

3. Pristranost potvrđivanja

<https://doi.org/10.31212/kogn.prist.2024.vale.03>

Traganje za dokazima koji potvrđuju da smo u pravu i ignoriranje informacija koje to negiraju

Pavle Valerjev¹ 

SAŽETAK

U ovom poglavlju opisana je kognitivna pristranost poznata pod nazivom *pristranost potvrđivanja*. U prvom dijelu teksta predstavljeni su primjeri i definicija za ovu pristranost. Nakon toga opisana je kao važan faktor u kontekstu filozofije znanosti gdje je artikuliran zahtjev za falsifiabilnost znanstvenih teorija. Pristranost potvrđivanja upravo otežava taj zahtjev. Slijedi opis glavnih istraživačkih paradigmi u psihologijskim istraživanjima u kojima se manifestira ova pristranost. To su, prije svega, problem 2 4 6 i Wasonov zadatak izbora te razne varijacije tih zadataka. Konačno, pristranost potvrđivanja je prikazana u širem kontekstu istraživanja misaonih procesa testiranja hipoteza. Unutar tog konteksta ova se pristranost naziva i *strategija pozitivnog testiranja*. Naznačeno je da određenim okolnostima ova strategija predstavlja konstruktivan dio istraživačkog procesa te se tada na nju ne može gledati kao na pristranost.

Ključne riječi: pristranost potvrđivanja, strategija pozitivnog testiranja, falsifiabilnost, induktivno rasuđivanje, Wasonov zadatak selekcije/izbora

Uvod

Zamislimo sljedeću situaciju. Imate prijateljicu koja voli čitati horoskope. Uvjeren je da kad razmišlja o zvijezdama i sudbini „tu ima nešto“. Uvjeren je i da njoj horoskopi daju dobre prognoze, tako da prati dnevne i tjedne horoskope. Ponekad pročita nešto što se izvrsno slaže s njenim utiskom i to pojača njenu uvjerenost u ispravnost

astrološke metode. Ponekad se i malo zamisli nad prognozom, no ubrzo ipak uvidi da je bila ispravna. Recimo, ovaj tjedan je pročitala u horoskopu da ju očekuje poboljšanje na financijskom i ljubavnom planu. To za financijski plan se sigurno odnosi na 2 eura koje je pro-našla na ulici čime si je platila kavu. Za poboljšanje na ljubavnom planu sigurno

¹ Odjel za psihologiju, Sveučilište u Zadru; email: valerjev@unizd.hr

se misli na činjenicu da je prekjučer na koncertu srela mladića koji joj se sviđa i koji joj se nasmiješio u prolazu. Te dvije okolnosti svakako predstavljaju svojevrsno poboljšanje na oba plana. Na ovaj način, ona svaki put kad su okolnosti u skladu s predviđanjima horoskopa potvrđuje valjanost horoskopa. S druge strane, svaki put kad evidencija ne prati prognoze dogodi joj se mala mentalna gimnastika tako da se negativna opažanja ili zanemare (jer, „zar ne postoje i slučajne oscilacije događaja?“) ili se pak reinterpreteraju kao pozitivna opažanja (kao u navedenom primjeru s 2 eura). Netko sa strane će možda lakše uočiti da je ona žrtva iskrivljenog razumijevanja svijeta. U skladu s takvim razumijevanjem, ona vjeruje da horoskopi predviđaju njenu sudbinu te nastoji održati to svoje uvjerenje čvrstim. Da li zbog toga ona selektivno opaža, procesira i interpretira informacije koje se tiču uvjerenja o ispravnosti horoskopa?

Istraživači kognitivnih pristranosti kažu – da. Osim astrologije, mogli bi se navesti još razni slučajevi. Na primjer, onaj o uvjerenju da je dijeta za mršavljenje uspješna kad vaga danima samo varira oko iste vrijednosti ili onaj o uvjerenju da je istraživačka hipoteza ispravna kad ju podaci tek „djelomično potvrđuju“. Ovaj tip pristranosti je svugdje prisutan u našoj svakodnevnicu. Primjerice, kod ljudi jakih religijskih, ideoloških ili filozofskih uvjerenja, kod ljudi sklonih

teorijama zavjere, među ekonomistima, novinarima, političarima i drugim skupinama čiji članovi se u pravilu vode svojim uvjerenjima. Također, među ljudima koji imaju paranormalna uvjerenja ili prakticiraju tzv. alternativnu medicinu. Zatim, u socijalnom kontekstu, kod održavanja slike o sebi, ili kod masovnih pojava poput „lova na vještice“ pa čak i kod polarizacije stavova među suprotstavljenim grupama koje su izložene istoj evidenciji. Internet, koji nudi nepregledna gledišta i mišljenja svih vrsta i za sve svjetonazore, iznimno je plodno tlo za gajenje ovakve pristranosti jer upravo ona upravlja odabirom informacija. Među daljnjim primjerima, tu su zastupnici svakakvih pseudoznanstvenih teorija, samopomoći, razni životni treneri i prodavači životnih mudrosti. Na žalost, ni medicinari niti znanstvenici nisu imuni na ove pogreške, iako bi utreniranost u znanstvenoj metodi i kritičkom mišljenju trebale biti od pomoći za odupiranje sklonosti nekritičkom potvrđivanju vlastite hipoteze.

Ova specifična kognitivna pristranost je **pristranost potvrđivanja** (engl. *confirmation bias*) i praktički nitko nije imun na nju. Evans (1989) definira pristranost potvrđivanja kao „temeljnu tendenciju ljudskih bića u traganju informacije koja je konzistentna s njihovim tekućim uvjerenjima, teorijama i hipotezama, i izbjegavanja skupljanja potencijalno falsificirajuće evidencije“. Drugim riječima,

tražimo evidenciju koja potvrđuje naša uvjerenja i hipoteze, dok istovremeno zanemarujemo ili reinterpreteramo evidenciju koja ih ne potvrđuje. Otuda i njen naziv. Pristranost potvrđivanja se odnosi na pristrano (selektivno) traženje informacija, pristranu interpretaciju informacija i pristrano dosjećanje informacija. U drugoj literaturi (npr. [Klayman & Ha, 1987](#); Manktelow, 2012) navodi se i drugo ime za ovu tendenciju: *strategija pozitivnog testiranja* (engl. *positive test strategy*).

U navedenoj definiciji se spominje izbjegavanje „falsificirajuće evidencije“. Filozof znanosti Karl Popper je u svojem djelu *Logika znanstvenog otkrića* (objavljenom na njemačkom jeziku 1932. i ponovno objavljenom 1959. na engleskom jeziku; Popper, 2002) argumentirao da se znanstvena metoda mora temeljiti na mogućnosti falsifikacije (falsifiabilnost – engl. *falsifiability*) kao temeljnom sudu za valjanost znanstvene teorije. Falsifikacija je metodološka mogućnost prikupljanja takve empirijske evidencije na temelju koje bismo (ukoliko postoji) prihvatili temeljne tvrdnje koje su kontradiktorne s teorijom, što bi za posljedicu moralo oboriti ili promijeniti teoriju. Takav falsificirajući efekt mora imati i mogućnost reprodukcije (ne smije se temeljiti na izoliranom slučaju za koji ne postoji mogućnost ponavljanja). Ukoliko znanstvena teorija prođe (rigorozni) test falsifikacije, imamo valjanu znanstvenu teoriju (barem za one temeljne

tvrdnje koje su testirane). Popper ističe da je glavna razlika između znanstvenih i neznanstvenih teorija upravo falsifiabilnost. Neznanstvene teorije ne dopuštaju mogućnost falsifikacije i zbog toga su prividno uvijek u pravu, bez obzira koja evidencija se obušila na njih. Na primjer, prema kreacionističkoj teoriji o „mladoj Zemlji“ naš je planet i sve na njemu stvorio abrahamski Bog u razdoblju od prije 6000–10000 godina. Iako postoje jasni geološki i paleontološki dokazi da je Zemlja stara više od 4,5 milijardi, a život na Zemlji više od 3,5 milijardi godina, po toj je teoriji Zemlja upravo tako i stvorena (s geološkim slojevima, fosilima i razlikama u koncentraciji radioaktivnih elemenata) te ne postoji tehnologija koja može izmjeriti stvarnu dob Zemlje. Upravo zbog toga je ta teorija neznanstvena jer ne nudi niti jednu temeljnu tvrdnju koja bi bila falsifiabilna. S druge strane, teorija evolucije života na Zemlji je prava znanstvena teorija jer nudi brojne temeljne tvrdnje koje su falsifiabilne. Na primjer, ako bi netko pronašao jasne dokaze da postoje fosili zečeva u prekambrijskom geološkom sloju, teorija evolucije bi pala u vodu. Nitko nikada nije pronašao takve fosile u tako starom sloju, niti je itko drugi našao drugu evidenciju koja bi falsificirala teoriju evolucije (iako postoje brojne temeljne tvrdnje koje se mogu falsificirati), tako da je ona i dalje živa i zdrava znanstvena teorija (i koja se, obzirom na nepreglednu

nagomilanu evidenciju koja ju potvrđuje, već smatra znanstvenom činjenicom). Carl Sagan (2000) je opisao važnost falsifiabilnosti na primjeru u kojem tvrdi da u svojoj garaži drži zmaja. Jedino, taj zmaj je nevidljiv, lebdi i nije u interakciji s okolinom te bljuje nevidljivu vatru „bez topline“ tako da nije vidljiv ni na infracrvenoj kameri. Dakle, ne postoji način da se testira i utvrdi postojanje zmaja (kao ni za sva mitska i religijska bića) te zaključuje da nema razlike između ovakvog zmaja i nepostojanja zmaja.

Dakle, falsifiabilnost je jedan od temeljnih zahtjeva za znanstvenu teoriju, te je testiranje temeljnih tvrdnji na

način da ih se pokuša falsificirati jedan od temelja za kritičko znanstveno mišljenje. Istovremeno, kako je navedeno prije, ljudi imaju tendenciju da manifestiraju pristranost potvrđivanja tako što izbjegavaju falsificirajuću evidenciju. Iz ovoga slijedi da ljudima kritičko i znanstveno mišljenje nije prirodno te da se moraju trenirati u njemu. Psiholozi su prepoznali važnost istraživanja ovog tipa mišljenja tako da danas postoje cijela istraživačka područja poput *znanstvenog rasuđivanja* unutar šireg područja *psihologije znanosti* u koja su uključena i istraživanja pristranosti potvrđivanja (vidjeti npr. knjigu o psihologiji znanosti, Feist & Gorman, 2013).

Zadaci induktivnog rasuđivanja za ispitivanje pristranosti potvrđivanja

Problem 2 4 6

Jedan od najranijih radova koji opisuje pristranost potvrđivanja je onaj britanskog kognitivnog psihologa Petera Wasona (1960; Wason & Johnson-Laird, 1972) u kojem je prvi put opisan problem 2 4 6. Taj se problem smatra klasičnim i vjerojatno prvim istraženim problem-skim zadatkom kojim se demonstrira i ispituje pristranost potvrđivanja. U tom problemu ispitanicima je prezentiran kratki niz brojeva, 2 4 6 (moguće je i druge nizove prezentirati) te im je rečeno da se taj niz uklapa u jednostavno pravilo po kojem su ti brojevi poredani. Zadatak ispitanika je da otkriju to pravilo tako

da samostalno prave nove proizvoljne nizove od tri broja i da pri tome napišu razlog za upotrebu tog niza. Za svaki pokušaj bi dobili povratnu informaciju od eksperimentatora o tome uklapaju li se novi nizovi u pravilo ili ne. Na taj način ispitanici su mogli testirati svoju hipotezu o pravilu. Također im je rečeno da, iako nema ograničenja, pokušaju otkriti pravilo sa što manje testnih pokušaja. Postupak se nastavljao dok ne bi otkrili pravilo, što bi bilo vidljivo na temelju navedenog razloga za korištenje novog niza, ili dok ne bi prekoračili trajanje od 45 minuta ili pak odustali. Netočna rješenja koja bi ispitanici često ponudili

u ovom zadatku jesu sljedeća pravila: „Rastući niz parnih brojeva“, „Intervali koji rastu za dva“, „Aritmetička progresija“, „Višekratnici prvog broja“. Testirajući svoje hipoteze ispitanici bi često zadavali razne varijacije rastućih nizova vjerujući da trebaju otkriti uzorak po kojem nizovi rastu. Međutim, točan odgovor je „Rastući niz brojeva“. Dakle, bilo koji rastući niz brojeva, parnih, neparnih, s pravilnim ili nepravilnim intervalima. Zadatak se pokazao relativno teškim za ispitanike jer su mnogi ispitanici generalizirali i testirali uglavnom potvrđujuću evidenciju, a sva evidencija koja je potvrđivala njihovo hipotetsko pravilo je ujedno potvrđivala i općenitije stvarno pravilo (ciljno pravilo), tako da je vrlo teško moglo doći do diskriminacije pravila. Većini ispitanika je trebalo više pokušaja, a neki ni nakon toga nisu otkrili ciljno pravilo. Ukoliko netko vjeruje da niz 2 4 6 reprezentira pravilo „rastući parni brojevi“, pokušao bi s nizom „10 12 14“ i dobio odgovor eksperimentatora „da – niz se uklapa u pravilo“ jer se uklapao ne samo u ispitanikovu hipotezu, već i u ciljno pravilo. Time ispitanik nije saznao mnogo ni o pravilu ni o nevaljanosti svoje hipoteze (koju je nesvjesno ili svjesno nastojao potvrditi). Kritički nastrojen ispitanik bi za hipotezu o rastućim parnim brojevima barem pokušao ponuditi neparne brojeve. Ali pravi kritički korak je isprobavanje silaznog niza. Eksperimenta-

torov odgovor „ne – niz se ne uklapa u pravilo“ omogućuje brz dolazak do rješenja. Ispitanicima je zadatak 2 4 6 težak jer je najčešće njihova hipoteza podskup općenitijeg pravila koje treba otkriti (čime se niz uklapa u oba pravila) i onda je nemoguće diskriminirati hipotezu od pravila potvrđujućom evidencijom. Analiza odnosa između ispitanikovih hipoteza i ciljnog pravila detaljno je opisana u radu autora Klaymana i Ha (1987). Taj rad je značajan za daljnji razvoj istraživanja hipotetskog mišljenja zbog toga što su autori pokazali da nije svako pozitivno testiranje pristranost potvrđivanja, te da postoje situacije u kojima je pozitivno testiranje efikasnije od negativnog. To ovisi o tome kakav je logički odnos između hipoteze i ciljnog pravila. U situacijama koje nisu takve da je ispitanikova hipoteza podskup ciljnog pravila, nego se radi o presjeku ili odvojenim skupovima, ili pak o situaciji gdje je ciljno pravilo podskup hipoteze, potvrđujuća evidencija je korisna.

No bez obzira na to, zadatak 2 4 6 je pokazao da ljudi imaju poteškoća čak i u jednostavnim zadacima ukoliko su žrtve pristranosti potvrđivanja. Dobelli (2013) u vezi s tim navodi kako je pristranost potvrđivanja *majka svih miskoncepcija* zbog toga što nas ona navodi da svaku novu informaciju interpretiramo tako da bude kompatibilna s našim postojećim teorijama, znanjima i uvjerenjima. Mi

isključujemo informaciju koja je kontradiktorna našim postojećim gledištima i to ponekad može imati ozbiljne posljedice, pogotovo ako utječe na životne ili poslovne odluke. Nadalje, pristranost potvrđivanja je neka vrsta općenite pristranosti ili metapristranosti. Ona je donekle nalik nekim drugim pristranostima, poput pristranosti uvjerenja, pristranostima koje proizlaze iz heuristike dostupnosti, Barnum–Forerovom efektu, zatim povezana je sa selektivnošću obrade informacija (npr. sljepoća zbog nepažnje, iluzija vijesti i efekt istaknutosti), iluzijom introspekcije, iluzornom korelacijom i kognitivnom disonancijom. Sve ove kognitivne iluzije imaju tendenciju da filtriraju informaciju na pogodan način ili da daju reinterpretaciju događaja. No, za razliku od većine njih, pristranost potvrđivanja se ne očituje u jednoj izoliranoj paradigmatškoj aktivnosti, već se čini da je ona sveprisutna i izmiješana sa svim drugim aspektima kognitivnog procesiranja. Stoga nije lako napraviti specifične zadatke za izolirano ispitivanje pristranosti potvrđivanja. Na primjer, kad ciljano ispitujemo selektivnost pažnje koristimo paradigmu pretraživanja gdje se zadaju mete i distraktori. Kad ispitujemo pristranosti utemeljene u heuristici dostupnosti koristimo odgovarajuće zadatke pamćenja i dosjećanja u kojima variramo dostupnost učenog materijala. No, kad ispitujemo valjanost zaključivanja s uvjerljivim i neuvjerljivim zaključcima, u stvari, u konkretnom

smislu ispitujemo pristranost uvjerenja, a općenitije gledajući zahvaćamo i pristranost potvrđivanja.

Rezultati koji sugeriraju da bi pristranost potvrđivanja mogla biti pokrovni termin za nekoliko povezanih procesa su dobiveni u radu Teovanovića i Bertheta (2024). Autori su koristili tri zadatka koji manifestiraju pristranost potvrđivanja i to svaki u tri kognitivna procesa (pretraživanje informacija, odvagivanje evidencije i dosjećanje) te su analizirali faktorsku strukturu dobivenih rezultata. Dobiveni model pokazuje tri procesna faktora prvog reda (za tri tipa kognitivnih procesa) koji međusobno koreliraju, ali nedovoljno za indicaciju jedinstvenog faktora drugog reda.

Wasonov problem 2 4 6 je zadatak koji kod rješavača aktivira induktivno rasuđivanje i po tome je nalik procesu znanstvene metode gdje se također prikuplja evidencija koja se onda generalizira u opće zakone ili efekte. Daljnja poveznica s indukcijom je i kategoričko/konceptualno učenje. Stoga je ovaj zadatak donekle nalik zadacima kategorizacije. U tim zadacima ispitanici kategoriziraju pojedine primjerke na temelju vlastite hipoteze o pravilu koje moraju naučiti. U tom zadatku također dobivaju povratnu informaciju (na primjer, Wisconsin test sortiranja karata), ali s tom bitnom razlikom što u problemu 2 4 6 ispitanici sami generiraju pojedine primjerke kojima testiraju hipoteze. Dakle, uključen je proces aktivnog testiranja hipoteza. Dobiveni rezultati

jasno demonstriraju da ljudi nikad ne idu prazna uma u testiranje evidencije, već u pravilu imaju generirane hipoteze ili uvjerenja koja onda utječu na odabir načina testiranja i na generiranje zaključaka.

Drugi zadaci s induktivnim rasuđivanjem

Daljnja istraživanja ovog efekta su uključivala i varijacije ovog zadatka. Smatrajući problem 2 4 6 preapstraktnim, Maynatt i suradnici (1977; 1978) su za svoja istraživanja smislili „konkretne“ zadatke za testiranje hipoteza koji su više nalikovali stvarnim situacijama znanstvenog istraživanja. U tim zadacima se tražilo od ispitanika da u simuliranoj grafičkoj okolini određenog računalnog programa odrede uzrok „apsorbiranja“ čestica. Ispitanici su testirali hipoteze ispaljujući čestice u raznim smjerovima. Trebalo je odrediti ovisi li to o obliku (kriva hipoteza) ili o razini svjetline geometrijskih oblika. Oba svojstva likova su varirala. U prvom istraživanju (Maynatt et al., 1977) su im prikazani ekrani s grafičkim preprekama koje su služile za inicijalno generiranje hipoteza i koje su mogli sami birati za kasnije testiranje. Grupe ispitanika su također dobile različite upute o tome da potvrde, odbace ili testiraju hipoteze. Rezultati su potvrdili pristranost potvrđivanja i u ovom zadatku. Ispitanici su radije birali okoline koje potvrđuju njihovu hipotezu o obliku nego one koji omogućuju

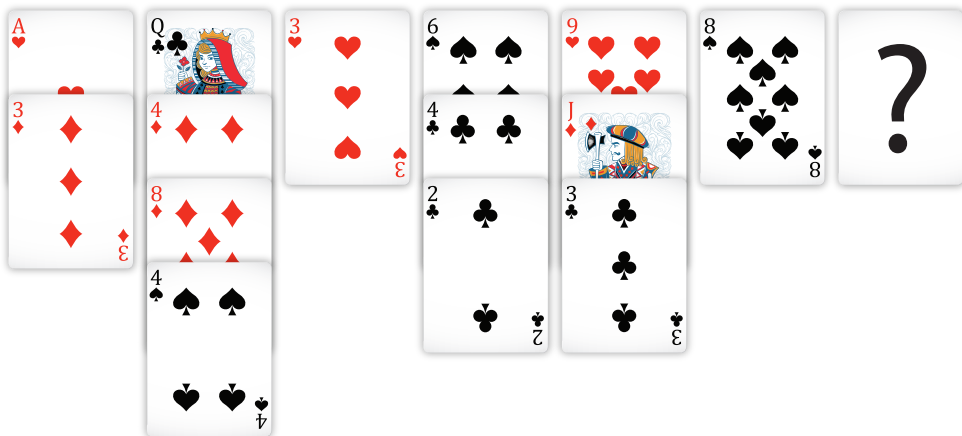
testiranje alternativne hipoteze. Efekt pristranosti je ostao postojan bez obzira na varijacije u uputama. Ispitanici su odbacili netočnu hipotezu tek kad su bili suočeni s falsificirajućim evidencijom. U drugom istraživanju (Maynatt et al., 1978) zadatak je bio kompleksniji, a ispitanici su pokazali znatno manju sklonost da napuste svoju pogrešnu hipotezu, čak i kad im je bila prezentirana falsificirajuća evidencija. Dakle, ne samo da nisu tragali za takvom evidencijom nego ju nisu uspjeli ni iskoristiti. Autori pretpostavljaju da je to stoga što u ovim uvjetima ni pogrešna hipoteza nije bilo dovoljno konsolidirana da bi bila uspješno testirana. Nadalje, autori su istaknuli moguću pozitivnu ulogu pristranosti potvrđivanja jer ona omogućuje da se u kompleksnim problemima stvori čvrsti temelj induktivnih generalizacija što je onda preduvjet za primjenu falsifikacijskih strategija u kasnijim koracima.

Razni induktivni zadaci mogu poslužiti za ispitivanje pristranosti potvrđivanja i sposobnosti za falsifikacijsko mišljenje. Jedan od takvih problema je razvio i poznati matematičar i promotor znanstvenog skepticizma Martin Gardner, koji je od 1957. do 1982. godine pisao kolumnu Matematičke igre za časopis *Scientific American*. Za tu kolumnu je razvio brojne igre i zadatke. Problemski zadatak koji je opisao je prikladan za ispitivanje pristranosti potvrđivanja. Riječ je o igri Eleusis (engl. *Game of Eleusis*), koja se igra

sa standardnim igračim kartama (Gardner, 1977). U toj igri četvero do osmero igrača sjedi za stolom i podijeljene su im igraće karte. Koriste se dva do tri špila standardnih igračih karata (ovisno o broju igrača). Djelitelj određuje tajno pravilo. Igrači nižu jednu po jednu kartu na stolu pri čemu djelitelj navodi uklapa li se karta u pravilo ili se ne uklapa i tu informaciju čuju svi igrači. Karte koje se ne uklapaju se nižu bočno od glavnog niza. Zadatak igrača je da otkriju pozadinsko pravilo i prvi koji to otkrije je pobjednik. Pravila mogu biti proizvoljna, no najbolja su ona koja nisu ni prejednostavna ni presložena. Recimo, Gardner navodi ovaj primjer dobrog pravila: „Ukoliko je zadnja legalno odigrana karta neparna, onda igraj crnu kartu. U suprotnom, igraj crvenu.“ Igra ima i svoj sustav bodovanja i opisana je u Gardnerovu (1977) radu.

Ovu igru su koristili Gorman i suradnici (1984) u istraživanju utjecaja suradnje i upute na pristranost potvrđivanja. Dobiveno je da suradnja nije poboljšala rezultate, dok je uputa o nepotvrđivanju hipoteze imala prednost pred uputom o potvrđivanju, što je donekle u suprotnosti s prijašnjim rezultatima (uputa, dakle, ipak djeluje na ispitanike). Čini se da tip i težina zadatka, te tip upute mogu utjecati na pristranost potvrđivanja, odnosno na pokušaje korištenja alternativne ili falsificirajuće evidencije.

Još jedan zadatak koji se svrstava u domenu znanstvenog rasuđivanja i u kojem se, između ostalog, manifestira pristranost potvrđivanja, jest zadatak detekcije kovarijacije (engl. *covariation detection task*). U tom zadatku ispitanici dobivaju kontingencijsku tablicu s rezultatima fiktivnog eksperimenta u kojem



SLIKA 3.1. Primjer igre Eleusis. Tajno pravilo je: „Poredak neparnih i parnih karata na kojemu se izmjenjuju crne i crvene“. (Ilustracija je preuzeta iz Walkup & Key (2020), str. 160.)

su uspoređene dvije situacije na način da su prikazane frekvencije pozitivnih i negativnih ishoda za obje situacije. Zadatak ispitanika je da procijeni koja od dvije situacije ima bolji efekt, odnosno da procijeni je li došlo do promjene uslijed primjene novog postupka ili tretmana (npr. da li je došlo do poboljšanja primjenom novog lijeka u odnosu na stari). Više varijacija tog zadatka mogu se vidjeti u radu Valerjeva i Dujmovića (2019). U ovim zadacima ispitanici manifestiraju pristranost potvrđivanja tako što često donose svoju procjenu samo usporedbom frekvencija pozitivnih ishoda dvije situacije i zanemaruju frekvencije negativnih ishoda.

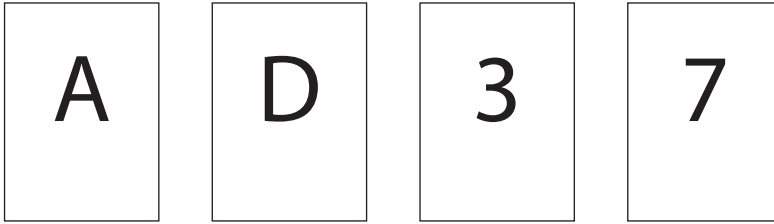
Dedukcija i pristranost potvrđivanja – Wasonov zadatak izbora

Pristranost potvrđivanja se može manifestirati i u zadacima koji su temeljeni na pravilima logike sudova koji pored induktivnog aktiviraju i deduktivno rasuđivanje. Najpoznatiji takav zadatak je Wasonov zadatak izbora poznat i kao problem s četiri karte (engl. *Wason selection task*; *Four card problem*), u kojem se zadaje kondicionalno pravilo koje sadrži operaciju implikacije. U osnovi, opet je riječ o zadatku s testiranjem hipoteza, s tim da ovdje hipotezu koju valja testirati nije generirao ispitanik već je eksplicitno navedena. Wasonov zadatak izbora je razvio Peter Wason

(1966) te je prvi put opisan u poglavlju koje je objavljeno u uredničkoj knjizi *New Horizons in Psychology*, te uskoro i u samostalnom znanstvenom članku (Wason, 1968). Wasonov zadatak izbora je postao jedan od najkorištenijih zadataka psihologije rasuđivanja, na kojem je demonstriran čitav niz zanimljivih efekata i koji je poslužio i kao bojište za argumentaciju i testiranje valjanosti različitih teorija deduktivnog mišljenja s najistaknutijim trima skupinama: teorije formalnih pravila zaključivanja, teorije specifičnih pravila zaključivanja i teorija mentalnih modela (na primjer, Braine et al., 1984; Cheng & Holyoak, 1985; Johnson-Laird & Byrne, 1991; Rips, 1983). Nadalje, zadatak je poslužio i kao paradigma za testiranje evolucijski utemeljene teorije socijalne izmjene o specijaliziranim i adaptivnim modulima rasuđivanja koji su osjetljivi na kršenje socijalnog ugovora (Cosmides, 1989; Cosmides & Tooby, 1989). Osim toga, Sperber i suradnici (1995) su rezultate iz ovog zadatka uklopili u teoriju relevantnosti po kojoj rezultati Wasonova zadatka proizlaze iz jezične pragmatike i socijalnog konteksta. U ovom tekstu nemamo prostora ni potrebe ulaziti u sve primjene i varijacije ovog zadatka. Stoga ćemo se uglavnom držati najčešćih varijanti koje se tiču pristranosti potvrđivanja.

Na Slici 3.2 je prikazana jedna od standardnih verzija Wasonova izbornog

Na stolu leže četiri karte. Svaka ima veliko slovo na jednoj strani i jednu znamenku na drugoj strani. Izložene strane su prikazane ovdje:



Pravilo koje je prikazano dolje se odnosi na ove četiri karte i ono može biti istinito ili neistinito.

Ako je A na jednoj strani karte, onda je 3 na drugoj strani karte.

Vaš je zadatak da odaberete one karte koje bi trebale biti okrenute kako biste otkrili je li pravilo istinito ili neistinito.

SLIKA 3.2. Standardni Wasonov zadatak izbora karata.

zadatka. Ponekad se ovakav tip zadatka naziva i apstraktna verzija zadatka jer ne sadrži nikakvo konkretno znanje, već samo opisuje odnose između slova i brojki. U ovom zadatku potrebno je testirati valjanost kondicionalnog pravila koje se odnosi na prikazane karte. Ispitanici mogu okrenuti bilo koji broj karata, iako je jasno da je okretanje sve četiri karte besmisleno. Trebamo se podsjetiti da se svako kondicionalno pravilo može logički zapisati kao $P \rightarrow Q$ („P implicira Q“), iako se često navodi i rečenična forma *Ako P, onda Q*. Ovaj sud je uvijek valjan osim ako je antecedens istinit i konsekvens neistinit. Kad tako razumijemo i pravilo navedeno gore u primjeru Wasonova zadatka, onda treba uočiti da četiri predložene karte uvijek sadrže stanja P , $ne-P$, Q i $ne-Q$ (dakako, redosljed se u eksperimentima može rotirati ili randomizirati). Točan odgovor je odabir karata

P i $ne-Q$. Dakle, u gornjem slučaju to bi bile karte A i 7. Odabir slučaja P (karta „A“) je jasan. Za valjano pravilo očekuje se da iza „A“ bude „3“, a u suprotnom pravilo nije valjano. Time se primjenjuje tip zaključka iz logike sudova koji se zove afirmacija antecedensa ili *modus ponens*.

Ako P, onda Q.

P .

Dakle, Q.

Međutim, odabir slučaja $ne-Q$ (karta „7“) je za ispitanike često težak. Iza karte „7“ ne smije biti „A“ da bi pravilo bilo točno. Ukoliko se ipak nalazi „A“, pravilo je nevaljano. Dakle, odabirom ove karte, po Wasonu, traži se negativna, falsificirajuća evidencija za zadano pravilo. Ovaj odabir se temelji na tipu zaključka iz logike sudova koji se zove negacija konsekvensa ili *modus tollens*.

Ako P, onda Q.

Nije Q.

Dakle, nije P.

Izbor preostale dvije karte (slučajevi *ne-P* i *Q*) je neispravan i nije informativan. Iza svake od te dvije karte može biti bilo koji znak i time ne saznajemo ništa o valjanosti kondicionalnog pravila. Odbiri tih pravila predstavljaju logičke forme koje se u logici sudova zovu negacija antecedensa i afirmacija konsekvensa i one su nevaljani zaključci.

Vrlo mali broj ispitanika uspijeva točno riješiti ovaj zadatak u apstraktnoj formi. Wason (1968) je dobio da su dva ispitanika od 26 (5.8%) točno riješili zadatak birajući *P* i *ne-Q*. Većina ispitanika je u tom eksperimentu (46%) birala netočnu kombinaciju *P* i *Q* i time su, prema Wasonu, manifestirali pristranost potvrđivanja. Ovi odnosi točnih i netočnih odgovora su replicirani velik broj puta pokazujući značajnu robusnost ovog efekta s određenim varijacijama u veličini koje su bile povezane s načinom zadavanja i odgovaranja na zadatak. Na primjer, u raznim varijacijama ovog eksperimenta (Valerjev, 2000; Valerjev & Pedisić, 2001) omjer točnih i pristranih odgovora za apstraktnu verziju Wasonova zadatka izbora je bio 14% naprema 74%, odnosno 7.8% naprema 46.9%. Izvjesne varijacije mogu proizaći u načinu odabira – ukoliko se od ispitanika traži da izabere proizvoljan broj karata ili izričito dvije karte. Na ovaj drugi način

se izbjegavaju kombinacije odgovora od jedne i tri karte te se dobivaju odgovori koji su preciznije fokusirani na mjerenje stupnja pristranosti i točnosti.

Standardno objašnjenje rezultata je da ispitanici tijekom rješavanja provode verifikaciju hipoteze, ali ne i falsifikaciju i time su pod utjecajem pristranosti potvrđivanja. Biranje karata *P* i *Q* je rezultat namjere da se pravilo potvrdi. Ne slažu se svi istraživači oko efekata koji su zapaženi u izvedbi Wasonova zadatka izbora. Prema prije navedenim autorima raznih teorija rasuđivanja, uzroci slabog učinka u zadatku su pojedine implikacije pojedinih teorija (koje su navedene gore), a prema drugima (Evans, 1972; Evans & Lynch, 1973; Evans, 1989) uzrok slabog učinka je jedan drugi tip kognitivne pristranosti – **pristranost podudaranja** (engl. *matching bias*). Po Evansu, pristranost podudaranja je tendencija ispitanika da preferiraju odgovore koji se podudaraju s vrijednostima koje su imenovane u pravilu. Po tome, ispitanici biraju karte *P* i *Q* jer se te karte podudaraju s vrijednostima u pravilu. Ako *P*, onda *Q*. Svoju pretpostavku je dokazivao koristeći varijacije pravila koje uključuju negacije. Na primjer, u zadatku s pravilom *Ako P onda nije Q* (na primjer, „Ako je na jednoj strani karte S, onda na poleđini nije 9.“), dio ispitanika je i dalje birao karte *P* i *Q* što je u novoj situaciji ispravan odgovor. Pristranost podudaranja vjerojatno ima utjecaj na odgovore u Wasonovu zadatku izbora,

no to ne isključuje utjecaj pristranosti potvrđivanja. Pristranost potvrđivanja je vjerojatno barem jedan od značajnih faktora koji se manifestiraju u ovom zadatku.

Do izrazito dramatične promjene u odgovorima u Wasonovu zadatku izbora dolazi kad se promjeni sadržaj kondicionalnog pravila te se umjesto apstraktnog sadržaja koristi konkretno poznato pravilo. Zamislimo pravilo: *Ako osoba pije pivo, onda mora biti starija od 19 godina*. I onda su tu prezentirane četiri karte/osobe koje možemo provjeriti kako bismo utvrdili poštuje li se ili krši pravilo. Te karte su [Pije pivo], [Pije kolu], [22 godine starosti] i [16 godina starosti]. Kako vidimo, u ovom zadatku, karte opet reprezentiraju slučajeve *P*, *ne-P*, *Q* i *ne-Q*. Takav su zadatak zadavali Griggs i Cox u svojim eksperimentima (Cox & Griggs, 1982; Griggs & Cox, 1982) te su ih nazivali tematskim zadacima. Točnost ispitanika skočila je s manje od 10% u apstraktnom zadatku na preko 70% u tematskom. To je jedan od najjačih efekata u eksperimentalnoj psihologiji i naziva se *efekt tematskog materijala*. Postavilo se pitanje povećava li bilo koji konkretni sadržaj uspješnost u Wasonovu zadatku izbora. Valerjev i Pedišić (2001) su uspoređivali apstraktne, konkretne-neutralne i konkretne-tematske zadatke koji uključuju socijalni ugovor (poput zadatka s pivom i punoljetnošću). Dobiveno je da tek konkretni-tematski zadaci znatno povećavaju uspješnost dok su konkretni-neutralni zadaci polučili slabe rezultate, slične apstraktnima.

Također je dobiveno da korištenje bikondicionala (*Samo ako P, onda Q*) ne pospješuje uspješnost, već ju naprotiv ograničava. Nadalje, u manipulaciji redoslijeda zadavanja zadataka dobiveno je i da nema transfera s tematskog na apstraktni zadatak. Objašnjenja uzroka efekta tematskog materijala su raznolika, no jedno od objašnjenja je da tematski materijal u znatnoj mjeri eliminira pristranost potvrđivanja, bilo zbog poznatosti pravila s kojima ispitanici već imaju iskustvo, bilo zbog situacije socijalnog ugovora na koju ljudi imaju evolucijsku adaptaciju.

Nadalje, pokazalo se i da efekt tematskog materijala nije skroz postojan i da nestaje u situaciji pod vremenskim pritiskom. U kronometrijskom eksperimentu sa Wasonovim zadacima koji su proveli Valerjev i Dujmović (2017) ispitanici su morali odabrati karte što brže, tj. izabrati inicijalni odgovor koji im intuitivno padne na pamet. Pokazalo se da su ispitanici u svim situacijama (apstraktnim, konkretnim-neutralnim i socijalnim) davali gotovo isključivo netočan *P* i *Q* odgovor, uz visoku razinu sigurnosti u taj odgovor, dok je postotak točnih odgovora bio manji od 5%. Ti rezultati pokazuju da je pristranost potvrđivanja, kao oblik heurističkog odgovora, značajno izraženija u situacijama gdje su vremenski resursi ograničeni.

Glavni zaključak ovog dijela teksta glasi: ukoliko se Wasonov zadatak izbora koristi kao paradigma ispitivanja pristranosti potvrđivanja, onda se mogu koristiti različite varijacije toga zadatka kojima

se može manipulirati sadržajem, socijalnim ugovorima, relevantnošću, formom zadatka, tipom kondicionala (kondicional/bikondicional), tipom logike (klasično kondicionalno pravilo ili deontičko pravilo), uputom za biranje karata, uputom

o brzini odgovaranja te vježbom ili nagovještajem prije zadatka. Pri tome treba imati na umu da ovaj zadatak može aktivirati različite kognitivne procese, a ne samo proces koji je vođen pristranošću potvrđivanja.

Istraživanja testiranja hipoteza

Još jedan oblik istraživanja kojim se mogu proučavati pristranosti potvrđivanja i, u općenitijem smislu, procesi testiranja hipoteza, jest proučavanje kronologije stvarnih otkrića u povijesti znanosti. Poznati takav primjer je detaljno proučavanje istraživačkog procesa znanstvenika Michaela Faradaya u domeni elektromagnetizma i elektrokemije (Tweney et al., 2005). Proučena je znatna količina njegovih istraživačkih dokumenata te je opažen obrazac u njegovom pristupu. Čini se da se Faradayeva eksperimentalna strategija temeljila na tzv. dvostupanjskom procesu. U prvom koraku je Faraday koristio usko pretraživanje kako bi našao potvrđujuću evidenciju za nove ideje u razvoju, a ignorirajući nepotvrđujuću evidenciju. Iako djeluje kao da je ovaj korak manifestacija pristranosti potvrđivanja (možda i je), ona u stvari ima pozitivan efekt – smanjuje problemski prostor i omogućuje istraživaču da se fokusira na jasne korake. Kasnije, u drugom koraku se uzima u obzir nepotvrđujuća evidencija i čak se aktivno koristi kako bi se nastojale odbaciti ideje nastale u ranoj fazi.

To, zapravo, predstavlja korak pokušaja falsifikacije hipoteze koji je po Popperu neophodan u znanstvenom procesu. Ova je strategija, nazvana „rano potvrdi kasnije odbaci“ (engl. *confirm early-disconfirm late*), uočena i u radu drugih znanstvenika, npr. kod Alexandera Grahama Bella. U proučavanju rada laboratorijskih biologa ([Fugelsang et al., 2004](#)) pokazalo se da znanstvenici u stvarnom znanstvenom radu nisu skloni olako odbaciti hipotezu ili teoriju uslijed negativne evidencije. Stvarni znanstvenici revidiraju svoje teorije na postepeni, a ne nagli način, koji više nalikuje statističkom rasuđivanju nego logičkom zaključku tipa sve-ili-ništa. Potrebna su ponovljena opažanja nepotvrđujuće evidencije da se ona prihvati. Za više detalja o istraživanjima znanstvenog rasuđivanja vidjeti djelo Gorman i suradnika (2005).

Iz navedenog slijedi da pristranost potvrđivanja nije uvijek negativna pojava koja rezultira rigidnim mišljenjem. Kada traganje za potvrđujućom evidencijom stavimo u širi kontekst testiranja hipoteza, može se vidjeti da ona u određenim

fazama ili okolnostima može imati pozitivan učinak. U tom smislu vjerojatno je bolje za odgovarajuće situacije usvojiti noviji termin koji nema negativnu konotaciju: *strategija pozitivnog testiranja*. Na primjer, ukoliko ljudi sumnjaju u grešku u podacima testiranje hipoteze je otežano. Nadalje, baratanje hipotezama i evidencijom opterećuje radno pamćenje. Za to vrijeme teško je aktivirati i testirati alternativnu hipotezu. To pokazuju eksperimenti s irelevantnim sekundarnim zadatkom koji dodatno opterećuje radno pamćenje te je za to vrijeme nemoguće promijeniti hipotezu ili razmatrati više od jedne (Dunbar & Klahr, 2012).

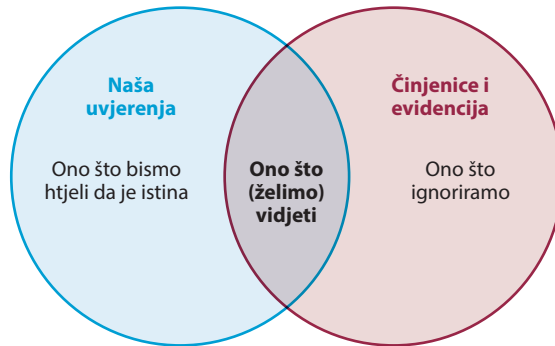
Istraživanja koja su 1960-ih počela s problemom 2 4 6 i istraživanjem pristranosti potvrđivanja, evoluirala su tijekom vremena u istraživanja testiranja hipoteza i istraživanja znanstvenog rasuđivanja. Nadovezujući se na spoznaje iz takvih istraživanja, te nadahnut i vlastitom revidiranom teorijom dvojnog procesiranja (opisana u Evans, 2008), Evans je razvio teoriju hipotetičnog mišljenja (2006, 2007). Hipotetično mišljenje² (engl. *hypothetical thinking*) je mišljenje koje zahtjeva zamišljanje mogućih stanja svijeta te uključuje razne aspekte testiranja hipoteza, predviđanja, donošenja odluka, protučinjeničnog mišljenja (engl. *counterfactual thinking*), deduktivnog

rasuđivanja i pretpostavljenog rasuđivanja (engl. *suppositional reasoning*). Prema teoriji hipotetičnog mišljenja ljudi u raznim formama mišljenja djeluju na temelju tri načela – (1) singularnosti, (2) relevantnosti i (3) zadovoljavajućeg cilja. Prema načelu singularnosti, ljudi mogu razmatrati jednu po jednu hipotetsku mogućnost. Prema načelu relevantnosti, ljudi razmatraju hipotezu koja je najrelevantnija u tekućem kontekstu. A prema načelu zadovoljavajućeg cilja (engl. *satisficing principle*) mogućnosti su razmatrane u odnosu na tekući cilj i prihvaćene su dok su zadovoljavajuće. Dvojno procesiranje se manifestira u tome da heuristički procesi sudjeluju u konstrukciji relevantnog modela i izvođenja zaključaka iz njega, a analitički procesi ocjenjuju je li model zadovoljavajući. To znači da, prema ovoj teoriji, ljudi na temelju raspoloživih informacija stvaraju hipotetski model i većinu vremena mogu razmatrati tek jedan model koji im se čini relevantan (načela 1 i 2). Takav model se održava sve dok je zadovoljavajući. U skladu s time, evidencija se nastoji uklopiti u model koji djeluje relevantno i zadovoljavajuće. Kad prestaje biti dovoljno zadovoljavajući, traži se alternativna evidencija.

Zaključno, traganje za potvrđujućom evidencijom može biti dio heurističke pristranosti. Ona proizlazi iz predstavljenog

² Termin „hipotetsko mišljenje“ bi zvučao kao da je sam proces mišljenja upitan, pod hipotezom. Umjesto toga, odabrao sam termin „hipotetično mišljenje“ jer se radi o mišljenju o hipotezama.

PRISTRANOST POTVRĐIVANJA



Tendencija da tražimo, interpretiramo i pamtimo informaciju koja potvrđuje naša uvjerenja

(Ilustraciju konceptualno osmislio Pavle Valerjev.)

dvoprocenog modela, pri čemu se hipoteza održava dok se dogodi suočavanje s nepotvrđujućom evidencijom (a nekad i nakon toga). No, s druge strane, obzirom da se pokazalo da testiranje hipoteza uglavnom ne funkcionira po „poperovskim“ načelima, već prije po „bajesovskim“ (temeljenim na vjerojatnosti i postepenom mijenjanju snage uvjerenja), i kako je prema teoriji hipotetičnog mišljenja način testiranja hipoteza vođen načelom relevantnosti, pozitivno

testiranje nije nužno štetno. Prema tom objašnjenju, kad zamišljamo određenu mogućnost onda stvaramo mentalnu reprezentaciju s plauzibilnim sadržajem za koji je vjerojatno da je istinit slučaj ukoliko je hipoteza istinita. I onda se prema tome bira relevantna evidencija koja je pozitivna. Dakle, na pozitivno testiranje može se gledati i kao na prirodan proces koji u realnim situacijama testiranja hipoteza dobavlja zadovoljavajuće hipoteze i objašnjenja o svijetu.

Literatura

- Braine, M.D.S., Reiser, B.J., & Romain, B. (1984). Some empirical justification for a theory of propositional logic. U G.H. Bower (Ur.), *The Psychology of Learning and Motivation*. Vol. 18. New York: Academic Press. [https://doi.org/10.1016/S0079-7421\(08\)60365-5](https://doi.org/10.1016/S0079-7421(08)60365-5)
- Cheng, P.W., & Holyoak, K.J. (1985). Pragmatic reasoning schemas. *Cognitive Psychology*, 17(4), 391-416. [https://doi.org/10.1016/0010-0285\(85\)90014-3](https://doi.org/10.1016/0010-0285(85)90014-3)
- Cosmides, L. (1989). The logic of social exchange: Has natural selection shaped how

- humans reason? Studies with the Wason selection task. *Cognition*, 31(3), 187–276. [https://doi.org/10.1016/0010-0277\(89\)90023-1](https://doi.org/10.1016/0010-0277(89)90023-1)
- Cosmides, L., & Tooby, J. (1989). Evolutionary psychology and the generation of culture, Part II. Case study: A computational theory of social exchange. *Ethology and Sociobiology*, 10, 51–97. [https://doi.org/10.1016/0162-3095\(89\)90013-7](https://doi.org/10.1016/0162-3095(89)90013-7)
- Cox, J.R., Griggs, R.A. (1982). The effects of experience on performance in Wason's selection task. *Memory & Cognition*, 10, 496–502 <https://doi.org/10.3758/BF03197653>
- Dobelli, R. (2013). *The Art of Thinking Clearly*. Sceptre.
- Dunbar, K. N., & Klahr, D. (2012). Scientific thinking and reasoning. In K. J. Holyoak & R. G. Morrison (Eds.), *The Oxford Handbook of Thinking and Reasoning*. Oxford University Press. <https://psycnet.apa.org/doi/10.1093/oxfordhb/9780199734689.013.0035>
- Evans, J. St. B. T. (1972). Interpretation and matching bias in a reasoning task. *Quarterly Journal of Experimental Psychology*, 24(2), 193–199. <https://doi.org/10.1080/0033557243000067>
- Evans, J. St. B. T. (1989). *Bias in Human Reasoning*. Lawrence Erlbaum Associates, Ltd.
- Evans, J. St. B. T. (2006). The heuristic-analytic theory of reasoning: Extension and evaluation. *Psychonomic Bulletin & Review*, 13(3), 378–395. DOI: [10.3758/bf03193858](https://doi.org/10.3758/bf03193858)
- Evans, J. St. B. T. (2007). *Hypothetical Thinking: Dual Processes in Reasoning and Judgement*. Taylor & Francis Routledge.
- Evans, J. St. B. T. (2008). Dual-processing accounts of reasoning, judgment, and social cognition. *Annual Review of Psychology*, 59, 255–78. DOI: [10.1146/annurev.psych.59.103006.093629](https://doi.org/10.1146/annurev.psych.59.103006.093629)
- Evans, J. St. B. T., & Lynch, J. S. (1973). Matching bias in the selection task. *British Journal of Psychology*, 64(3), 391–397. <https://doi.org/10.1111/j.2044-8295.1973.tb01365.x>
- Feist, G. J., & Gorman, M. E. (Eds.). (2013). *Handbook of the Psychology of Science*. Springer Publishing Company.
- Fugelsang J. A., Stein C. B., Green A. E., Dunbar K. N. (2004). Theory and data interactions of the scientific mind: evidence from the molecular and the cognitive laboratory. *Canadian Journal of Experimental Psychology*, 58(2), 86–95. <https://doi.org/10.1037/h0085799>
- Gardner, M. (1977). Mathematical games. *Scientific American*, 237(4), 18–25. <http://www.jstor.org/stable/24953955>
- Gorman, M. E., Gorman, M. E., Latta, R. M., & Cunningham, G. (1984). How disconfirmatory, confirmatory and combined strategies affect group problem solving. *British Journal of Psychology* 75(1), 65–79. <https://psycnet.apa.org/doi/10.1111/j.2044-8295.1984.tb02790.x>
- Gorman, M. E., Tweney, R. D., Gooding, D. C., & Kincannon, A. E. (Eds.). (2005). *Scientific and Technological Thinking*. Lawrence Erlbaum Associates, Publishers.
- Griggs, R. A., & Cox, J. R. (1982). The elusive thematic-materials effect in Wason's selection task. *British Journal of Psychology*, 73(3), 407–420. <https://doi.org/10.1111/j.2044-8295.1982.tb01823.x>
- Johnson-Laird, P. N., & Byrne, R. M. J. (1991). *Deduction*. Lawrence Erlbaum Associates, Inc.

- Klayman, J., & Ha, Y.-w. (1987). Confirmation, disconfirmation, and information in hypothesis testing. *Psychological Review*, 94(2), 211-228. <https://doi.org/10.1037/0033-295X.94.2.211>
- Manktelow, K. (2012). *Thinking and Reasoning: An Introduction to the Psychology of Reason, Judgment and Decision Making*. Psychology Press.
- Mynatt, C. R., Doherty, M., & Tweney, R. (1977). Confirmation bias in a simulated research environment: An experimental study of scientific inference. *Quarterly Journal of Experimental Psychology*, 29(1), 85-95. <https://doi.org/10.1080/0033557743000053>
- Mynatt, C. R., Doherty, M. E., & Tweney, R. D. (1978). Consequences of confirmation and disconfirmation in a simulated research environment. *Quarterly Journal of Experimental Psychology*, 30(3), 395-406. <https://doi.org/10.1080/0033557843000007>
- Popper, K. (2002). *The Logic of Scientific Discovery*. Routledge.
- Rips, L. J. (1983). Cognitive processes in propositional reasoning. *Psychological Review*, 90(1), 38-71. <https://doi.org/10.1037/0033-295X.90.1.38>
- Sagan, C. (2000). *Svijet proganjen demonima: Znanost kao svijeća u tmini*. Naklada Jesenski i Turk. [Originalno izdanje Sagan, C. (1995) *The Demon-Haunted World: Science as a Candle in the Dark*. Random House]
- Sperber, D., Cara, F., & Girotto, V. (1995). Relevance theory explains the selection task. *Cognition*, 57(1), 31-95. [https://doi.org/10.1016/0010-0277\(95\)00666-M](https://doi.org/10.1016/0010-0277(95)00666-M)
- Teovanović, P., & Berthet, V. (2024). On a generality of confirmation bias: Individual differences perspective. In P. Valerjev, & I. Tucak Junaković (Eds.), *23rd Psychology Days in Zadar: Book of Selected Proceedings* (pp. 113-119). University of Zadar. <https://doi.org/10.15291/9789533314792>
- Tweney, R. D., Mears, R. P. & Spitzmüller, C. (2005). Replicating the practices of discovery: Michael Faraday and the interaction of gold and light. In M. E. Gorman, R. D. Tweney, D. C. Gooding, & A. E. Kincannon (Eds.), *Scientific and Technological Thinking*. Lawrence Erlbaum Associates, Publishers.
- Valerjev, P. (2000). Wasonov izborni zadatak: Stabilnost efekata pristranosti i tematskog materijala. *Radovi: Razdio filozofije, psihologije, sociologije i pedagogije*, 39, 16, 97-111. <https://doi.org/10.15291/radovifpsp.2537>
- Valerjev, P., & Dujmović, M. (2017). Metacognitive judgments during solving of Wason selection task. In N. Đapo, S. Zvizdić, A. Dautbegović & M. Marković (Eds.), *Zbornik radova: Četvrti sarajevski dani psihologije* (pp. 83-93). Filozofski fakultet Univerziteta u Sarajevu.
- Valerjev, P. & Dujmović, M. (2019). Performance and metacognition in scientific reasoning: The covariation detection task. *Psihologijske teme*, 28(1), 93-113. <https://hrcak.srce.hr/psihologijske-teme>
- Valerjev, P., & Pedisić, A. (2001). Wasonov izborni zadatak: Utjecaj upute, tipova kondicionala i tematskog materijala. *Razdio filozofije, psihologije, sociologije i pedagogije*, 40, 17, 45-64. https://www.researchgate.net/publication/326353880_Wasonov_izborni_zadatak_Utjecaj_upute_tipova_kondicionala_i_tematskog_materijala

- Walkup, J. R., & Key, R. (2020). Speed Eleusis: Variation on an old educational card game. *The Physics Teacher*, 58, 160-163. <https://doi.org/10.1119/1.5145403>
- Wason, P. C. (1960). On the failure to eliminate hypotheses in a conceptual task. *Quarterly Journal of Experimental Psychology* 12(3), 129-140. <https://doi.org/10.1080/17470216008416717>
- Wason, P. C. (1966). Reasoning. In B. Foss (Ed.), *New Horizons in Psychology*. Penguin Books. pp. 135-151.
- Wason, P. C. (1968). Reasoning about a rule. *Quarterly Journal of Experimental Psychology*, 20(3), 273-281. <http://dx.doi.org/10.1080/14640746808400161>
- Wason, P. C., & Johnson-Laird, P. N. (1972). *Psychology of Reasoning: Structure and Content*. Harvard University Press.

ABSTRACT

This chapter describes a cognitive bias known as confirmation bias. The first part of the text presents examples and a definition for this bias. Following this, we discuss this bias as an important factor in the context of the philosophy of science, where the demand for the falsifiability of scientific theories was articulated. Confirmation bias makes that requirement very difficult. We present the main research paradigms in psychological research in which this bias is manifested. These are primarily the 2 4 6 problem and Wason's choice task and variations of these tasks. Finally, confirmation bias is presented in the broader context of research into the thought processes of hypothesis testing. Within this context, this bias is also called the positive test strategy. It is indicated that under certain circumstances this strategy represents a constructive part of the research process, and then it cannot be seen as bias.

Keywords: confirmation bias, positive test strategy, falsifiability, inductive reasoning, Wason selection task