

Replikovatelnost českého psychologického výzkumu

Psychologické dny Olomouc 2016

8. 9. 2016



Českomoravská psychologická společnost



UNIVERZITA
PALACKÉHO
V OLOMOUCI

Kontakt:

vit.gabrhel@mail.muni.cz

Autorský tým

Hynek Cígler,
Karolína Bernatová,
Hana Blažková,
Kitti Farkasová,
Světlana Farkašová,
Vít Gabrhel,
Jaroslav Gottfried,

Edita Chvojková,
Anna Janů,
Stanislav Ježek,
Lenka Knapová,
Tomáš Kratochvíl,
Vojtěch Mýlek,
Ondřej Novák,

Petr Palíšek,
Kateřina Prachařová,
Karel Rečka,
Anna Marie Rosická,
Tomáš Rudecký,
Klára Šmejkalová,
David Vašek



MASARYKOVA UNIVERZITA
FAKULTA SOCIÁLNÍCH STUDIÍ



FILOZOFICKÁ FAKULTA
UNIVERZITY KARLOVY
V PRAZE

Východiska

Počet publikací českých docentů a profesorů psychologie jako hodnocení kvality jejich práce:

- *Bezdíček, Preiss a Dočkalová (2009)*
- *Preiss a Mačudová (2012)*

Počet studií však nic neříká o jejich kvalitě

- Jedním z ukazatelů kvality výzkumu je jeho **replikovatelnost**

Replikovatelnost v kontextu naší studie:

- Počet statistických testů, které by při replikaci vyšly jako statisticky signifikantní

Replikační krize?

SCIENCE
The New York Times
Many Psychology Findings Not as Strong as Claimed, Study Says

PSYCHOLOGY IS IN CRISIS OVER
WHETHER IT'S IN CRISIS WIRED

Speaking of Science
Errors riddled 2015 study showing
replication crisis in psychology research,
scientists say

Psychology's replication crisis
sparks new debate ScienceNew
MAGAZINE OF THE SOCIETY FOR SCIENCE & THE P

The Washington Post
What psychology's crisis means for the future
of science Vox

Co znamená replikační krize pro psychologii?

Social Priming (dle Aronson, Wilson, & Akert, 2014)

Proces, při kterém nedávná zkušenost zvyšuje pravděpodobnost aktivace určitého *schématu* či *konceptu*. S větší pravděpodobností bude aktivováno *schéma*, které se vyskytlo nedávno, nebo již v minulosti bylo použito. Jde o příklad automatického uvažování a dochází k němu neúmyslně.

- **Změnu chování** či **úsudku** mohou vyvolat i relativně subtilní vodítka jako *výška židle* (Sanna et al., 2011), *teplota hrnku s kávou* (Williams & Bargh, 2008) nebo *barva vytištěného slova* (Bertrams et al., 2014).

Opravdu tomu tak je?

- Očekávaný efekt nebyl replikován (Klein et al., 2014)

Implikace - terapie, manželské a rodinné poradenství:

- *Pocit závazku* ve vztahu **(NE)vede** k *odpouštění* (Cheung et al., accepted)

Open Science, 2015

"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*

Původní

Průměrná velikost účinku:

$M_r = 0,403; SD = 0,188$

Statistická signifikance:

97 % studií $p < 0,05$

Replikace

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.

Výzkumná otázka

Jaká je replikovatelnost českého psychologického výzkumu?

Design

Vyhledání autorů/studií s afiliací k ČR (Československo...)

- Provázání publikací s autory, aby bylo možné zjistit citační parametry autorů (IF) jednotlivých studií (snaha o provázání co nejvíce studií).

Identifikace klíčových hypotéz a zápis relevantních údajů

- Např. $t(45.7) = 4.23, p = .001; d = .3$

Analýza vybraných indikátorů

- *R-index, statistická síla, velikosti účinků* atd.

Vzorek

Kritéria pro zařazení

- „Czech Republic“, „Czechoslovakia“, „article“, „psychology“ (zahrnuje i různé podoblasti).

Klíčová slova

- *social, applied, clinical, biological, educational, experimental, mathematical, developmental, psychoanalysis a multidisciplinary psychology*

1301 studií s celkem 3469 autory

- Roky: **1996-2016**
- Před změnou ve WoS (1996) bylo dohledatelných méně než **50 % autorů**
- Alespoň jeden autor byl dohledán u **986 studií (76 %)**

Metody

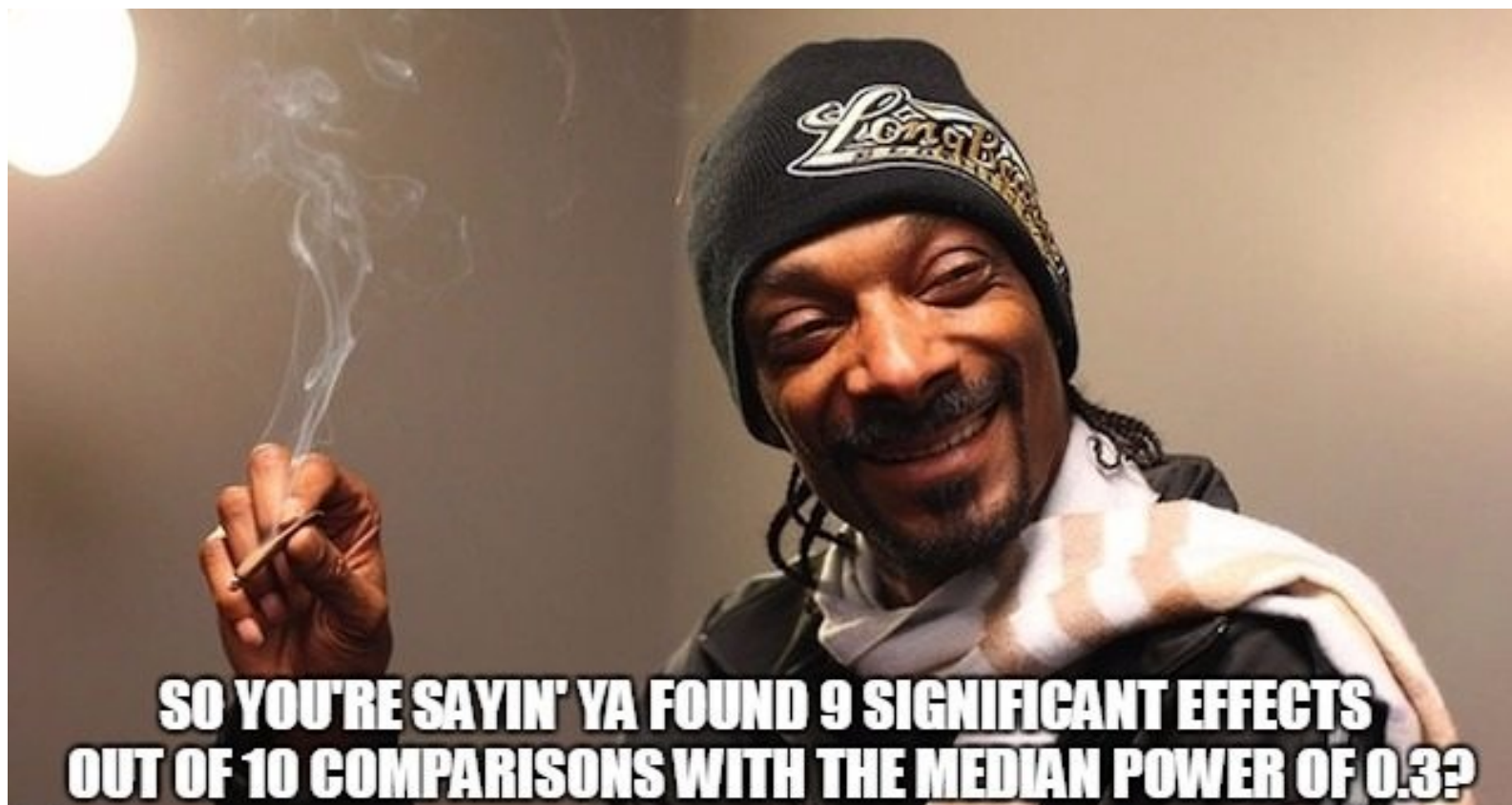
Replikovatelnost

Statistická síla

- Pravděpodobnost, že kdybychom výzkum **replikovali** stejně, **identifikovali** bychom **signifikantní výsledek**.
 - Pakliže by měl být hledaný efekt stejně velký jako ten v původní studii.
- Studie se statistickou silou **50 %** by v případě 100 provedených analýz měla vést k **50 signifikantním** a **50 nesignifikantním** výsledkům
- **Publikační zkreslení** - publikovány jsou v naprosté většině případů signifikantní výsledky

Replicability index (R-index) (Schimmack, 2014)

- **Ukazatel** (nikoli přímý odhad) přesnější než síla testu, protože **kompensuje publikační zkreslení**
- Kombinuje informace o statistické síle v původní studii spolu s inflací statistické síly vzhledem k publikačnímu zkreslení (tzv. publication bias).
 - *Inflation = Percentage of Significant Results – Median (Estimated Power)*
 - **R-Index = Median Observed Power – Inflation**



**SO YOU'RE SAYIN' YA FOUND 9 SIGNIFICANT EFFECTS
OUT OF 10 COMPARISONS WITH THE MEDIAN POWER OF 0.3?**

Metody

Scientometrie

Impact Factor (IF)

"Průměrný" počet citací "průměrného" článku v daném žurnálu ve vybraný rok.
Takto definovaný IF je hodnocením dopadu **žurnálu jako celku**, ne však jednotlivých článků.

- **Author Impact Factor (AIF)** = *"averages the number of citations received by all papers published by an author in a given time window"* (Pan & Fortunato, 2014, p. 2)

Hirschův index (h-index) (Hirsch, 2005)

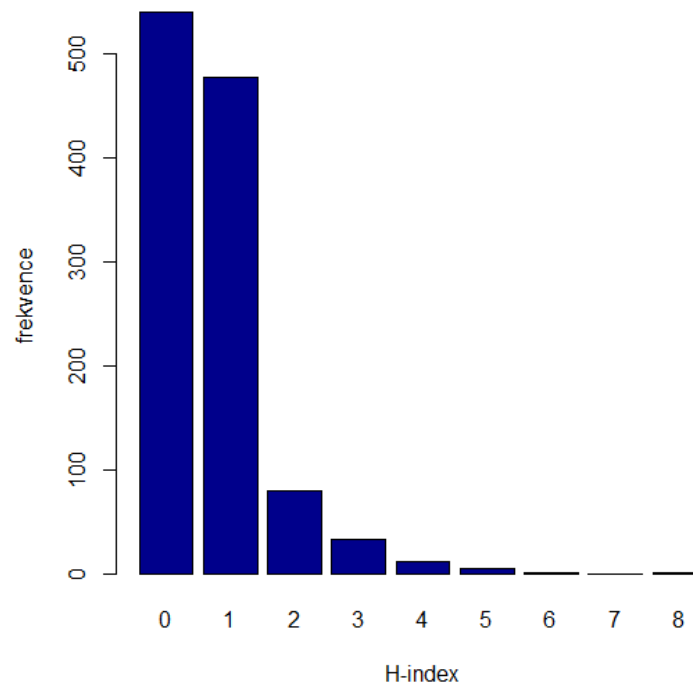
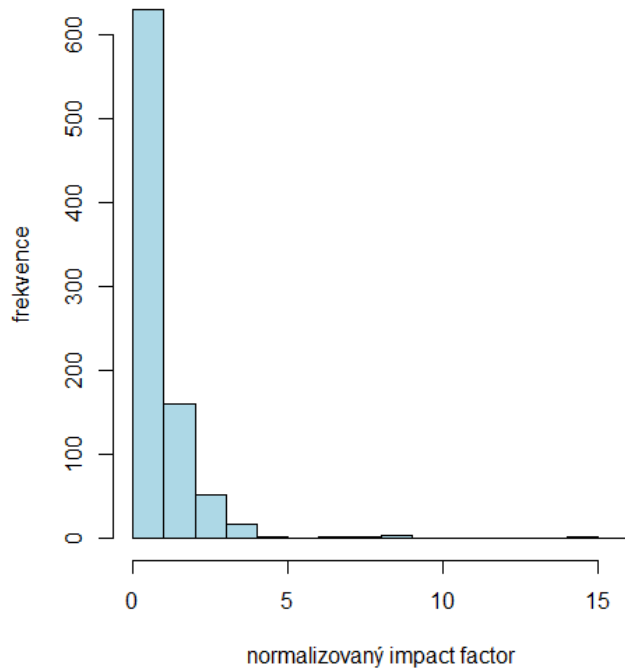
Nástroj pro současné určení *produktivity* a *citačního dopadu* daného autora
Kolik článků daného autora dosahuje citovanosti vyšší, než je pořadové číslo článku dle počtu citací.

$$f(A)=10, f(B)=8, f(C)=5, f(D)=4, f(E)=3 \rightarrow h\text{-index}=4$$

$$f(A)=25, f(B)=8, f(C)=5, f(D)=3, f(E)=3 \rightarrow h\text{-index}=3$$

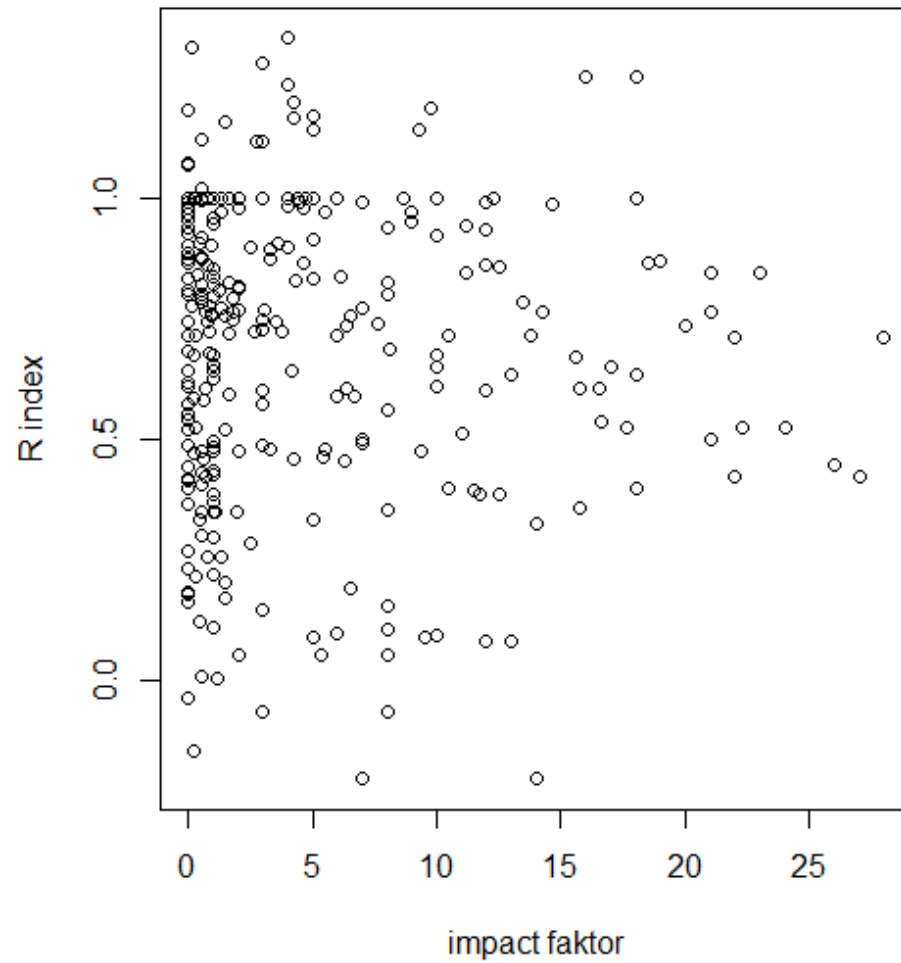
Výsledky

Scientometrie



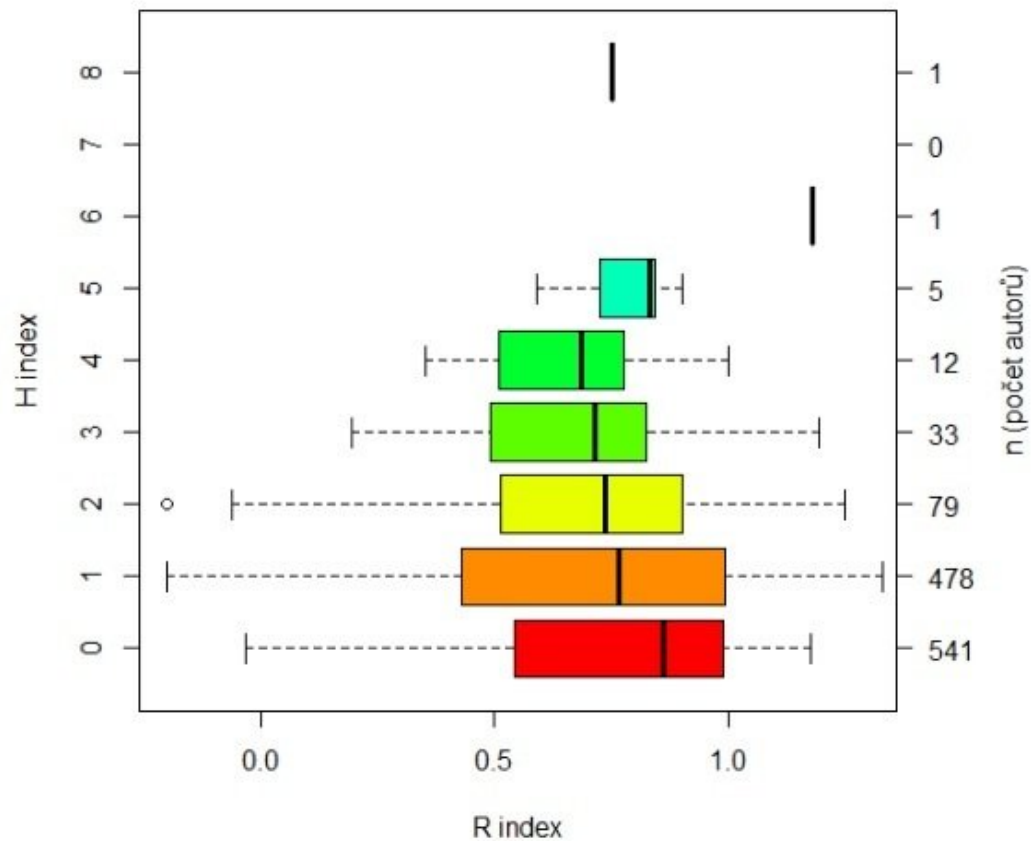
Výsledky

Souvislost R-indexu a IF



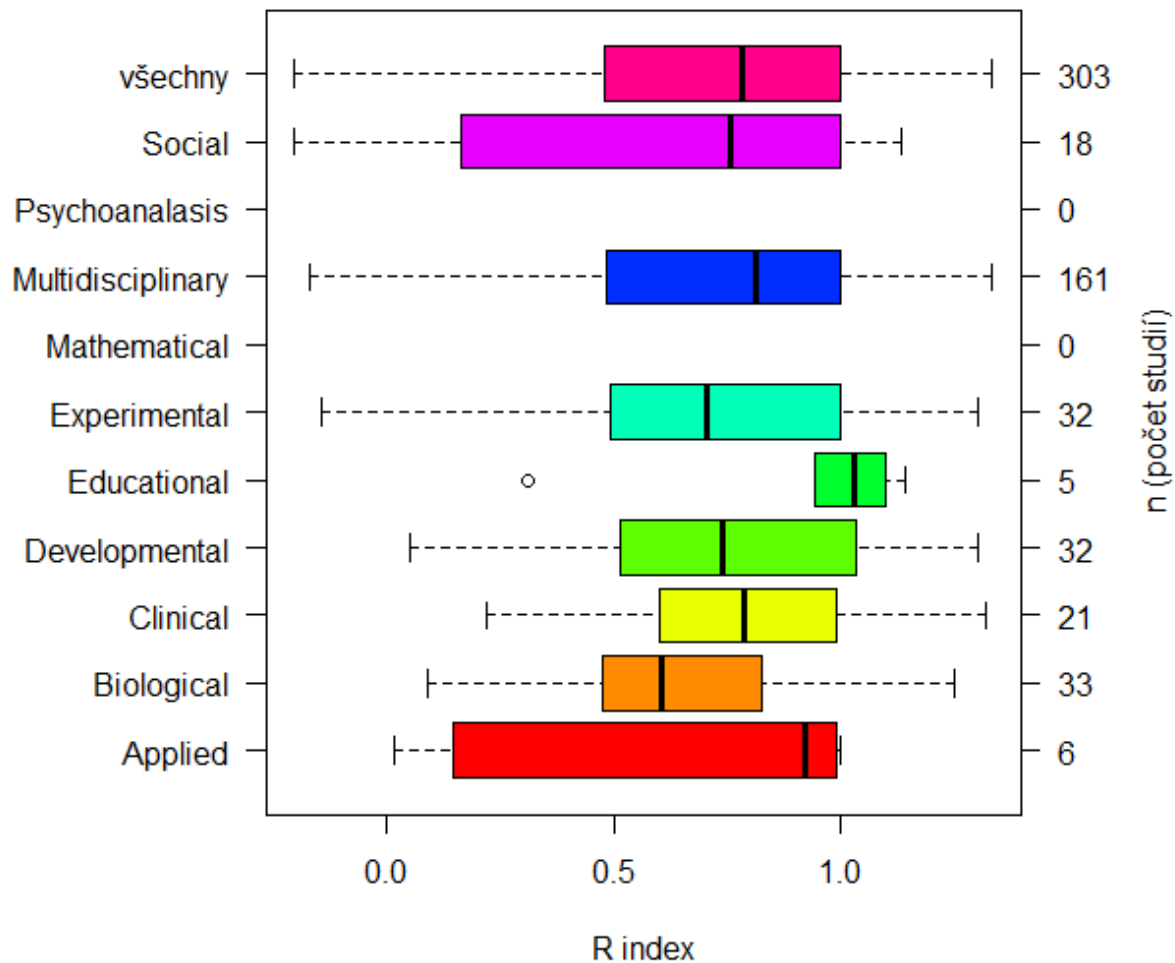
Výsledky

Souvislost R-indexu a h-indexu



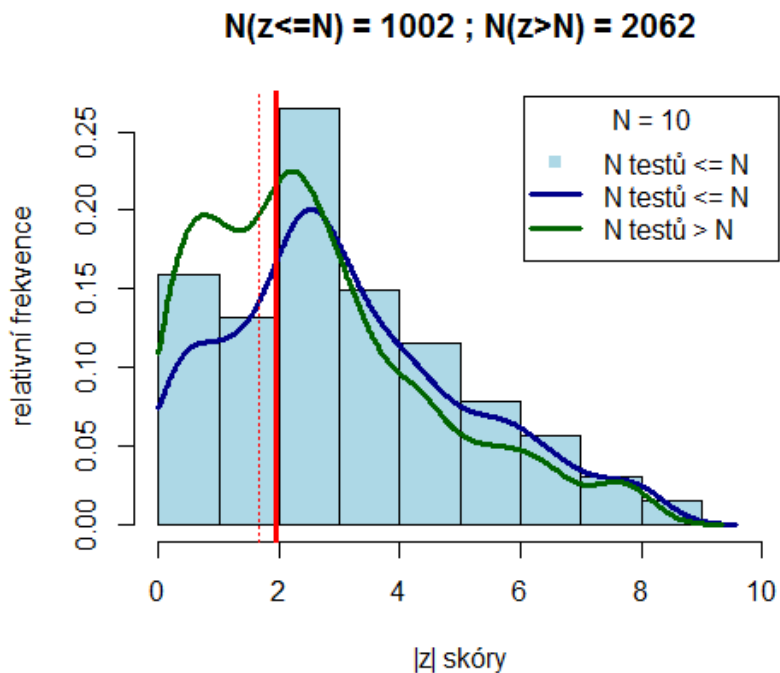
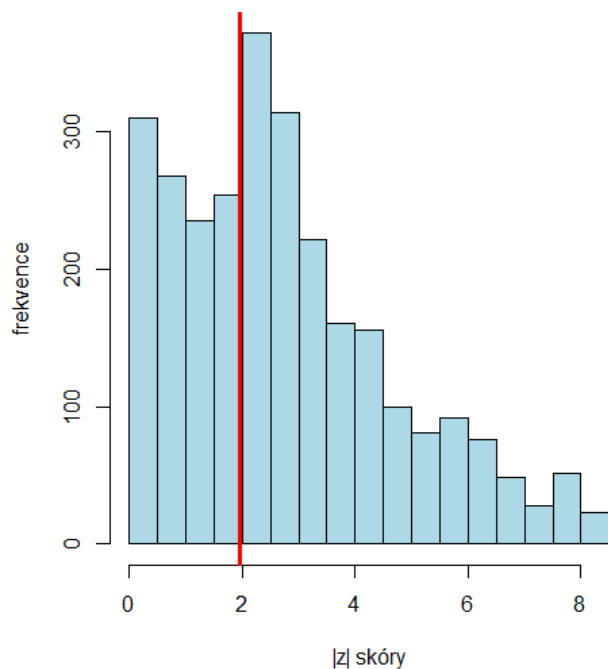
Výsledky

R-index dle oblasti psychologie



Výsledky

Provedené testy s příslušným z-skórem



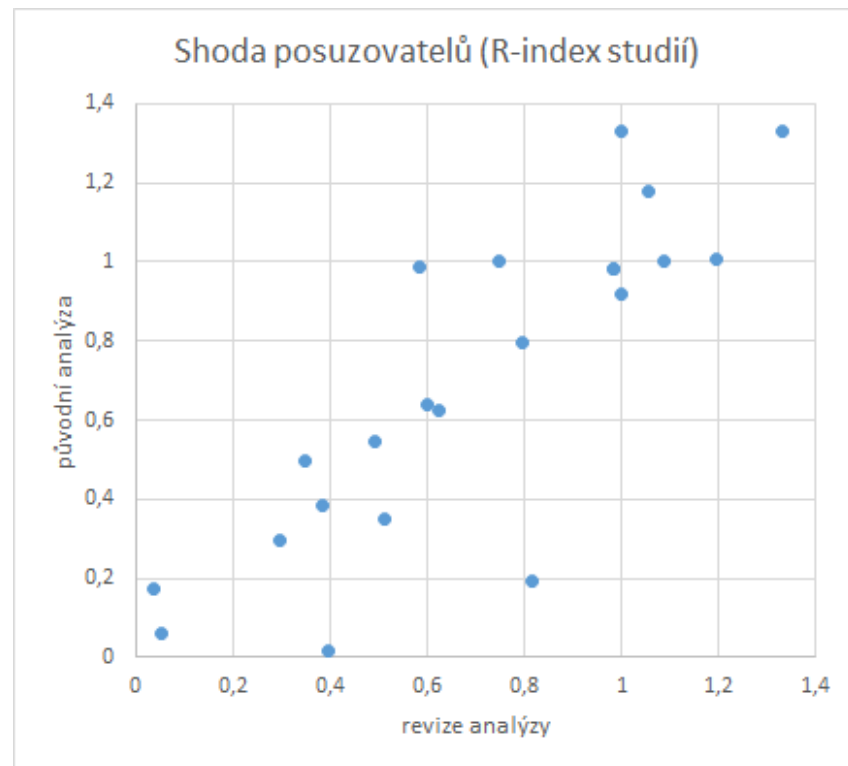
Interní validita: Shoda posuzovatelů

Shoda posuzovatelů byla ověřena na 27 náhodně vylosovaných zařazených studií, které ohodnotil další hodnotitel podruhé.

Z těchto 27 druhý hodnotitel 6 vyřadil (4 netestovaly hypotézy, 1 byla "zmatená", 1 "jiné důvody").

R-index zbylých 21 studií **Spearmanovo rho = 0,832**, 95% CI [0,60; 0,92] (bootstrap)

Nelišil se ani průměrný R-index u posuzovaných studií (**Wilcoxonův test p = 0,796**).



Diskuze

Výsledky

Podstatná část českého psychologického výzkumu není citovaná vůbec nebo jen minimálně.

Celkový R-index = 0,7. Replikovali bychom tedy 70 % studií.

U českých psychologů citační indexy (AIF, h-index) nesouvisí s R-indexem.

Výzkumníci nedodržují standardy při prezentaci statistik.

Diskuze

Reflexe sběru dat

Klíčové nedostatky

- Chybějící informace o...
 - Přesné hodnotě "p", zejména při non-signifikantním výsledku (kupř. "n.s."),
 - **Velikost účinku** (d, r, atd.),
 - **Výsledek testu** (F = ...),
 - **Počet případů** (N) či **stupně volnosti** (df)
- Nesnadná **identifikace klíčových hypotéz**
- I přes deklarování **explorační** povahy bylo prováděno **testování signifikance**, doprovázené **reportem** vybraných (**signifikantních**) výsledků
- **Absence korekce** při vícenásobném srovnání na týchž datech (kupř. **Bonferroniho metoda**)
- **Deskriptivní statistiky** chyběly či obsahovaly jen část relevantních údajů

Řešení?

Dodržování APA formátu

Take home message

**Replikovatelnost výzkumu
jednotlivých autorů se liší, ale ne v
závislosti na počtu jejich publikací,
respektive na dopadu těchto
publikací.**

Děkuji za pozornost

Zdroje

- American Psychological Association. (2001). *Publication manual of the American Psychological Association* (6th ed.). Washington, DC: Author.
- Aronson, E., Wilson, T. D., & Akert, R. M. (2014). *Social psychology*. Harlow: Pearson.
- Bertrams, A., Baumeister, R. F., Englert, C., & Furley, P. (2015). Ego depletion in color priming research: self-control strength moderates the detrimental effect of red on cognitive test performance. *Personality & Social Psychology Bulletin*, 41(3), 311–22. doi:10.1177/0146167214564968
- Bezdíček, O., Preiss, M., & Dočkalová, E. (2009). D Isregulation and R Esponse To O Piate B Lockers. *Psychiatrie*, 13(2-3), 108–113.
- Hirsch, J. E. (2005). An index to quantify an individual's scientific research output. *Proc Natl Acad Sci U S A*, 102(46), 16569–16572. doi:10.1073/pnas.0507655102
- Cheung, I., Campbell, L., LeBel, E.,...Yong, J. C. (accepted). Registered Replication Report: Study 1 from Finkel, Rusbult, Kumashiro, & Hannon (2002). *Perspectives on Psychological Science*
- Klein, R. A., Ratliff, K. A., Vianello, M., Adams Jr, R. B., Bahník, Š., Bernstein, M. J., ... & Nosek, B. A. (2014). Data from Investigating Variation in Replicability: A “Many Labs” Replication Project. *Journal of Open Psychology Data*, 2(1), e4.
- Morgan, S. E., Reichert, T., Harrison, T. R. (2002). *From numbers to words. Reporting statistical results for the social sciences*. Allyn & Bacon.
- Open Science (2015). Estimating the reproducibility of psychological science. *349* (6251), <http://doi.org/10.1126/science.aac4716>
- Pan, R. K., & Fortunato, S. (2014). Author Impact Factor: tracking the dynamics of individual scientific impact. *Scientific Reports*, 4, 4880. doi:10.1038/srep04880
- Preiss, M., & Mačudová, G. (2012). Publikační Aktivita a Citovanost Českých Docentů a Profesorů V Oboru Psychologie. *E-Psychologie*, 6(4).
- Sanna, L. J., Chang, E. C., Miceli, P. M., & Lundberg, K. B. (2011). Rising up to higher virtues: Experiencing elevated physical height uplifts prosocial actions. *Journal of Experimental Social Psychology*, 47(2), 472–476. doi:10.1016/j.jesp.2010.12.013
- Schimmack, U. (2014). Quantifying Statistical Research Integrity: The Replicability Index, 1–31. Retrieved from http://r-index.org/uploads/3/5/6/7/3567479/introduction_to_the_r-index__14-12-01.pdf
- Williams, L. E., & Bargh, J. A. (2008). *Science*. October, 322(5901), 606–607. doi:10.1126/science.1162548.Experiencing

Otázky?

