Sveučilište u Zagrebu

Filozofski fakultet

Odsjek za sociologiju

Luka Mandić

Testiranje utjecaja efekata metode na rezultate anketnih istraživanja

Zagreb, 2020. godina

Ovaj rad izrađen je na Odsjeku za sociologiju Filozofskog fakulteta Sveučilišta u Zagrebu pod vodstvom doc. dr. sc. Ksenije Klasnić i predan je na natječaj za dodjelu Rektorove nagrade u akademskoj godini 2019./2020.

Sadržaj

[Uvod 1](#_Toc48822572)

[Metode 5](#_Toc48822573)

[*Anketni upitnik* 5](#_Toc48822574)

[*Uzorak* 5](#_Toc48822575)

[*Analitički pristup* 5](#_Toc48822576)

[*Formulacija iskaza* 6](#_Toc48822577)

[*Potvrdna pristranost* 7](#_Toc48822578)

[*Nemarno odgovaranje* 7](#_Toc48822579)

[*Pristranost slaganja s iskazima* 8](#_Toc48822580)

[*Statistička analiza* 9](#_Toc48822581)

[Rezultati 10](#_Toc48822582)

[*Utjecaji formulacije iskaza i potvrdne pristranosti* 10](#_Toc48822583)

[*Utjecaji nemarnog odgovaranja* 14](#_Toc48822584)

[*Utjecaji pristranosti slaganja s iskazima* 15](#_Toc48822585)

[Rasprava i zaključci 18](#_Toc48822586)

[Sažetak 20](#_Toc48822587)

[Literatura 22](#_Toc48822588)

[Zahvale i životopis 26](#_Toc48822589)

**Popis kratica**

PIM- provjera instruktivne manipulacije (engl. *instructional manipulation check*)

NARS- neto stil odgovora na pristranost slaganja (engl. *net acquiescence response style)*

# Uvod

Jedan od glavnih ciljeva društvenih istraživanja jest osmisliti nacrt koji omogućuje dobivanje validnih rezultata, koji se potom mogu reproducirati koristeći iste metodološke pristupe na drugim uzorcima iste populacije. Ukoliko se rezultati uspješno reproduciraju ovim putem, znanstvena disciplina kumulativno napreduje, a ujedno se dobivaju i snažniji temelji za argumente o predmetu proučavanja, koje su bazirane na empirijskim rezultatima. Kako bi se validni rezultati postigli putem anketnih upitnika, istraživači većinom nastoje slijediti neke općeprihvaćene smjernice: korištenje reprezentativnog uzorka, objektivna interpretacija rezultata utemeljena na statističkim kriterijima, jasan i iscrpan prikaz metoda i rezultata istraživanja, ali i razne druge pristupe koji omogućuju postizanje validnih rezultata. Budući da je proces pribavljanja reprezentativnog uzorka često težak i skup, uzorci korišteni u istraživanjima često nisu reprezentativni. Navedeno se smatra jednim od razloga zbog kojega generalizacija na populaciju iz koje je uzorak pribavljen nije moguća, jer postoji mogućnost da rezultati prikazuju iskrivljenu sliku populacije zbog nereprezentativnosti uzorka. No, neki od fenomena koji potencijalno mogu utjecati na rezultate istraživanja nisu tako očiti, nisu nužno ni toliko spominjani, a k tome ih je teško otkriti koristeći tradicionalnije metode društvenih istraživanja. Takvi efekti često se nazivaju i efektima metode (engl. *method effect*). Za efekte metode se često misli da predstavljaju neželjenu varijancu u prikupljenim podatcima, koja je nepovezana s varijancom korištenog instrumenta, što posljedično utječe na rezultate (Maul, 2013). Generalno, naziv efekta metode može se odnositi na bilo koji utjecaj za koji se može pretpostaviti da utječe na rezultate mjerenja, tako da rezultat mjerenja ne predstavlja samo rezultat s obzirom na korišteni instrument, već i s obzirom na neki dodatni efekt. Unutar domene anketnih upitnika, neki efekti metode koji mogu utjecati na rezultate su: a) utjecaj formulacije iskaza (engl. *item wording*), b) utjecaj potvrdne pristranosti (engl. *confirmatory bias*), c) utjecaj nemarnog odgovaranja (engl. *careless responding*), d) utjecaj pristranosti slaganja (engl. *acquiescence bias*).

Sve veći broj istraživanja proučava moguće utjecaje koje formulacija iskaza ima na rezultate. Bez obzira na moguće utjecaje, skale korištene u raznim područjima društvenih znanosti često se sastoje od podjednakog broja pozitivnih (engl. *regular* *item*) i negativnih (engl. *reversed item*) iskaza. Negativni iskazi su oni iskazi koje je potrebno rekodirati, odnosno prije obrade podataka učiniti takozvanu refleksiju skale, kako bi svi iskazi uključeni u skalu imali isti smjer u odnosu na mjereni konstrukt. Na primjer, iskazi poput „*Unosim živost u neku zabavu*“ i „*Započinjem razgovore*“, trebali bi se smatrati pozitivnim iskazima, a iskazi poput „*Držim se u pozadini*“ i „*Ne pričam puno*“ trebali bi se smatrati negativnim iskazima, ukoliko se mjeri konstrukt *ekstraverzije*. Čak i ako iskaz u sebi sadrži negaciju, on se i dalje može smatrati pozitivnim iskazom, jednako kao što se i iskazi bez negacije mogu smatrati negativnim iskazima. Na primjer, iskaz „*Ne smeta mi kada sam centar pozornosti*“ započinje negacijom, ali se i dalje može smatrati pozitivnim izrazom, budući da ga nije potrebno rekodirati prilikom mjerenja *ekstraverzije*.

Korištenje mjernih skala koje su sastavljene od istog broja pozitivnih i negativnih iskaza, često nazivane uravnoteženim skalama, prvobitno je uvedeno s ciljem smanjivanja pristranosti stila odgovaranja (engl. *response style bias*) prilikom ispunjavanja upitnika (Nunally, 1978; Paulhus, 1991). Pristranost stila odgovaranja često se definira kao svaka individualna tendencija koja navodi ispitanike na odgovor neovisan o iskazu, čime se posljedično iskrivljuju rezultati (Cronbach, 1946). Smatralo se da spomenuta pristranost manje utječe na uravnotežene skale, nego na skale sastavljene od samo pozitivnih ili samo negativnih iskaza. Iako recentna istraživanja pronalaze sve veći broj potencijalnih mana korištenja negativnih iskaza, njihovo korištenje u mjernim skalama se i dalje preporuča, jer se smatra kako unaprjeđuju validnost konstrukta koji se mjeri, jer daju širu sliku uvjerenja na kojima se zasniva odgovor sudionika (Weijters i Baumgartner, 2012). Bez obzira na ovu moguću prednost, istraživanja pokazuju kako uravnotežene skale, u usporedbi sa skalama sastavljenim od samo pozitivnih ili samo negativnih iskaza, često imaju slabiju diskriminacijsku snagu te rezultiraju modelima koji slabije pristaju podatcima (Suárez-Alvarez et al., 2018; Marsh, 1986). Generalno, za uravnotežene skale se može tvrditi da, u usporedbi sa skalama koje su sačinjene od samo pozitivnih ili samo negativnih iskaza, češće dovode do problema u analizi, poput odbijanja jednodimenzionalnih modela u korist višedimenzionalne solucije, zbog pojavljivanja artificijelnog faktora koji je stvoren zbog utjecaja formulacije iskaza (Suárez-Alvarez et al., 2018; Marsh, 1986). To znači da bi uravnotežena skala mogla rezultirati naizgled dvofaktorskom strukturom, koja je dvofaktorska isključivo zbog toga što se sastoji i od pozitivnih i od negativnih iskaza. Formulacija iskaza samo je jedan od mnogih efekata metode koji mogu utjecati na rezultate i iskriviti ih, a koje je često teško detektirati unutar istraživanja, što dovodi do toga da ti efekti potencijalno mogu biti zanemareni prilikom interpretacije rezultata.

Potvrdna pristranost (engl. *confirmatory bias*) je efekt metode koji ima mogućnost utjecaja na rezultate, a čiji smo utjecaj testirali u ovoj studiji. Kada sudionici odgovaraju na iskaz, aktiviraju se uvjerenja koja su u skladu s načinom na koji je iskaz postavljen (Kunda, Fong, Santioso, i Reber, 1993). Na primjer, ukoliko se sudionike pita jesu li sretni, veća je šansa da će se aktivirati uvjerenja i sjećanja o situacijama koje su u skladu sa srećom, što može utjecati na odgovore na iskaze koji slijede. Također, ukoliko se sudionike pita jesu li tužni, veća je šansa da će se aktivirati uvjerenja i sjećanja o situacijama koje su u skladu s tugom. Neka provedena istraživanja, u kojima su sudionicima nasumično dodijeljene dvije inačice anketnog upitnika koje se međusobno razlikuju s obzirom na tip iskaza (slično kao u prethodnoj rečenici), pokazala su kako su odgovori pristrani s obzirom na formulaciju iskaza (Kunda et al., 1993; McClendon, 1991; Schuman i Presser, 1981). Dakle, kada se sudionike pitalo jesu li ekstroverti, u prosjeku su rezultati tendirali prema ekstrovertiranosti, a kada ih se pitalo jesu li introverti, rezultati su tendirali prema introvertiranosti.

Nemarno odgovaranje (engl. *careless responding*) je termin koji je dominantno korišten za opisivanje tipa odgovaranja u kojemu sudionici ne obraćaju dovoljno pozornosti na sadržaj iskaza na koji odgovaraju (Schmitt i Stults, 1985; Woods, 2006). Neki istraživači to pripisuju nedostatku pažnje, motivacije, ili tendenciji sudionika da stvore očekivanja o iskazima koji slijede na temelju prethodnih iskaza (Weijters et al., 2013). Bez obzira na uzrok, takav pristup ispunjavanja anketnih upitnika može dovesti do kontradiktornih odgovora i nemarnog odgovaranja, što posljedično može iskriviti rezultate istraživanja. Neke simulacijske studije pokazale su da ako samo 10% sudionika nemarno odgovara na pitanja u anketnom upitniku, rezultati ukazuju na mogućnost pojave dodatnog artificijelnog faktora (Schmitt i Stults, 1985; Woods, 2006). U takvom slučaju, jedan faktor će sačinjavati pozitivni iskazi, a drugi faktor negativni iskazi. Zapravo, ovakvi rezultati mogu navesti istraživače na odbacivanje jednofaktorskog modela nekih konstrukata, jer je nemarno odgovaranje teško za detektirati koristeći tradicionalnije statističke pristupe poput eksploracijske faktorske analize. Postoji osjetna diskrepancija u načinu na koji se nemarno odgovaranje operacionalizira, što neporecivo utječe na izvještene rezultate o tom fenomenu (Kam i Meyer, 2015). Jedan od najkorištenijih pristupa istraživanja nemarnog odgovaranja jest provjera instruktivne manipulacije (PIM, engl. *instructional manipulation check*), koji predstavlja specifičan tip iskaza koji sadrži direktnu instrukciju postavljenu od strane istraživača, te se pomoću nje utvrđuje jesu li sudionici odgovarali nemarno prilikom ispunjavanja anketnog upitnika. Na primjer, takav iskaz može biti formuliran na sljedeći način: „*Odaberite kako se u potpunosti slažete s ovom tvrdnjom*“. Iskaz može biti formuliran i na sljedeći način: „*Ispunite prazan kvadratić na gornjem desnom kutu ove stranice*“. Sudionici koji ne slijede ove upute klasificiraju se nemarnima. Sudionici istraživanja anketni upitnik mogu ispunjavati pažljivo na početku, da bi nakon nekog vremena počeli odgovarati nemarno (posebice ako je upitnik podulji). Upravo iz tog razloga, potrebno je napomenuti kako se iskazi koji služe za provjeravanje nemarnog odgovaranja moraju postaviti prije ključne skale od interesa, kako bi se moglo odrediti u kojoj su mjeri njezini rezultati pod utjecajem nemarnog odgovaranja.

Pristranost slaganja s iskazima bez obzira na njihov sadržaj (engl. *acquiescence bias*), često se opisuje kao preferencija prema pozitivnoj strani skale, jer upravo opisuje nemarno slaganje s iskazima koje zanemaruje značenje samog iskaza (Bentler, Jackson, i Messick, 1971; Weijters, Baumgartner, i Schillewaert, 2013). Do pristranosti slaganja može doći prilikom nemarnog odgovaranja od strane sudionika, koji se nedovoljno angažiraju oko sadržaja iskaza te posljedično daju nedovoljno promišljen odgovor (Knowles i Condon, 1999). Smatra se kako pristranost slaganja iskrivljuje korelacije među konstruktima (Bentler et al., 1971). Preciznije, ono može rezultirati višom korelacijom među konstruktima istoga smjera, a smanjiti korelacije među konstruktima suprotnog smjera (Kam i Mayer, 2015). Informacija koja bi trebala posebno interesirati društvene znanstvenike jest to da su skale Likertova tipa, to jest najčešće korištene skale u društvenim istraživanjima, izrazito sklone utjecaju pristranosti slaganja (McClendon, 1991). Upravo iz tog razloga smatramo važnim istražiti potencijalne utjecaje koje ono može imati na rezultate. Generalno, na pristranost slaganja se gleda kao na osobnu karakteristiku sudionika istraživanja, koja je odvojena i većinom van utjecaja konstrukta mjerenja (Weijters, Geuens, i Schillewaert, 2010). Pristranost slaganja se može mjeriti na razne načine, a jedan od učestalijih načina temelji se na stupnju u kojemu se sudionici slažu s iskazima skale sačinjene od heterogenih iskaza, s pretpostavkom da iskazi nisu sadržajno slični. Ovakva metoda, koja se naziva neto stil odgovora na pristranost slaganja (NARS, engl. *net acquiescence response style*), uklapa pristranost slaganja u jedinstveni rezultat, koji je predstavljen aritmetičkom sredinom koju sudionik ostvari odgovorom na sve heterogene iskaze skale (npr., Weijters et al., 2013).

Cilj ovog istraživanja je pokazati moguće utjecaje opisanih efekata metode: formulacije iskaza, potvrdne pristranosti, nemarnog odgovaranja, te pristranosti slaganja s iskazima. Navedeni efekti metode imaju potencijal navesti istraživače na krive zaključke, a s obzirom na to da ih je teško detektirati korištenjem tradicionalnih metoda, validnost rezultata osjetnog broja istraživanja može se ozbiljno dovesti u pitanje. Svrha istraživanja je proširiti metodološke spoznaje o načinima mjerenja u društvenim znanostima, kako bi se opisani efekti metode mogli bolje kontrolirati u budućim istraživanjima.

# Metode

## *Anketni upitnik*

Podatci su prikupljani putem dvije inačice online anketnog upitnika. Poveznica za pristup anketnom upitniku distribuirana je putem raznih društvenih mreža. Anketni upitnik napisan je na hrvatskom jeziku te su se podatci prikupljali tijekom ožujka 2020. Provedbu istraživanja odobrilo je nadležno etičko povjerenstvo Odsjeka za sociologiju, Filozofskog fakulteta Sveučilišta u Zagrebu. Prvo pitanje u anketnom upitniku služilo je kao filter varijabla pomoću koje je sudionicima dodijeljena jedna od inačica anketnog upitnika. Sudionici kojima je rođendan u parnom mjesecu dodijeljena je jedna inačica, a onima kojima je rođendan u neparnom mjesecu druga.

## *Uzorak*

Istraživanje je provedeno na prigodnom uzorku korisnika društvenih mreža u Hrvatskoj. Uzorak sudionika koji su rođeni na parni mjesec sastojao je od N= 391 sudionika, a uzorak onih koji su rođeni na neparni mjesec od N= 400 sudionika. 82,3% od svih sudionika odgovorilo je na sva pitanja u anketnom upitniku. 97% od svih sudionika nije imalo više od 3,4% praznih odgovora. Iz analize su isključeni sudionici kojima je nedostajalo više od 10% odgovora. Nisu pronađene statističke značajne razlike među uzorcima s obzirom na varijable *spol*, *dob*, i *obrazovanje*. Prosječna dob svih sudionika bila je 30 godina [Min= 15; Max= 83; SD= 12,9). 71,5% sudionika bili su pripadnici muškog spola. 45,3% sudionika ostvarilo je neki oblik visokog obrazovanja, gdje visoko obrazovanje označava ono obrazovanje koje je iznad sekundarnog nivo (srednja stručna sprema). Ovaj postotak uključuje i studente preddiplomskog studija koji još nisu završili svoje obrazovanje.

## *Analitički pristup*

Odlučili smo testirati utjecaj svih efekata metode uključenih u anketni upitnik koristeći Mini-IPIP skale ličnosti (Donellan, Oswald, Baird, i Lucas, 2006). Navedene skale odabrane su zbog velikog broja empirijskih istraživanja u području ličnosti, ali i zbog njihove praktičnosti za upotrebu. Naime, skale su kratke te se svaka od pet skala sastoji od samo četiri iskaza. Skale smo adaptirali prema potrebama istraživanja, a adaptacije će biti opisane zajedno s opisom metoda. Svi iskazi bodovali su se na skali procjene slaganja od 1 do 5, kako bismo osigurali usporedive varijance.

## *Formulacija iskaza*

Kako bismo testirali utjecaje formulacije iskaza odnosno utjecaj promjene pozitivnog iskaza u negativni (ili obrnuto), izabrali smo dvije facete iz Mini-IPIP-a, *savjesnost* i *ugodnost*. Navedene facete izabrane su iz razloga što mislimo kako preformuliranje njihovih iskaza rezultira iskazima koji zvuče prirodnije, nego da se ista stvar napravi s nekom od preostalih faceta. Na primjer, iskaz „*Empatičan sam*“ u drugoj će inačici postati „*Nisam empatičan*“. Tako se u našoj studiji jedna inačica skala *savjesnosti* i *ugodnosti* sastojala od četiri pozitivna iskaza po skali, a druga od četiri negativna iskaza po skali. Testirali smo razlike u aritmetičkim sredinama nakon što smo sve iskaze rekodirali u istom smjeru, budući da predviđamo da će formulacija iskaza imati statistički značajan utjecaj na aritmetičke sredine. Koristeći konfirmacijsku faktorsku analizu (CFA), testirali smo odgovara li jednodimenzionalni model pojedine skale (*savjesnost*/ *ugodnost*) podatcima. Navedeno smo učinili zasebno za obje inačice skale, kako bismo vidjeli kakav utjecaj na pristajanje podatcima ima formulacija iskaza. Nadalje, proveli smo testove invarijantnosti među dvjema inačicama istog konstrukta, u cilju dobivanja dodatnih saznanja o utjecaju formulacije iskaza. Osim skala ličnosti, uključili smo dvije skale koje se odnose na domenu rodne nejednakosti, koje u sebi uključuju jedan identičan iskaz (Inglehart i Norris, 2003; Tougas, Brown, Beaton, i Jolly, 1995). Jedna skala rodne nejednakosti u anketnom je upitniku bila smještena prije svih skala ličnosti, dok je druga smještena nakon njih. Odabrali smo iskaz koji je identičan u obje skale kako bismo imali indikator koji direktno testira utjecaje formulacije iskaza. Odabrani iskaz u prvoj je skali koja se nalazila prije skala ličnosti glasio „*Žena mora imati dijete kako bi se osjećala ispunjeno*“, dok je u drugoj skali koja se nalazila nakon skala ličnosti glasio „*Žena ne mora imati dijete kako bi se osjećala ispunjeno*“. Ukoliko se jedan iskaz može smatrati polarnom suprotnošću drugoga, te ukoliko ih sudionici istraživanja tako dožive, korelacija među njima trebala bi se približiti iznosu -1. Kako bismo osigurali dodatnu validnost glede ovog indikatora, u anketnom upitniku nije bio mogući povratak na prethodnu stranicu, kako ispitanici ne bi namještali, to jest usklađivali svoje odgovore.

## *Potvrdna pristranost*

Utjecaj potvrdne pristranosti testirali smo na isti način kao i utjecaj formulacije iskaza. Ovaj put smo izdvojili druge dvije facete ličnost: *emocionalnu stabilnost* i *ekstraverziju*, budući da su sastavljene od dva pozitivna (p) iskaza i dva negativna (n) iskaza u svojem originalnom formatu. U našoj studiji, jedna inačica anketnog upitnika sadržavala je skalu koja je počinjala s pozitivnim iskazom (p1, p2, n3, n4), a druga inačica skalu koja je započinjala s negativnim iskazom (n3, n4, p1, p2). Iskazi p1 i p2 predstavljaju pozitivno formulirane iskaze, a iskazi n3 i n4 negativno formulirane iskaze. Osim promjene redoslijeda, iskazi iz jedne inačice bili su identični iskazima u drugoj inačici anketnog upitnika. Rezultati svih iskaza u obradi su podataka rekodirani tako da budu u istom smjeru s konstruktom mjerenja. Za testiranje utjecaja potvrdne pristranosti, usporedili smo aritmetičke sredine istih iskaza među dvjema inačicama (npr. p1 u obje inačice). Nadalje, testirali smo kako jednofaktorski modeli navedenih faceta ličnosti pristaju podatcima zasebno za svaku skalu i svaki uzorak. Za kraj, proveli smo i testove invarijantnosti između pojedinih skala ličnosti korištenih u različitim inačicama anketnog upitnika (npr. test invarijantnosti *ekstraverzije* u dvije inačice anketnog upitnika).

## *Nemarno odgovaranje*

Kako bismo mogli mjeriti utjecaj nemarnog odgovaranja, uključili smo poseban tip iskaza (PIM) koji je postavljen prije svih skala ličnosti. S obzirom na relativno kratku duljinu anketnog upitnika, nismo se odlučili za više od jednog takvog iskaza, budući da u tom slučaju takvi iskazi mogli postati očiti ili bi mogli iritirati sudionike. Iskaz je izričito i jasno navodio sudionike da izaberu određenu opciju na ponuđenoj skali procjene: „*Izaberite kako se uopće ne slažete s ovom tvrdnjom*“. Sudionike koji nisu slijedili uputu klasificirali smo nemarnima u analizama koje su uslijedile. Testirali smo utjecaj nemarnog odgovaranja tako da smo uspoređivali rezultate nemarnih i pažljivih sudionika na sljedeće načine: usporedbe korelacije međusobno sadržajno obrnutih iskaza iz skala rodne nejednakosti (jedan pozitivan iskaz i jedan negativan iskaz), usporedbe korelacija skala ličnosti, te usporedba aritmetičkih sredina na indikatorima pristranosti slaganja s iskazima.

## *Pristranost slaganja s iskazima*

Kako bismo efikasno uključili pristranost slaganja s iskazima u model, specificiran je eksplicitni indeks pristranosti slaganja s iskazima bez obzira na njihov sadržaj odnosno NARS (engl. *net acquiescence response style*). Rezultat pojedinog sudionika na tom indeksu je aritmetička sredina svih odgovora na skali od 17 heterogenih iskaza korištenih u anketnom upitniku, za koje smo pretpostavili kako ne dijele zajednički sadržaj. Skala je blago modificirana verzija one koju je Greenleaf (1992) koristio kako bi mjerio ekstremni stil odgovaranja, te ju je sastavio s ciljem da iskazi budu heterogeni i da ne dijele zajednički sadržaj. Pristranost slaganja smo testirali koristeći svih pet faceta Mini-IPIP-a: *ugodnost*, *savjesnost*, *emocionalnu stabilnost*, *ekstraverziju*, te *intelekt*. Kao prvi korak provjerili smo pristajanje jednofaktorskih modela podatcima zasebno za svaku facetu svakog uzorka. Nakon toga, dodali smo pristranost slaganja u model te smo usporedili rezultate s modelima bez pristranosti slaganja. Kako bismo uspješno implementirali pristranost slaganja u model, dizajnirali smo dvofaktorski model koji u sebi uključuje konstrukt mjerenja ličnosti, ali i latentan faktor pristranosti slaganja. Kreirani model podsjeća na model koji su koristili Kam i Meyer (2015), s razlikom da naš model sadrži samo jedan sadržajni faktor (jedna faceta ličnosti). Nadalje, koristili smo četiri parcele iskaza kao indikatore pristranosti slaganja, kako bismo izbjegli procjenu razmjerno velikog broja dodatnih parametara (npr. Kam i Meyer, 2015; Weijters et al., 2010). Parcele su kreirane uprosječivanjem rezultata na indikatorima pristranosti slaganja. Postavili smo nultu korelaciju između sadržajnog faktora ličnosti i latentnog faktora pristranosti slaganja, te je jedna faktorska saturacija na sadržajnom faktoru ličnosti postavljena na 1, kako bismo definirali njegovu varijancu. Sve faktorske saturacije na faktoru pristranosti slaganja postavljene su na 1 te se zbog toga ovaj model može smatrati preprečnim faktorskim modelom (engl. *intercept factor model*) (Maydeu-Olivares i Coffman, 2006). Svi iskazi rekodirani su tako da budu u istom smjeru sa sadržajnim konstruktom mjerenja. Predloženi model prikazan je na slici 1.

Skala koja mjeri *intelekt* ostala je neizmijenjena u obje inačice anketnog upitnika te nam je služila kao kontrolni instrument, pomoću kojeg se jasnije vide utjecaji spomenutih efekata metode.



**Slika 1**. Grafički prikaz predloženog modela koji uključuje latentan faktor pristranosti slaganja. η1= sadržajni konstrukt jedne od skala ličnosti; η2= latentni faktor pristranosti slaganja; a1-a4= indikatori pristranosti slaganja stvoreni parcijalizacijom 17 iskaza koji mjere pristranost slaganja; y1-y4= indikatori sadržajnog konstrukta ličnosti; η1 i η2 su u nultoj korelaciji.

## *Statistička analiza*

Analize su provedene koristeći R sustav za statističko računanje 3.6.2. (R Development Core Team, 2019). Svi univarijatni testovi provedeni su koristeći dodatni statistički paket *psych*, koji je razvio Revelle (2019). Svi strukturalni modeli i svi testovi metrijske invarijantnosti provedeni su koristeći dodatni statistički paket *lavaan*, koji je razvio Rosseel (2012).

# Rezultati

## *Utjecaji formulacije iskaza i potvrdne pristranosti*

Testirali smo utjecaje formulacije iskaza i potvrdne pristranosti putem istih metoda. Skale *ugodnosti* i *savjesnosti* korištene su za testiranje utjecaja formulacije iskaza, a skale *emocionalne stabilnosti* i *ekstraverzije* za testiranje utjecaja potvrdne pristranosti. Prvo smo usporedili aritmetičke sredine iskaza koji su bili ili međusobne polarne suprotnosti (pozitivan iskaz u prvoj inačici postao je negativni iskaz u drugoj inačici) ili su bili drugačije raspoređeni u skali mjerenja (pozitivni iskaz koji je na prvom mjestu u prvoj inačici skale postaje pozitivni iskaz koji je na trećem mjestu druge inačice skale). Očekivali smo kako će formulacija iskaza i potvrdna pristranost imati statistički značajan efekt na aritmetičke sredine iskaza. Rezultati su prikazani u tablici 1. Jedina skala koja nije bila izmijenjena u odnosu na original je skala koja mjeri *intelekt*. Navedena skala pokazala je invarijantnost u vidu nepostojanja statistički značajnih razlika u aritmetičkim sredinama pojedinog iskaza u prvoj i drugoj inačici. Pronađena invarijantnost potvrđuje validnost sljedećih rezultata, jer je to jedina skala koja je u potpunosti identična u obje inačice anketnog upitnika. Sedam od osam parova iskaza koji su bili međusobno sadržajno obrnuti (pozitivni iskaz u negativni iskaz ili obrnuto) pokazuju statistički značajnu razliku u prosjecima. Ovo je jasan pokazatelj utjecaja koju formulacija iskaza ima na rezultate. Predlažemo istraživačima poseban oprez prilikom razmišljanja o preformulaciji iskaza unutar skale koju planiraju koristiti, budući da obrtanja mogu utjecati na rezultate i validnost skale. Utjecaj pozicijskih promjena iskaza odnosno promjene tipa prvog iskaza može posljedično utjecati na rezultate zbog prisutnosti potvrdne pristranosti (Kunda et al., 1993). Naši rezultati pokazuju kako je među šest od ukupno osam parova iskaza, promjena pozicije uzrokovala statistički značajne razlike među njihovim aritmetičkim sredinama, potvrđujući utjecaj potvrdne pristranosti. Određen utjecaj potvrdne pristranosti neizbježan je korištenjem skala Likertova tipa te je spomenuti utjecaj teško detektirati koristeći tradicionalne metode, pogotovo ukoliko ne postoji više inačica anketnog upitnika u kojima je promijenjen redoslijed iskaza. Čak i sofisticiranije metode koje uključuju potvrdnu pristranost u linearne strukturalne modele koriste više anketnih inačica (npr., Weijters et al., 2013).

**Tablica 1**. Utjecaj formulacije iskaza i potvrdne pristranosti na aritmetičke sredine

|  |
| --- |
| UTJECAJ FORMULACIJE ISKAZA |
|  | Uzorak parnih mjeseci (N = 400) | Način promjene iskaza iz uzorka parnih mjeseci u uzorak neparnih mjeseci | Uzorak neparnih mjeseci (N = 391) | p < ,05 |
|  |  |  |  |  |
| *Savjesnost, iskaz 1* | 3,01 | Negativni u pozitivni | 3,22 | **\*** |
| *Savjesnost, iskaz 2* | 3,42 | Negativni u pozitivni | 3,76 | **\*** |
| *Savjesnost, iskaz 3* | 3,80 | Negativni u pozitivni | 3,65 | **\*** |
| *Savjesnost, iskaz 4* | 4,02 | Negativni u pozitivni | 4,02 | - |
| *Ugodnost, iskaz 1* | 3,82 | Pozitivni u negativni | 4,13 | **\*** |
| *Ugodnost, iskaz 2* | 4,07 | Pozitivni u negativni | 4,28 | **\*** |
| *Ugodnost, iskaz 3* | 3,37 | Pozitivni u negativni | 3,95 | **\*** |
| *Ugodnost, iskaz 4* | 2,98 | Pozitivni u negativni | 3,69 | **\*** |
|  |  |  |  |  |
| UTJECAJ POTVRDNE PRISTRANOSTI |
|  | Uzorak parnih mjeseci (N= 400) | Pozicijska promjena iz uzorka parnih mjeseci u uzorak neparnih mjeseci | Uzorak neparnih mjeseci (N= 391) | p < ,05 |
|  |  |  |  |  |
| *Em. stabilnost, iskaz 1* | 3,36 | 1. u 3. | 3,03 | **\*** |
| *Em. stabilnost, iskaz 2* | 3,04 | 2. u 4. | 2,86 | **\*** |
| *Em. stabilnost, iskaz 3* | 3,08 | 3. u 1. | 3,36 | **\*** |
| *Em. stabilnost, iskaz 4* | 2,82 | 4. u 2. | 3,07 | **\*** |
| *Ekstraverzija, iskaz 1* | 3,26 | 1. u 3. | 2,54 | **\*** |
| *Ekstraverzija, iskaz 2* | 3,12 | 2. u 4. | 3,17 | - |
| *Ekstraverzija, iskaz 3* | 2,54 | 3. u 1. | 3,27 | **\*** |
| *Ekstraverzija, iskaz 4* | 3,19 | 4. u 2. | 3,32 | - |
|  |  |  |  |  |
| NEIZMJENJENA SKALA |
|  | Uzorak parnih mjeseci (N= 400) | Promjena između uzorka parnih mjeseci I uzorak neparnih mjeseci | Uzorak neparnih mjeseci (N= 391) | p < ,05 |
|  |  |  |  |  |
| *Intelekt, iskaz 1* | 3,66 | Nikakva | 3,73 | - |
| *Intelekt, iskaz 2* | 3,51 | Nikakva | 3,53 | - |
| *Intelekt, iskaz 3* | 3,69 | Nikakva | 3,74 | - |
| *Intelekt, iskaz 4* | 3,64 | Nikakva | 3,70 | - |
| Napomene: \* označava da je razlika aritmetičkih sredina dvaju iskaza statistički značajna uz p < ,05, a – označava da nije. Svi iskazi su rekodirani tako da budu u istom smjeru u odnosu na konstrukt mjerenja |

Kako bismo dobili dodatan uvid u utjecaje formulacije iskaza i potvrdne pristranosti na rezultate prikupljene anketnim upitnicima, proveli smo testove metrijske invarijantnosti među istim konstruktima. Ovo je donekle problematično, budući da neki jednofaktorski modeli konstrukata ličnosti koji su provedeni zasebno za svaki uzorak nisu pokazali zadovoljavajuće pristajanje podatcima (tablica 4). No, bez obzira na taj nedostatak, smatramo kako testovi metrijske invarijantnosti daju dodatne korisne uvide. Kako bismo proveli spomenute testove, bilo je potrebno spojiti sudionike iz uzorka parnih datuma i uzorka neparnih datuma, a da pritom zadržimo filter varijablu (1= uzorak parnih mjeseci, 2= uzorak neparnih mjeseci), to jest informaciju o inačici anketnog upitnika.

Prvo smo specificirali jednofaktorske modele za svih pet konstrukata ličnosti, na kojima su potom bili provedeni testovi invarijantnosti. Svaki model sastojao se od četiri indikatora njemu svojstvene skale. Rezultati testova multivarijatnog normaliteta pokazuju kako varijable korištene u modelima nisu bile normalno distribuirane. Zbog asimetrične prirode uključenih varijabli, ali i zbog prisutnosti određene količine nedostajućih vrijednosti, koristili smo procjenu maksimalne vjerojatnosti (engl. *maximum likelihood estimator*) s robusnim standardnim pogreškama (MLR; White, 1980) i skaliranim statistikom koji je asimptotičko jednak Yuan-Bentlerovom (2007) testovnom statistiku. Analize su provedene koristeći R sustav za statističko računanje 3.6.2. (R Development Core Team, 2019). Svi strukturalni modeli i svi testovi metrijske invarijantnosti provedeni su koristeći dodatni statistički paket *lavaan*, koji je razvio Rosseel (2012).

Prilikom specifikacije modela za skalu *ugodnosti* (skala u kojoj su iskazi izmijenjeni iz pozitivnih u negativne), dodatni statistički paket lavaan (Rosseel, 2012) je izvijestio o pojavljivanju negativnih varijanci. S obzirom da dobivanje negativnih varijanci generalno ukazuje na krivu specifikaciju modela (Browne, 2015), prisiljeni smo zaključiti kako je formulacija iskaza imala toliko jak utjecaj, da osnovni model nije uspio konvergirati (za usporedbu, svi ostali modeli provedeni na zasebnim uzorcima su bez problema konvergirali). Zbog pojave negativnih varijanci unutar specificiranog modela *ugodnosti*, odlučili smo ne provoditi testove metrijske invarijantnosti na toj skali. Tri od preostala četiri modela pokazali su nedovoljno dobro pristajanje podatcima. To ne čudi s obzirom na to da su to skale koje su korištene i za testiranje utjecaja formulacije iskaza i za testiranje utjecaja potvrdne pristranosti (*savjesnost*, *emocionalna stabilnost*, *ekstraverzija*), te da smo za navedene utjecaje pretpostavili postojanje plauzibilnog efekta na rezultate. Skala *intelekta*, koja je jedina ostala ista u obje inačice anketnog upitnika, pokazala je dobro pristajanje podatcima [χ 2= 3,88, df= 2, p= ,14, CFI=,995, TLI=,986, RMSEA= ,038, 90% CI= ,000- ,094, SRMR= ,019], što dodatno ide u prilog validnosti rezultata. Bez obzira na izmjene među inačicama, neke skale su pokazale određene razine metrijske invarijantnosti. Jedina skala za koju tvrdnja o konfiguralnoj metrijskoj invarijantnosti ne može biti legitimno opravdana jest skala *emocionalne stabilnosti*. Navedeno se može opisati kao rezultat utjecaja potvrdne pristranosti. Problem je u tome, što druga skala koja je korištena za testiranje potvrdne pristranosti, skala *ekstraverzije*, pokazuje konfiguralnu metrijsku invarijantnost. Rezultati postaju još zanimljiviji kada se usporede razine metrijske invarijantnosti između skale *ekstraverziju*, koja je korištena za mjerenje potvrdne pristranosti, i (neizmijenjene) skale *intelekta*, jer obje skale pokazuju istu razinu metrijske invarijantnosti. Iste rezultate pokazuje i skala *savjesnosti*. Naime, skala koja je korištena za testiranje formulacije iskaza rezultirala je istom razinom metrijske invarijantnosti kao i skala *intelekta*. Potrebno je spomenuti kako je skala *intelekta* koja uključuje oba uzorka predstavljala jedini model koji je dobro pristajao podatcima, te da je nezadovoljavajuće pristajanje drugih modela moglo utjecati na rezultate u multivarijatnom prostoru.

Stoga, možemo tvrditi kako formulacija iskaza i potvrdne pristranosti definitivno ima utjecaja na univarijatnoj razini glede aritmetičkih sredina rezultata pojedinih iskaza, na multivarijatnoj razini glede pristajanja modela podatcima, ali ne nužno i na multivarijatnoj razini glede metrijske invarijantnosti. Rezultati metrijske invarijantnosti prikazani su u tablici 2.

**Tablica 2.** Testovi metrijske invarijantnosti

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  | χ2 (df) | RMSEA [90% CI] | CFI | AIC | BIC | Δχ2 (*p*) |
| *EM. STABILNOST* |
| Konfiguralna invarijantnost | 54,89 (4) | ,199 [,154- ,247] | ,914 | 8698,40 | 8810,56 |  |
|  |  |  |  |  |  |  |
| *EKSTRAVERZIJA* |
| Konfiguralna invarijantnost | 16,06 (4) | ,102 [,053- ,156] | ,987 | 7938,85 | 8051,01 |  |
| Metrijska invarijantnosti | 73,47 (7) | ,169 [,135- ,205] | ,938 | 7998,82 | 8096,96 | 57,41 (,000) |
|  |  |  |  |  |  |  |
| *SAVJESNOST* |
| Konfiguralna invarijantnost | 12,15 (4) | ,078 [,030- ,130] | ,976 | 8058,83 | 8170,99 |  |
| Metrijska invarijantnosti | 16,28 (7) | ,062 [,022- ,102] | .974 | 8057,27 | 8155,41 | 4,13 (,000) |
|  |  |  |  |  |  |  |
| *INTELEKT* |
| Konfiguralna invarijantnost | 13,57 (4) | ,086 [,038- ,138] | ,969 | 8286,06 | 8398,22 |  |
| Metrijska invarijantnosti | 73,47 (7) | ,169 [,135- ,205] | ,938 | 7998,82 | 8096,96 | 5,9 (,000) |
| AIC= Akaike information criterion; BIC= Bayesian information criterion Napomene: metrijska invarijantnost označava fiksirane faktorske saturacije; H0= model konfiguralne invarijantnosti nije unaprjeđenje u odnosu na model metrijske invarijantnosti. Ako je p < ,05, model metrijske invarijantnosti ne unaprjeđuje pristajanje podatcima; ukoliko model ne pokazuju barem približno zadovoljavajuće pristajanje podatcima na konfiguralnoj razini, ne provodi se testiranje metrijske invarijantnosti  |

## *Utjecaji nemarnog odgovaranja*

74,5% sudionika iz oba uzorka uspješno je slijedilo uputu posebnog iskaza koji je služio za razlikovanje nemarnih i pažljivih sudionika, „*Izaberite kako se uopće ne slažete s ovom tvrdnjom*“. Pretpostavili smo kako će rezultati sudionika koji su klasificirani kao nemarni utjecati na korelacije konstrukata, što bi trebalo biti uočljivo prilikom usporedbe korelacija između nemarnih i pažljivih sudionika. Kako bismo to testirali, prvo smo usporedili korelacije između dva polarno suprotna iskaza : „*Ženi je potrebno dijete da se osjeća ispunjeno*“ i „*Ženi nije potrebno dijete da se osjeća ispunjeno*“, među nemarnim i među pažljivim sudionicima.

Proveli smo Williamsov test koji testira invarijantnost korelacija, a rezultati pokazuju kako se razlike u korelacijama rezultata nemarnih i pažljivih sudionika na dva navedena iskaza nisu pokazale statistički značajnima niti na jednoj razini analize: u dva odvojena poduzorka, kao ni u spojenom uzorku [*Uzorak parnih mjeseci*: rnemarni= -,644; rpažljivi= -,616; *Uzorak neparnih mjeseci*: rnemarni= -,569; rpažljivi= -,641; *Spojeni uzorak*: rnemarni= -,599; rpažljivi= -,632]. Značajnost korelacijskih razlika testirala se na razini p< ,05. Dodatno zabrinjava to što korelaciji nisu bile ni blizu -1, što ozbiljno testira diskriminacijsku snagu posebnog iskaza korištenog za razlikovanje nemarnih od pažljivih sudionika. Da je klasifikacija s obzirom na posebni iskaz bila validna, korelacije pažljivih sudionika trebale su se približiti -1, te su trebale biti pronađene statistički značajne razlike među korelacijama s obzirom na nemarne i pažljive sudionike.

Nadalje, testirali smo razlike aritmetičkih sredina nemarnih i pažljivih sudionika u rezultatima na NARS indeksu, za koje smo pretpostavili da će biti statistički značajne. Umjesto da testiramo samo razlike u aritmetičkim sredinama samo na ukupnom NARS indeksu, odlučili smo ih dodatno testirati za svaki indikator koji sačinjava taj indeks. Dakle, testirali smo ih između 17 indikatora zasebno, u svakom uzorku odvojeno. Od 34 ukupna indikatora NARS-a (17 indikatora po uzorku), t-testovi su ukazali na statistički značajne razlike u aritmetičkim sredinama samo kod dva indikatora. Nadalje, nije pronađena statistički značajna razlika niti u ukupnim prosječnim rezultatima nemarnih i pažljivih sudionika na NARS indeksu, što dodatno dovodi validnost posebnog iskaza u pitanje. Također, testirali smo imaju li nemarni sudionici drugačije rezultate u korelacijama među skalama ličnosti u odnosu na pažljive sudionike. Ovo je napravljeno zasebno na svakom uzorku, jer bi u slučaju spajanja uzoraka utjecaje na rezultate dodatno mogli imati i formulacije iskaza i potvrdna pristranost. Od 20 ukupnih korelacija između skala ličnosti, statistički značajne razlike među korelacijama rezultata nemarnih i pažljivih sudionika pronađene su samo kod tri koeficijenta korelacije. Razlike korelacija testirane su putem Williamsovog testa, na razini značajnosti p< ,05. Rezultati su prikazani u tablici 3.

Temeljem navedenih rezultata, zaključujemo da posebni iskaz čija je svrha bila razlikovanje nemarnih od pažljivih sudionika, nedovoljno dobro razlikuje nemarne od pažljivih sudionika. Relativno niske korelacije (s obzirom na kontekst) polarno obrnutih iskaza u intervalu [-,599 do -,644] sugeriraju da je nemarno odgovaranje prisutno u istraživanju, pod pretpostavkom da je jedan iskaz antipod drugog. Pretpostavljamo da posebni iskazi čija je svrha razlikovanje nemarnih od pažljivih sudionika, a koji su formulirani slično iskazu korištenom u ovome istraživanju, mogu navesti sudionike na uvjerenje da se radi o mjerenju poslušnosti ili konformizma. Iz tih razloga, sugeriramo korištenje drugačije formulacije posebnih iskaza kojima je svrha razlikovanje nemarnih od pažljivih sudionika. Takva formulacija mogla bi glasiti: „*Označite kvadratić u gornjem desnom kute ove stranice*“.

**Tablica 3.** Korelacije skala ličnosti između nemarnih i pažljivih sudionika

|  |
| --- |
| UZORAK PARNIH MJESECI (N = 391; Nnemarnih = 97; Npažljivih = 294) |
|  | 1. | 2. | 3. | 4. | 5. |
| 1. *Intelekt* | - |  |  |  |  |
| 2. *Em. stabilnost* | ,08 (,12\*) | - |  |  |  |
| 3. *Ekstraverzija* | ,04 (,19\*\*) | **-,05 (,18\*)** | - |  |  |
| 4. Savjesnost | ,09 (,09) | ,06 (,12\*) | ,0 (,03) | - |  |
| 5. *Ugodnost* | ,09 (,17\*\*) | ,04 (,03) | ,15 (,10) | ,01 (,09) | - |
|  |  |  |  |  |  |
| EVEN MONTH SAMPLE (N = 400; Nnemarnih = 103; Npažljivih = 297) |
|  | 1. | 2. | 3. | 4. | 5. |
| 1. *Intelekt* | - |  |  |  |  |
| 2. *Em. stabilnost* | -,02 (.04) | - |  |  |  |
| 3. *Ekstraverzija* | ,16 (,12\*) | **-,02 (,27\*)** | - |  |  |
| 4. Savjesnost | -,09 (-,07) | ,22\* (,16\*\*) | **-,25 (,09\*)** | - |  |
| 5. *Ugodnost* | ,32\*\* (,18\*\*) | -,25\* (-,12\*) | ,19 (,17\*\*) | -,07 (-,03) | - |
| \* označava korelacije značajne na razini p < ,05, \*\* označava korelacije značajne na razini p < ,01; Napomena: podebljani parovi označavaju statistički značajnu razliku u korelacijama koristeći Williamsov test na razini p< ,05 (invarijantnost korelacija, (Δ χ 2) |

## *Utjecaji pristranosti slaganja s iskazima*

Kako bismo testirali utjecaje pristranosti slaganja, prvo smo odlučili konstruirati jednofaktorske modele za svaku skalu ličnosti zasebno za svaki uzorak. Nakon toga, svakom modelu smo dodali latentni faktor pristranosti slaganja (slika 1), te smo zatim usporedili dvofaktorske modele koji kontroliraju pristranost slaganja i jednofaktorske modele koji ne kontroliraju pristranost slaganja. Prije prezentacije rezultata uključivanja latentnog faktora pristranosti slaganja u modele, podsjećamo kako su skale *savjesnosti* i *ugodnosti* bile izmijenjene za potrebe istraživanja, tako da se jedna inačica skala sastoji od samo pozitivnih iskaza, a druga od samo negativnih. S obzirom na izmjene, ne očekujemo nužno da njihovi jednofaktorski modeli pokažu zadovoljavajuće odgovaranje podatcima, jer ovom metodom ne možemo točno procijeniti utjecaj formulacije iskaza na pristajanje modela podatcima. Rezultati su prikazani u tablici 4.

**Tablica 4.** Pristajanje s i bez kontroliranja ACQ-a podatcima

|  |
| --- |
| UZORAK PARNIH MJESECI (N = 400) |
|  | **χ 2** | **df** | **p** | **CFI** | **TLI** | **RMSEA** | **90% CI** | **SRMR** |
| *Intelekt* | 6,07 | 2 | ,05 | ,967 | ,901 | ,084 | ,007- ,164  | ,034 |
| *Intelekt + ACQ* | 38,82 | 19 | ,00 | ,923 | ,887 | ,056 | ,030- ,073 | ,045 |
| *Savjesnost* | 6,41 | 2 | ,04 | ,966 | ,898 | ,082 | ,015- ,158 | ,033 |
| *Savjesnost + ACQ* | 35,98 | 19 | ,01 | ,931 | ,899 | ,051 | ,024- ,076 | ,050 |
| *Ekstraverzija* | 6,42 | 2 | ,04 | ,988 | ,963 | ,096 | ,017- ,184 | ,028 |
| *Ekstraverzija + ACQ* | 32,33 | 19 | ,03 | ,976 | ,965 | ,047 | ,017- ,065 | ,054 |
| *Ugodnost* | 77,07 | 2 | ,00 | ,764 | ,293 | ,379 | ,309- ,453 | ,115 |
| *Ugodnost + ACQ* | 176,16 | 19 | ,00 | ,811 | ,721 | ,124 | ,107- ,141 | ,070 |
| *Emocionalna stabilnost* | 39,47 | 2 | ,00 | ,894 | ,683 | ,229 | ,170- ,294 | ,058 |
| *Emocionalna stabilnost + ACQ* | 56,42 | 19 | ,00 | ,917 | ,877 | ,075 | ,053-,098 | ,044 |
|  |
| UZORAK NEPARNIH MJESECI (N = 391) |
|  | **χ 2** | **df** | **p** | **CFI** | **TLI** | **RMSEA** | **90% CI** | **SRMR** |
| *Intelekt* | 7,69 | 2 | ,02 | ,971 | ,912 | ,087 | ,029- ,156 | ,031 |
| *Intelekt + ACQ* | 49,56 | 19 | ,00 | ,890 | ,838 | ,066 | ,044- ,089 | ,060 |
| *Savjesnost* | 5,74 | 2 | ,06 | ,983 | ,948 | ,074 | ,035- ,148 | ,029 |
| *Savjesnost + ACQ* | 27,50 | 19 | ,09 | ,971 | ,957 | ,036 | ,000- ,063 | ,045 |
| *Ekstraverzija* | 10,48 | 2 | ,00 | ,987 | ,960 | ,107 | ,050- ,174 | ,024 |
| *Ekstraverzija + ACQ* | 37,38 | 19 | ,00 | ,973 | ,961 | ,052 | ,026- ,076 | ,048 |
| *Ugodnost* | 69,41 | 2 | ,00 | ,734 | ,202 | ,415 | ,335- ,502 | ,121 |
| *Ugodnost + ACQ* | 179,63 | 19 | ,00 | ,750 | ,632 | ,138 | ,120- ,157 | ,007 |
| *Emocionalna stabilnost* | 17,79 | 2 | ,00 | ,936 | ,809 | ,164 | ,100- ,238 | ,044 |
| *Emocionalna stabilnost + ACQ* | 43,46 | 19 | ,00 | ,929 | ,895 | ,062 | ,038- ,087 | ,049 |
| Napomena: “+ ACQ” označava da je model dvofaktorski te da uključuje latentni faktor pristranosti slaganja. Svi indikatori su robusne procjene izračunate koristeći MLR |

U većini slučajeva, rezultati pokazuju kako dvofaktorski modeli koji uključuju latentni faktor pristranosti slaganja u prosjeku bolje pristaju podatcima od jednofaktorskih modela. Skale *intelekta* i *savjesnosti* pokazuju dobro pristajanje podatcima u svojim jednofaktorskim varijantama, a isto se može i reći za skalu *ekstraverzije*, makar u toj skali postoji prostor za protuargument, s obzirom na to da je njihovo pristajanje podatcima nešto lošije od predloženih kriterija. Za navedene skale, modeli koji uključuju latentni faktor pristranosti slaganja pokazuju nešto niže vrijednosti usporednih indeksa pristajanja (engl. *comparative fit indices*) u odnosu na jednofaktorske modele. Vrijednosti usporednih indeksa pristajanja (CFI: Bentler, 1990; TLI: Tucker i Lewis, 1973) se u obje verzije modela većinom i dalje kreću unutar zadovoljavajućih intervala [CFI: ,966- ,988; CFIacq= ,890- ,976] [TLI: ,898- ,963; TLIacq= ,838- ,965]. Generalno, kada se govori o usporednim indeksima pristajanja, vrijednosti koje se kreću u intervalu [od ,90 do ,95] sugeriraju da model zadovoljavajuće pristaje podatcima (Bentler, 1990). Pokazatelj RMSEA (engl. *root mean square error of approximation*) jedan je od najkorištenijih parsimonično-korektivnih indeksa (Steiger i Lind, 1980). Hu i Bentler (1999) navode kako model zadovoljavajuće pristaje podatcima kada su vrijednosti pokazatelja RMSEA ispod ,06. Nadalje, ukoliko vrijednost pokazatelja RMSEA nije ispod ,06, potrebno je da barem njezin 90% interval pouzdanosti zahvaća vrijednosti ispod ,06. Neki autori navode kako su spomenuti kriteriji prestrogi za većinu modela koji se testiraju u društvenim znanostima (Marsh, Hau, i Wen, 2004). Vrijednosti pokazatelja RMSEA za jednofaktorske modele skala *intelekta*, *savjesnosti* i *ekstraverzije* pokazuju kako njihovo pristajanje podatcima nije nužno zadovoljavajuće, makar njihovi 90% intervali pouzdanosti zahvaćaju potrebne kriterije. Ali, nakon uključivanja latentnog faktora pristranosti slaganja, sve vrijednosti pokazatelja RMSEA i svi njezini 90% intervali pouzdanosti zadovoljavaju sve potrebne kriterije koji ukazuju na zadovoljavajuće pristajanje podatcima. Skale *emocionalne stabilnosti* i *ugodnosti* daju nešto drugačije rezultate. Kao što smo već spomenuli, skala *ugodnosti* je adaptirana u svrhu mjerenja utjecaja formulacije iskaza, što je posljedično moglo pogoršati njezino jednofaktorsko pristajanje podatcima. Posljedice adaptacije u obje inačice skale *ugodnosti* rezultiralo je modelima koji nipošto ne pokazuju zadovoljavajuće pristajanje podatcima. Bez obzira što se pristajanje popravilo nakon implementacije latentnog faktora pristranosti slaganja, i dalje smo bili prisiljeni odbaciti sve jednofaktorske i dvofaktorske modele *ugodnosti*. Modifikacijski indeksi pokazuju kako je uvođenje latentnog faktora pristranosti slaganja generalno poboljšalo pristajanje modela podatcima, ali ipak nedovoljno da bi model zadovoljio postojeće kriterije. Odbacivanje modela može biti rezultat toga da skala ugodnosti nije primjerena za polarno obrtanje iskaza na hrvatskom jeziku. Naime, druga skala koja se koristila za mjerenje utjecaja formulacije iskaza, skala *savjesnosti*, rezultirala je modelom koji zadovoljavajuće odgovara podatcima u oba uzorka. Skale *emocionalne stabilnosti* rezultirale su jednofaktorskim modelima koji loše odgovaraju podatcima u oba uzorka, te bi ih se kao takve trebalo odbaciti. No, nakon uvođenja latentnog faktora pristranosti slaganja, dvofaktorski modeli pokazuju bolje pristajanje podatcima, te bi argument za njihovo prihvaćanje bio validan. Bez obzira na to što je većina vrijednosti indikatora pristajanja nešto niža od poželjnih, treba imati na umu da su postavljeni kriteriji prestrogi za većinu modela koji se koriste u društvenim znanostima (Marsh et al., 2004).

# Rasprava i zaključci

Rezultati ovog istraživanja potvrdili su većinu početnih hipoteza. Prvo, pretpostavili smo da će utjecaji formulacije iskaza i potvrdne pristranosti biti osjetni barem na univarijatnoj razini, što je potvrđeno rezultatima. Naime, među većinom parova iskaza pronađene su statistički značajne razlike u njihovim aritmetičkim sredinama, čime su navedeni utjecaji potvrđeni. Dodatno, nakon spajanja dvaju uzoraka, specifikacija jednofaktorskog modela *ugodnosti* rezultirala je negativnim varijancama, koje obično sugeriraju krivo specificiran model (Browne, 2015). S obzirom da su svi jednofaktorski modeli, čije se pristajanje testiralo zasebno za oba uzorka, bez problema uspjeli konvergirati, zaključujemo kako formulacija iskaza može iznimno utjecati na i iskriviti rezultate. Kako bismo vidjeli do koje mjere utjecaji formulacije iskaza i potvrdne pristranosti sežu, proveli smo testove invarijantnosti konstrukata ličnosti. Rezultati su pokazali da spomenuti fenomeni utječu na univarijatnoj razini glede statistički značajnih razlika u aritmetičkim sredinama i na multivarijatnoj razini glede pristajanje jednofaktorskih modela podatcima. No, ne moraju nužno utjecati na multivarijatnoj razini glede mjerne invarijantnosti konstrukata. Na primjer, skale *intelekta*, koje nisu ni na koji način adaptirane, pokazale su istu razinu invarijantnosti kao i skale korištene za testiranje utjecaja potvrdne pristranosti (*ekstraverzija*) i skale koja je bila korištena za testiranje utjecaja formulacije iskaza (*savjesnost*), ali je bitno napomenuti kako skala intelekta pruža jedini jednofaktorski model sačinjen od oba uzorka koji dobro pristaje podatcima. Ograničenje ovakvog pristupa je u tome što početna adaptacija mjernih skala može rezultirati odbacivanjem jednfoaktorskih modela, koji bi bez takvih adaptacija bili prihvaćeni. Navedeno može utjecati na validnost rezultata dobivenih testovima invarijantnosti.

Drugo, posebni iskaz koji se koristio kako bi razlikovao nemarne od pažljivih sudionika, pokazao je slabu diskriminacijsku snagu, čime zaključujemo da nije validan niti uspješan u razlikovanju navedenih grupa. Polarni par iskaza nije pokazao statistički značajnu razliku u aritmetičkim sredinama rezultata nemarnih i pažljivih sudionika, a validnost su pod upitnik stavljali i rezultati razlika indeksa NARS i koeficijenta korelacija konstrukata ličnosti. Naime, u slučaju validnosti posebnog iskaza očekivale su se statistički značajne razlike među navedenim elementima, koje u većini slučaja nisu pronađene. Pretpostavljamo da je mogući razlog toga način na koji je iskaz formuliran, to jest koji navodi sudionike da izaberu jednu opciju na peterotomnoj skali slaganja. Vjerujemo kako je ovo moglo navesti sudionike da pomisle kako se radi o testiranju poslušnosti ili konformizma, što bi moglo znatno utjecati na validnost iskaza. Zbog navedenog, preporučamo korištenje drugačijih tipova posebnih iskaza kojima je svrha razlikovanje nemarnih od pažljivih sudionika. Takav iskaz mogao bi biti formuliran na sljedeći način: „*Označite kvadratić u gornjem desnom kute ove stranice*“. Smatramo da je u budućim metodološkim studijama potrebno dodatno istražiti vezu između sklonosti konformizmu i stila odgovaranja na iskaze koji služe za identifikaciju nemarnog odgovaranja te ne temelju njihovih rezultata utvrditi koja je inačica iskaza za mjerenje nemarnog odgovaranja najprihvatljivija.

Treće, pretpostavili smo da je pristranost slaganja s iskazima fenomen koji je u nekoj mjeri prisutan u većini društvenih istraživanja, koji može utjecati na rezultate. Kada smo usporedili pristajanje jednofaktorskih modela koji ne uključuju pristranost slaganja i dvofaktorskih modela koji uključuju latentan faktor pristranosti slaganja, postalo je jasno kako kontroliranje tog fenomena poboljšava pristajanje modela podatcima, posebice ako je model prije kontroliranja bio na rubu odbacivanja. Moguće ograničenje ovakvog pristupa je veličina korištenog instrumenta koji mjeri NARS, a koji se sastoji od 17 heterogenih iskaza koji ne dijele zajednički sadržaj. Preporučamo pokušaje implementacije kraćih skala, jer bi se smanjivanjem instrumenata mogla povećati učestalost njegovog korištenja, a time i pospješiti kontrola pristranosti slaganja u društvenim istraživanjima.

Cilj ovog istraživanja bio je prikazati razne fenomene koji imaju mogućnost utjecati na rezultate anketnih upitnika, a koji se rijetko uzimaju u obzir prilikom konstrukcije anketnih upitnika. Ukoliko se mogući utjecaj ovdje proučavanih efekata metode ne uzme u obzir, validnost rezultata dobivenih anketnim upitnikom može se znatno umanjiti. Prilikom konstrukcije anketnog upitnika, preporučeno je imati u vidu sve vanjske utjecaje koji mogu utjecati na rezultate, osim korištenog instrumenta. Mjerenje nemarnog odgovaranja moglo bi se vršiti putem samo jednog posebnog iskaza, te se iz takvih iskaza mogu saznati razne korisne informacije, koje se inače ne bi mogle uvidjeti.

# Sažetak

Autor: Luka Mandić

Naslov rada: Testiranje utjecaja efekata metode na rezultate anketnih istraživanja

Često se misli da su rezultati anketnih istraživanja isključivo odraz korištenih mjernih instrumenata u anketnom upitniku. No, postoje razni fenomeni koji mogu utjecati na rezultate i koji ih mogu iskriviti, a čiji se utjecaji često zanemaruju pri provedbi anketnih istraživanja. Ova studija bavila se ispitivanjem utjecaja raznih efekata metode koji mogu iskriviti rezultate anketnih upitnika. Testirali su se utjecaji formulacije iskaza (engl. *item wording*), potvrdne pristranosti (engl. *confirmatory bias*), nemarnog odgovaranja (engl. *careless responding*) te pristranosti slaganja (engl. *acquiescence bias*). Koristeći dvije inačice online anketnog upitnika, prikupili smo rezultate 791 korisnika društvenih mreža. Testirali smo jesu li prethodno navedeni efekti metode imali utjecaja na aritmetičke sredine čestica, korelacije čestica, korelacije konstrukata, pristajanje metrijskih modela podatcima te na invarijantnost mjerenja. Instrumenti putem kojih su efekti metode bili testirani bili su iz sfere ličnosti i rodne nejednakosti, te su njihove čestice bile izmijenjene shodno potrebama mjerenja utjecaja pojedinog efekta metode. Svi testirani efekti metode, izuzev nemarnog odgovaranja, pokazali su statistički značajne utjecaje na rezultate na barem jednoj jedinici analize. Formulacija iskaza i potvrdna pristranost utjecali su na aritmetičke sredine, na pristajanje modela podatcima te na invarijantnost mjerenja. Kontroliranje pristranosti slaganja rezultiralo je modelima koji su bolje pristajali podatcima. Ovaj rad potvrđuje da istraživane efekte metoda treba uzeti u obzir prilikom provedbe istraživanja metodom ankete, ujedno dajući konkretne preporuke istraživačima.

*Ključne riječi*: efekti metode, formulacija iskaza, nemarno odgovaranje, potvrdna pristranost, pristranost slaganja

**Summary**

Author: Luka Mandić

Title: Testing the influence that method effects have on survey results

It is often thought that survey results reflect only the underlying measuring instruments that were used in the survey. However, there are various phenomena that can affect and skew survey results, whose influences are often overlooked in survey research. This aim of this study was to test the influences which various method effects have on survey results. We tested the influences of the following method effects: item wording, confirmatory bias, careless responding, and acquiescence bias. Using a split-ballot study of an online questionnaire, we collected data from 791 participants. We tested if the mentioned method effects had an influence on mean values, item correlations, construct correlations, model fits, and construct measurement invariance. The instruments used to test these influences were from the domain of personality and gender inequality, and their items were changed based on the method effect being tested. All tested method effects, with the exception of careless responding, had a statistically significant effect on at least one component of analysis. Item wording and confirmatory bias affected mean values, model fit, and measurement invariance. Controlling for acquiescence bias improved the fit of the model. This paper confirms that the tested method effects should be carefully considered when using surveys in research, and suggests guidelines to do so.

*Keywords*: method effects, item wording, careless responding, confirmatory bias, acquiescence bias

# Literatura

Bentler, P. M., Jackson, D. N., i Messick, S. (1971). Identification of content and style: A two-dimensional interpretation of acquiescence. *Psychological Bulletin*, 76(3), 186-204. doi: https://doi.org/10.1037/h0031474

Bentler, P. M. (1990). Comparative fit indexes in structural models. *Psychological Bulletin*, 107(2), 238-246. doi: 10.1037/0033-2909.107.2.238

Brown, T. A. (2015). *Methodology in the social sciences. Confirmatory factor analysis for applied research* (2nd ed.). The Guilford Press.

Cronbach, L. J. (1946). Response sets and test validity. Educational and Psychological Measurement, 6, 475-494. doi:10.1177/001316444600600405

Davies, M. F. (2003). Confirmatory bias in the evaluation of personality descriptions: Positive test strategies and output interference. *Journal of Personality and Social Psychology*, 85(4), 736–744. doi:10.1037/00223514.85.4.736

Donnellan, M. B., Oswald, F. L., Baird, B. M., i Lucas, R. E. (2006). The Mini-IPIP scales: Tiny-yet-effective measures of the Big Five factors of personality. *Psychological Assessment*, 18(2), 192-203. doi: 10.1037/1040-3590.18.2.192

Greenleaf, E. A., Measuring Extreme Response Style. *Public Opinion Quarterly*, 56(3), 328–351. doi: <https://doi.org/10.1086/269326>

Hu, L., i Benter, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling*, 6(1), 1–55. doi: <https://doi.org/10.1080/10705519909540118>

Inglehart, R., i Norris, P. (2010). *Rising tide: Gender equality and cultural change around the world*. Cambridge: Cambridge University Press. doi: 10.1017/CBO9780511550362

Kam, C. C. S., i Meyer, J. P. (2015). How careless responding and acquiescence response bias can influence construct dimensionality: the case of job satisfaction. *Organizational Research Methods*, 18(3), 512–541. doi: 10.1177/1094428115571894

Knowles, E. S., i Condon, C. A. (1999). Why people say "yes": A dual-process theory of acquiescence. *Journal of Personality and Social Psychology*, 77(2), 379–386. doi: <https://doi.org/10.1037/0022-3514.77.2.379>

Kunda, Z., Fong, G. T., Santioso, R., i Reber, E. (1993). Directional questions direct self-conceptions. *Journal of Experimental Social Psychology*, 29(1), 63–86. doi: 10.1006/jesp.1993.1004

Marsh, H. W. (1986). Negative item bias in ratings scales for preadolescent children: A cognitive-developmental phenomenon. *Developmental Psychology*, 22, 37–49. doi:10.1037/0012-1649.22.1.37

Marsh, H., Hau, K., i Wen, Z. (2004). In Search of Golden Rules: Comment on Hypothesis-Testing Approaches to Setting Cutoff Values for Fit Indexes and Dangers in Overgeneralizing Hu and Bentler's (1999) Findings. *Structural Equation Modeling: A Multidisciplinary Journal* 11(3), 320-341. doi: 10.1207/s15328007sem1103\_2

Maul, A. (2013). Method Effects and the Meaning of Measurement. Frontiers in Psychology, 4, Article 169. doi: https://doi.org/10.3389/fpsyg.2013.00169

Maydeu-Olivares, A., i Coffman, D. L. (2006). Random intercept item factor analysis. *Psychological Methods*, 11(4), 344–362. doi: https://doi.org/10.1037/1082-989X.11.4.344

McClendon, M. J. (1991). Acquiescence: Tests of the cognitive limitations and question ambiguity hypotheses. *Journal of Official Statistics*, 7, 153–166.

Nunnally, J. C. (1978). *Psychometric theory* (2nd ed.). New York, NY: McGraw-Hill.

Paulhus, D. L. (1991). Measurement and control of response bias. U J. P. Robinson, P. R. Shaver, & L. S. Wrightsman (Ur.), *Measures of personality and social psychological attitudes* (17–59). San Diego, CA: Academic Press. doi: https://doi.org/10.1016/B978-0-12-590241-0.50006-X

R Development Core Team. (2019). R: *A language and environment for statistical computing* [Computer software manual]. Vienna, Austria: Author.

Revelle, W. (2019). *psych: Procedures for Psychological, Psychometric, and Personality Research*. Northwestern University, Evanston, Illinois. R package version 1.9.12, https://CRAN.R-project.org/package=psych.

Rosseel, Y. (20129). lavaan: An R package for structural equation modeling. *Journal of Statistical Software*, 48, 1-36. doi: 10.18637/jss.v048.i02

Schmitt, N., i Stults, D. M. (1985). Factors defined by negatively keyed items: The result of careless respondents? *Applied Psychological Measurement*, 9(4), 367–373. doi: https://doi.org/10.1177/014662168500900405

Schuman, H., i Presser, S. (1981*). Questions and answers in attitude surveys:Experiments on question form, wording and context*. SanDiego, CA: Academic Press.

Steiger, J.H. i Lind, J.C. (1980, May 30, 1980*). Statistically-based tests for the number of common factors*. Paper presented at the Annual Spring Meeting of the Psychometric Society, Iowa City.

Suárez-Álvarez, J., Pedrosa, I., Lozano, L. M., García-Cueto, E., Cuesta, M., i Muñiz, J. (2018). Using reversed items in Likert scales: A questionable practice. *Psicothema* 30(2), 149-158. doi: 10.7334/psicothema2018.33

Tougas, F., Brown, R., Beaton, A. M., i Joly, S. (1995). Neosexism: Plus Ça Change, Plus C'est Pareil. *Personality and Social Psychology Bulletin*, 21(8), 842-849. doi: <https://doi.org/10.1177/0146167295218007>

Tucker, L.R., i Lewis, C. A. (1973). reliability coefficient for maximum likelihood factor analysis. *Psychometrika,* 38(1), 1–10. doi: <https://doi.org/10.1007/BF02291170>

Weijters, B., Geuens, M., i Schillewaert, N. (2010). The stability of individual response styles. *Psychological Methods*, 15(1), 96-110. doi: 10.1037/a0018721.

Weijters, B., i Baumgartner, H. (2012). Misresponse to Reversed and Negated Items in Surveys: A Review. *Journal of Marketing Research*, 49(5), 737-747. doi: https://doi.org/10.1509/jmr.11.0368

Weijters, B., Baumgartner, H., i Schillewaert, N. (2013). Reverse item bias: An integrative model. *Psychological Methods*, 18(3), 320-334. doi: 10.1037/a0032121

White, H. (1980). A Heteroskedasticity-Consistent Covariance Matrix Estimator and a Direct Test for Heteroskedasticity. *Econometrica*, 48(4), 817–38.

Yuan, K., i Bentler, P. M. (2007). Multilevel Covariance Structure Analysis By Fitting Multiple Single‐Level Models. *Sociological Methodology*, 37(1), 53-82. doi: https://doi.org/10.1111/j.1467-9531.2007.00182.x

Woods, C. M. (2006). Careless responding to reverse-worded items: Implications for confirmatory factor analysis. *Journal of Psychopathology and Behavioral Assessment*, 28(3), 186-191.

# Zahvale

Provođenje ovog istraživanja potvrdilo mi je čime se želim baviti nakon završetka studija, te mi time dodatno obogatilo iskustvo studiranja. Prije (a i za vrijeme) provođenja istraživanja morao sam naučiti nove oblike statističkih analiza, što je predstavljao mukotrpan proces brojnih pokušaja i još brojnijih pogrešaka. No, od kada se finalni produkt počeo nazirati, shvatio sam da je metodologija područje kojim se želim nastavit baviti.

Ovim putem bih htio zahvaliti svojoj mentorici, doc. dr. sc. Kseniji Klasnić, koja je unatoč svojem prenatrpanom rasporedu prihvatila i podržala ideju, te me ohrabrivala i savjetovala tokom cijelog procesa. Nacrt istraživanja prošao je kroz mnogobrojne izmjene, a upravo njen pristup (nikada nije odbila niti jednu moju ideju) dao mi je samopouzdanja da nastavim s provođenjem i oblikovanjem istraživanja. Zahvaljujem svima koji su anketu dijelili i rješavali, pogotovo s obzirom na to da je za ovo istraživanje bio potreban veliki broj sudionika.

Zahvaljujem svojoj baki, koja je prvi (ujedno i najstariji) sudionik koji je sudjelovao u istraživanju. Ali posebice, neizmjerno, zahvaljujem Tii, koja se zadnjih godinu dana uspješno pravi da ju zanima o čemu pričam i pruža mi bezuvjetnu podršku, čak i kada se osjećam izgubljeno u enormnoj količini nerazumljivih podataka. Bez njenih savjeta i strpljenja ovaj rad ne bi bio ostvariv.

**Životopis**

Luka Mandić rođen je u Zagrebu, a dolazi iz Velike Gorice, u kojoj je završio opću gimnaziju i Umjetničku školu Franje Lučića (gitara). Trenutno je student diplomskog studija sociologije, gdje se osim interesa za metodologiju i dalje teorijski usmjerava. Dobitnik je nagrade za izvrsnost u (preddiplomskom) studiju, te mu je 2020. dodijeljena stipendija za izvrsnost Sveučilišta u Zagrebu. Bivši je član predsjedništva Kluba studenata sociologije Diskrepancija, unutar kojeg je sudjelovao u organiziranju simpozija i radu sekcija. Kroz studij je radio razne poslove na brojnim mjestima: Visage Technologies (anotacije na softveru za prepoznavanje lica), IPSOS (izvještaji za marketinške svrhe), HKO (unos podataka), Američka ambasada (vozač), Fleet rent-a-car (vozač), ali se najviše zadržao u Koncertnoj dvorani Vatroslav Lisinski, u kojoj i dalje ponosno radi kao biljeter. U slobodno vrijeme neobavezno se bavi glazbom (gitara, klavir) i šahom.