: Různé postupy pro dosažení statisticky významného efektu.

- Vyzkoušení více statistických analýz, zařazování a vyřazování podmínek...

HARKing: Hypothesizing After the Results are Known

Phishing: Selektivní reportování dat a proměnných.

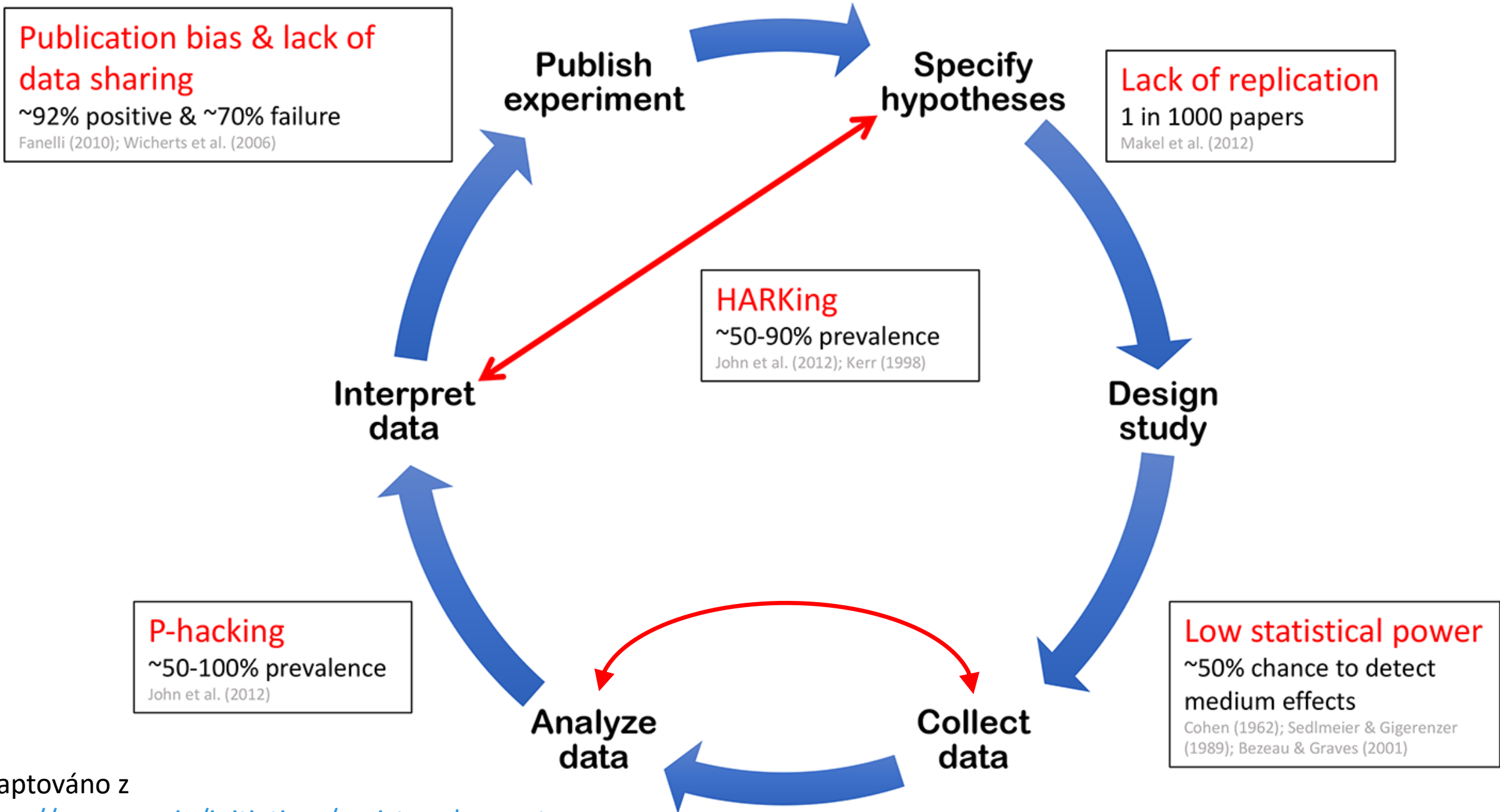
Zamlčení dalších podmínek.

Dosbírávání dat podle průběžných výsledků.

Ex-post manipulace s designem experimentu.

Manipulace s odlehlými hodnotami a významnými případy.

Vyzkoušejte si sami: <https://shinyapps.org/apps/p-hacker/>



Adaptováno z

<https://www.cos.io/initiatives/registered-reports>

Preregistrace: Proč?

Výzkumník si „dobrovolně sváže ruce“ před zahájením výzkumu.

Snaha zabránit, aby byla explorace a náhodná zjištění vydávána jako konfirmační test hypotéz.

Snaha předcházet QRP.

Uložena ideálně na důvěryhodném repozitáři.

Typicky kvantitativní a hlavně experimentální studie.

- Ale i korelační design, dokonce i kvalitativní výzkum.

Preregistrace: Co?

Hypotézy.

Design studie.

Způsob tvorby dat a velikost vzorku.

- Ideálně včetně zdůvodnění velikosti (power analýza).

Všechny proměnné, měřicí nástroje.

Postupy statistické analýzy.

- Úpravy dat, tvorba odvozených proměnných, testování hypotéz.

Registrovaný report

Specifický formát výzkumného článku.

První peer-review ještě před sběrem dat:

- Úvod, hypotézy, způsob sběru dat a analýz, případně pilotní data.

Někdy ještě krok 0: Posouzení samotného výzkumného záměru editorem.



Registrovaný report

Snaha zabránit publikačnímu zkreslení.

- Publikace přijata neohledně na výsledky.

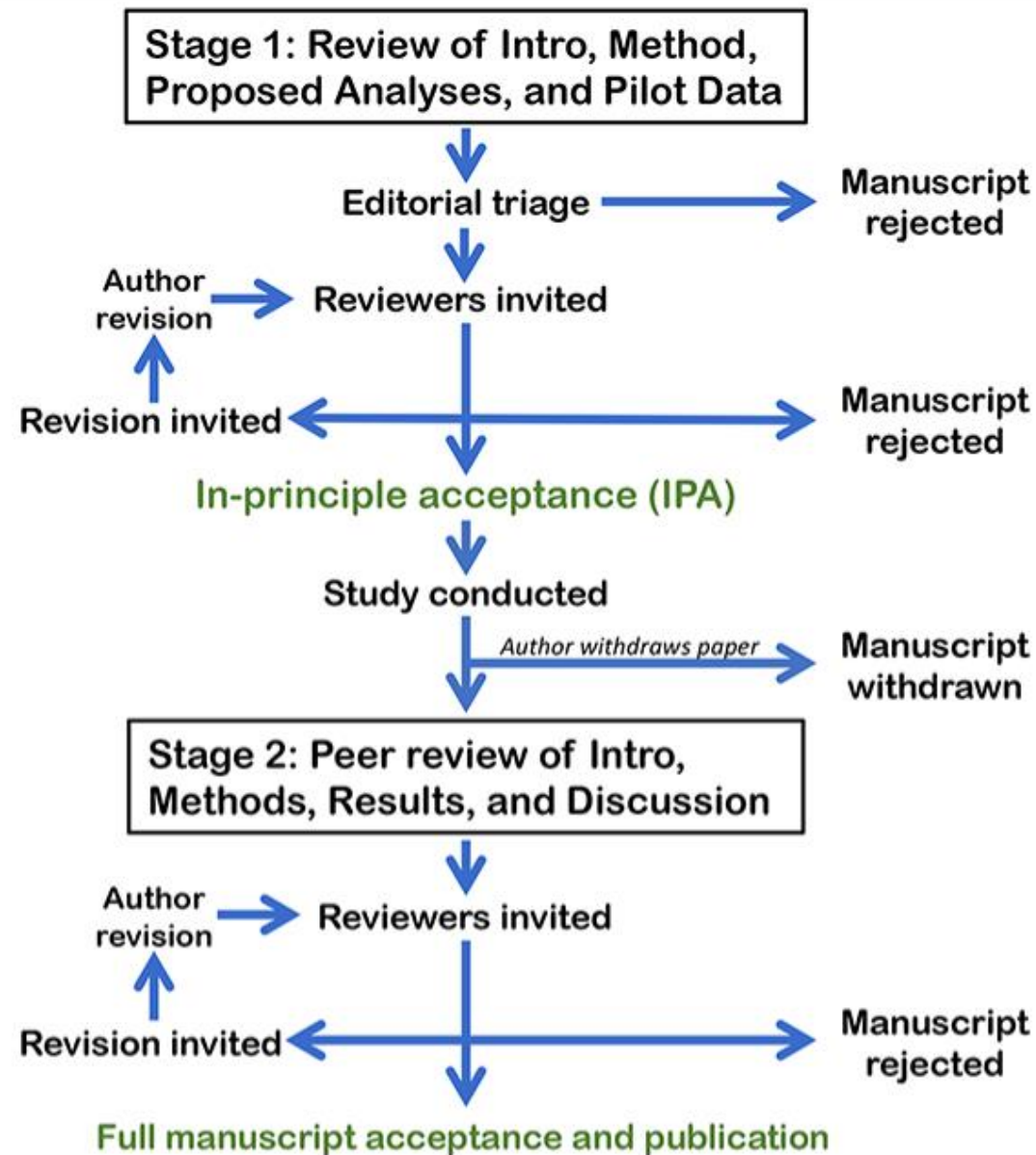
Možnost opravit design studie ještě před sběrem dat.

Úspora času autorů i recenzentů.

Více než 300 časopisů.

Přehledné informace:

<https://www.cos.io/initiatives/registered-reports>



<https://www.cos.io/initiatives/registered-reports>

Případová studie: Dobronauti: Zpátky v čase

Zakázkový výzkum INPSY FSS MUNI pro Dobronauty,
s.r.o.

Ověření preventivního programu pro zlepšování
komunikace (a prevence šikany).

Rizika:

- Naše motivace „vyjít vstříc“.
- Důvěryhodnost celého ověření.
- Komplikovaný analytický design.

Řešení: Preregistrace celého výzkumu.



www.dobronauti.cz

Případová studie: Dobronauti: Zpátky v čase

Program nemá vliv na self-esteem dětí, ale má vliv na jejich znalosti, jak se v krizové situaci zachovat.

Výzkumná studie:

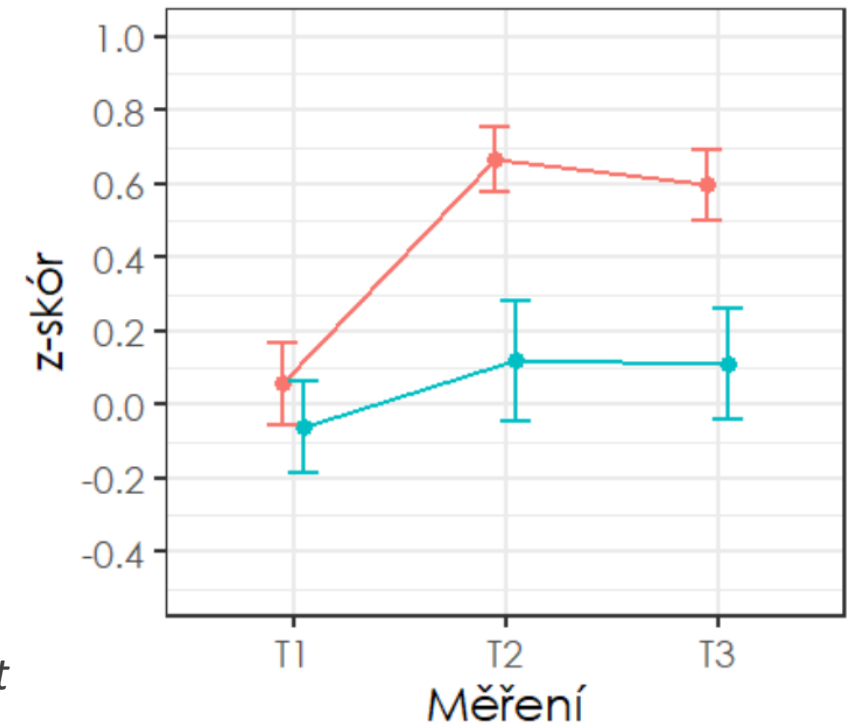
- Cígler, H., Fikrlová, J., & Tancoš, M. (2022). *Ověřovací studie preventivního programu Dobronauti: Zpátky v čase* [výzkumná zpráva]. Masarykova univerzita. <https://is.muni.cz/go/dobronauti>

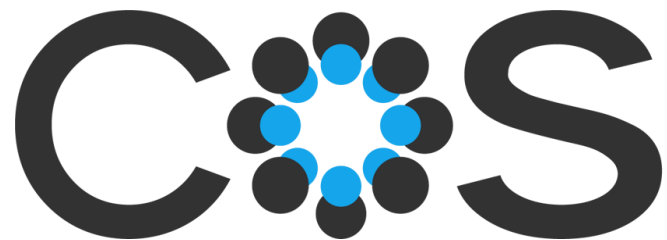
[Popularizační studie.](#)

Supplement: <https://doi.org/10.17605/OSF.IO/G2Z3N>

Preregistrace:

- Cígler, H., Tancoš, M., & Fikrlová, J. (2021, November 4). *Assessment of the effectiveness of the Dobronauti prevention program: Back in time.* <https://doi.org/10.17605/OSF.IO/9HZYC>





— CENTER FOR —
OPEN SCIENCE
SCIENCE WORKS BEST IN THE OPEN

www.osf.io

Pojďme si to vyzkoušet!

Děkuji za pozornost!

Hynek Cígler

INPSY a Katedra psychologie

Fakulta sociálních studií, Masarykova univerzita

Kontakty:

- E-mail: cigler@fss.muni.cz
- ORCID: <http://orcid.org/0000-0001-9959-6227>
- GitHub: <https://github.com/hynekcigler>
- OSF: <https://osf.io/t6ufg/>

Další materiály a prezentace
[dostupné online](#) (QR kód).

