



<https://doi.org/10.5559/di.34.2.04>

UTJECAJ ZNAČAJNOSTI REZULTATA NA PROCJENU TOČNOSTI REPLIKACIJE: EKSPERIMENTALNO ISTRAŽIVANJE

Ivan BULJAN
Filozofski fakultet, Sveučilište u Splitu, Split, Hrvatska

UDK: 001.891.5:159.9

Izvorni znanstveni rad

Primljeno: 17. lipnja 2024.

Cilj istraživanja bio je ispitati utjecaj redoslijeda prikaza istraživanja koja daju suprotne zaključke na procjenu točnosti istraživanja. U ovom registriranom paralelnom eksperimentalnom istraživanju sudionicima iz opće populacije nasumično su dodjeljivani opisi istraživanja u kojima su replikacije psiholoških istraživanja pronašle rezultate koji pokazuju značajan ili neznačajan učinak intervencije, a zatim su zamoljeni da procijene točnost zaključaka opisanih istraživanja. Istraživanje je provedeno u virtualnom okruženju. Ukupno je analizirano 194 sudionika, 97 u svakoj eksperimentalnoj skupini. Intervenciju su u ovom istraživanju predstavljali različiti zaključci istraživanja koje je program za testiranje nasumično dodjeljivao sudionicima. Rezultati upućuju na to da je veći postotak sudionika kojima je prikazana replikacija sa značajnim učinkom procjenjivao replikaciju točnijom u odnosu na početno istraživanje ($n = 71$, 73,20 %) u odnosu na postotak sudionika kojima je prikazana replikacija koja nije pokazala statistički značajan učinak ($n = 53$, 54,64 %). Rezultati upućuju na potencijalnu pristranost redoslijeda prilikom interpretacije replikacijskih istraživanja, jer je procjena točnosti ovisila o nalazima istraživanja. Buduća kvalitativna istraživanja trebala bi istražiti čimbenike povezane s interpretacijom replikacija.

Ključne riječi: ponovljivost znanstvenih istraživanja, replikacijska kriza, razumijevanje znanstvenih istraživanja, eksperimentalno istraživanje



Ivan Buljan, Odsjek za psihologiju, Filozofski fakultet,
Sveučilište u Splitu, Poljička cesta 35, 21000 Split, Hrvatska.
E-mail: ibuljan@ffst.hr

Ponovljivost znanstvenih rezultata jedan je od temeljnih principa znanstvenoga pristupa (Alston i Rick, 2021). Metode, analize i znanstveni nalazi trebali bi biti objektivno provjerljivi od drugih istraživača, koji bi, ako je istraživanje dobro provedeno, trebali dobiti iste ili slične rezultate (Taylor i Kuyatt, 1994). Ipak, u zadnjem se desetljeću pojavio koncept krize ponovljivosti u znanstvenim istraživanjima kao reakcija na neponovljivost velikoga dijela znanstvenih rezultata u pojedinim disciplinama. Pokret za poboljšanja ponovljivosti izazvao je val novih istraživanja, koja imaju za cilj povećati vjerodostojnost dobivenih rezultata (npr. Baker, 2016; Ioannidis, 2005; Open Science Collaboration, 2015). Novije su analize metodom strojnog učenja na više od 14 tisuća znanstvenih radova iz psihologije iznijele podatak da je tek oko 40 % istraživanja u psihologiji ponovljivo, što može predstavljati problem u interpretaciji psihologijskih fenomena i njihovoj primjeni u praksi (Youyou i sur., 2023). Neki od razloga za nisku ponovljivost u psihologijskim istraživanjima jesu mali uzorci, nedostupnost istraživačkih podataka te mijenjanje početne hipoteze nakon što se analiziraju rezultati (Munafò i sur., 2017). Iz krize su nastale jako pozitivne međunarodne inicijative, poput *Psychological Science Accelerator*, koja pokušava ponoviti prethodno provedena istraživanja na različitim populacijama i u različitim kulturama kako bi provjerila stabilnost izvorno pronađenih učinaka (Psychological Science Accelerator, 2023).

S druge strane, pojedinci su kritizirali pristup istraživanju ponovljivosti u kojem se statistička značajnost i mjere učinka izvornog istraživanja postavljaju kao glavni kriterij u interpretaciji nalaza. Stoga možemo reći da sam kriterij uspješno ponovljenog istraživanja još uvijek nije ustanovljen, s obzirom na to da odluka o tome je li istraživanje ponovljivo ovisi o interpretaciji znanstvenika i/ili čitatelja, koja se pritom može zasnivati na različitim kriterijima (statistička značajnost, veličina učinka, Bayes faktor) (Held i sur., 2022). Primjerice, nedavno istraživanje na uzorku novinara pronašlo je da veličina uzorka utječe na procjenu vjerodostojnosti istraživanja, za razliku od p vrijednosti, reprezentativnosti uzorka i institucije autora (Bottesini i sur., 2023). Navedeni i drugi čimbenici utječu na procjenu vjerodostojnosti istraživanja, pa sukladno s time donose pristranosti koje utječu na interpretaciju. Dodatan problem u današnjem vremenu jest izloženost čitatelja velikoj količini informacija koja im otežava njihovo kritičko sagledavanje, pa se može očekivati da će pribjegavati nizu mentalnih kratica prilikom zaključivanja o vjerodostojnosti istraživanja. Jedna od najpoznatijih pristranosti u kontekstu znanosti jest pristranost u objavljivanju (engl. *publication bias*), prema kojoj će istraživanja s pozitivnim (značajnim) ishodima prije biti objavljena ne-

go ona s negativnim ishodima, što u konačnici dovodi do njihova nerazmjera (Fanelli, 2012). Istodobno, istraživanja su pokazala da su i čitatelji skloniji vjerovati statistički značajnim nego neznačajnim rezultatima (Jankowski i sur., 2022). U isto vrijeme postoji mogućnost da, osim istraživačkih kriterija, postoje i drugi vanjski faktori na temelju kojih čitatelji donose odluke (npr. Sherif i sur., 1958). Primjerice, postoji mogućnost da je interpretacija kasnijih istraživanja (replikacija) pod utjecajem informacija iz istraživanja koja su napravljena inicijalno, odnosno da sam redoslijed dobivanja informacija utječe na interpretaciju točnosti istraživanja. Iz toga razloga se pristupilo provođenju ovoga eksperimentalnog istraživanja, kako bi se odgovorilo na pitanje kako sudionici interpretiraju rezultate ponovljenih istraživanja kada oni nisu u skladu s rezultatima inicijalnih istraživanja. Uobičajen način ispitivanja uvjerenja o točnosti zaključaka istraživanja uključivao bi primjenu kontinuirane skale na kojoj bi se označavao stupanj uvjerenja (npr. Buljan i sur., 2020). Tomu unatoč, u ovom istraživanju odlučili smo sudionicima dati mogućnost binarnog odgovora u procjeni slaganja s mogućim ishodima istraživanja, kako da u stvarnom svijetu moraju odlučiti vjeruju li ili ne vjeruju zaključku istraživanja koje su pročitali, to jest moraju dati binaran odgovor. Primarni cilj ovog istraživanja bio je ispitati hoće li izmjena redoslijeda prikaza identičnih istraživanja koja daju suprotne zaključke utjecati na procjenu točnosti svakoga pojedinačnog istraživanja. Također, s obzirom na nedostatak istraživanja o svjesnosti o problemu ponovljivosti istraživanja u hrvatskoj populaciji, sudionicima je postavljen niz pitanja vezan za ponovljivost istraživanja i sintezu znanstvenih dokaza, kako bi se ispitala razina svjesnosti o navedenim konceptima te koliko je pritom uzorak prikupljen u ovom istraživanju uopće reprezentativan za opću populaciju.

Hipoteza 1: Veći postotak sudionika odabrat će inicijalno istraživanje kao točnije u usporedbi s replikacijom.

Hipoteza 2: Većina sudionika (više od 50 % uzorka) izjavit će da nije upoznata s pojmovima ponovljivosti istraživanja i sustavnih pregleda, s obzirom na to da je riječ o općoj populaciji.

METODE

Nacrt istraživanja

Provedeno je randomizirano eksperimentalno istraživanje s dvjema usporednim nezavisnim eksperimentalnim skupinama koje je bilo registrirano na javno dostupnoj platformi *Open Science Framework (OSF)*: <https://osf.io/b3et7>.

Uzorak i prikupljanje podataka

U istraživanju su mogle sudjelovati punoljetne osobe, a poziv za sudjelovanje bio je oglašavan na stranicama Hrvatske mreže za odgovornu i ponovljivu znanost (<https://crorin.hr/>), na društvenim mrežama i putem osobnih kontakata, od svibnja do kraja srpnja 2023. Istraživanje je provedeno putem internetske platforme za anketno testiranje *SurveyMonkey* (SurveyMonkey Inc., San Mateo, California, USA). Sukladno Zakonu o provedbi opće uredbe o zaštiti podataka (NN 42/2018), sudionicima su nakon otvaranja poveznice bile prikazane informacije o istraživanju, nakon kojih su, ako se slažu sa sudjelovanjem, mogli odabrati opciju da žele sudjelovati u istraživanju, pri čemu bi bili odvedeni na sljedeću stranicu. Program je bio ugođen tako da ne prikuplja IP adrese sudionika, kako bi se osigurala njihova potpuna anonimnost.

Materijali i intervencije

U svrhu istraživanja izmišljena su dva scenarija istraživanja koja su predstavljala intervenciju u ovom istraživanju. U oba slučaja tema istraživanja bila je učinkovitost psihoedukacije, terapijske tehnike koja se u opisanim scenarijima rabila u tretmanu perfekcionizma. U scenarijima je inicijalno napravljeno istraživanje koje je testiralo učinkovitost spomenute tehnike (broj sudionika u opisanom istraživanju $n = 250$), a zatim je napravljeno replikacijsko istraživanje s istim brojem sudionika ($n = 250$) koje je pokušalo ponoviti rezultate inicijalnog istraživanja. Rezultati obaju sljedećih (replikacijskih) istraživanja bili su u suprotnosti s onima inicijalnih istraživanja.

Opis skupina

Intervencijska skupina A: Rezultat prvog istraživanja pokazivao je da je psihoedukacija učinkovita u smanjenju simptoma perfekcionizma, drugo (replikacijsko) istraživanje pokazalo je da nije.

Intervencijska skupina B: Rezultat prvog istraživanja pokazivao je da psihoedukacija nije učinkovita u smanjenju simptoma perfekcionizma, drugo (replikacijsko) istraživanje pokazivalo je da jest.

Na ovaj način pokušala se stvoriti situacija u kojoj sudionici najprije pronalaze istraživanje s nalazom u jednom smjeru, a zatim ono sa suprotnim nalazom.

Oba scenarija dostupna su u dodatku A. Nije bilo vremenskog ograničenja za čitanje scenarija.

Randomizacija

Randomizacija se provela pomoću mrežnoga sustava za testiranje (SurveyMonkey), zatim A/B metodom za testiranje (<https://help.surveymonkey.com/en/surveymonkey/create/ab-tests/>). Sustav je dodjeljivao sadržaj prema slučajnom rasporedu u omje-

ru 50:50. Primjenom metode A/B testiranja u SurveyMonkey programu polovici sudionika bio je prikazan format A (intervencijska skupina A), a polovici format B (intervencijska skupina B).

Varijable

a) Primarni ishodi: zaključci o provedenim istraživanjima

Nakon čitanja scenarija, sudionici su zamoljeni da odgovore na pitanja vezana uz njihove zaključke o scenarijima i da procijene točnost provedenih istraživanja. Pritom su trebali izraziti svoje slaganje označivši "Da" ili "Ne" uz svaku od pojedinih tvrdnji:

- 1) Prva studija je točna.
- 2) Druga studija je točna.
- 3) Obje studije su točne.
- 4) Ni jedna studija nije točna.
- 5) Ne mogu donijeti zaključak.

S obzirom na to da su scenariji sadržavali vrlo malo informacija na temelju kojih se mogu donijeti odluke, za svaki odgovor trebalo je označiti i stupanj sigurnosti od 0 (nimalo siguran/a) do 100 (potpuno siguran/a).

b) Sekundarni ishodi: demografske informacije i znanja o ponovljivosti rezultata

Osim zaključaka o scenarijima, sudionicima je postavljen niz pitanja vezanih uz ponovljivost istraživanja. Prikupljanjem ovih informacija pokušala se: a) ispitati opća informiranost i mišljenje sudionika o ponovljivosti znanstvenih istraživanja i b) provjeriti da informiranost o ponovljivosti znanstvenih istraživanja nije bila značajno veća u jednoj skupini u odnosu na drugu. Sukladno tomu, sudionicima su postavljena sljedeća pitanja:

- a) Poznavanje samoga koncepta ponovljivosti (reproducibilnosti): Jeste li upoznati s pojmom reproducibilnosti (ponovljivosti) znanstvenih istraživanja? (odgovori "Da" i "Ne")
- b) S ciljem ispitivanja koliki bi postotak istraživanja trebao biti ponovljiv, sudionici su, ako je odgovor na prethodno pitanje bio potvrđan, zamoljeni da unesu procjenu u kolikom bismo postotku (raspon od 0 do 100 %) prilikom provođenja znanstvenih istraživanja svaki put trebali dobiti iste ili slične rezultate.
- c) Dva su pitanja bila usmjerena na ispitivanje mišljenja vezanih uz sintezu dokaza, koja je zapravo logično rješenje situacije u kojoj imamo suprotne nalaze istraživanja:
 - 1) Jeste li upoznati s pojmom sustavnih pregleda? (odgovori "Da" i "Ne")

- 2) Ako da, prema Vašem mišljenju, što je znanstveno vrjednije za određeno istraživačko pitanje:
 - a) Kvalitetan eksperiment s velikim brojem osoba.
 - b) Kvalitetan sustavni pregled.
 - c) Obje vrste istraživanja podjednako su važne.
- d) Konačno, mišljenje sudionika o ponovljivosti u znanstvenom radu ispitano je jednom česticom, s mogućim odgovorima "Da" i "Ne": "Smatrate li da je potrebno provoditi ponovljena istraživanja fenomena koji su dokazani prije više od desetljeća?"

Od demografskih informacija prikupljeni su spol i stupanj obrazovanja.

Etička privola

Sve postupke opisane u ovom istraživanju odobrilo je Etičko povjerenstvo Filozofskog fakulteta u Splitu (Klasa: 029-06/22-03/0001, Ur. broj: 2181-190-23-00020).

Izračun veličine uzorka

S obzirom na to da se radi o istraživanju koje je usmjereno na relativno nov koncept ponovljivosti, nisu pronađena eksperimentalna istraživanja na temelju kojih bismo mogli napraviti izračun veličine uzorka. Stoga je pretpostavljena potencijalna razlika u frekvencijama odgovora "Da" na pitanja o točnosti istraživanja između skupina od 20 % (60 % u skupini A i 40 % u skupini B), pa je sa snagom istraživanja od 80 % i razinom alfa pogreške od 5 % izračunano da treba minimalno 95 osoba po skupini da bi se uočila značajna razlika. Izračun je napravljen upotrebom programa *Select* (<https://select-statistics.co.uk/calculators/sample-size-calculator-two-proportions/>).

Obradba podataka

Sve su analize prethodno planirane u registraciji istraživanja. Kategorijske varijable prikazane su kao frekvencije i postotci, dok su numeričke varijable testirane na odstupanja od normalnosti i prikazane su kao medijani s interkvartilnim rasponima ili rasponima pouzdanosti. Skupine su uspoređene hi-kvadrat testom za kategorijske varijable ili Mann Whitney testom za numeričke varijable. Kao veličine učinka uzeti su omjer rizika za kategorijske varijable i rang-biserijalni koeficijent korelacije za kontinuirane varijable. Značajne razlike ušle su u logističku regresiju kao prediktori intervencijskih skupina kao kriterija. Snaga prediktora izražena je kao omjer izgleda s rasponom pouzdanosti. Usporedba stupnja sigurnosti u odgovor između različitih pitanja napravljena je Friedmanovim

testom za ponovljena mjerenja. Sve su analize napravljene u programskom paketu Jamovi (The jamovi project, 2024), osim izračuna medijana s pripadajućim 95 % rasponima pouzdanosti, koji su napravljeni u R jeziku za statističku analizu (R Core Team, 2021). Svi podatci i analize napravljene u ovom istraživanju dostupni su na OSF stranici projekta.

REZULTATI

Iz analize su isključeni sudionici koji su odustali odmah pri dobivanju scenarija ($n = 29$) ili koji su naknadno odustali ($n = 1$), stoga je analizirano 194 sudionika, 97 u svakoj eksperimentalnoj skupini.

Varijabla		Grupa	
		Replikacija nije statistički značajna ($n = 97$)	Replikacija je statistički značajna ($n = 97$)
Spol ($n, \%$)	Muški	33 (34,02)	17 (17,53)
	Ženski	63 (64,95)	77 (79,38)
	Ne želim se izjasniti	1 (1,03)	3 (3,09)
Stupanj obrazovanja ($n, \%$)	Završena osnovna škola	0 (0,00)	1 (1,03)
	Završena srednja škola	5 (5,15)	4 (4,12)
	Trenutačno studiram	64 (65,98)	69 (71,13)
	Završen stručni studij	2 (2,06)	1 (1,03)
	Završen fakultet	24 (24,74)	11 (11,34)
	Završen doktorat znanosti	2 (2,06)	11 (11,34)
Jeste li upoznati s pojmom reproducibilnosti (ponovljivosti) u istraživanjima? (Da, $n, \%$)		69 (71,13)	67 (69,07)
Ako ste odgovorili potvrdno na prethodno pitanje, označite u kojem postotku bismo trebali dobiti iste ili slične rezultate kao i u inicijalnom istraživanju, ako bismo ponovili znanstveno istraživanje na drugim podatcima (0-nijednom, 100-svaki put) (Md, 95 % CI)		75 (70-80)	74 (63-76)
Jeste li upoznati s pojmom sustavnih pregleda? (Da, $n, \%$)		77 (79,38)	69 (71,13)
Ako da, što je po Vama vrijednije za određeno istraživačko pitanje (Da, $n, \%$): ^a	kvalitetan eksperiment s velikim brojem osoba	10 (10,31)	9 (9,28)
	kvalitetan sustavni pregled obje vrste istraživanja su podjednako važne	15 (15,57)	16 (16,49)
		57 (58,76)	52 (53,61)
Smatrate li da je potrebno provoditi ponovljena istraživanja fenomena koji su dokazani prije više od desetljeća? (Da, $n, \%$)		92 (94,85)	88 (90,72)

📌 **TABLICA 1**
Demografske karakteristike sudionika i stavovi prema ponovljivosti istraživanja prikazani prema eksperimentalnim skupinama

Kratice: Md – medijan; CI – engl. *confidence interval*, raspon pouzdanosti.

^a 15 odgovora nedostaje u skupini *Replikacija nije statistički značajna* i 20 odgovora u skupini *Replikacija je statistički značajna*.

U istraživanju je ukupno sudjelovalo 140 (72,16 %) osoba ženskoga spola (Tablica 1). Sudionici su većinom bili studenti ($n = 133, 68,56 \%$) te su pretežito bili upoznati s pojmom po-

novljivosti u istraživanjima ($n = 136, 70,10\%$). Prema sudionicima koji su bili upoznati s pojmom ponovljivosti, za većinu istraživanja (medijan 75% , interkvartilni raspon: $51-90\%$) trebali bismo uvijek dobiti iste ili slične rezultate. Većina sudionika ($n = 146, 75,26\%$) izjavila je kako su upoznati s pojmom sustavnih pregleda, a polovica ($n = 109, 55,86\%$) smatrala je kako su i sustavni pregled i eksperiment s velikim brojem osoba podjednako znanstveno vrijedni. Konačno, čak 180 sudionika ($92,78\%$) navelo je da treba provoditi ponovljena istraživanja (replikacije) fenomena koji su dokazani prije više desetljeća.

Usporedbom skupina pronađeno je da su sudionici najčešće odgovarali kako ne mogu donijeti zaključak na temelju prikazanih informacija, a najrjeđe da ni jedna od opisanih studija nije točna (Tablica 2, Slike 1 i 2). Ipak, u slučaju kad je replikacija inicijalnog istraživanja pokazala da je intervencija statistički značajna, sudionici su u značajno većem postotku vjerovali da je replikacija točna nego u odnosu na slučaj kad replikacija nije uspjela ponoviti značajne rezultate inicijalnog istraživanja.

● **TABLICA 2**
Odgovori sudionika o vjerojatnosti točnosti rezultata podijeljeni prema eksperimentalnim skupinama

Varijabla	Replikacija nije statistički značajna ($n = 97$)	Replikacija je statistički značajna ($n = 97$)	p^a	Veličina učinka ^b
Prva studija je točna (Da, $n, \%$)	50 (51,55)	52 (53,61)	0,892	0,96
Sigurnost u odgovor (Md, 95 % CI)	60 (51-64)	60 (55-68)	0,732	0,03
Druga studija je točna (Da, $n, \%$)	53 (54,64)	71 (73,20)	0,012	0,46
Sigurnost u odgovor (Md, 95 % CI)	56 (50-70)	64 (55-70)	0,563	0,05
Obje studije su točne (Da, $n, \%$)	42 (43,30)	51 (52,58)	0,246	0,72
Sigurnost u odgovor (Md, 95 % CI)	75 (60-80)	70 (60-75)	0,728	0,03
Niti jedna studija nije točna (Da, $n, \%$)	19 (19,59)	15 (15,46)	0,429	1,35
Sigurnost u odgovor (Md, 95 % CI)	73 (60-81)	71 (60-81)	0,925	0,01
Ne mogu donijeti konačan zaključak (Da, $n, \%$)	68 (70,10)	63 (64,95)	0,505	1,23
Sigurnost u odgovor (Md, 95 % CI)	82 (75-90)	79 (73-91)	0,517	0,05

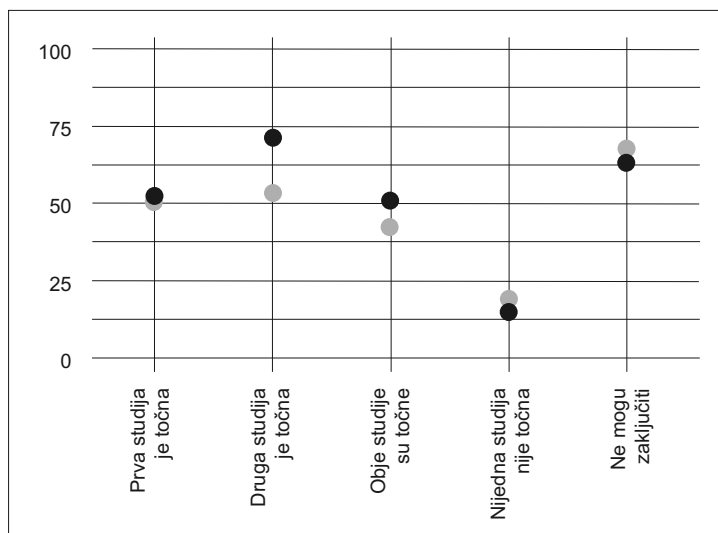
Md – medijan; CI – engl. *confidence interval*, raspon pouzdanosti.

^a Hi kvadrat za usporedbu kategorijskih varijabli i Mann Whitney U test za usporedbu kontinuiranih varijabli.

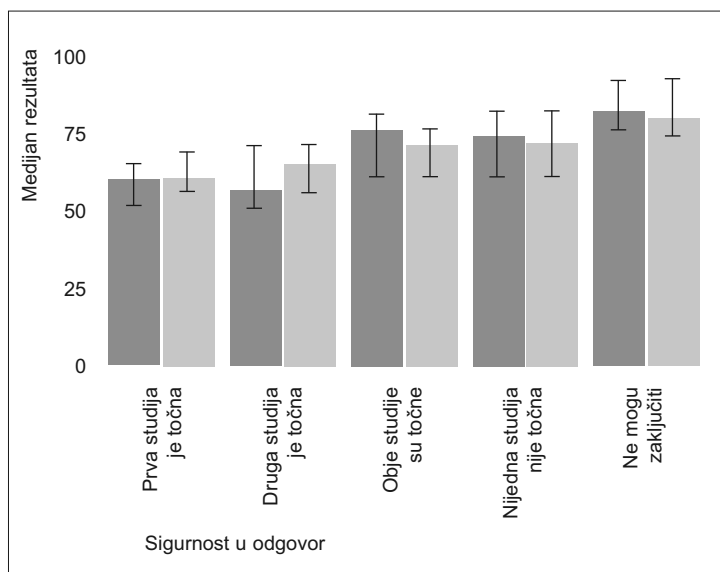
^b Omjer rizika za kategorijske varijable i rang-biserijski koeficijent korelacije za kontinuirane varijable.

U logističkoj regresiji, u koju je kao prediktor ušla samo frekvencija odgovora "Da" na pitanje "Druga studija je točna" te kojom se posljedično pokušalo utvrditi predviđaju li frekvencije odgovora o točnosti istraživanja eksperimentalnu skupinu, izračunano je da je omjer izgleda 2,16, odnosno da je toliko puta veća vjerojatnost da će sudionik izjaviti kako je istraživanje točno ako je ponovljeno istraživanje (replikacija) bilo značajno (95 % CI 1,18-3,96) (Tablica 3).

➔ SLIKA 1
Broj sudionika u istraživanju koji se složio s izjavama o istraživanju prema intervencijskoj skupini A (siva boja): "Replikacija nije statistički značajna" i intervencijskoj skupini B (crna boja): "Replikacija je statistički značajna"



➔ SLIKA 2
Usporedba procjene stupnja sigurnosti od 0 – nimalo siguran/a do 100 – potpuno siguran/a (medijan i 95 % rasponi pouzdanosti) u odgovor između intervencijske skupine A (tamnosiva boja): "Replikacija nije bila značajna" i intervencijske skupine B (svijetlosiva boja): "Replikacija je bila značajna"



🔗 TABLICA 3
Logistička regresija u predviđanju intervencijske skupine A

Prediktor	Omjer izgleda (95 % raspon pouzdanosti)	<i>p</i>	McFadden R ²
Točka presjeka	0,35 (0,15-0,82)	0,016	0,02
Druga studija je točna.	2,16 (1,18-3,96)	0,012	

U prediktorima je odgovor "Da" bio označen sa 1, a odgovor "Ne" označen sa "2". Referentna razina kriterijske varijable "grupa" je bila intervencijska skupina B: "Replikacija je statistički značajna".

Friedmanovim testom pokazano je kako se stupanj sigurnosti na različitim odgovorima razlikovao između ponuđenih odgovora ($\chi^2 = 54,86$, $df = 4$, $p < 0,001$). U odnosu na ostale

odgovore, stupanj sigurnosti bio je značajno niži za odgovore koji su navodili da je isključivo prva ili druga studija točna. Stupnjevi sigurnosti sudionika koji su navodili da ni jedna studija nije točna ili da su obje točne nisu se međusobno razlikovali, dok se stupanj sigurnosti u odgovor koji navodi da se ne može donijeti konačan zaključak nije razlikovao od zaključka da ni jedna studija nije točna.

RASPRAVA

U ovom eksperimentu provedenom u virtualnom okruženju pokušalo se ispitati hoće li manipulacija redosljeda prikaza identičnih istraživanja koja daju suprotne zaključke utjecati na procjenu točnosti svakoga pojedinačnog istraživanja. Iako su sudionici na temelju dobivenih informacija izjavljivali da ne mogu sa sigurnosti donijeti konačan zaključak te su se u situaciji kada su morali donijeti odluku o točnosti originalnog istraživanja ili replikacije osjećali najmanje sigurni u svoje procjene ipak je pronađen obrazac donošenja odluka u takvim nejasnim situacijama. Hipoteza da će sudionici više vjerovati prvim istraživanjima koja su prikazana nije se pokazala točnom. Značajno veći broj sudionika vjerovao je da je replikacija točna kada je ona pokazala značajan rezultat (a inicijalno istraživanje neznačajan rezultat) nego kad je pokazala suprotno (to jest kada je inicijalno istraživanje pokazalo značajan rezultat). Ovo upućuje na mogućnost da pristranost prema značajnim rezultatima postoji čak i u slučaju repliciranih i nejasnih rezultata, što ima važne implikacije za interpretaciju znanstvenih nalaza. Osim toga, značajno je više sudionika izjavljivalo da su upoznati s temama ponovljivosti i sustavnim pregledima, iz čega se može zaključiti da ni druga hipoteza nije potvrđena. Ovo implicira da je tema ponovljivosti znanstvenih rezultata prisutna u javnom kontekstu, ali je moguće i to da su sudionici u istraživanju bile osobe koje su zainteresiranije za teme.

Glavni nalaz ovog istraživanja jest taj da su sudionici više vjerovali u točnost nalaza replikacije koja je bila statistički značajna u odnosu na neznačajnu inicijalnu studiju, naspram obrnute situacije. Nalaz da se u psihologijskim istraživanjima češće nalaze značajni rezultati nije novina (npr. Kühberger i sur., 2014), ali ovo istraživanje ide korak dalje i pokazuje kako postoji mogućnost da čak postoji pristranost čitatelja da značajne rezultate procjenjuju točnijima naspram neznačajnih. U istraživanju Jankowskog i suradnika (2022), sudionici su bili randomizirani u jednu od četiri eksperimentalne skupine, u kojima su u samo jednom scenariju mijenjani statistička značajnost te sudionici nisu mogli uspoređivati vjerodostojnost

različitih opisa. Naši su rezultati u skladu s rezultatima spomenutog istraživanja, jer se pokazalo da će sudionici vjerovati značajnom rezultatu, ne samo kada vide samostalno istraživanje nego i kad im je prikazano identično istraživanje s potpuno suprotnim nalazima. Nadalje, osim same pristranosti u procjeni točnosti istraživanja, postoji mogućnost da čitatelji čekaju značajan rezultat. Dodatan problem u svakodnevnom radu jest pitanje definicije uspješne replikacije zbog izostanka složnosti oko potrebnih kriterija (Muradchian i sur., 2021). Stoga se u scenarijima u ovom istraživanju nije navodio ni jedan ustaljeni kriterij značajnosti istraživanja (p vrijednost ili veličina učinka), nego je samo navedeno da učinak postoji, bez navođenja kvantitativnih podataka. U intervencijskoj skupini A psihoedukacija se inicijalno pokazala značajnom, no taj je nalaz bio doveden u pitanje naknadnom replikacijom. S druge strane, kad se inicijalno istraživanje nije prikazano kao značajno u skupini B, ako se replikacija pokazala značajnom, sudionici su procjenjivali da je druga studija točnija. Ni u jednom od opisanih scenarija ne može se donijeti konačan zaključak, jer su oba istraživanja imala iste veličine uzorka i metodološke karakteristike. Sljedeći korak u testiranju ove hipoteze bilo bi ponavljanje ovog istraživanja kako bi se moglo utvrditi jesu li ovi nalazi slučajni.

Sudionici su bili najsigurniji u odabiru odgovora da se ne može donijeti konačan zaključak o učinkovitosti provedenih istraživanja, što implicira da je većina sudionika (oko 70 %) shvaćala da opisani nalazi nisu dovoljni za donošenje zaključaka o učinkovitosti, jer su nalazi konceptualno istih istraživanja potpuno suprotni. Nedostatak dokaza može predstavljati problem u slučajevima u praktičnom radu, gdje pojedinac treba donijeti odluku na temelju provođenja određenoga testiranja, terapije ili intervencije. U ovom istraživanju sudionici su se očito oslanjali na podatke o značajnosti istraživanja kao mentalnu kraticu za njegovu točnost, to jest procjenjivali replikaciju značajno točnijom ako je tretman bio značajno učinkovit u odnosu na situaciju kada replikacija nije bila točna. Nedavno istraživanje pokazalo je kako postoji velik broj medicinskih intervencija za koje se ne može izvesti snažan zaključak s obzirom na dostupnu razinu dokaza (Banić i sur., 2022). Stoga, s obzirom na složenost i prirodu istraživanih konstrukata, valja očekivati da bi na velikim uzorcima ista situacija s nedostatkom snažnih i jasnih zaključaka bila i u psihologijskim istraživanjima (u teorijskima i praktičnima). Zbog toga bi bilo važno poznavati i prepoznavati kognitivne pristranosti povezane s procesom interpretacije znanstvenih nalaza te recenzije i objavljivanja znanstvenih radova da bi se utvrdilo kako

profesionalci reagiraju kada su suočeni s nejasnim informacijama. Naši rezultati sukladni su s onima Bottesini i suradnika (2023), jer pokazuju da, uz čimbenike poput veličine uzorka, redosljed prikazivanja istraživanja može utjecati na interpretaciju točnosti istraživanja. Uz to je važno i naglasiti da je razmjerno malen broj sudionika izjavio kako ni jedna studija nije točna ($n = 34, 17,44 \%$), što je također bilo moguće, s obzirom na to da nismo imali detalje o provedenim istraživanjima, pa su obje studije mogle biti netočne. Postoji i mogućnost da je većina sudionika smatrala kako od prikazanih istraživanja barem jedno mora biti točno, što bi također upućivalo na prisutnost pristranosti uvjerenja (Trippas i sur., 2014), što bi pak u praktičnom kontekstu potencijalno dovelo do pogrešnih odluka. Međutim, ovaj nalaz trebao bi biti bolje i sustavnije istražen u budućnosti prije donošenja snažnijih zaključaka.

Svim je sudionicima postavljen niz pitanja o ponovljivosti znanstvenih istraživanja i sintezama dokaza. Sekundarna hipoteza ovog istraživanja – da većina sudionika neće biti upoznata s konceptima sustavnih pregleda i ponovljivosti – također nije potvrđena. S obzirom na to da dosad nisu bila provedena istraživanja o tome koliko je hrvatska populacija upoznata s praksama otvorene znanosti, ovo istraživanje započelo je s pretpostavkom da će koncepti biti zanemareni. Visok postotak osoba koje su izjavile da su upoznate s konceptima može upućivati na to da su istraživanju pristupile osobe koje su zainteresiranije za temu, što dovodi u pitanje poopćivost rezultata na opću populaciju. Osim toga, upoznatost s konceptima ispitana je pitanjima s binarnim odgovorima ("Da" ili "Ne"), koji neprecizno mjere konstrukt. Buduća istraživanja trebala bi rabiti detaljnija i preciznija pitanja, ali uključiti i kvalitativan pristup. U buduću treba provesti i istraživanja koja bi detaljnije ispitala upoznatost hrvatske populacije s konceptima otvorene znanosti, pa je važno naglasiti da takvi pokušaji već postoje (Pale i sur., 2024). Većina sudionika složila se s tim da je provođenje replikacija potrebno u kontinuitetu te da treba provoditi kvalitetna eksperimentalna istraživanja (primarna istraživanja) i kvalitetne sustavne preglede (sekundarna istraživanja).

Kada su upitani u kolikom bismo postotku istraživanja trebali dobiti iste ili slične nalaze, medijan odgovora sudionika bio je oko 75 %, što je još uvijek značajno više od dosad dobivenih 40 % u prijašnjim procjenama psihologijskih istraživanja (Open Science Collaboration, 2015; Youyou i sur., 2023). Iako je promjena u pristupu provođenja psihologijskih istraživanja nužna, do nje dolazi izrazito sporo, s obzirom na to da je još vrlo malen broj istraživanja preregistriran (Thibault i sur.,

2023) te da jako malo istraživača dijeli svoje podatke (Teder-
soo i sur., 2021). Kako bi pratili svjetske znanstvene standarde
u slučajevima kada imaju ograničene resurse i ograničen pri-
stup ispitanicima, istraživači bi se trebali usredotočiti na pro-
vođenje kvalitetnih i metodološko rigoroznih sustavnih pre-
gleda ili se uključivati u međunarodne istraživačke skupine.
Ovaj pristup poboljšao bi iskorištenost resursa te pridonio
međunarodnoj prepoznatljivosti manjih istraživačkih skupi-
na u globalnom kontekstu. Pritom bi u procesu racionalne upo-
trebe resursa značajno pomogli časopisi uvođenjem preregi-
striranih izvještaja (engl. *registered reports*) (Purgar i sur. 2024),
u kojima protokol evaluiraju nezavisni stručnjaci, pri čemu se
istraživanje provedeno na osnovi protokola objavljuje neovi-
sno o smjeru rezultata dokle god je ono provedeno prema spo-
menutom protokolu. Tako je moguće unaprijed uočiti istraži-
vanja s metodološkim problemima te istraživanja koja obra-
đuju već istražene teme, kao i da se unaprijed mogu dogovo-
riti kriteriji značajnosti rezultata. Osim toga, u dogledno je vri-
jeme cilj napraviti preregistraciju istraživanja i dijeljenje po-
dataka normom tijekom provođenja znanstvenih istraživanja,
što bi znatno pridonijelo prepoznavanju stabilnih i validnih
zaključaka u psihologiji, naspram onih koji su rezultat slučaj-
nosti i statističkih manipulacija (Flis, 2023). Usprkos velikom
zamahu istraživanja koja su se usredotočila na tehničke nači-
ne poboljšanja ponovljivosti – poput preregistracije znanstve-
nih istraživanja (Strømmland, 2019) ili dijeljenja podataka (Merz
i sur., 2020) – psihološki faktori razumijevanja reproducibil-
nosti poput kognitivnih pristranosti i iskustava još su uvijek
relativno neistraženi. Stoga je preporuka da buduća istraživa-
nja stave veći naglasak na ispitivanje koliko i kako su nave-
deni faktori povezani s donošenjem odluka o interpretaciji
ponovljivosti istraživanja.

Prilikom interpretacije rezultata provedenog istraživanja
treba uzeti u obzir neka ograničenja. Istraživanje je provede-
no u virtualnom okruženju i u pozivu je oglašavano kao is-
traživanje o interpretaciji znanstvenog istraživanja, stoga je
vjerojatnije kako se sudionici koji su pristali sudjelovati više
zanimaju za teme razumijevanja znanstvenih istraživanja. Sa-
mi scenariji bili su napisani kao sažetci iz svakodnevnoga
tiska, a ne nalazi znanstvenog istraživanja, pa bi rezultate pro-
vedenog eksperimenta trebalo promatrati uz dozu opreza ako
bi ih se željelo poopćiti za svakodnevno čitanje znanstvenih
članaka. Uz to, među osobama s doktoratom postojao je ma-
nji nerazmjer između skupina podijeljenih prema slučaju jer
su sudionici raspoređeni u skupine putem mrežnoga progra-
ma, no zbog malobrojnosti malo je vjerojatno da je to utjeca-

lo na rezultate. Nadalje, uzorak se pretežno sastojao od studenata, što je jedan od razloga zašto nije poopćiv na opću populaciju. S druge strane, uzorci u kojima sudjeluju uglavnom studenti relativno su česti za psihologijska istraživanja. Buduća istraživanja trebala bi ispitati i smjer studija uključenih studenata i uključivanje šire populacije, s ciljem dobivanja bolje reprezentativnosti uzorka. Osim toga, nisu se uspjele prikupiti informacije o sudionicima koji su napustili istraživanje prije randomizacije u intervencijske skupine, ali taj otpad ispitanika nije bio velik (12,8 %). Konačno, u ovom se istraživanju kao scenarij rabilo istraživanje jedne intervencije (utjecaj psihoedukacije na smanjenje perfekcionizma) koju su sudionici mogli otprije bolje poznavati te znati da je učinkovita u smanjenju simptoma perfekcionizma. Ova se pristranost pokušala kontrolirati slučajnim rasporedom u skupine, s obzirom na to da nisu ispitana znanja sudionika o psihoedukacijama prije provođenja istraživanja. Za čvršću snagu dokaza ipak bi trebalo ponoviti istraživanje upotrebom različitih tema kao podražaja. U analizi podataka korišteni su neparametrijski postupci koji imaju manju statističku snagu, pa je i manja vjerojatnost da će se opaziti razlika između skupina negoli upotrebom parametrijskih postupaka. Ipak, s obzirom na to da pretpostavke za parametrijske postupke nisu bile zadovoljene, njihovo provođenje ne bi bilo opravdano. Konačno, sudionici su pitani treba li ponavljati istraživanja koja su već dokazana. Iako se ovim putem pokušalo procijeniti potrebe za replikacijama, ispitanici su potencijalno mogli biti zbunjeni, jer se u društvenim istraživanjima fenomeni prikazuju, a ne dokazuju. Stoga je preporuka u budućim istraživanjima izbjegavati formulaciju "dokaza" u kontekstu reproducibilnosti te staviti veći naglasak na predstavljanje fenomena.

Naknadna istraživanja ili replikacije ovog istraživanja trebala bi stremiti većem omjeru ekološke valjanosti, jer bi se istraživanje provedeno u ovakvim uvjetima moglo smatrati "umjetnim". Kao primjer za ovu tvrdnju jest intervencijska skupina, u kojoj se pokušava ponoviti istraživanje koje nije inicijalno pronašlo značajan učinak, što se u praktičnom radu gotovo nikad ne događa. Buduća istraživanja mogu dodati i treću skupinu, u kojoj bi učinak postojao i u originalnom istraživanju i u replikaciji, kako bi se poboljšala unutarnja valjanost i uvela kontrolna skupina. Ipak, moguće je da osobe u pretraživanju literature najprije naiđu na istraživanje koje navodi da nema efekta, a zatim i na sljedeće koje navodi da ima, što smo pokušali preslikati u ovom istraživanju.

S obzirom na niz pitanja koja proizlaze iz ovog istraživanja, preporuka je da buduća istraživanja idu u (barem) dva

smjera. Prvi smjer odnosio bi se na istraživanje donošenja odluka o točnosti istraživanja, za što bi bio potreban kvalitativni pristup. Kroz niz intervjua i/ili fokusnih skupina istraživači mogu pokušati dobiti uvid u čimbenike na koje se sudionici oslanjaju kada donose odluku o točnosti znanstvenih istraživanja, odnosno kada trebaju donijeti odluku je li istraživanje uspješno ponovljeno ili ne. Nadalje, potrebno je ponovno provesti ovo istraživanje ispitivanjem niza različitih tema i različitih populacija kako bi se utvrdila održivost pronađenoga fenomena. Moguće je da rezultati ne bi bili ponovljivi ako bi u istraživanje bili uključeni samo stručnjaci, koji bi se vjerojatno suzdržavali od donošenja bilo kakve odluke na temelju male količine podataka. Također je moguće da bi rezultati bili različiti ako bi u istraživanje bile uključene samo one osobe kojima je tema izrazito važna (npr. osobe s izraženim perfekcionizmom koji narušava svakodnevno funkcioniranje). Potrebno je sve nove čimbenike za koje se potencijalno utvrdi da utječu na pristranost u interpretaciji rezultata znanstvenih istraživanja testirati upotrebom složenijih eksperimentalnih nacrtava, provjeravajući interakcije čimbenika, s ciljem stvaranja smjernica što objektivnije interpretacije nalaza znanstvenih istraživanja.

Konačno, praktična implikacija ovog istraživanja jest isticanje potrebe za uključivanjem teme ponovljivosti u metodološke kolegije. Da bi se povećalo razumijevanje znanstvenih istraživanja u općoj populaciji, moguće je uz objavljene radove imati i sažetke istraživanja napisane jednostavnijim jezikom, pri čemu bi se možda smanjio jaz između proizvodnje znanstvenih dokaza i njihove upotrebe u općoj populaciji.

ZAKLJUČAK

Sudionici u ovom istraživanju procjenjivali su rezultate točnijima kad je replikacija bila statistički značajna u odnosu na suprotni slučaj. Ovaj nalaz implicira da čitatelji vjeruju novijim i značajnim istraživanjima, iako se ona ne moraju metodološki značajno razlikovati od istraživanja u kojima nije dobio značajan rezultat. Za potvrdu ovoga nalaza trebalo bi ponoviti ovo istraživanje na drugom uzorku i kvalitativno istraživanje o donošenju odluka prilikom interpretacije rezultata znanstvenih istraživanja, kako bismo bolje razumjeli moguće pristranosti u tom procesu.

Financiranje

Ovaj je rad nastao u okviru znanstvenoga projekta *Hrvatska mreža za ponovljivost znanstvenih rezultata i znanstveno-istraživačku čestitost* u okviru programa Znanje i Otkrića 2022, koji financira Zaklada Adris.

Anketni upitnik

Obavijest o istraživanju

Poštovana/i, pred Vama je upitnik koji se koristi kao instrument istraživanja u sklopu projekta Hrvatska mreža za ponovljivost znanstvenih istraživanja i znanstveno-istraživačku čestitost, koje je financirano od Zaklade Adris u sklopu natječaja Znanje i otkrića 2022.

Istraživanjem se pokušava otkriti kako osobe interpretiraju i donose odluke na temelju sažetaka znanstvenih informacija. Vaši odgovori su potpuno anonimni, bit će upotrijebljeni samo u istraživačke svrhe i analizirani na grupnoj razini, te će samo istraživači imati pristup Vašim podacima. Anketa je dobrovoljna i namijenjena svim punoljetnim građanima Republike Hrvatske. Ukoliko ne želite sudjelovati u istraživanju, slobodni ste prestati ispunjavati anketu u svakom trenutku.

Ispitivanje je odobrilo Etičko povjerenstvo Filozofskog fakulteta u Splitu (Broj odobrenja: Klasa: 029-06/22-03/0001, Ur. broj: 2181-190-23-00020). Za daljnje obavijesti o istraživanju, slobodni ste kontaktirati istraživača putem emaila: ibuljan@ffst.hr. Molimo Vas da pročitate pitanja i odgovorite na njih odabirom ponuđenih odgovora ili upisivanjem odgovora na predviđeno mjesto. Iskreno cijenimo Vašu povratnu informaciju i zahvaljujemo na strpljenju i susretljivosti. Ukoliko pristajete sudjelovati u istraživanju, molimo potvrdite odabirom opcije "Slažem se" koja se nalazi ispod Suglasnosti za sudjelovanje.

Suglasnost za sudjelovanje

1. Potvrđujem da sam pročitao/pročitala ovu obavijest za gore navedeno znanstveno istraživanje.
2. Razumijem da je moje sudjelovanje dobrovoljno te se mogu povući u bilo koje vrijeme, bez navođenja razloga i bez ikakvih posljedica po zdravstvenom ili pravnom pitanju.
3. Razumijem da mojim anonimiziranim informacijama imaju pristup odgovorni pojedinci, tj. glavni istraživač i članovi Etičkog povjerenstva koje je odobrilo ovo znanstveno istraživanje. Dajem dozvolu tim pojedincima za prikupljanje i obradu podataka.
4. Razumijem da će moje anonimizirane informacije biti javno dostupne. Razumijem da to znači da će se moći koristiti za potrebe istraživanja osim onih navedenih u obavijesti o istraživanju. Također razumijem da se moje anonimizirane informacije mogu koristiti u državama izvan Europe te da uredbe o zaštiti podataka u tim zemljama mogu biti drugačije od onih unutar Europske unije.
5. Želim sudjelovati u navedenom znanstvenom istraživanju.

Format A Kognitivno bihevioralna terapija je jedan od pristupa u psihoterapiji koji je usmjeren na rješavanje trenutnih psiholoških problema kod pojedinca, kroz promjenu misli i ponašanja, uz strukturiran plan i vremenski ograničene ciljeve. Perfekcionizam je skup osobina koje

karakterizira želja osobe za savršenstvom i ispunjavanjem vrlo visokih standarda, te je osoba često sklona sebe strogo procjenjivati i brinuti se da ju drugi vide pozitivno. Istraživači su pokušali utvrditi je li kognitivno bihevioralna terapija učinkovita u smanjenju perfekcionizma. Primijenjena je tehnika psihoedukacije, u kojoj su iskusni terapeuti osobe s perfekcionizmom obrazovali o definiciji perfekcionizma, povezanosti s drugim problemima, potencijalnim uzrocima i podržavajućim čimbenicima te su pacijenti pročitali ulomke knjige za samopomoć. Pacijenti su prema slučaju raspoređeni u skupinu koja je dobivala psihoedukaciju i koja nije. Ukupan broj pacijenata koji su sudjelovali u istraživanju je bio 250. Razina perfekcionizma je mjerena upitnikom prije psihoedukacije te četiri tjedna nakon, kako bi se utvrdilo je li došlo do smanjenja perfekcionizma nakon intervencije, te su rezultati uspoređeni s drugom skupinom osoba koja nije dobila psihoedukaciju. Pacijenti su se pridržavali terapije te je utvrđeno kako je došlo do značajnog smanjenja perfekcionizma nakon četiri tjedna. Zaključak istraživanja je da je psihoedukacija učinkovita u tretmanu perfekcionizma.

Druga skupina istraživača je odlučila ponoviti spomenuto istraživanje, te su prikupili jednak broj pacijenata (u drugom istraživanju je ukupan broj osoba s perfekcionizmom bio opet 250). Primijenili su iste postupke u istraživanju kao i prethodna studija; rasporedili su osobe u skupine prema slučaju, mjerili su perfekcionizam na početku istraživanja, sa sudionicima su razgovarali iskusni terapeuti, dobili su identične materijale za čitanje te je perfekcionizam mjereno četiri tjedna nakon psihoedukacije. Kao rezultat istraživanja je dobiveno kako nema razlike u razini perfekcionizma prije i četiri tjedna nakon intervencije. Istraživači su zaključili kako psihoedukacija nije učinkovita u tretmanu perfekcionizma.

Format B

Kognitivno bihevioralna terapija je jedan od pristupa u psihoterapiji koji je usmjeren na rješavanje trenutnih psiholoških problema kod pojedinca, kroz promjenu misli i ponašanja, uz strukturiran plan i vremenski ograničene ciljeve. Perfekcionizam je skup osobina koje karakterizira želja osobe za savršenstvom i ispunjavanjem vrlo visokih standarda, te je osoba često sklona sebe strogo procjenjivati i brinuti se da ju drugi vide pozitivno. Istraživači su pokušali utvrditi je li kognitivno bihevioralna terapija učinkovita u smanjenju perfekcionizma. Primijenjena je tehnika psihoedukacije, u kojoj su iskusni terapeuti osobe s perfekcionizmom obrazovali o definiciji perfekcionizma, povezanosti s drugim problemima, potencijalnim uzrocima i podržavajućim čimbenicima te su pacijenti pročitali ulomke knjige za samopomoć. Pacijenti su prema slučaju raspoređeni u skupinu koja je dobivala psihoedukaciju i koja nije. Ukupan broj pacijenata koji su sudjelovali u istraživanju je bio 250. Razina perfekcionizma je mjerena upitnikom prije psihoedukacije te četiri tjedna nakon, kako bi se utvrdilo je li došlo do smanjenja perfekcionizma nakon intervencije, te su rezultati uspoređeni s drugom skupinom osoba koja nije dobila psihoedukaciju. Pacijenti su se pridržavali terapije te je utvrđeno kako nije došlo do značajnog smanjenja perfekcionizma nakon četiri tjedna. Zaključak istraživanja je da psihoedukacija nije učinkovita u tretmanu perfekcionizma.

Druga skupina istraživača je odlučila ponoviti spomenuto istraživanje, te su prikupili jednak broj pacijenata (u drugom istraživanju je ukupan broj osoba s perfekcionizmom bio opet 250). Primijenili su iste postupke u istraživanju kao i prethodna studija; rasporedili su osobe u skupine prema slučaju, mjerili su perfekcionizam na početku istraživanja, sa sudionicima su razgovarali iskusni terapeuti, dobili su identične materijale za čitanje te je perfekcionizam mjeren četiri tjedna nakon psihoedukacije. Kao rezultat istraživanja je dobiveno značajno smanjenje u razini perfekcionizma prije i četiri tjedna nakon intervencije. Istraživači su zaključili kako je psihoedukacija učinkovita u tretmanu perfekcionizma.

U sljedećem pitanju Vas molimo da za svaku od ponuđenih izjava označite slažete li se ili ne odabirom jedne od ponuđenih opcija. Također molimo, označite koliko ste sigurni u svoj odgovor od 0 (nimalo siguran/a) do 100 (potpuno siguran/a).

Prva studija je točna.	Da	Ne	Sigurnost u odgovor:
Druga studija je točna.	Da	Ne	Sigurnost u odgovor:
Obje studije su točne.	Da	Ne	Sigurnost u odgovor:
Ni jedna od studija nije točna.	Da	Ne	Sigurnost u odgovor:
Nije moguće donijeti konačan zaključak.	Da	Ne	Sigurnost u odgovor:

Jeste li upoznati s pojmom reproducibilnosti (ponovljivosti) znanstvenih istraživanja? Da Ne

Ako da, molimo označite u kolikom postotku bismo prilikom provođenja znanstvenih istraživanja svaki put trebali dobiti iste ili slične rezultate (postotak)

Jeste li upoznati s pojmom sustavnih pregleda? Da Ne

Ako da, što je, prema Vašem mišljenju, znanstveno vrijednije za određeno istraživačko pitanje:

- a) Kvalitetan eksperiment s velikim brojem osoba.
- b) Kvalitetan sustavni pregled.
- c) Obje vrste istraživanja su podjednako važne.

Smatrate li da je potrebno provoditi ponovljena istraživanja fenomena koji su dokazani prije više od desetljeća? Da Ne

Molimo označite svoj spol: M Ž

Vaš stupanj obrazovanja:

- a) Završena osnovna škola
- b) Završena srednja škola
- c) Trenutno studiram
- d) Završen stručni studij
- e) Završen fakultet
- f) Završen doktorat znanosti

- Alston, J. M. i Rick, J. A. (2021). A beginner's guide to conducting reproducible research. *The Bulletin of the Ecological Society of America*, 102(2), e01801. <https://doi.org/10.1002/bes2.1801>
- Baker, M. (2016). 1,500 scientists lift the lid on reproducibility. *Nature News*, 533(7604), 452–454. <https://doi.org/10.1038/533452a>
- Banić, A., Fidahić, M., Šuto, J., Roje, R., Vuka, I., Puljak, L. i Buljan, I. (2022). Conclusiveness, linguistic characteristics and readability of Cochrane plain language summaries of intervention reviews: A cross-sectional study. *BMC Medical Research Methodology*, 22(1), Article no. 240. <https://doi.org/10.1186/s12874-022-01721-7>
- Bottesini, J. G., Aschwanden, C., Rhemtulla, M. i Vazire, S. (2023). How do science journalists evaluate psychology research? *Advances in Methods and Practices in Psychological Science*, 6(3), Article no. 25152459231183912. <https://doi.org/10.1177/25152459231183912>
- Buljan, I., Tokalić, R., Roguljić, M., Zakarija-Grković, I., Vrdoljak, D., Milić, P., Puljak, L. i Marušić, A. (2020). Comparison of blogshots with plain language summaries of Cochrane systematic reviews: A qualitative study and randomized trial. *Trials*, 21(1), Article no. 426. <https://doi.org/10.1186/s13063-020-04360-9>
- Fanelli, D. (2012). Negative results are disappearing from most disciplines and countries. *Scientometrics*, 90(3), 891–904. <https://doi.org/10.1007/s11192-011-0494-7>
- Flis, I. (2023). Scientific reform and replication crisis in psychology. *Psiholozijske teme*, 32(2), 227–256. <https://doi.org/10.31820/pt.32.2.1>
- Held, L., Micheloud, C. i Pawel, S. (2022). The assessment of replication success based on relative effect size. *Annals of Applied Statistics*, 16(2), 706–720. <https://doi.org/10.1214/21-AOAS1502>
- Hrvatska mreža za odgovornu i ponovljivu znanost (2023). Pristupljeno 17. lipnja 2024. <https://crorin.hr/>
- Ioannidis, J. P. (2005). Why most published research findings are false. *PLoS Medicine*, 2(8), e124. <https://doi.org/10.1371/journal.pmed.1004085>
- Jankowski, S., Boutron, I. i Clarke, M. (2022). Influence of the statistical significance of results and spin on readers' interpretation of the results in an abstract for a hypothetical clinical trial: A randomised trial. *Bmj Open*, 12(4), e056503. <https://doi.org/10.1136/bmjopen-2021-056503>
- Kühberger, A., Fritz, A. i Scherndl, T. (2014). Publication bias in psychology: A diagnosis based on the correlation between effect size and sample size. *PLoS ONE*, 9(9), Article e105825. <https://doi.org/10.1371/journal.pone.0105825>
- Merz, K. M. Jr., Amaro, R., Cournia, Z., Rarey, M., Soares, T., Tropsha, A., Wahab, H. A. i Wang, R. (2020). Editorial: Method and data sharing and reproducibility of scientific results. *Journal of Chemical Information and Modeling*, 60(12), 5868–5869. <https://doi.org/10.1021/acs.jcim.0c01389>
- Munafò, M. R., Nosek, B. A., Bishop, D. V. M., Button, K. S., Chambers, C. D., Percie du Sert, N., Simonsohn, U., Wagenmakers, E.-J., Ware, J. J. i Ioannidis, J. P. A. (2017). A manifesto for reproducible science. *Nature Human Behaviour*, 1(1), Article no. 0021. <https://doi.org/10.1038/s41562-016-0021>

Muradchanian, J., Hoekstra, R., Kiers, H. i van Ravenzwaaij, D. (2021). How best to quantify replication success? A simulation study on the comparison of replication success metrics. *Royal Society Open Science*, 8(5), 201697. <https://doi.org/10.1098/rsos.201697>

Open Science Collaboration. (2015). Estimating the reproducibility of psychological science. *Science*, 349(6251). <https://doi.org/10.1126/science.aac4716>

Pale, U., Hoić, M., Patarcic, I., Stojanovski, J., Buljan, I., Marusic, A. i Culina, A. (2024). *Survey on perception and practices of open science in Croatia*. Open Science Framework. Pristupljeno 9. travnja 2025. na <https://osf.io/2fmyz/>

Psychological Science Accelerator: A distributed laboratory network. (2023). <https://psysciacc.org>

Purgar, M., Glasziou, P., Klanjscek, T., Nakagawa, S. i Culina, A. (2024). Supporting study registration to reduce research waste. *Nature Ecology & Evolution*. Advance online publication. <https://doi.org/10.1038/s41559-024-02433-5>

R Core Team (2021). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria. <https://www.R-project.org>

Sherif, M., Taub, D. i Hovland, C. I. (1958). Assimilation and contrast effects of anchoring stimuli on judgments. *Journal of Experimental Psychology*, 55(2), 150–155. <https://doi.org/10.1037/h0048784>

Strömmland, E. (2019). Preregistration and reproducibility. *Journal of Economic Psychology*, 75, Part A. <https://doi.org/10.1016/j.joep.2019.01.006>

SurveyMonkey (2024). *SurveyMonkey software*. Momentive. Pristupljeno 17. lipnja 2024. na <https://www.surveymonkey.com>

Taylor, B. N. i Kuyatt, C. E. (1994). *Guidelines for evaluating and expressing the uncertainty of NIST measurement results (Vol. 1297)*. US Department of Commerce, Technology Administration, National Institute of Standards and Technology. <https://doi.org/10.6028/NIST.TN.1297>

Tedersoo, L., Küngas, R., Oras, E., Köster, K., Eenmaa, H., Leijen, Ä., Pedaste, M., Raju, M., Astapova, A., Lukner, H., Kogermann, K. i Sepp, T. (2021). Data sharing practices and data availability upon request differ across scientific disciplines. *Scientific Data*, 8(1), Article no. 192. <https://doi.org/10.1038/s41597-021-00981-0>

The jamovi project (2024). *jamovi (Version 2.3.28)* [Computer Software]. <https://www.jamovi.org>

Thibault, R. T., Pennington, C. R. i Munafò, M. R. (2023). Reflections on preregistration: Core criteria, badges, complementary workflows. *Journal of Trial & Error*, 4(1). <https://doi.org/10.36850/mr6>

Trippas, D., Verde, M. F. i Handley, S. J. (2014). Using forced choice to test belief bias in syllogistic reasoning. *Cognition*, 133(3), 586–600. <https://doi.org/10.1016/j.cognition.2014.08.009>

Youyou, W., Yang, Y. i Uzzi, B. (2023). A discipline-wide investigation of the replicability of Psychology papers over the past two decades. *Proceedings of the National Academy of Sciences of the United States of America*, 120(6), e2208863120. <https://doi.org/10.1073/pnas.2208863120>

The Impact of Result Significance on the Assessment of Replication Accuracy: An Experimental Study

Ivan BULJAN
Faculty of Humanities and Social Sciences,
University of Split, Split, Croatia

The study aimed to examine the impact of manipulating the order of presentation of studies with opposing conclusions on the assessment of the accuracy of research. In this registered parallel experimental study, participants from the general population were randomly given descriptions of studies in which replications of psychological research produced results that showed either an effect or no effect for interventions, and were asked to assess the accuracy of the conclusions of the described studies. The study was conducted in a virtual environment. Analyses included 194 individuals, 97 in each experimental group. The intervention in this study consisted of different research conclusions randomly assigned to participants by the testing program. The results suggest that a higher percentage of participants who were shown a replication with a significant effect assessed the replication as more accurate compared to the initial research ($n = 71$, 73.20%), than participants who were shown a replication that did not show a statistically significant effect ($n = 53$, 54.64%). The results suggest a potential bias in the order of presentation when interpreting replication studies because the accuracy assessment depended on the findings of the research. Future qualitative research should investigate factors associated with the interpretation of replications.

Keywords: reproducibility of scientific research, replication crisis, understanding of scientific research, experimental research



Međunarodna licenca / International License:
Imenovanje-Nekomercijalno / Attribution-NonCommercial