

Radovi Znanstvenog centra “Berislav Žarnić”

**Uredili: Gabriela Bašić Hanžek,
Ljudevit Hanžek, Dario Škarica**



IZDANJA FILOZOFSKOG FAKULTETA
SVEUČILIŠTA U SPLITU
EDITIONES FACULTATIS PHILOSOPHICAE
UNIVERSITATIS SPALATENSIS

Nakladnik

Sveučilište u Splitu, Filozofski fakultet
Poljička cesta 35, 21000 Split

Odgovorni urednik

prof.dr.sc. Ina Reić Ercegović,

Recenzenti knjige

prof. dr. sc. Hrvoje Relja
dr. sc. Ivana Skuhala Karasman, viša znanstvena suradnica

Recenzenti pojedinačnih priloga

dr. sc. Josip Balabanić (hon. prof., znanstveni savjetnik), izv. prof. dr. sc. Tomislav Bracanović, dr. sc. Dušan Dožudić (znanstveni suradnik), izv. prof. dr. sc. Aleksandra Golubović, prof. dr. sc. Filip Grgić, prof. dr. sc. Kristijan Krkač, prof. dr. sc. Stipe Kutleša, prof. dr. sc. Davor Lauc, dr. sc. Ivan Restović, doc. dr. sc. Ines Skelac, prof. dr. sc. Zvonimir Šikić, prof. dr. sc. Danilo Šuster, prof. dr. sc. Jure Zovko

Korektura i priprema za tisak

Božo Dujmović

ISBN (Tisak) 978-953-352-081-0

ISBN (Online) 978-953-352-192-3

Online izdanje odobrilo vijeće Filozofskog fakulteta
odlukom od 22. 9. 2026.



RADOVI ZNANSTVENOG CENTRA “BERISLAV ŽARNIĆ”

urednici

Gabriela Bašić Hanžek

Ljudevit Hanžek

Dario Škarica



Sveučilište u Splitu,
Filozofski fakultet

Split, 2026.

Riječ urednika

Znanstveni centar “Berislav Žarnić” osnovan je 2016. godine kao ustrojbeni jedinicu Filozofskog fakulteta u Splitu, s ciljem promocije, širenja i poticanja izvornog znanstvenog rada na području teorijske filozofije, u prvom redu logike, epistemologije i filozofije znanosti. Među aktivnostima koje su primjerene ostvarivanju tih ciljeva svoje mjesto svakako nalazi i objavljivanje znanstvenih radova kako članova, tako i suradnika Centra. Radovi sabrani u ovom zborniku svojom tematskom raznorodnošću čine svojevrsan kolaž izabranih problema u teorijskoj filozofiji, što je sasvim u skladu s idejom na kojoj je Centar utemeljen, kao mjesto razvoja teorijske filozofije u najširem smislu te riječi, i odgovara planu koji je osnivač centra, pokojni profesor Berislav Žarnić (1959.–2017.) imao prilikom pokretanja centra.

Srećko Kovač (Institut za filozofiju, Zagreb) u članku “O izvorima i temeljima formalizacije” analizira odrednice formalnih sustava i formalizacije koje je u filozofiji moguće pratiti već od Aristotela, unatoč činjenici da su standardi moderne formalizacije precizno opisani tek krajem 19. i početkom 20. stoljeća. K tomu, autor opisuje i način na koji je egzaktnost dokazivanja uvjetovana osjetilnom zornošću i slikovnošću, izlažući i o strojnom karakteru formalnih sustava na temelju Turing-Gödelove definicije formalnoga sustava kao Turingova stroja. Naposljetku, na temelju Gödelovih i Wittgensteinovih refleksija autor ocrtava teleološki karakter temeljnoga pojma dokazljivosti te elaborira smisao u kojem se “životne forme”, iz Wittgensteinove druge faze, mogu shvatiti kao izvorište dokazljivosti.

Goran Lojkić (Institut za filozofiju, Zagreb) u radu “Russellov Russellov paradoks” analizira Russellove formulacije po njemu imenovanog glasovitog paradoksa u naivnoj teoriji skupova: ako svaki predikat određuje skup svih predmeta koji ga zadovoljavaju, što s predikatom “nije član sa-

moga sebe” i skupom koji je određen tim predikatom? Zadovoljava li taj skup taj predikat? Ako da, onda nije član samoga sebe; ali ako nije član samoga sebe, onda ne zadovoljava taj predikat, jer njegovi su članovi svi entiteti koji zadovoljavaju taj predikat. Autor ekstenzivno dokumentira Russellove pokušaje analize tog paradoksa i njegova rješavanja, što uključuje i analizu pojmova usko povezanih s paradoksom: stavačne funkcije i razreda. Autor ističe Russellovu frustraciju nezadovoljavajućim rješenjima paradoksa, koja na koncu vodi do teorije tipova u remek-djelu *Principia Mathematica*.

Nenad Smokrović (Sveučilište u Rijeci) u tekstu “Može li se naučiti kritički misliti?” raspravlja o fenomenu kritičkog mišljenja, koji se posljednjih desetljeća našao u središtu brojnih debata u edukativnim krugovima. Smokrović ističe da suvremeni diskurs o kritičkom mišljenju uključuje dvije tvrdnje: prvo, kritičko je mišljenje specifična sposobnost zaključivanja koja se jasno može razlikovati od i koja ne uključuje druga epistemička svojstva, prije svega epistemičke vrline, i drugo, kritičko je mišljenje sposobnost koju može steći svaki normalni, prosječni subjekt, odnosno, to je sposobnost koja se edukacijom može (i treba) u takvom subjektu razviti. Autor uočava da takvo shvaćanje kritičkog mišljenja implicira da je subjekt koji je uvježban za kritičko mišljenje također sposoban kritički misliti u svakoj domeni realnosti, kako u onima o kojima posjeduje činjenično znanje, tako i u onima za koje nije stručan. Autor takav stav kritizira i smatra da je jedini način na koji je kritičko mišljenje moguće naučiti onaj u kojem se kritičko mišljenje konceptualizira kao epistemička vrlina koju je moguće razvijati, a ne kao poseban oblik zaključivanja.

Luca Malatesti i Predrag Šustar (obojica sa Sveučilišta u Rijeci) u tekstu “Naturalizam ‘quattro stagioni’” raspravljaju o odnosu filozofije i znanosti, tj. o tome kakav bi taj odnos trebao biti. U debati između diskontinuiteta i kontinuiteta, autori se odlučuju za verziju kontinuiteta, ali umjerenog tipa. Svoj pristup pojašnjavaju na primjeru dviju debata: prvo, rasprava o genetičkoj informaciji u biologiji, u okviru koje autori primjećuju da se pojmovna analiza svojstvena filozofiji pokazala korisnom u razlikovanju između pojmova “informacije” i “uzroka”, te je na taj način pridonijela jasnoći u znanstvenoj debati. Drugi je primjer

važnost neuroznanstvenih spoznaja o ljudskom mozgu i ponašanju u razrješenju sudskih procesa, posebno pri procjeni nečije neubrojivosti. Malatesti i Šustar ističu da takva praksa zahtijeva jasan i praktičan pojmovni okvir, koji omogućava pripisivanje mentalnih stanja subjektima, u čemu vide ulogu i filozofije. Na koncu, autori naglašavaju da se njihova analiza navedenih slučajeva dobro slaže s umjerenim metafilozofskim naturalizmom, prema kojemu postoji kontinuitet i interakcija između filozofije i znanosti, koja omogućava filozofski doprinos znanstvenim debatama, uz jasno ograničenje rezultatima dotične znanstvene discipline.

Tonći Kokić (Sveučilište u Splitu) u radu "On the concept of design as a very general idea in philosophy of science" raspravlja o često kontroverznoj ideji dizajna u biologiji, ili znanosti općenito. Kokićev je stav da pitanje dizajna nije isključivo metodološki problem, neovisno o metafizičkim ili epistemološkim stavovima znanstvenika. Danas uobičajeni negativan stav prema bilo kakvom pozivanju na dizajn u znanstvenom objašnjenju ima svoju metafizičku i epistemološku pozadinu, koja je prešutno prevedena u metodološki naputak. Autor također primjećuje da je dizajn kao moguće svojstvo fizičkih procesa često nejasno definiran i donosi neke prijedloge, utemeljene na ideji procesa koji se odvijaju suprotno uobičajenom prirodnom tijeku takvih i sličnih procesa. Uvođeci analogiju crnih monolita i ljudskih organizama u filmu *2001: Odi-seja u svemiru*, autor brani tezu da su objašnjenja s pomoću dizajna često prerano odbačena, zaključujući da pojam dizajna može imati znanstveni legitimitet kao vrlo općenita ideja u znanstvenoj praksi, slična uvjerenju u uniformnost prirode ili zahtjevu provjerljivosti za znanstvene teorije.

Boran Berčić (Sveučilište u Rijeci) u svome tekstu "Novi relativizam: Može li relativist reći da Hitler nije trebao narediti istrebljenje Židova?" raspravlja o moralnom relativizmu, promatrajući ga kroz prizmu "novog relativizma", ili analitičkog relativizma u filozofiji jezika. Berčić ističe važnost kontekstualističkih elemenata u formulacijama novog relativizma (npr. Ističe važnost konteksta izricanja moralnih sudova i konteksta njihova vrednovanja), te iznosi ozbiljne prigovore takvom kontekstualističkom pristupu: prvo, kontekstualizam ne može dobro objasniti fenomen moralnog neslaganja među različitim subjektima; drugo, kontekstualizam ne može dobro objasniti fenomen revizije mo-

ralnog stava, tj. nečije promjene stava o nekom moralnom pitanju. Zbog toga Berčić dovodi u pitanje i samu mogućnost koherentne formulacije novog relativizma.

Marin Biondić (Sveučilište u Rijeci) u tekstu "Metafizika vremena i šteta smrti" razmatra ulogu metafizičkih teorija vremena u suvremenoj analitičkoj raspravi o vrijednosti smrti za osobu koja umre. U prvom dijelu iznosi pregled i implikacije dviju temeljnih metafizičkih teorija vremena, eternalizma i prezentizma. U drugom dijelu ukratko iznosi osnovne probleme rasprave o vrijednosti smrti te problem nepostojećeg subjekta, a zatim prelazi na rješenja problema nepostojećeg subjekta u eternalističkom i prezentističkom metafizičkom okviru. U završnom dijelu Biondić iznosi implikacije i probleme eternalističkog i prezentističkog rješenja nepostojećeg subjekta te analizira tvrdnju da je opravdanje vrijednosti smrti neutralno prema fundamentalnoj metafizici. U konačnici, autor zagovara tezu o neraskidivom odnosu fundamentalnih metafizičkih teorija i vrijednosti u raspravi o šteti smrti.

Filip Čeč (Sveučilište u Rijeci) u tekstu naslovljenom "Je li problem slobode volje pseudoproblem?" tematizira ideju, blisku klasičnim kompatibilistima (Hume, Hobart, Schlick, Ayer), da je problem slobode volje zapravo pseudoproblem, tek puka verbalna rasprava, a ne sadržajna debata. Čeč analizira argumente Marka Balaguera, koji u sklopu svoje libertarijaniističke pozicije postavljene na naturalističkim temeljima argumentira da su kompatibilističko pitanje kao i pitanje "Što je sloboda volje?" trivijalna, semantička pitanja lišena metafizičkog naboja. Čeč nastoji pobliže oslikati u čemu se sastoji shvaćanje da problem slobode volje predstavlja pseudoproblem te nudi argumente kojima je za cilj dokazati da se u raspravi o slobodi volje ne iznose besmisleni ili trivijalni iskazi nego naprotiv, da se vodi smisljena, punokrvna metafizička rasprava.

Davor Pećnjak (Institut za filozofiju, Zagreb) u članku "Alternativne mogućnosti i odgovornost: diskusija o načelu alternativnih mogućnosti" raspravlja o principu alternativnih mogućnosti, jednom od središnjih pojmova u debati o slobodi volje. Princip alternativnih mogućnosti tvrdi da je neka radnja slobodna samo ako je djelatnik imao mogućnost

učiniti drugačije, tj. ne izvršiti tu radnju. Harry Frankfurt iznio je poznati protuprimjer za tu tezu: Black na neki način osigura postojanje mehanizma koji će natjerati Jonesa da izvrši radnju A, ako Jones sam ne donese odluku za vršenje A. Međutim, u danoj situaciji, Jones bez prisile donese tu odluku i mehanizam se ne aktivira. Frankfurt tvrdi da je Jonesova radnja slobodna, i on je za nju moralno odgovoran, unatoč tome što uvjeti iz principa alternativnih mogućnosti nisu zadovoljeni. Pećnjak komentira Ginetovu kritiku Frankfurta, u kojoj Ginet uvođenjem pojma temporalne specifičnosti pokušava pokazati da se intuicije o moralnoj odgovornosti u tim slučajevima ipak mogu objasniti na način konzistentan s principom alternativnih mogućnosti. Pećnjak se ne slaže sasvim s Ginetom, i misli da je Frankfurt uspješno pokazao da princip alternativnih mogućnosti ne vrijedi generalno, ali misli da može vrijediti u nekim posebnim slučajevima, ovisno o detaljima njihova opisa.

Gabriela Bašić Hanžek
Ljudevit Hanžek
Dario Škarica

Sadržaj

SREĆKO KOVAČ	
O izvorima i temeljima formalizacije.....	13
GORAN LOJKIĆ	
Russellov Russellov paradoks.....	29
NENAD SMOKROVIĆ	
Može li se naučiti kritički misliti?.....	61
LUCA MALATESTI, PREDRAG ŠUSTAR	
Naturalizam “quattro stagioni”.....	77
TONČI KOKIĆ	
On the concept of design as a very general idea in philosophy of science.....	95
BORAN BERČIĆ	
Novi relativizam: može li relativist reći da Hitler nije trebao narediti istrjebljenje Židova?.....	117
MARIN BIONDIĆ	
Metafizika vremena i šteta smrti.....	135
FILIP ČEČ	
Je li problem slobode volje pseudoproblem?.....	153
DAVOR PEĆNJAK	
Alternativne mogućnosti i odgovornost: diskusija o načelu alternativnih mogućnosti.....	179

O izvorima i temeljima formalizacije

Srećko Kovač

Moderni standardi egzaktnosti znanja ostvarene u obliku formaliziranih teorija uvjetovali su bitne promjene i postavili nove logičko-metodološke zahtjeve u filozofiji i temeljima matematike: ne samo u smislu podizanja standarda egzaktnosti spoznaja nego i u smislu osvješćivanja povijesnih izvora kao i dubljih strukturnih temelja logike i formalizacije. Najprije ćemo pokazati u kojem se smislu bitne odrednice formalnih sustava i formalizacije mogu naći već u Aristotela, iako su standardi moderne formalizacije postavljeni tek kao rezultat promjena u logici krajem 19. i početkom 20. stoljeća. U idućem odjeljku opisujemo na koji je način egzaktnost dokazivanja uvjetovana osjetilnom zornošću i slikovnošću. Treći se odjeljak bavi strojnim obilježjem formalnih sustava na temelju Turing-Gödelove definicije formalnoga sustava kao Turingova stroja. Naposljetku, na temelju Gödelovih i Wittgensteinovih refleksija o crtavamo teleološki karakter temeljnoga pojma dokazljivosti te pokazujemo u kojem se smislu “životne forme” mogu shvatiti kao izvorište dokazljivosti.

1 Formalizacija i povijesno podrijetlo logike¹

Prema modernim standardima sigurnosti i egzaktnosti spoznaje, ne možemo biti potpuno zadovoljni nekom danom teorijom sve dok nije formalizirana, tj. prikazana u obliku formalnoga sustava. Formalizirana teorija treba precizno definirati svoj jezik (skup S rečenica) i izričito

¹ Odjeljci 1–3 prijevod su s manjim izmjenama većega dijela autorova priloga “Remarks on the origin and the foundations of formalisation”, Leiden, Boston: Brill | Rodopi, 2020., str. 163–179, © Brill, Leiden. Prijevodi su na hrvatski iz citirane literature autori ili modificirani.

navesti svoje aksiome; u logici je riječ o logičkim aksiomima i pravilima zaključivanja (relacija $\vdash$ izvedljivosti rečenica iz skupa rečenica), te se stoga logika može definirati kao uređen par $\langle S, \vdash \rangle$.² Takve su standarde ustanovili utemeljitelji moderne logike krajem 19. i početkom 20. stoljeća. U tom je kontekstu 1922. godine Łukasiewicz primijetio:

Nismo više zadovoljni običnim matematičkim dedukcijama, koje obično počinju negdje “u sredini”, pokazuju česte praznine i neprekidno se pozivaju na zor. ... Želimo znati aksiome na kojima je svaki sustav osnovan i pravila zaključivanja koja rabi.

Łukasiewicz 1967, 20

Osvrćući se na filozofiju svojega vremena, Łukasiewicz je još kritičniji:

Filozofija se mora rekonstruirati iz samih svojih temelja; treba se nadahnuti znanstvenom metodom i osnivati se na novoj logici.

Łukasiewicz 1967, 21

Također dodaje: “To je rad za generacije i umove mnogo snažnije od onih koji su već rođeni” (str. 21). No ti moderni standardi imaju svoje povijesne začetak još u starome vijeku. Stoga se posebno osvrćemo na rad Łukasiewicza i Bocheńskoga, koji su se podrobno i više nego velika većina drugih logičara u ranoj fazi razvoja moderne logike posvetili i povijesti logike u njezinu suodnosu sa suvremenom logikom, mijenjajući perspektivu gledanja na povijesne logičke teorije, raspolažući i potrebnim filološko-povijesnim kompetencijama. Na taj su se način mogli modernom logičkom metodologijom nadovezati na pojedina otvorena pitanja iz povijesti logike i aktualizirati ih. Łukasiewicz i Bocheński pokazali su da je Aristotel razvio deduktivni sustav “formalne logike” te postavio opću teoriju aksiomatizacije. Doduše, zaključili su da Aristotelova formalna logika nema obilježje formalizirane logičke teorije (“formalnoga sustava” u strogoj smislu), opisane u rigorozno definiranom jeziku, držeći da je to postignuće stoičkih logičara.³ Međutim, istakli su

² V. npr. Tarskiovo rano shvaćanje u (Tarski 1983). Usp. Béziau (2005).

³ Łukasiewicz (1957) katkad gotovo isključivo naglašuje jezični izraz logičkih oblika te nije čuo da je Aristotel uspio postaviti logički formalizam zbog njegova neispravnoga supstituiranja konkretnih pojmova u logičke oblike i sinonimne uporabe logičkih pojmova (npr. “pripada” i “prirodi se”, Łukasiewicz 1957, 17–18, v. Bocheński 2002, 113). Łukasiewicz priznaje da je Aristotel razvio aksiomatski sustav silogizama (npr., Łuka-

da su Aristotelov pojam i izgradnja “formalne logike” zasnovani ponajprije na njegovoj uporabi pojmovnih slova (varijabla), što mu je omogućilo da apstrahira od pridruženoga značenja rečenica i usredotoči se isključivo na oblik (rečenice), koji je bitan za deduktivni zaključak. Kako je naglasio Bocheński, Aristotelova je formalna logika obilježena porabom varijabla i bavi se “formulama”, tj. rečenicama (uključujući zakone) u kojima su opisni pojmovi (“konstante”) zamijenjeni varijablama.⁴ Štoviše, kako nastavlja Bocheński, tu je i sljedeći “prijelomni” Aristotelov prinos logici: aksiomatizacija (“usprkos njezinoj slabosti”) silogistike “kategoričnih” iskaza (tj. oblika podmet-prirok) (Bocheński 2002, npr. str. 3–5, 74, 84–87). Prema Bocheńskom, Aristotelova opća teorija “kategoričnoga” silogizma, kako je prikazana u *Prvoj analitici*, povijesna je paradigma onoga što treba razumjeti kao formalnu logiku.

Neovisno o nedostacima u Aristotelovu prikazu i formulacijama,⁵ očito je da je njegovo zasnivanje formalne logike i opće teorije aksiomatizacije smjeralo na rigorozne norme (osobito ako imamo na umu vrijeme kada Aristotel djeluje) i da je predstavljalo znatan skok prema postizanju potpuno formalizirane logike. Iako je Aristotel tvrdio da je silogizam stvar “unutrašnjega govora” (ἔσω λόγος, *An. Post.* A10 76b 27), želio je implementirati svoju teoriju silogizma i aksiomatske teorije općenito u prikladan, umjetni jezik, ne tek rabeći pojmovna slova (“varijable”) nego također izražavajući ustroj rečenica i zaključaka na nedvosmislen način, pa i u slučaju kad taj ustroj odstupa od običnoga načina govora.⁶ Stoga je u svojim *Analitikama* jasno razlikovao logičke riječi (‘svi’, ‘ne’, ‘neki’, ‘pripada’) i izražavao logički ustroj rečenica. Običan način kazivanja “Svi B jesu A” (koji je rabio ranije, npr. u *O tumačenju*, još uvijek bez pojmovnih slova) preoblikovao je u “A pripada svakomu B” (ili, metateorijski, “A se pririče svakomu B”); analogno vrijedi i za druge rečenice takozvanoga “logičkoga kvadrata”. Često je silogizam izražavao

siewicz 1957, str. 44, 73; osobito v. Bocheński 2002, 74, 81, 84, 86), premda ne zadovoljavajući moderne “formalističke” standarde.

⁴ “[P]oraba slova umjesto konstantnih riječi rodila je formalnu logiku” (Bocheński 2002, 80–81).

⁵ V. npr. u Łukasiewicz (1957, 16–19).

⁶ Aristotelovo “strukturno razilaženje između apstraktnih i konkretnih oblika silogizama”, koje ističe Łukasiewicz, (1957, 17), može se smatrati upravo potvrdom umjetnoga Aristotelova idioma kako bi se egzaktno izrazili logički oblici.

u obliku pogodbene rečenice. Tim je sredstvima Aristotel mogao strogo izraziti nužan “slijed” zaglavka iz njegovih premisa (*An. Pr.* A1 24b 18–21), osobito prijelaznost priricanja: Ako se A pririče svakomu B, a B svakomu C, onda se A pririče svakomu C. To je Aristotelu omogućilo da učini izričitim sve implicitne pretpostavke (npr. obrat u drugoj i trećoj figuri silogizma) kako bi se zaključivanje svelo na “savršene” silogizme, u kojima nema takvih implicitnih pretpostavaka (v. npr. *An. Pr.* A4 26b 28–30, A5 27a 15–18). Sama temeljna načela silogizma (načela protuslovlja i isključenoga srednjega), neovisno o “radnoj” formalnoj logici (teorija silogizma), Aristotelova su tema u *Metafizici* (Aristotel 1973), posebice u knjizi Γ. Aristotel stoga nije bio daleko od strogoga definiranja umjetnoga, formalističkoga jezika, što bi zajedno s logičkim aksiomima i pravilima pružilo formalni sustav logike (ili neke druge, posebne teorije).⁷ Usprkos Aristotelovoj uporabi izraza naravnoga jezika zajedno s pojmovnim slovima i umjetnim frazama, Aristotelov je pristup sukladan zahtjevima formalnoga sustava kako ih je formulirao Frege: (a) formalni jezik (“pojmomopis”, *Begriffsschrift*) treba izraziti samo ono što je bitno za zaključivanje, isključujući svaku prešutnu konotaciju koja se može izvesti iz konteksta naravnoga jezika i diskursa (“pojmomopis” sadrži samo posebne, umjetne simbole, ne “riječi” naravnoga jezika); (b) zaključivanje treba biti opisano bez praznina (*lückenlos*), s izričitim svim nužnim aksiomima, pravilima zaključivanja i definicijama (Frege 1988, x, 3; Frege 1998, v–vii).

2 Formalizam između osjetilne zamjedbe i pojmova

Formalni sustav (formalizam) kako ga Frege shvaća, bitno je usmjeren na apstraktna, neosjetilna bitstva kao što su misli, pojmovi, sudovi i zaključci (usp. Fregeove “čiste misli” ili “čiste pojmove”). Međutim formalni se sustav definira tek “preslikavanjem” tih apstraktnih bitstava na osjetilno zamjetljive, pisane znakove (“formalni jezik”, “pojmomopis”). Prema Fregeu se dokazi formalnoga sustava prikazuju našem “oku”

⁷ Intenzionalna semantika Aristotelova silogističkoga sustava predložena je u (Kovač 2013).

kao niz formula (Frege 1998, v).⁸ To upućuje na to da temelji formalnoga sustava uključuju uvjet osjetilne danosti, koja je pak nešto s onu stranu logike ako se predmet logike doslovno razumije kao razumska radnja formalnoga zaključivanja.⁹

Valja primijetiti da zapisi sami po sebi, samo kao predmeti osjetilnoga iskustva, ne otkrivaju nužno ništa o onom što zapisuju. Samo ako unaprijed znamo da je dani zapis udružen s određenim izrazom, napose s određenim pojavkom izraza (npr. s prvim pojavkom 'A' u 'A v non-A'), možemo čitati i gramatički ili dokaznoteorijski provjeriti, na primjer, dane zapise nekoga dokaza. Štoviše, logički razumijemo izraz kao što je 'A v ¬A' samo ako razumijemo da je to tek jedan način (uz 'p v ¬p', 'ApNp' itd.) na koji možemo izabrati formalni jezik kako bismo izrazili neki logički zakon kojemu će taj izraz biti strukturno sličan. Stoga, premda egzaktnost, strogoća i (gramatička i dokazna) provjerljivost u formalizmu potječu, s jedne strane, od osjetilne evidencije zapisanih izraza, one, s druge strane, pretpostavljaju neko apstraktno i "idealno" ili "pojmovno" predrzumijevanje *tipova* izraza i samih izraženih "forma" (usp. na primjer Fregeove ili Gödelove "pojmove" i "misli", kao i Aristotelov $\epsilon\sigma\omega$ λόγος).

Vezano uz "pojmovnu" sastavnicu razumijevanja formaliziranih zapisa postavlja se i pitanje dizajna i izbora formalnoga sustava (njegovih

⁸ Vizualna dostupnost pisanoga jezika preduvjet je "strogoće dokazivanja" (Strenges der Beweisführung) i "oštrine razlikovanja" (Frege 1998, vi; Frege 1988, xi). V. ovdje Kantovu refleksiju u bilješci 11.

⁹ Osjetilna zorna danost izrazâ u dokazu zapisanom u formalnom jeziku osobito je istaknuta u Hilberta, neovisno o svim razlikama prema Fregeovu filozofskom stajalištu: "Ako logički zaključak treba biti pouzdan, mora biti moguće potpuno pregledati te predmete u svim njihovim dijelovima, a činjenica da se oni [= ti predmeti] pojavljuju, da se međusobno razlikuju i da slijede jedan za drugim ili su ulančani, izravno je *zorno* dana, zajedno s predmetima, kao nešto što se ne može svesti na nešto drugo niti zahtijeva svođenje" (Hilbert 1967, 376, naš kurziv). Kant tu igra znatnu filozofsku ulogu: "[U] suglasnosti smo s filozofima, osobito s Kantom. Već je Kant naučavao ... da matematika ima na raspolaganju sadržaj neovisno o svoj logici i stoga se ne može priskrbiti samo utemeljenjem pomoću logike. ... [N]ešto već mora biti dano našoj moći predočavanja, određeni izvanlogički konkretni predmeti koji su zorno nazočni kao neposredno iskustvo prije svake misli" (Hilbert 1967, 376). Hilbertova je razmatranja o Kantovoj ulozi dalje produbio Gödel na osnovi svojih poučaka o nepotpunosti (v. zadnja dva odlomka u 1995b).

aksioma, definicija i pravila). Kad je riječ o formalizaciji empirijskoga znanja, jasno je da dizajn sustava treba uzeti u obzir empirijske rezultate i osnivati se na njima. S druge strane, može se postaviti pitanje kriterija našega izbora logičkih načela koja se ugrađuju u formalni sustav i prihvaćaju se bez formalnoga dokazivanja. Neovisno o tom kako bi se moglo odgovoriti na to pitanje, odluku treba očito donijeti barem djelomice na temelju pojmovnih razmatranja i samoevidentnosti (za logičkoga činitelja) s onu stranu dokaznih postupaka. Jedan se primjer, za ilustraciju, može naći u Łukasiewiczu istraživanju zakona o isključenom srednjem i zakona dvovrijednosti u kontekstu Aristotelove diskusije u *De interpretatione* (v. Aristotel 1974). Kako Łukasiewicz ističe, u temeljima logike treba se osloniti na osobnu "samoevidenciju":

Kako ono [načelo dvovrijednosti] leži u samim temeljima logike, to načelo, o kojem se raspravlja, ne može se dokazati. Netko samo može vjerovati u nj, i samo onaj koji ga drži *samoevidentnim*, vjeruje u nj. Meni se, *osobno*, načelo dvovrijednosti ne čini samoevidentnim. Prema tome, ovlašten sam ne priznati ga te prihvatiti pogled da uz istinu i neistinu opstoje i druge istinitosne vrijednosti, uključujući barem još jednu, treću istinitosnu vrijednost.

Łukasiewicz 1967, 36–37, naš kurziv¹⁰

Dok s jedne strane osjetilna danost znakova olakšava egzaktnost, s druge strane, ako se uzme doslovno, tj. bez dostatna apstraktna razmatranja znakova, ona vodi antinomijama u koje su upletena ne tek prividna, nezbiljska "bitstva" (kao Aristotelov "jarac-jelen"), nego i logički, sebe-niječući "podmeti" i rečenice (usp. "ova rečenica" u formulaciji Lažljivca i Curryeva paradoksa).

¹⁰ V. u njegovu ranijem tekstu: "Ipak, u svem zaključivanju ima neko inherentno formalno kreativno zaključivanje: logičko načelo zaključivanja. ... Logika je znanost a priori. Njezini su poučci istiniti snagom definicija i aksioma izvedenih iz uma, a ne iz iskustva. Ta je znanost sfera čiste mentalne činitbe" (Łukasiewicz 1970a, 11). I poslije, u 1930-ima: "Bavimo se tim značenjem, mišlju i idejama izraženim znakovima čak i kad ne znamo što su ta značenja, a ne znakovima kao takvim" (Łukasiewicz 1970b, 241). Ipak, Łukasiewicz je snažno naglašivao jezične izraze logičkih zakona ("nominalistička maska" logike, Łukasiewicz 1970c, 222), a katkad se radikalno odrekao "mišljenja" kao predmeta logike: "Nije, međutim, istina da je logika znanost mišljenja. Nije predmet logike istraživati kako zbilja mislimo ili kako bismo trebali misliti. Prva zadaća pripada psihologiji, a druga nekomu praktičnomu umijeću sličnomu mnemonici" (Łukasiewicz 1957, 12, usp. i Łukasiewicz 1970c).

3 Stroj kao simbol formalizma

Prethodno spomenuta tarskijevska opća definicija formalnoga sustava uzima pojmove rečenice i izvedljivosti (“posljedica” ranoga Tarskoga) kao primitivne i aksiomatski ih opisuje u pojmovima teorije skupova, tj. ona definira formalni sustav pomoću posebno izgrađenoga meta-teorijskoga aksiomatskoga sustava, koji pak ima vlastite predformalne pretpostavke (v. Tarski 1983). Pojam se formalnoga sustava može precizirati u njegovu “apstraktnom” (“apsolutnom”) smislu neovisno o bilo kojem formalizmu (tj. neovisno o danom formalnom sustavu), i ujedno potpuno egzaktno i strogo. Kako je poznato, prema Gödelu, koji slijedi Turinga (1965), opći pojam formalnoga sustava neovisan je o nekom posebnom formalizmu (ali nije neformalan, v. Crocco 2019) i dan je pomoću jasne i precizne zamjedbe pojma mehaničkoga postupka, definiranoga Turingovim strojem koji ispisuje sve poučke formalnoga sustava (Gödel 1995a, 308): “formalni sustav jest mehanički postupak za proizvodnju formula”. Prema Gödelu, “bit” je formalnoga sustava “da je zaključivanje potpuno zamijenjeno mehaničkim radnjama na formulama” (1986b, 195; 1986d, 370), tj. istovrijedno Turingovu stroju (“mehaničkoj proceduri”, “algoritmu”), kako je spomenuto, za ispisivanje poučaka toga sustava (Gödel 1995a, 308, v. Crocco 2019, usp. Kennedy 2014).¹¹

U definiciji formalnoga sustava pomoću Turingova stroja pojam je formalnoga sustava sveden na mehaničke (i stoga uzročne) pojmove i objektiviran.¹² Takva definicija ne bi trebala biti iznenađujuća. Tvrdimo da je mehanička (ili barem općenito uzročna) perspektiva bila u bit-

¹¹ Kantova “tvornica silogizama” svojevrsna je anticipacija pojma formalnoga sustava kao stroja: “O tvornici umskih zaključaka [= silogizama]. Svaki se put traži da um naposljetku učinimo tehničkim kako bismo se, podvrgavajući um obradi osjetila, osigurali od pogrešaka” (“Von der fabric der Vernunftschlüsse. Man sucht jederzeit die Vernunft zulezt technisch zu machen, damit, indem man sie der Behandlung der Sinne unterwirft, man wegen der Fehler gewiß sey”, Kant 1924, 742, refl. 3256). Kant je ocrtao, za svoje standarde, dokažljivo potpun “sustav” formalne logike (proširujući aristotelovsku logiku hipotetičnim i disjunktivnim oblikom iskaza) na kojem se ta “tvornica” treba osnivati (usp. na primjer Kant 1911, B viii–ix, 94–101, 131–142, 359–361; Kant 1924).

¹² Objektivnost inherentne “mehaničke” misli logičkoga formalizma istaknuo je J. Salamucha (2003). On drži da je “logičko ‘mehaniziranje’, kultivirajući jasnoću, preciznost i objektivnost, bitno za razvoj ‘zdrave individualnosti’” (str. 68).

nome uključena u izgradnji i definiranju središnjih logičkih pojmova (ne tek u smislu pomoćnih tehnika i sredstava za izvođenje zaključaka) od početaka povijesti logike. Može se pokazati npr. da je Aristotelovo razumijevanje silogizma u osnovi ovisno o formalno shvaćenim uzročnim pojmovima (npr. premise kao uzroci silogizma, Aristotel, *An. Pr.* A1 24b 18–22, *Met.* Δ2, 1013b 20–21). “Savršen” silogistički zaključak (silogizam), u skladu s Aristotelovim pristupom, nije drugo nego (mehanički) “izračun” (συλλογίζειν) iz pojmova silogizma prema kakvoći veće (opće) premise i kolikoći manje (potvrdne) premise na osnovi prijelaznosti priricanja.¹³ “Nužnost” se u Aristotelovoj definiciji silogizma može shvatiti kao mehaničko (automatsko) uzrokovanje zaglavka iz premise kao njegovih “uzroka”. Iako se zaglavak može razumjeti kao svrha (τέλος) silogizma, prema obziru na koju se silogizam uopće postavlja i biraju premise, kad su premise jednom već postavljene zaglavak slijedi (συμβαίνει) zahvaljujući samo njima, bez ikakva dodatna, vanjskoga uzroka (“time što premise jesu”, διὰ τὸ ταῦτα εἶναι; *An. Pr.* A1 24b 18–20).

Dodatno, nerefleksivnost, nemonotonost i prijelaznost silogističkoga zaključivanja kao i nevaljanost slijeda P iz ne-P (v. Kovač 2013) također su u suglasju sa shvaćanjem silogizma barem formalno po uzoru na uzročnost. Ukratko, nerefleksivnost isključuje samouzrokovanje, a nemonotonost uvrštavanje uzročno irelevantnih sastavnica u uzrok. Argumenti koji se u literaturi mogu naći protiv prijelaznosti uzrokovanja obično se služe eliptičnim ili višeznačnim pojmom uzroka. Također, nije razumljivo kako bi izostanak nekoga stanja mogao biti uzrokom toga stanja (podrobnije o navedenim karakteristikama zaključivanja i uzročnosti u Kovač 2013).

Ono što na poseban način pridonosi mehanističkoj “slici” Aristotelove silogistike jest uloga ekteze (“ekspozicije”), koja se pretpostavlja u temeljnom sloju te silogistike. Kako je primijetio Hintikka (1967),¹⁴ Aristotelova je ekteza slična Euklidovoj ektezi s priključenom konstrukcijom (κατασκευή, “priprava”, “mašinerija”). Pomoću Aristotelove ekteze

¹³ O posebnoj, “aritmetičkoj interpretaciji silogistike” u Leibniza, usp. u Łukasiewiczza (1957, 126–129).

¹⁴ V. Hintikkinu diskusiju o ektezi u povezanosti s euklidovskom konstrukcijom kao pripremom za dokaz.

konstrukcija se oprimjerujućega pojma X inicira kao svojevrsno umjetno, mehaničko sredstvo koje oprimjeruje pojmove kako su zadani premisama te potom “automatski” pokazuje pripada li X-u neko svojstvo P (iz veće premise) o kojem se pita.¹⁵ Pod pretpostavkom da je ektetički pojam “niži” od svih pojmova razmatranoga silogizma, taj se pojam može razumjeti kao singularan, kako bi se predočio neki pojedinačan predmet. Čini se da u mogućem svođenju silogistike na njezine temelje – načela neprotuslovlja i isključenoga srednjega (*Met.* Γ, in Aristotel 1973) – ekteza treba imati ključnu ulogu (usp. ektetički stil Aristotelove definicije općega iskaza u *An. Pr.* A1 24b 28–30), bez obzira na činjenicu da u svojoj “radnoj logici” Aristotel rabi ektezu samo prigodice, kao na primjer u dokazivanju obrata ili trećega lika silogizma (Kovač 2013).¹⁶

Turingov stroj kao i konstruirani ektetični pojmovi imaju vrlo općenit, simbolični karakter. Kako bi se bolje razumjela njihova bit upućujemo na neke posebno zanimljive Wittgensteinove refleksije o strojevima.

Stroj ili sliku stroja rabimo kao simbol određenoga načina djelovanja. Primjerice, dajemo nekomu tu sliku i pretpostavljamo da on iz nje izvodi pojave gibanja dijelova.

PU §193

Mogli bismo reći da je stroj ili njegova slika početak niza slika koje smo naučili izvoditi iz nje.¹⁷

PU §193

U skladu s Wittgensteinom stroj se (ili slika stroja) može uporabiti kao simbol za određen način rada (uzrokovanje) ili gibanja (radnja), u razlici prema stvarnom stroju, koji dodatno uključuje prigodna svojstva kao što je izobličenje dijelova stroja. U slučaju stroja kao simbola, stroj ima svoj određeni način rada, svoje moguće pokrete “u sebi” (isključujući

¹⁵ O Aristotelovoj ektezi v. npr. Żarnecka-Biały (1993).

¹⁶ Sljedeća je aristotelovska uzročna analiza silogizma predložena u (Kovač 2013): premise kao tvarni uzrok (*Met.* Δ2, 1013b 20–21), lik silogizma (položaj srednjega pojma) kao njegov formalni uzrok, ekteza kao pokretački uzrok i zaglavak kao svršni uzrok.

¹⁷ “Wir gebrauchen eine Maschine, oder das Bild einer Maschine, als Symbol für eine bestimmte Wirkungsweise. Wir teilen z.B. Einem dieses Bild mit und setzen voraus, daß er die Erscheinungen der Bewegung der Teile aus ihm ableitet” (PU §193). “Wir könnten sagen, die Maschine, oder ihr Bild, sei der Anfang einer Reihe von Bildern, die wir aus diesem Bild abzuleiten gelernt haben” (PU §193).

njegovo dodatno ponašanje kao danoga zbiljskoga stroja). Wittgenstein sugerira da simboličan stroj (njegova slika), sadržeći moguće pokrete, ne tek odslikava te moguće pokrete nego, kao simbol, ima neki bliži, neempirijski odnos prema njima¹⁸ (u TLP, mogućnost bi gibanja, i stoga samoga stroja, doista bila razumljena kao slika gibanja). Wittgensteinov simboličan stroj dijeli svoju apstraktnost s Turingovim strojem.

Turingov stroj (ili njegova slika koja predočava glavu koja čita i piše, vrpcu, uz dodatak tokovnika stroja) možemo u posebnom slučaju shvatiti kao simbol specifične dokazljivosti (mogući koraci u dokazima) odgovarajućega formalnoga sustava (ne tek kao spravu, makar i idealnu, koja obavlja zadani posao). U tom slučaju Turingov stroj simbolizira taj formalni sustav, sadržavajući “u sebi” sve “pokrete” dokazljivosti u sustavu. Strojevi i slike strojeva sami kao takvi čine univerzalan jezik, neovisan o danim formalizmima, ali ništa manje potpuno egzaktni i strog, posjedujući očitu rigoroznost formalizma.¹⁹ Primijetimo da se u Wittgensteinovu smislu Aristotelov ektetički pojam X može razumjeti kao simbol međuodnosa pojmova kako su postavljeni premisama silogizma, gdje apstrahiramo od svih drugih “stvarnih” svojstava koja bi bilo koji predmet mišljen pod X mogao posjedovati, tj. sam pojam X možemo razumjeti kao simbol određenoga silogizma (zajedno s obratima i drugim silogizmima koje bi on mogao logički uključivati).

¹⁸ Usp. “[D]akle mogućnost gibanja stoji u jedinstvenu odnosu prema samom gibanju; užem nego što je onaj slike prema njezinu predmetu”; jer se može dvojiti o tom je li slika ovoga ili onoga predmeta. ... [A]li ne kažemo ‘Iskustvo će poučiti je li to mogućnost toga gibanja...’ (PU §194).

¹⁹ Wittgensteinovo se filozofsko zanimanje za strojeve može prirodno povezati s njegovim studijem aeronautike i posebice s djelima koja je posjedovao u svojoj privatnoj knjižnici uključujući djela Leonarda da Vincia (*Les manuscrits de Léonard de Vinci*, sv. 1–6, Paris, 1881–1891) i “knjige strojeva” (*theatra machinarum*) kao što su djela Fausta Vrančića (*Machinae novae*, Venetiis, 1615/16, v. Vrančić 1993), Georga Andreasa Böcklera (*Theatrum machinarum novum*, Nürnberg, 1661) i šest djela Jacoba Leupolda (npr. *Theatrum machinarum hydraulicarum: Schau-Platz der Wasser-Künste*, 1, 1724; *Theatrum machinarum: Schau-Platz der Heb-Zeuge*, 1725.) Usp. Spadoni (1985, 25–45), Hide (2004) i <http://digitalcollections.mcmaster.ca/russell-lib/media/machinae-novae-fausti-verantii-siceni>. Navedene su knjige mogle pridonijeti Wittgensteinovoj inspiraciji za njegovu ranu “filozofiju slike” (znak i stroj kao slike), kao i za njegovu kasniju filozofiju jezičnih igara. Wittgenstein vjerojatno nije znao da je Vrančić bio i pisac male logike (Vrančić 2018), koja u svojem drugom izdanju sadrži argumentaciju protiv mogućnosti metafizike kao zasebne discipline uz logiku, fiziku i teologiju, te etiku.

4. Apstraktni pojam dokazljivosti i “životne forme”

Wittgensteinovo shvaćanje stroja kao simbola može pridonijeti razumijevanju “apstraktnih” logičkih pojmova kao što su “dokazljivost” (“demonstrabilnost”) i “definirljivost” (Gödel 1990a), u smislu u kojem su ti pojmovi pretpostavka svakoga pojedinoga logičkoga formalizma. Razumijevanje je tih pojmova, koji trebaju obuhvatiti dinamiku nadogradnje i izgradnje novih i novih formalnih sustava, zadaća koju je Gödel (uz neke nacрте i napomene) ostavio otvorenom. Iako Wittgensteinov pristup, usredotočen na jezik i njegovu porabu, nije srodan Gödelovu pojmovnomu i matematičkomu realizmu i kauzalizmu (Kovač 2020), držimo da on ipak može pripomoći u pristupu apstraktnim logičkim pojmovima i njihovu razumijevanju, bez obzira na spomenuta različita filozofijsko-metodologijska stajališta.

Sam je Gödel već u svojoj disertaciji o potpunosti logike prvoga reda imao na umu, iako kao upitan, pojam dokazljivosti “bilo kojim uopće izmišljivim pomoćnim sredstvima” (*alle überhaupt erdenkliche Hilfsmittel*, Gödel 1986a, str. 62–64), tj. ono što će poslije nazvati “apsolutnim” pojmom dokazljivosti, nesvedljivim na dokazljivost unutar jednoga određenoga formalnoga sustava. Iz perspektivne toga “apsolutnoga” pojma dokazljivosti, potpunost dokaznoga sustava s obzirom na modele nije neočekivana (Gödel 1990a, usp. Gödel 1986a). Gödel naznačuje da bi takav pojam dokazljivosti trebala karakterizirati shematska sličnost korak po korak nadograđenih aksioma s postojećim aksiomima na koje se nadograđuju, a otkrivanje novih i novih primitivnih pojmova usporedio je s konstrukcijom rednih brojeva (Gödel 1990a). S druge je strane Gödel (1986c) uočio da standardna shematizacija pojma nužnosti (modalni djelatelj $\Box$), ako se primijeni na dokazljivost (kao djelatelj B), vodi pojmu dokazljivosti koji izlazi iz okvira formalnih sustava koji mogu formalizirati aritmetiku. Primijetio je da se isto događa i kad se djelatelj nužnosti = dokazljivosti zamijeni konkretnim opravdanjima ali se sačuva *de dicto* pokoličavanje na njima (kao npr. u rečenicama oblika “svako je opravdanje t takvo da ... t ...”) (usp. Kovač 2019).²⁰ Naposljetku, Gödel je došao (1990b) do dinamičnoga pojma razvoja Turingovih

²⁰ Artemov je pokazao kako se logika opravdanja može definirati za pojam aritmetičke dokazljivosti (v. Artemov 2001).

strojeva i formalnih sustava, koji otvara perspektivu stalnoga povećanja broja “mentalnih” stanja (strojnih i ljudskih) s mogućnošću sve veće preciznosti te novih i novih pojmovnih razlikovanja. Postavlja se pitanje gdje je temelj, “uvjet mogućnosti”, takva dinamična pojma dokazljivosti, koji bi ujedno osiguravao i potpunost, “apsolutnu” cjelinu dokazljivosti i istine, tj. formalnih sustava u njihovu razvoju i stvarnoga svijeta?

U Wittgensteinovoj perspektivi, takav temelj, koji treba osiguravati i cjelovitost, treba tražiti u “oblicima života”. Iako time Wittgenstein nije ponudio neko razrađeno tehničko sredstvo razvoja dokazne evidencije i novih pojmova, upozorio je na horizont i način gledanja koji može pomoći u razumijevanju i opisivanju pretpostavljene “tehnike”.²¹ S Wittgensteinova motrišta Gödelovi bi se dinamični pojmovi demonstrabilnosti i definirljivosti trebali sastojati u načinu kako rabimo jezik i u “obiteljskoj sličnosti” predložaka mišljenja (zaključivanja, izgradnje pojmova) koji pripadaju raznim “oblicima života”. “Svaki se znak *sam* čini mrtvim... On živi u porabi” (PU §432). Na primjer, “Strjelica pokazuje samo u primjeni koju živo stvorenje njom ostvaruje” (PU §454). Slično i “sustav” i “sigurnost” nisu drugo nego oblici života.²² U tom su smislu život i jezične igre temeljno izvorište uzoraka i paradigma, kao i (naposljetku “nemehaničkoga”) prepoznavanja i razvoja novih i novih aksioma i načina dokazivanja, kao i novih i novih pojmova. Dokaz, kako napominje Wittgenstein, postavlja nove paradigme, tvori nove veze, mijenja i stvara pojmove, mijenja gramatiku (BGM str. 161, 166, 172, usp. str. 139, 178, 248, 297) na osnovi uzoraka koji su već tu. Apstraktni pojmovi vode prema “obitelji tvorba”, umjesto prema “formalnom jedinstvu” (PU §108). Tautologije, dokazi i pojmovi utemeljeni su naposljetku u jezičnim i predjezičnim predlošcima i “oblicima života”, iz kojih po-

²¹ Ono što za Gödela nije obećavajuće u Wittgensteinovu pristupu jest orijentacija na jezik, o kojem, kako Gödel drži, “nemamo nikakve primitivne intuicije”. “Smiješno je [očekivati] da bismo trebali imati ikakve primitivne intuicije o jeziku, koji je tek udruživanje simbola s pojmovima i drugim bitstvima” (Wang 1996, str. 180 5.5.7–8). No s druge strane, Gödelov je stav da je “svakodnevno znanje, kad se analizira u sastavnice, relevantnije [nego znanost] u davanju podataka za filozofiju. Sama znanost ne će dati filozofiju; ona je neobvezujuća prema obziru na ono što stvarno [jest]” (Wang 1996, str. 297 9.2.3). Kako je poznato, svakodnevna praksa, običaji i jezik dobivaju središnju ulogu u Wittgensteinovoj filozofiji nakon TLP.

²² “Sustav nije toliko polazište koliko životni element argumenata” (UG 105); usp. sigurnost “kao (neki) oblik života” (UG 358).

tječu sve jezične igre i logički predlošci (Moyal-Sharrock 2015). Umjesto težnje prema “kristalnoj čistoći” logika se treba referirati na “naše prave potrebe kao čvrstu točku” (PU §108) i utemeljiti se u “uzorcima na vrpici [= sponi, njem. *Band*] života” (PP xi §362) (kao što se analogno svaki pojedini formalni sustav očituje u uzorcima zapisa na vrpici Turingova stroja koji ispisuje poučke toga sustava). Prauzori i zakoni mišljenja, koje opisuje logika, jesu institucija, običaj (*Denkgewohnheit*, usp. BGM str. 89, I–131) i kao takvi su oblik života. Premda Wittgenstein nije studirao Aristotela, takve refleksije, ostavljajući po strani Wittgensteinove razlike prema Aristotelu, možemo shvatiti i kao daleku reminiscenciju na “mišljenje mišljenja”, koje je prema Aristotelu (*Met.* A7, Aristotel 1973) život i izvor svega gibanja.

Referencije

- Aristotel (1964). *Analytica Priora et Posteriora*. Ur. D. Ross i L. Minio-Paluello. Oxford: Oxford University Press. – Hrvatski: Aristotel, *Druga Analitika*, prijevod i komentar F. Grgić, Zagreb: Matica hrvatska, 2020.
- Aristotel (1973). *Metaphysica*. Ur. W. Jaeger. Oxford: Oxford University Press. – Hrvatski: Aristotel, *Metafizika*, prijevod i komentar T. Ladan, Zagreb: SN Liber, 1985.
- Aristotel (1974). *Categoriae et Liber de Interpretatione*. Ur. L. Minio-Paluello. Oxford: Oxford University Press. – Hrvatski: Aristotel, *Kategorije*, prijevod i komentar F. Grgić, Zagreb: Hrvatska sveučilišna naklada, 1992; Aristotel, *O tumačenju*, prijevod i komentar J. Talanga, Zagreb: Latina et Graeca, 1989.
- Artemov, S. (2001). “Explicit provability and constructive semantics”. *Bulletin of Symbolic Logic* 7, 136.
- Béziau, J.-Y. (2005). “From Consequence Operator to Universal Logic”. U: *Logica Universalis*, ur. J.-Y. Béziau, Basel: Birkhäuser, 3–17.
- Bocheński, I.M. (2002). *Formale Logik*. Freiburg, München: Alber.
- Crocco, G. (2019). “Informal and Absolute Proofs: Some Remarks from a Gödelian Perspective”. *Topoi* 38, 561–575 (online 11. studenoga 2017).
- Frege, G. (1988). *Begriffsschrift und andere Aufsätze*. 2. izdanje. Hildesheim etc.: Olms.
- Frege, G. (1998). *Grundgesetze der Arithmetik*. Hildesheim etc.: Olms.
- Gödel K. (1986–2003). *Collected Works*. Ur. S. Feferman et al. Oxford: Oxford University Press.

- Gödel, K. (1986a). "Über die Vollständigkeit des Logikkalküls" (1929). U: Gödel 1986–2003, sv. 1, 60–101.
- Gödel, K. (1986b). "Über formal unentscheidbare Sätze der Principia mathematica und verwandter Systeme I" (1931). U: Gödel 1986–2003, sv. 1, 144–195. – Hrvatski u: E. Nagel i J. R. Newman, *Gödelov dokaz*, prev. V. Kirin, Zagreb: Kruzak, 2001.
- Gödel, K. (1986c). "Eine Interpretation des intuitionistischen Aussagenkalküls" (1933). U: Gödel 1986–2003, sv. 1, 300–303.
- Gödel, K. (1986d). "On Undecidable Propositions of Formal Mathematical Systems" (1934). U: Gödel 1986–2003, sv. 1, 346–371. (Postscriptum 1964.).
- Gödel, K. (1990a). "Remarks before the Princeton Bicentennial Conference on Problems in Mathematics" (1946). U: Gödel 1986–2003, sv. 2, 150–153.
- Gödel, K. (1990b). "Some Remarks on the Undecidability Results" (1972). U: Gödel 1986–2003, sv. 2, 305–306.
- Gödel, K. (1995a). "Some Basic Theorems on the Foundations of Mathematics and their Implications". U: Gödel 1986–2003, sv. 3, 304–323.
- Gödel, K. (1995b). "The Modern Development of the Foundations of Mathematics in the Light of Philosophy". U: Gödel 2003, sv. 3, 374–387.
- Hide, Ø. (2004). "Wittgenstein's Books at the Bertrand Russell Archives and the Influence of Scientific Literature on Wittgenstein's Early Philosophy". *Philosophical Investigations* 27: 68–91.
- Hilbert, D. (1967). "On the Infinite" (1926). U: *From Frege to Gödel*, ur. J. van Heijenoort, Cambridge (MA): Harvard University Press, 367–392.
- Hintikka, J. (1967). "Kant on the Mathematical Method". *Monist* 51, 3: 352–375.
- Kant, I. (1924). *Logik. Kant's handschriftlicher Nachlaß*. U: *Kant's gesammelte Schriften*, ur. Königlich Preußische Akademie der Wissenschaften, sv. 16. Berlin, Leipzig: de Gruyter.
- Kant, I. (1911). *Kritik der reinen Vernunft*. 2. izdanje 1787. *Kant's gesammelte Schriften*, ur. Königlich Preußische Akademie der Wissenschaften, sv. 3. Berlin: Reimer. – Hrvatski: I. K., *Kritika čistoga uma*, prev. V. Sonnenfeld, Zagreb: Nakladni zavod Matice hrvatske, 1984.
- Kovač, S. (2013). "Causation and Intensionality in Aristotelian Logic". *Studia Philosophiae Christianae* 49, 2: 117–136.
- Kovač, S. (2020). "On Causality as the Fundamental Concept of Gödel's Philosophy". *Synthese* 197: 1803–1839. Obj. online 2018.
- Kovač, S. (2019). "Proofs, necessity and causality". U: *Aventuras en el Mundo de la Lógica: Ensayos en Honor a María Manzano*, ur. E. Alonso, A. Huertas i A. Moldovan, London: College Publications, 239–263.

- Łukasiewicz, J. (1957). *Aristotle's Syllogistic: From the Standpoint of Modern Formal Logic*. 2. proš. izdanje. Oxford: Oxford University Press.
- Łukasiewicz, J. (1967). "On Determinism" (1961). U: *Polish Logic*, ur. S. McCall, Oxford: Oxford University Press, 19–39.
- Łukasiewicz, J. (1970a). "Creative Elements in Science" (1912). U: Łukasiewicz 1970d, 1–15.
- Łukasiewicz, J. (1970b). "In Defence of Logistic" (1937). U: Łukasiewicz 1970d, 236–249.
- Łukasiewicz, J. (1970c). "Logistic and Philosophy" (1936). U: Łukasiewicz 1970d, 218–235.
- Łukasiewicz, J. (1970d). *Selected Works*. Ur. L. Borkowski. Amsterdam, London, Warszawa: North-Holland.
- Moyal-Sharrock, D. (2015). "Wittgenstein on forms of life, patterns of life, and ways of living". *Nordic Wittgenstein Review* (special issue): 21–42.
- Salamucha, J. (2003). "On the 'Mechanisation' of Thinking" (1937). U: J. Salamucha, *Knowledge and Faith*, ur. K. Świętorzecka i J. J. Jadacki, Amsterdam, New York: Rodopi.
- Spadoni, C. i D. Harley. (1985). "Bertrand Russell's Library". *The Journal of Library History* (1974–1987) 1: 25–45.
- Tarski, A. (1983). "On Some Fundamental Concepts of Metamathematics" (1930). U: A. Tarski, *Logic, Semantics, Metamathematics*, prev. J. H. Woodger, 2. izdanje, Indianapolis (IN): Hackett, 30–37.
- Turing, A. (1965). "On Computable Numbers, with an Application to the Entscheidungsproblem" (1936/1937). U: *The Undecidable*, ur. M. Davis, New York: Raven, 116–154.
- Vrančić, F. (1993). *Machinae novae*. Zagreb, Šibenik: Novi Liber, Gradska knjižnica "Juraj Šišgorić". Reprint izdanja Venecija, 1615/1616.
- Vrančić, F. (2018). *Nova logika*. Ur. i prev. L. Boršić. Šibenik, Zagreb: Gradska knjižnica "Juraj Šišgorić", Institut za filozofiju. Sadrži: *Logica nova suis ipsius instrumentis formata et recognita*, Venetiis 1616.
- Wang, H. (1996). *A Logical Journey: From Gödel to Philosophy*. Cambridge (MA), London: The MIT Press.
- Wittgenstein, L. (2019) [BGM]. *Bemerkungen über die Grundlagen der Mathematik*. Frankfurt: Suhrkamp.
- Wittgenstein, L. (2009) [PU, PP za drugi dio "Philosophie der Psychologie – Ein Fragment"]. *Philosophische Untersuchungen / Philosophical Investigations*. Izmijenjeno 4. izdanje. Prev. G. E. M. Anscombe, P. M. S. Hacker i J. Schulte. Oxford: Blackwell. – Hrvatski: L. W., *Filozofijska istraživanja*, prev. I. Mikecin, Zagreb: Nakladni zavod Globus, 1998.

- Wittgenstein, L. (1972) [UG]. *Über Gewißheit / On Certainty*. G. E. M. Anscombe i G. H. von Wright ur., prev. D. Paul i G. E. M. Anscombe. New York etc.: Harper and Row. – Hrvatski: L.W., *O izvjesnosti*, prev. A. Periša, Zagreb: Nakladni zavod Globus, 2007.
- Wittgenstein, L. (1976) [TLP]. *Tractatus Logico-Philosophicus*. Frankfurt: Suhrkamp. – Hrvatski: L. Wittgenstein, *Tractatus Logico-Philosophicus*, prev. G. Petrović, Zagreb: Moderna vremena, 2003.
- Żarnecka-Biały, E. (1993). "Aristotle's Proofs by Ecthesis". *Bulletin of the Section of Logic* 22, no. 1: 40–44.

Russellov Russellov paradoks

Goran Lojkić

Russellov paradoks razreda svih razreda koji nisu svoji članovi najpoznatija je antinomija tzv. naivne teorije skupova. Paradoks je izravno motivirao nastanak Russellove teorije tipova i, zajedno s drugim antinomijama naivne teorije, aksiomatizaciju teorije skupova. Međutim, puno je manje poznato na koji je način sam Russell, u svjetlu svojih onovremenih metafizičkih pogleda na narav razreda, isprva interpretirao paradoks koji je otkrio i rješenje koje je na paradoks ponudio. Cilj je ovoga rada izložiti Russellovo razumijevanje paradoksa i odgovora na nj u kontekstu njegove rane filozofije matematike.

Russellov paradoks

U svojoj filozofijskoj autobiografiji *My Philosophical Development* [31] Bertrand Russell piše da je u ljeto 1900. u Parizu, sudjelujući na prvom Međunarodnom filozofijskom kongresu, postao "svjestan važnosti logičke reforme za filozofiju matematike" [31, str. 65].¹ Kao presudan utjecaj navodi Peana, koji je "u svakoj raspravi pokazao više preciznosti i logičke strogosti nego što je pokazao bilo tko drugi" [31, str. 65]. Zamolio je Peana za primjerke njegovih radova i, kako kaže, "smjesta ih sve pročitao" [31, str. 65]. Prema Russellovu sjećanju, bio je kraj srpnja kada je upoznao Peana, a prvu radnu inačicu knjige *The Principles of Mathematics* [30] dovršio je do kraja iste godine (usp. [31, str. 72–73]).²

¹ Nije li navedeno drukčije, sve citate preveo je s engleskoga autor rada.

² Grattan-Guinness u [17, str. 57] Russellovu priču opisuje kao "apsurdno pogrješnu" te tvrdi da Russell većinu radova tzv. peanista nije zaprimio još mjesec dana od zahtjeva, a da se iz sačuvanih rukopisa vidi da je pisanje *Principles* trajalo do u 1902. godinu. Posebno, napominje da su ključni dijelovi s logikom relacija i definicijama ciljelih brojeva nastali tek 1901. (usp. [17, str. 57–58]).

To razdoblje glasovito opisuje kao “intelektualni medeni mjesec kakav ni prije ni poslije nije nikada iskusio” [31, str. 73].

Peano u svojoj aksiomatizaciji aritmetike pomoću nove simboličke notacije u *Arithmetices principia* [26], maloj knjižici na latinskom u kojoj se prvi puta pojavljuju njegovi slavni aksiomi, oštro razlikuje matematičke od logičkih simbola (usp. [27, str. 85]). Posebno, tvrdi Peano, izrazi za brojeve pripadaju aritmetičkim, a ne logičkim simbolima. Russell, s druge strane, u *Principles* predlaže definiciju kojom se brojevi svode na određenu vrstu razreda. Preciznije, definira kardinalne brojeve kao razrede ekvipotentnih ili, držeći se Russellova nazivlja, sličnih (*similar*) razreda. Prema toj je definiciji broj 0 razred svih razreda bez ijednoga člana (tj. razred koji kao svoj jedini element sadrži prazan razred (*null-class*)), broj 1 razred je svih razreda sličnih razredu 0 (tj. razred svih razreda sa samo jednim članom), broj 2 pak razred svih razreda koji su slični razredu koji sadrži samo 0 i 1 (dakle, razred svih razreda koji sadrže upravo dva člana) itd. Za tu se definiciju uobičajilo ime ‘Frege-Russellova definicija brojeva’.³ Redni su brojevi definirani na sličan način, kao razredi sličnih dobro uređenih razreda. Unatoč odstupanju od Peana, te su definicije postale dostupne upravo zahvaljujući njemu: razlikovanje jediničnoga razreda od njegova jedinoga elementa, o kojem predložene definicije brojeva ovisе, Russell iznosi kao jedno od najvažnijih tehničkih rezultata koje je naučio od Peana (usp. [31, str. 66–67]).⁴

³ Pritom, Fregeu pripada prednost i vjerojatno je upravo on, preko Peana, posredno utjecao na Russella. Definicija slična Russellovoj nalazi se u *Grundgesetze der Arithmetik* iz 1893. [10], ali Frege istu predlaže još 1884. u *Die Grundlagen der Arithmetik* [9], dakle barem šesnaest godina prije nego što Russell dolazi do iste ideje. Iako Russell još u listopadu 1900. posjeduje *Grundgesetze* i sa skromnim ih uspjehom pokušava čitati (usp. [2, str. 65]), prije predaje *Principles* u tisak u svibnju 1902. nije uspio “shvatiti njezinu važnost ili razumjeti njezin sadržaj” [30, str. xvi]. Tek u lipnju 1902. pažljivo čita Fregea i naknadno piše prvi dodatak knjige, naslovljen “The Logical and Arithmetical Doctrines of Frege”, gdje za Fregeovo djelo kaže da “obiluje finim distinkcijama i izbjegava sve uobičajene pogriješke”, a za njegov simbolizam da je, iako nezgrapan za praktičnu primjenu, “zasnovan na analizi logičkih pojmova koja je mnogo dublja od Peanove” i da je ktomu “filozofijski jako nadmoćan” Peanovu [30, str. 501]. Svoju definiciju brojeva Russell, čini se, izravno duguje Peanu, jer u *Principles*, odgovarajući na prigovore definiciji, citira Peana, koji tu definiciju odbacuje (usp. [30, str. 115]). No Peano tu definiciju razmatra vjerojatno imajući na umu Fregeove *Grundgesetze*, prikaz kojih je objavio još 1895. u *Rivista di matematica* (usp. [33, str. xxvii]).

⁴ Puna priča o jediničnim razredima u *Principles* dosta je složenija, no više o tome kazat ćemo u trećem odsječku.

Međutim, smatra Russell, s definicijom brojeva u terminima razreda nestalo je potrebe razlikovati logičke od aritmetičkih simbola (ili logičke od aritmetičkih pojmova), kako to čini Peano. Pojam razreda pojam je koji pripada (također i) logici. Naime, Russell o razredima, u skladu s Booleovom logičkom tradicijom i pod utjecajem Whiteheadove *A Treatise on Universal Algebra* [36], razmišlja prije svega kao o opsezima pojmova (usp. [23, str. 635]).⁵ Znači to da se svi aritmetički pojmovi mogu definirati u terminima pojmova same logike. Svođenje aritmetičkoga pojma broja na pojam razreda tako postaje osnova Russellova pogleda na matematiku, prema kojemu se cjelokupna “čista” matematika u konačnici svodi na logiku. Malo preciznije govoreći, svi se pojmovi matematike mogu svesti na temeljne pojmove logike, a sve matematičke tvrdnje na logičke istine: cjelokupna je matematika izvedljiva iz općih logičkih aksioma. U predgovoru *The Principles of Mathematics* kaže Russell da je jedan od glavnih ciljeva knjige ponuditi “dokaz da se sva čista matematika bavi isključivo pojmovima definirljivim pomoću vrlo maloga broja temeljnih logičkih pojmova, a da su svi njezini stavci [*propositions*] izvedljivi iz vrlo maloga broja temeljnih logičkih načela” [30, str. xv]. Slično, na samom početku knjige, u prvoj rečenici prvoga poglavlja, naslovljenoga “Definition of Pure Mathematics”, Russell piše:

Čista je matematika razred svih stavaka oblika “ p implicira q ”, gdje su p i q stavci koji sadrže jednu ili više varijabla, istih u obama stavcima, a ni p ni q ne sadrže nijednu konstantu osim logičkih konstanti. A sve su logičke konstante pojmovi definirljivi u terminima sljedećih: implikacije, relacije oznake [*term*] prema razredu kojega je član, pojma *takav da*, pojma relacije i takvih daljnjih pojmova koji mogu biti uključeni u pojam stavaka gore spomenuta oblika. Dodatno, matematika *rabi* pojam koji nije sastavnica [*constituent*] stavaka na koje se odnosi, naime pojam istine [30, str. 3].

Za stajalište prema kojemu su bilo sva matematika bilo samo određene njezine tradicionalne grane ili discipline svedljive na logiku kasnije se u filozofiji matematike ustalio naziv ‘logicizam’, a Russell je, uz Fregea,

⁵ Russell u [30, str. 18–19] piše da, za razliku od Peana, koji pojmove razreda i članstva u razredu uzima kao temeljne ili primitivne pojmove, u svojoj teoriji razreda polazi od tri nedefinirljiva pojma (*indefinables*): relacije članstva ili pripadnosti, stavne funkcije – tj. svojevrsnoga Russellova poopćenja pojma *pojma* – i pojma “takav da” (*such that*). Malo dalje piše da “razred može biti definiran kao sve one oznake koje zadovoljavaju neku stavnu funkciju” [30, str. 20]. Preciznije značenje te Russellove tvrdnje postat će jasnije u narednim odsječcima.

ostao povijesno najznačajniji njegov predstavnik.⁶ Iako će kasnije sam odbaciti gore citiranu definiciju čiste matematike i znatno promijeniti svoje početne metafizičke poglede o naravi predmeta logike i matematike, to osnovno razumijevanje odnosa logike i matematike Russell nikada nije napustio.⁷

Nažalost, kako se pokazalo, Russellov intelektualni medeni mjesec nije dugo potrajao i uskoro ga je zamijenila “intelektualna tuga” [31, str. 73]. Razmatrajući slavni Cantorov dokaz o nepostojanju najvećega kardinalnog broja, Russell u proljeće 1901. otkriva protuslovlje koje danas znamo pod imenom ‘Russellov paradoks’, iako isprva ne shvaćajući posve važnost paradoksa i misleći da je možda riječ tek o “trivijalnoj pogrešci u [njegovu] rasuđivanju” [31, str. 76].⁸ Iduće godine, nekoliko tjedana nakon što je rukopis *Principles* već predao u tisak, u kratkom pismu od 16. lipnja 1902. – pismu koje će tijekom narednoga stoljeća postati vjerojatno najcitiranije pismo u povijesti filozofije – Russell o otkrivenom paradoksu obavještava Fregea:

Neka je *w* prirok: biti prirok koji se ne može priricati sam sebi. Može li se *w* priricati sam sebi? Iz bilo kojega odgovora, suprotno slijedi. Stoga moramo zaključiti da *w* nije prirok. Slično tomu, ne postoji razred (kao sveukupnost)

⁶ Definicija logicizma prema kojoj je logicizam stajalište da se *sva* matematika može svesti na logiku bila bi prestroga (usp. [16, str. 171]). Russellov je logicizam radikalniji od Fregeova utoliko što Frege *nije* smatrao da se geometrija, za razliku od aritmetike, svodi na logiku. Russell jest, već stoga što je, za razliku od Fregea, prihvatao aritmetizaciju geometrije. Sažet ali informativan prikaz razlike između Russellove i Fregeove koncepcije logicizma može se naći u [24, str. 13–16]. Sam izraz ‘logicizam’ nije ni Russellov ni Fregeov, već je kasnijega podrijetla i dugujemo ga Carnapu u [5].

⁷ Dobar dio uvoda drugomu izdanju *Principles* iz 1938. posvećuje Russell kritičkoj analizi svoje definicije čiste matematike s početka knjige, ali ujedno naglašava da nikada nije vidio nikakva razloga izmijeniti “temeljnu tezu [knjige], da su matematika i logika istovjetne” [30, str. v]. Koliko su se s vremenom Russellovi metafizički pogledi na logiku promijenili govori to što u vrijeme pisanja *My Philosophical Development*, izdane 1959., zakone logike smatra tek jezičnim konvencijama: “Više ne mislim da su zakoni logike zakoni stvari; naprotiv, sada ih smatram čisto jezičnima” [31, str. 102].

⁸ Griffin u [19] detaljno opisuje povijest i kontekst otkrića Russellova paradoksa, a Klement u [22] analizira odnos između Cantorova poučka o partitivnom skupu i različitih paradoksa koje je Russell bilo sam otkrio bilo samo razmatrao u svojim djelima, uključujući paradoks razreda koji nisu svoji članovi i paradoks priroka koji ne oprimjeruju sami sebe. Sam Russell u *Principles* otkriveni paradoks dosljedno naziva jednostavno ‘protuslovlje’ (*the contradiction*).

onih razreda koji, svaki uzet kao sveukupnost, ne pripadaju sami sebi. Iz toga zaključujem da pod određenim okolnostima definirljiva kolekcija [Menge] ne tvori sveukupnost [20, str. 125].⁹

Važno je napomenuti da Russell, kada govori o priroku, ne misli na jezični izraz, već otprilike na ono što bismo nazvali 'svojstvo', 'atribut' ili 'pojam'. Više ćemo o tome kazati u narednom odsječku pa je za sada ta napomena dovoljna. Govoreći pak o razredu kao sveukupnosti (*totality*), Russell na umu ima kolekciju ili množinu entiteta koji, uzeti zajedno, tvore sjedinjenu *cjelinu*, tj. kolekciju entiteta koja se i sama može razumjeti kao *jedna* stvar ili kao zaseban *entitet*. Puno značenje toga izraza postat će nam jasno tek nakon analize u trećem odsječku, no Russellov jezik pritom vjerojatno odražava Cantorov način konceptualizacije problema tzv. paradoksalnih razreda. Cantor, naime, razlikuje dvije vrste razreda ili, vjernije njegovu nazivlju, dvije vrste mnoštvenosti (*Vielheiten*): suvisle, koje se mogu shvatiti kao neko *jedno*, kao jedna stvar (*Einheit*), i nesuvisle, koje se tako shvatiti ne mogu, npr. razred svih rednih brojeva, razred svih transfinitnih kardinalnih brojeva i sl.¹⁰ To je razlikovanje blisko razlici koju danas povlačimo u nekim aksiomatskim teorijama skupova (npr. u NBG) između skupova (tj. onih razreda koji su i sami elementi drugih razreda) i pravih razreda (tj. razreda koji nisu elementi drugih razreda, koji su "preveliki" da bi bili skupovi).¹¹ Russellov ćemo paradoks stoga za početak opisati na danas standardan način, govoreći jednostavno o skupovima. Kako ćemo vidjeti, to standardno izlaganje paradoksa prikriva važnu filozofijsku distinkciju, ključnu za razumijevanje Russellove vlastite interpretacije otkrivenoga protuslovlja, njegova odgovora na nj predloženoga u dodatku *Principles*, kao i njegovih razloga za nezadovoljstvo tim odgovorom. Raspravu o tome moramo odgoditi dok se malo detaljnije ne upoznamo s pozadinskom ontologijom *Principles*, no vratit ćemo joj se u trećem odsječku.

⁹ Pismo je pisano na njemačkom jeziku, no prijevod je prema engleskomu prijevodu u [20], koji sadržava riječ 'Menge' unutar uglatih zagrada. U kratkoj napomeni prije engleskoga prijevoda pisma stoji da je Russell prijevod pregledao i odobrio ga.

¹⁰ Vidi Cantorovo pismo Dedekindu u [20, str. 113–117]. Cantini u [4, str. 880] (u podrubnici 5) navodi da se slično razlikovanje može naći još ranije kod Schrödera u [34].

¹¹ Kriterij razlikovanja između skupova i pravih razreda u teoriji NBG može se shvatiti kao preciziranje Cantorova kriterija: razred je neko jedno ako i samo ako je element nekoga razreda.

Pismo Fregeu sadrži dvije inačice paradoksa, skupovnu i priročnu – ili, kazano malo drukčijim riječima, ekstenzionalnu i intenzionalnu – no njihova je struktura ista. Russell polazi od naizgled nesporne i očigledno istinite pretpostavke o skupovima, odnosno prirocima, pokazujući da je iz te pretpostavke sasvim elementarnim načinom zaključivanja takoreći umah izvedljivo protuslovlje.

“Očigledno istinita” pretpostavka tzv. naivne teorije skupova jest ta da za svaki dobro definiran uvjet φ , iskazan npr. formulom s jednom slobodnom varijablom, postoji neki skup (možda prazan) koji kao svoje članove sadrži upravo one predmete koji uvjet u pitanju zadovoljavaju. Recimo, ako je φ “ x je ljudsko biće”, onda postoji skup koji kao članove sadrži sva ljudska bića i ništa što nije ljudsko biće nije njegov član. Ako je φ “ x je ljudsko biće i x nije ljudsko biće”, postoji skup koji sadrži sve i samo one predmete koji ujedno i jesu i nisu ljudska bića – naime, prazan skup. Ta se pretpostavka danas obično naziva načelom sadržaja (komprehenzije) ili načelom apstrakcije (ili, u kontekstu aksiomatizacije teorije skupova, aksiomatskom shemom sadržaja). U priročnoj inačici paradoksa imamo analogno načelo sadržaja za priroke, prema kojemu za svaki uvjet φ postoji odgovarajuće svojstvo koje se pririče (ili, kažemo još, koje je oprimjereno) upravo onim predmetima koji taj uvjet zadovoljavaju. Primjerice, ako je φ “ x je ljudsko biće”, postoji svojstvo (“ljudskost” ili “biti ljudsko biće”, recimo), koje pripada svim i samo onim predmetima koji zadovoljavaju φ .

Načelo sadržaja (bilo za skupove bilo za priroke) formalno možemo prikazati na sljedeći način:

$$\exists y \forall x (x \in y \leftrightarrow \varphi).$$

Neformalno čitano, postoji skup y kojega su članovi upravo oni predmeti x koji zadovoljavaju uvjet φ . Odnosno, u priročnoj inačici, prilagodimo li neformalno tumačenje znaka ‘ $\in$ ’, postoji prirok y koji se odnosi upravo na one predmete x koji zadovoljavaju uvjet φ . Radi jednostavnosti ograničimo vrijednosti varijabla na skupove, odnosno priroke, i uzmimo sada da je naš uvjet φ “biti skup koji ne pripada sam sebi” (tj. “ x nije element od x ”) ili “biti prirok koji se ne pririče sam sebi” (tj. “ x ne oprimjeruje x ”). Iz načela sadržaja tada slijedi:

$$\exists y \forall x (x \in y \leftrightarrow x \notin x).$$

Taj iskaz kaže da postoji skup y kojega su članovi upravo oni skupovi koji nisu sami svoji članovi. Odnosno, da postoji prirok y koji je oprimjeren upravo onim prirocima koji nisu oprimjereni samima sobom. Nazovemo li taj skup ili prirok ‘ S ’, imamo:

$$\forall x (x \in S \leftrightarrow x \notin x).$$

Sada je lako vidjeti problem. Prema gornjemu iskazu, za *svaki* skup x vrijedi da je član skupa S ako i samo ako skup x nije sam svoj član. Vrijedi to za svaki skup x pa tako onda i za sam skup S . U priročno čitanju, gornji nam iskaz kaže da za svaki prirok x vrijedi da oprimjeruje prirok S ako i samo ako prirok x ne oprimjeruje sam sebe. Kako to vrijedi za svaki x , vrijedi i za prirok S . Stoga slijedi:

$$S \in S \leftrightarrow S \notin S.$$

Skup koji sadrži upravo one skupove koji sami sebe ne sadrže ujedno i jest i nije svoj član. Prirok koji se pririče upravo onim prirocima koji se ne pririču samima sebi ujedno i jest i nije oprimjeren samim sobom. Izvedeno protuslovlje jasno pokazuje da nešto nije u redu s naivnim načelom sadržaja – drugim riječima, da nešto nije u redu s našim “intuitivnim” ili “predteorijskim” razumijevanjem razreda i priroka. Što točno nije u redu, puno je teže za vidjeti.

Za razliku od Russella samoga, Frege odmah uočava važnost njegova paradoksa.¹² U odgovoru Russellu piše da ga je njegovo otkriće iznenadilo i zapanjilo i uzdrimalo temelje na kojima se nadao izgraditi aritmetiku – štoviše, da se boji da paradoks prijete ne samo njegovu već bilo kojemu mogućem utemeljenju aritmetike. No, koliko god se otkriće na prvi pogled čini neželjenim, kaže, možda će iz njega proizaći važan napredak u logici.¹³ U vrijeme kada prima Russellovo pismo, drugi sve-

¹² Iako se ponegdje tvrdi da je Russell postao svjestan važnosti paradoksa tek s Fregeovim dramatičnim odgovorom, Grattan-Guinness u [18, str. 324] upozorava da protiv toga govori već sama činjenica da Russell započinje korespondenciju s Fregeom *nakon* dovršetka *Principles* i poglavlja o paradoksu u prvom dijelu knjige. Prije znatnih izmjena teksta u korekturi, Russell u otkrivenom protuslovlju vidi ozbiljan problem za definiciju brojeva kao razreda sličnih razreda i čini mu se da bi se brojevi morali ipak prihvatiti kao nedefinirajući (usp. [2, str. 66]).

¹³ Engleski se prijevod Fregeova odgovora može naći u [20, str. 127–128]. Frege upozorava da se, strogo gledano, paradoks u obliku u kojem ga je opisao Russell ne

zak njegova glavnoga djela *Grundgesetze der Arithmetik* nalazi se u tisku i Frege knjizi uspijeva samo priložiti dodatak u kojem opisuje paradoks i iznosi – kako će se pokazati, neuspješan – prijedlog kako paradoks izbjeći.¹⁴ Od planiranoga trećeg sveska Frege odustaje, a kasniji njegovi radovi pokazuju da ga je Russellovo otkriće uvjerilo da je logicizam neistinit (usp. [16, str. 182]).

O paradoksu – ili protuslovlju, kako ga jednostavno naziva – Russell opširno piše u *The Principles of Mathematics*, posvećujući mu, između ostaloga, završno poglavlje prvoga dijela knjige, naslovljeno jednostavno “The Contradiction”. Naravno, određeni su paradoksi skupova bili poznati matematičarima i prije Russellova. Burali-Forti je paradoks najvećega rednog broja objavio još 1897., ali njime nije privukao veću pažnju.¹⁵ Cantor, doduše, paradoks najvećega kardinalnog broja, koji danas nosi njegovo ime, nije objavio, ali obznanio ga je u pismima Dedekindu 1899., a čini se da zna za antinomiju već 1895., možda čak i 1883. (usp. [19, str. 350] i [4, str. 879]).¹⁶ Štoviše, paradoks skupa skupova koji nisu svoji članovi – dakle, paradoks koji znamo kao ‘Russellov’ – od 1902., neovisno o Russellu i Fregeu, poznat je Hilbertu, Zermelu i Husserlu (usp. [4, str. 881–882]). Međutim, do Russella i Fregea paradoksi

može formulirati u njegovu sustavu – za Fregea, prirok se ne može priricati prirocima, tj. funkcija ne može kao svoj argument imati funkciju – ali da Russellovu formulaciju nije problem prikladno preinačiti.

¹⁴ U dodatku Frege piše: “Znanstvenom piscu jedva se može dogoditi nešto nepoželjnije negoli da mu se nakon dovršetka jednog rada uzdrma jedna od osnova njegove građevine. Ja sam stavljen u taj položaj pismom gospodina Bertranda Russella kad se tiskanje ovog sveska približavalo kraju...” (Prijevod Vladimira Devidéa preuzet je iz [8, str. 190].) Poznata analiza Fregeova prijedloga Quineov je “On Frege’s Way Out” [28], no Quine napominje, pozivajući se na Sobocińskog u [35], da je Leśniewski još 1938. pokazao da je Fregeov prijedlog nesuvisao.

¹⁵ Prijevod Burali-Fortijeva članka na engleski, “A Question on Transfinite Numbers”, nalazi se u [20, str. 105–112]. Svoj paradoks najvećega rednog broja Burali-Forti nije vidio kao *paradoks*, već tek kao *reductio* tvrdnje da je skup rednih brojeva linearno uređen (usp. [20, str. 104] i [4, str. 878–879]). Heijenoort u [20, str. 104] piše da je paradoks odmah izazvao zanimanje među matematičarima, ali Garciadiego u [13, str. 25–26] i, pozivajući se na nj, Griffin u [19, str. 350] tvrde da je, naprotiv, isprva potpuno ignoriran.

¹⁶ Engleski je prijevod Cantorova pisma Dedekindu od 28. srpnja 1889. u [20, str. 113–117]. U pismu Cantor razmatra i paradoks najvećega rednog broja i paradoks najvećega kardinalnog broja.

nisu prepoznati kao ozbiljan problem, bilo za matematiku bilo za logiku samu. Razloži su različiti: s jedne strane, Cantor u paradoksima vidi otkriće da postoje nesuvisle mnoštvenosti, što smatra pozitivnim doprinosom svojoj teoriji; s druge strane, Hilbert i krug matematičara oko njega paradokse skupova vide jednostavno kao prirodnu i očekivanu pojavu u rastućoj mladoj disciplini kakva je u to vrijeme teorija skupova (usp. [19, str. 351–352] i [4, str. 883]). Tek Russell i Frege u paradoksima vide dubok problem, koji nije vezan samo uz pojedinu matematičku disciplinu, već logiku i mišljenje općenito. Kako Russell kasnije piše, paradoks je pokazao da “iz premisa koje su svi logičari, bilo koje škole, prihvaćali još od vremena Aristotela, mogu se izvesti protuslovlja” [31, str. 75]. Važnost je Russellova paradoksa, četiri desetljeća nakon njegova otkrića ocjenjuje Gödel, upravo u tome što je pokazao da su naše logičke intuicije protuslovne:

Analizirajući paradokse do kojih je dovela Cantorova teorija skupova, oslobodio ih je [Russell] svih tehničkih matematičkih detalja, izvevši tako na vidjelo zapanjuću činjenicu da su naše logičke intuicije (tj. intuicije u pogledu takvih pojmova kao: istina, pojam [*concept*], biće [*being*], razred, itd.) protuslovne [15, str. 124].

U svega nekoliko godina nakon Russellova pojavljuju se u literaturi i drugi paradoksi, što, u uskom smislu, skupovnoteorijski, što, kako ih danas nazivamo, semantički.¹⁷ Kao jedan od odgovora na skupovnoteorijske paradokse, koji je u međuvremenu postao standardnim, razvijene su aksiomatske teorije skupova, prije svega ZF(C): ograničenjem aksiomatske sheme sadržaja – ili, drugim riječima, njezinom zamjenom aksiomatskom shemom odijeljenosti (separacije) – onemogućena je konstrukcija paradoksa, a dodatni aksiomi osiguravaju opstojnost pojedinih vrsta skupova, primjerice partitivnoga skupa, unije skupova i sl. Russellov je odgovor, s druge strane, logički sustav kojemu je dao ime ‘teorija tipova’.¹⁸

¹⁷ Russell mnoge od njih opisuje na početku članka “Mathematical Logic as Based on the Theory of Types” [29, str. 222–224] i u uvodu *Principia Mathematica* [37, str. 60–61]. Copi u [7, str. 1–9] daje dobar uvodni opis i analizu paradoksa, a Cantini u [4] iznosi noviji povijesni prijedlog odnosa između paradoksa i razvoja matematičke logike.

¹⁸ Zermelo u [38] aksiomatizira teoriju skupova 1908., kada je Russell većim dijelom već razvio svoju razgranatu teoriju tipova. No, aksiomatska teorija skupova za logicista Russellova tipa ne bi bila posve zadovoljavajuće rješenje, ma koliko bila tehnički pri-

Russellov odgovor

Russellova teorija tipova, kao formalni okvir njegova logicističkoga projekta svodenja matematike na logiku, svoj će konačan oblik dobiti tek u trima svescima monumentalnoga djela *Principia Mathematica* [37], pisanim u suradnji s Whiteheadom i objavljenima od 1910. do 1913. No prvi okvirni nacrt teorije tipova Russell objavljuje već 1903. u drugom dodatku *The Principles of Mathematics*, naslovljenom “The Doctrine of Types”.¹⁹ Osnovnu ideju teorije tipova opisuje na sljedeći način:

Svaka stavačna funkcija [*propositional function*] $\varphi(x)$ [...] ima, uz svoje područje istinitosti, područje značenja [*range of significance*], tj. područje unutar kojega x mora ležati da bi $\varphi(x)$ uopće bila stavak [*proposition*], bilo istinit bilo neistinit. To je prva točka u teoriji tipova; druga je točka da područja značenja tvore tipove, tj. ako x pripada području značenja funkcije $\varphi(x)$, onda postoji razred predmeta, tip x -a, od kojih svi moraju također pripadati području značenja funkcije $\varphi(x)$, kako god φ varirao; a područje značenja uvijek je ili jedan tip ili zbroj nekoliko cijelih tipova [30, str. 523].

Da bismo razumjeli Russellov prijedlog, valja pojasniti što zapravo ima na umu kada u *Principles* govori o stavcima i stavačnim funkcijama. Stavci, naime, za nj nisu ni jezični ni mentalni entiteti, ali nisu – preciznije, nisu *svi* ili barem nisu na očekivan način – ni potpuno apstraktni ili idealni predmeti. Kako je za očekivati, u onome što slijedi mnoga ćemo suptilna Russellova filozofijska razmatranja morati zanemariti ili pojednostaviti.

Russell u predgovoru kaže da riječ ‘stavak’ definira na način koji odstupa od njezine (tada) uobičajene uporabe, prema kojoj se stavci “obično shvaćaju kao (1) istiniti ili neistiniti, (2) mentalni” [30, str. xix]. Slaže se da je ‘stavak’ primjereno ime za “ono istinito ili neistinito kao takvo” [30, str. xix], ali niječe da su stavci mentalni entiteti jer “ono što je istinito ili neistinito općenito nije mentalno” [30, str. xix]. Jednako tako, stavci nisu puke rečenice: “[S]tavak, osim ako nije jezični [*lingui-*

kladna: aksiomatski *stipulirati* koji skupovi postoje nije za logicista prihvatljiv način utemeljenja.

¹⁹ Dodatak je napisan krajem 1902., vjerojatno u prosincu, no rudimentarnu zamisao teorije tipova iznosi Russell već u pismu Fregeu od 8. kolovoza, u kojem kaže da je teorija analogna Fregeovu razlikovanju funkcija različitih razina (usp. [33, str. xxxviii–xxxix]).

stic], sam ne sadrži riječi: on sadrži entitete naznačene [*indicated*] riječima” [30, str. 47]. Stavci, dakle, za Russella nisu ni jezične ni mentalne reprezentacije – činjenica ili stanja stvari, recimo – već otprilike ono što bismo smatrali samim činjenicama ili stanjima stvari, posebna vrsta složenih entiteta, općenito ontologijski neovisnih o jeziku kojim ih se izražava i umu koji ih spoznaje. Kada mislimo, vjerujemo ili tvrdimo neki stavak, taj stavak nije nešto (npr. rečenica ili misao) što tek imenuje ili znači neko stanje stvari, već upravo to stanje stvari samo. Tvrdimo li da je Mont Blanc viši od 4000 metara, ne tvrdimo misao, koja je, prema Russellu, “privatna psihološka stvar” [11, str. 169], već sâm predmet te misli, složeni entitet kojega je sastavnica upravo planina o kojoj je riječ, a ne tek njezino ime, predodžba ili bilo kakav drugi posrednik.²⁰ Slično, sastavnicom toga složenog entiteta nije izraz ‘viši od 4000 metara’, nije ni mentalna predodžba koja odgovara tomu izrazu, ali jest *pojam* koji je tim izrazom naznačen.

Posljednja rečenica zahtijeva dodatno pojašnjenje. Za početak, važno je napomenuti da Russellovi metafizički pogledi u *Principles* uključuju svojevrstu vrlo jaku inačicu ontologijskoga realizma:

Sve što može biti predmet mišljenja, ili što se može pojaviti u bilo kojem istinitom ili neistinitom stavku, ili računati kao *jedno*, nazivam *oznakom* [*term*]. To je tada najšira riječ u filozofijskom rječniku. Rabiću kao njoj istoznačne riječi jedinica [*unit*], pojedinačnost [*individual*] i biće [*entity*]. Prve dvije naglašavaju činjenicu da je svaka oznaka *jedno*, dok je treća izvedena iz činjenice da svaka oznaka ima bitak [*has being*], tj. u nekom smislu *jest*. Čovjek, trenutak, razred, relacija, himera, ili bilo što drugo što može biti spomenuto, svakako je oznaka; a zanimljivo da je takva i takva stvar oznaka uvijek mora biti neistinito [30, str. 43].

Sve što uopće može biti predmetom mišljenja *jest*, ima bitak, oznaka je. Oznake Russell dijeli na pojmove (*concepts*) i stvari (*things*). Pojmovi su ono što je naznačeno pridjevima i glagolima, tj. ono što naziva ‘priroci’ i ‘relacije’. Pojmovi u stavcima mogu imati dvije uloge. Osim kao podmet, ono o čemu je stavak, pojam se u stavku može “pojaviti i na drukčiji način” [30, str. 45] – naime, pojam može biti prirican oznakama, tj. odnositi se na podmete stavka, sâm biti o oznakama u stavku. Stvari su pak one oznake koje nisu pojmovi, tj. koje se ne mogu priricati drugim

²⁰ Primjer je iz pisma Fregeu od 12. prosinca 1904. Engleski se prijevod može naći u [11, str. 166–170].

oznakama – ili, drugim riječima, oznake koje u stavku mogu biti *samo* podmeti (usp. [30, str. 44–45]).²¹ Pojmovi pritom za Russella nisu, ništa više nego planine, tek nekakvi mentalni entiteti, predodžbe, psihološke sastavnice misli, već zasebni, o našim mislima i tvrdnjama neovisni idealni predmeti, jednako stvarni koliko i planine same.²²

Oznake su ili jednostavne ili složene, tj. cjeline (*wholes*). Neke cjeline ne pretpostavljaju nikakvu dodatnu uzajamnu povezanost svojih dijelova – cjelina je u tom slučaju “potpuno specificirana kada su sve njezine jednostavne sastavnice specificirane” [30, str. 139]. Takve cjeline naziva Russell ‘agregati’. Druge pak cjeline, one koje među svojim sastavnicama sadrže barem jedan pojam u njegovoj priričućoj ulozi, koji time strukturira cjelinu povezujući njezine dijelove na način na koji nisu međusobno povezani u agregatima – i koje stoga “nisu potpuno specificirane kada su svi njihovi dijelovi znani” [30, str. 140] – naziva Russell ‘jedinstva’ (*unities*) ili ‘stavci’. U primjeru koji daje, imamo li proizvoljne oznake *A* i *B* i pojam razlike, te tri oznake zajedno tvore jedan jedini agregat, ali sastavnice su dvaju različitih stavaka, “*A* se razlikuje od *B*” i “*B* se razlikuje od *A*” (usp. [30, str. 140]). Općenito, stavak da neka oznaka *A* ima svojstvo *P* složeni je entitet ili cjelina koje su sastavnice oznaka *A* i pojam *P*; stavak da je oznaka *A* u relaciji *R* s oznakom *B* složeni je entitet sastavljen od oznaka *A* i *B* i pojma *R* itd., no u svim tim slučajevima relevantan pojam nije tek jedna od oznaka u cjelini, već se u njoj javlja “kao povezujući ili kvalificirajući” [30, str. 140]. Oznake *A* i *B* pritom također mogu biti pojmovi ili pak mogu biti stvari. Kako su stavci, bili istiniti ili neistiniti, cjeline – složeni entiteti koji i sami mogu biti predmeti mišljenja – i istiniti i neistiniti stavci *jesu*, imaju bitak, oznake su.²³

²¹ Maloj terminološkoj zbrici pridonosi to što Russell ponegdje rabi ‘pridjev’ i ‘glagol’ za riječi koje naznačuju (*indicate*) pojmove, ponegdje pak za same pojmove, tj. za priroke i relacije. Također, ponegdje ‘oznaka’ rabi u užem smislu samo za podmet stavka pa govori o “*oznakama* stavka kao onim *oznakama*, koliko god brojne, koje se javljaju u stavku i mogu se shvatiti kao podmeti o kojima je stavak” [30, str. 45]. Tako, primjerice, u istom odsječku u kojem kaže da pridjevi naznačuju priroke, a glagoli relacije, ujedno piše da su priroci “pojmovi, osim glagola, koji se javljaju u stavicima koji imaju samo jednu oznaku ili podmet” [30, str. 45].

²² Prema Russellu, teorija po kojoj bi “pridjevi ili atributi ili idealne stvari, ili kako god ih zvali” bili nekako ontologijski manjkavi “čini se potpuno pogrešnom i lako svedljivom na protuslovlje” [30, str. 46].

²³ Teorija se usložnjava time što imati bitak *ne* znači ujedno i postojati. Russell strogo razlikuje bitak i opstojnost: “Razlika između bitka [*Being*] i opstojnosti [*existence*]

Primijetimo da Russell u iznad citiranom tekstu izričito tvrdi da sve što uopće *jest* jest neko *jedno* – ‘entitet’ i ‘pojedinačnost’ za nj su isto- značnice, koje se jednako odnose i na fizičke predmete i na razrede i na stavke. Imati na umu Russellovo prijanjanje u *Principles* uz tu “staru maksimu da ‘što god jest, jedno je’” [32, str. 189] važno je za razumijevanje finih nijansi njegove analize paradoksa i odgovora koji na nj daje.

Znajući sada što su za Russella stavci, možemo približno odgovoriti na pitanje što su stavačne funkcije. Pojam stavačne funkcije za Russella je jedan od temeljnih pojmova (*fundamental notions*), koji se stoga može objasniti, ali ne i definirati u pravom smislu (usp. [30, str. 19]). Objasnjavajući pojam, Russell kaže: “ $\varphi(x)$ je stavačna funkcija ako je, za svaku vrijednost x -a, $\varphi(x)$ stavak, određen [*determinate*] kada je dan x ” [30, str. 19]. Međutim, to je objašnjenje nedovoljno precizno – nije *svaka* funkcija koja svojim argumentima pridružuje stavke već samim time stavačna funkcija. Primjerice, funkcija koja Sokratu pridružuje stavak “Kiša pada”, Platonu stavak “Ulice su mokre”, a broju dva Pitagorin poučak nije stavačna funkcija. Stavačna funkcija svakomu argumentu pridružuje stavak koji kao svoju *sastavnicu* sadržava taj isti argument. Štoviše, dva različita stavka koji su vrijednosti iste stavačne funkcije $\varphi(x)$ za dva različita argumenta a i b moraju biti takvi da se međusobno razlikuju *samo* utoliko što stavak $\varphi(a)$ (tj. stavak koji je vrijednost funkcije $\varphi(x)$ za argument a) kao svoju sastavnicu sadržava oznaku a , a stavak $\varphi(b)$ (tj. stavak koji funkcija $\varphi(x)$ pridružuje argumentu b) sadržava oznaku b . Russellov je primjer stavačne funkcije “ x je čovjek”:

U svakom stavku, koliko god složenom, koji ne sadržava slobodne [*real*] varijable, možemo zamisliti da je jedna od oznaka, ne glagol niti pridjev, zamijenjena drugom oznakom: umjesto “Sokrat je čovjek” možemo staviti “Platon je čovjek”, “broj 2 je čovjek” i tako dalje. Time dobivamo uzastopne stavke koji se slažu u svemu osim u jednoj promjenjivoj [*variable*] oznaci. Stavivši x za promjenjivu oznaku, “ x je čovjek” izražava tip svih takvih stavaka [30, str. 19–20].²⁴

važna je, a dobro se ilustrira postupkom brojanja. Što se može brojati, mora biti nešto i svakako mora *biti*, iako nipošto ne mora posjedovati daljnju povlasticu postojanja” [30, str. 71]. Time stvarima, pa onda i podmetima stavaka, mogu biti i “mnoge oznake koje ne postoje, primjerice točke u neeuclidskom prostoru i pseudopostojeći entiteti [*pseudo-existents*] nekoga romana” [30, str. 45]. Razlika između opstojnosti i bitka ključna je za razlikovanje istinitih i neistinitih stavaka, ali nama ta razlika ovdje nije toliko bitna.

²⁴ Izrazi ‘glagol’ i ‘pridjev’ ne odnose se pritom na riječi koje naznačuju pojmove (relacije i priroke), nego na pojmove same (vidi bilj. 21). Upozoriti na Russellovu

Pojedina se stavačna funkcija, prema Russellu, može shvatiti kao “tip” stavaka, tj. kao “razred svih stavaka koji nastaju varijacijom jedne oznake” [30, str. 93] ili, drugim riječima, kao razred svih stavaka koji dijele određenu postojanost oblika (*constancy of form*) (usp. [30, str. 89]). Primjerice, stavačnu funkciju “ x je čovjek” može se shvatiti kao razred stavaka {“Sokrat je čovjek”, “Platon je čovjek”, “Broj dva je čovjek”, ...}. Russell, međutim, napominje da je pojam stavačne funkcije već pretpostavljen u definiciji takvih razreda i da stoga to pojašnjenje ne treba razumjeti kao *definiciju* pojma stavačne funkcije (usp. [30, str. 89, 509]).²⁵ To će nam pojašnjenje morati dostajati: pokušati kazati više o stavačnim funkcijama, odvelo bi nas predaleko.²⁶

Vidimo, dakle, da tako međusobno različite stvari kao što su filozofi i brojevi mogu biti argumentima iste stavačne funkcije. Russell, naime, logiku razumije kao potpuno univerzalnu znanost: logičke su istine istinite jednako o filozofima koliko i o brojevima. Logika stoga, za razliku od drugih znanosti, nema neko zasebno, ograničeno predmetno područje: ništa što u bilo kojem smislu jest ne leži izvan njezina dohvata. Takav se pogled na logiku već na samom početku *Principles* odražava u Russellovu zahtjevu da logika (i time matematika) mora imati samo jednu vrstu “istinskih” varijabla, vrijednosti kojih mogu biti bilo koji entiteti. U protivnom, ako bi bilo što – Sokrat, recimo, ili broj dva, ili stavak da se Sokrat razlikuje od broja dva – ležalo nekako izvan područja argumenata stavačnih funkcija poput “ako je x čovjek, onda je x čovjek”, “ x je istovjetan x -u” i sl., ne bi doista bila riječ o potpuno univerzalnim istinama. Dopušteno je, dakako, po potrebi ograničiti moguće vrijednosti varijabla ili uvoditi posebne vrste varijabla (npr. za brojeve ili za

nemarnost u razlikovanju onoga što izraz označava od samoga izraza – drugim riječima, njegovu nepažljivost u razlikovanju između uporabe (*use*) i spominjanja (*mention*) izrazâ – standardna je napomena u literaturi (usp. [24, str. 34] i [23, str. 634]).

²⁵ Kako smo spomenuli u bilj. 5, Russell opći pojam razreda definira pomoću pojma stavačne funkcije te bi stoga definicija stavačne funkcije u terminima razredâ bila cirkularna.

²⁶ Kako točno valja razumjeti Russellove stavačne funkcije stari je interpretacijski problem. Primjerice, Geach je još sedamdesetih napisao da je način na koji Russell u *Principia Mathematica* rabi izraz ‘stavačna funkcija’ “beznadno zbrkan i nesuvisao” [14, str. 217], a Cartwright, tridesetak godina kasnije, da “pokušaji kazati što točno russellovska stavačna funkcija jest, ili što bi trebala biti, obvezno završavaju u frustraciji” [6, str. 915].

razrede), ali varijable koje se javljaju u rečenicama logike i matematike kada su precizno i u cijelosti izložene kao svoje vrijednosti mogu imati bilo što:

[U] svakom stavku čiste matematike, kada je u cijelosti izložen, varijable imaju apsolutno neograničeno područje [*field*]: bilo koji pomišljiv [*conceivable*] entitet može se supstituirati za bilo koju od naših varijabla bez štete za istinitost našega stavka [30, str. 7].

To stajalište tvori važan dio Russellova shvaćanja logike iz vremena *Principles* (ali i daleko nakon njih) i postat će kasnije poznato pod imenom ‘nauk o neograničenoj varijabli’ (*doctrine of the unrestricted variable*).²⁷ Prema nauku o neograničenoj varijabli, svaka stavačna funkcija *svako*j oznaci pridružuje odgovarajući stavak: *bilo što* što jest, moguća je vrijednost varijable bilo koje stavačne funkcije.

Međutim, kako vidimo iz citata na početku odsječka, čini se da Russell, odgovarajući u dodatku knjizi na otkriveni paradoks, napušta nauk o neograničenoj varijabli. Prema predloženoj teoriji tipova, ne može bilo koji predmet biti argument bilo koje stavačne funkcije. Za neke vrijednosti varijable stavačna funkcija nije definirana, tj. argumentu ne pridružuje ni istinit ni neistinit stavak. Moguće vrijednosti varijable dane stavačne funkcije ograničene su na ono što Russell naziva ‘područjem značenja’ te funkcije.

Na koji način ograničenje mogućih vrijednosti varijabla stavačnih funkcija pomaže Russellu izbjeći paradoks? Prisjetimo se, u našoj smo analizi s početka ovoga rada kazali da je izvor paradoksa pretpostavka da za svaki uvjet φ postoji skup kojega su elementi upravo oni predmeti koji zadovoljavaju φ – odnosno, dijelom sada preuzevši Russellovo nazivlje, da za svaku stavačnu funkciju $\varphi(x)$ postoji skup upravo onih predmeta kojima $\varphi(x)$ pridružuje istinit stavak. Međutim, u izvodu protu-

²⁷ Što se točno tvrdi naukom o neograničenoj varijabli donekle je sporno. Nakon Heijenoortova utjecajna članka “Logic as Calculus and Logic as Language” [21] postalo je uobičajeno smatrati da nauk o neograničenoj varijabli povlači da logika mora sadržavati svoju metateoriju, ali Landini u [25, str. 268–269] piše da je nauk jednostavno sintaktički zahtjev da svaka formalna logička teorija mora usvojiti samo jednu vrstu varijabla (individualne ili entitetske varijable) i upozorava da je “važno ne pomiješati Russellov sintaktički nauk o neograničenoj varijabli sa semantičkim pitanjima koja se tiču dopustivoga područja varijabla”.

slovlja implicitno smo pretpostavili također to da je *skup* svih predmeta koji zadovoljavaju uvjet φ i sam jedan od predmeta za koje je legitimno pitati zadovoljava li φ – odnosno, da je taj skup i sam jedan od legitimnih argumenta stavačne funkcije $\varphi(x)$. Bez potonje se pretpostavke očito ne može izvesti protuslovlje da dotični skup ujedno i zadovoljava i ne zadovoljava φ , tj. da mu funkcija $\varphi(x)$ ujedno pridružuje i istinit i neistinit stavak. No ta pretpostavka trivijalno slijedi iz nauka o neograničenoj varijabli. Odbacivši nauk o neograničenoj varijabli i prikladno ograničivši moguće argumente stavačne funkcije, Russell onaj predmet koji je skup svih predmeta kojima dana funkcija pridružuje istinit stavak ostavlja izvan njezina područja značenja. Problematičan je skup uklonjen iz legitimnih argumenata te iste funkcije – njemu ona ne pridružuje nikakav stavak i protuslovlje je time izbjegnuto.

Russellov bismo prijedlog mogli opisati na sljedeći način. Suprotno nauku o neograničenoj varijabli, ne pridružuje svaka stavačna funkcija svakomu predmetu neki stavak, nego samo onim predmetima koji se nalaze u njezinu području značenja. Ta područja značenja tvore tipove – odnosno, ona nisu takoreći nasumično pridružena pojedinim funkcijama, nego postoji sustavan hijerarhijski način na koji su moguće vrijednosti varijabla stavačnih funkcija ograničene. S jedne strane, za svaki predmet x postoji neki maksimalan skup predmetâ T takav da je *svaka* stavačna funkcija $\varphi(x)$ koja je definirana za x definirana ujedno i za *svaki* član skupa T .²⁸ Takve skupove Russell naziva minimalnim tipovima. Ako je stavačna funkcija definirana za *neki* predmet iz nekoga minimalnog tipa – drugim riječima, ako taj predmet sadržava u svojem području značenja, tj. pridružuje mu bilo istinit bilo neistinit stavak – onda je ona definirana i za *svaki* predmet iz toga minimalnog tipa. Uzmemo li neku pojedinačnost, primjerice Sokrata, postoji neki maksimalan skup predmetâ – ovom je slučaju to skup svih pojedinačnosti – takav da je svaka stavačna funkcija koja Sokratu pridružuje bilo istinit bilo neistinit stavak definirana i za svaki član toga skupa, tj. za svaku pojedinačnost. Skup pojedinačnosti stoga je minimalan tip. Skup skupova pojedinačnosti također je minimalan tip: ako neka stavačna

²⁸ U tekstu citiranom na početku odsječka Russell ne navodi eksplicitno uvjet maksimalnosti, ali on se čini nužnim: u protivnom bi i skup $\{x\}$ bio tip, a to nije u skladu s Russellovim opisom teorije tipova dalje u tekstu.

funkcija ima vrijednost za neki skup pojedinačnosti kao svoj argument, onda ima i za svaki skup pojedinačnosti. Na temelju ideje da područja značenja tvore tipove, opisuje dalje Russell hijerarhiju tipova (i time hijerarhiju njima odgovarajućih stavačnih funkcija): skup svih oznaka ili pojedinačnosti najniži je minimalan tip, idući je minimalan tip skup svih skupova pojedinačnosti, zatim skup kojega su članovi svi skupovi skupova pojedinačnosti, i tako dalje. Općenito, piše Russell, skup od svojih elemenata “mora uvijek biti za jedan višega tipa” [30, str. 525]. S druge strane, ako je predmet određenoga minimalnog tipa argument neke stavačne funkcije, onda predmet nekoga drugog minimalnog tipa općenito nije legitiman argument te iste funkcije. Primjerice, ako neka funkcija skupu {Sokrat, Platon} pridružuje odgovarajući stavak, onda ona tipično ne će biti definirana ni za pojedinačnost poput Sokrata ni za skup poput {{Sokrat, Platon}, {Platon, Aristotel}}.²⁹ Skup svih argumenata kojima $\varphi(x)$ pridružuje istinit stavak, kako je taj skup različitoga minimalnog tipa od svojih članova, tj. od argumenata funkcije $\varphi(x)$, tipično ne će biti i sam jedan od legitimnih argumenata te iste funkcije. Time teorija tipova izbjegava paradokse poput onoga o skupu svih skupova koji nisu svoji članovi. Pored opisane hijerarhije tipova, zasebni su tipovi i skup svih uređenih parova pojedinačnosti, skup svih skupova uređenih parova pojedinačnosti, tj. skup dvomjesnih relacija (ekstenzionalno shvaćenih) između pojedinačnosti, skup dvomjesnih relacija između uređenih parova i sl. Zanimarimo li to što Russell dopušta da se područje značenja *nekih* stavačnih funkcija sastoji od zbroya više minimalnih tipova, dobivamo tako standardnu tipsku strukturu onoga što danas nazivamo ‘jednostavna teorija tipova’.

Međutim, opis Russellova prijedloga u prethodnim odlomcima ne govori nam cijelu priču. Pažljivijem čitatelju dva su detalja mogla zapeti za oko. Najprije, predmeti najnižega tipa nazvani su ‘oznake’ i ‘pojedinačnosti’. No vidjeli smo da Russell izraze ‘oznaka’ i ‘pojedinačnost’ razumije kao istoznačnice izrazu ‘entitet’ – dakle, kao

²⁹ Postoje iznimke. Russell smatra da nisu samo minimalni tipovi tipovi: zbroj nekoliko minimalnih tipova također može biti tip, tj. područje značenja neke stavačne funkcije. Štoviše, kaže da mu se čini da “zbroj bilo kojega broja minimalnih tipova jest tip” [30, str. 525]. Razlog za to su stavačne funkcije poput “ x je istovjetan x -u”, koja daje (istinit) stavak za bilo koju vrijednost varijable – tj. za bilo koji predmet, bilo kojega minimalnog tipa – i koja stoga ima posve neograničeno područje značenja.

izraze koji se odnose na sve što uopće može biti predmetom mišljenja, na sve što ima bitak. Štoviše, vidjeli smo da i *razrede* Russell ubraja među pojedinačnosti. Kako to pomiriti s teorijom prema kojoj su pojedinačnosti samo neke od mogućih vrijednosti varijabla, tj. s teorijom prema kojoj skupovi *nisu* pojedinačnosti? Što Russell ima na umu kada u izlaganju teorije tipova govori o pojedinačnostima? Kao drugo, Russellov smo odgovor na protuslovlje opisali govoreći o *skupovima*. No, kako vidimo iz dijelova *Principles* koje smo citirali, termin koji Russell redovito rabi je ‘razred’, ne ‘skup’. Ima li kakve relevantne razlike između skupova i Russellovih razreda? Iako se riječ ‘skup’ (*set*) pojavljuje na mnogim mjestima u *Principles*, ona nema tehničko značenje kakvo imaju ‘razred’ (*class*) i s njim povezani izrazi poput ‘kolekcija’ (*collection*) i ‘agregat’ (*aggregate*) (usp. [18, str. 269]). Russellovi razredi pritom doista *ne* odgovaraju u potpunosti onomu što danas standardno razumijemo pod terminom ‘skup’. Strogo gledano, ontologija *Principles* – barem do samoga prijedloga teorije tipova u dodatku knjizi – uopće ne sadržava skupove.

Za razumijevanje Russellova odgovor na paradoks – i time načina na koji je on sam razumio paradoks koji je otkrio – moramo sada malo detaljnije pogledati o čemu Russell zapravo govori kada govori o razredima.

Russellovi razredi

Opisati Russellove filozofijske poglede na narav razredā i kako su se ti pogledi mijenjali tijekom pisanja *The Principles of Mathematics* nije jednostavan zadatak. Knjiga je objavljena u svibnju 1903., ali Russell na njoj radi od 1897., tijekom godina često i znatno revidirajući svoja stajališta. Različiti dijelovi objavljenoga teksta napisani su u različitim razdobljima, omeđenim trima ključnim događajima, i mnogi su ostatci u međuvremenu napuštenih pogleda pronašli put u završni tekst knjige. Prvi je ključan događaj Russellovo upoznavanje s Peanom početkom kolovoza 1900. i oduševljenje njegovim djelima. Drugi je otkriće paradoksa u travnju ili svibnju 1901. Treći je pažljivo čitanje Fregeovih *Grundgesetze* u lipnju 1902. (nakon što su *Principles* već predani u tisak) i korespondencija s Fregeom koja je uslijedila nakon toga. Od sedam glava od kojih se sastoje *Principles* većina je napisana do kraja 1900., za

trajanja peanovskoga medenog mjeseca, no prva i druga – koje su ujedno filozofijski najznačajniji dio knjige – dovršene su kasnije.³⁰ Završnu inačicu druge glave (“Number”) napisao je Russell u lipnju 1901., a prve (“Indefinables of Mathematics”) tek u svibnju 1902., no uključivši u nju gotovo sav tekst prethodne inačice napisane godinu dana ranije. Međutim, u obje se glave nalaze i reciklirani dijelovi ranijega pretpeanovskog rukopisa. Također, obje glave – uključujući šesto poglavlje prve glave (“Classes”), u cijelosti posvećeno raspravi o razredima, i posebice deseto poglavlje (“The Contradiction”), koje uključuje nekoliko različitih formulacija Russellova paradoksa, pokušaj logičke dijagnoze i naznake predloženih rješenja – sadržavaju znatne naknadne izmjene unesene tijekom korekture, često s uočljivim Fregeovim utjecajem. Dva dodatka, važna za razumijevanje Russellovih, uvjetno govoreći, konačnih pogleda na razrede u *Principles*, napisana su krajem 1902. Prvi dodatak (“The Logical and Arithmetical Doctrines of Frege”), u kojem Russell, uspoređujući ih sa svojim, izlaže osnovne pojmove i ideje Fregeova pogleda na logiku i aritmetiku i modificira svoj nauk o razredima, dovršen je u studenom, a drugi (“The Doctrine of Types”), u kojem iznosi teoriju tipova kao prijedlog odgovora na paradoks, možda tek u prosincu. Korektura knjige, s većim ili manjim izmjenama teksta, trajala je do u veljaču 1903.³¹ Russellova se razmatranja o razredima stoga na mno-

³⁰ Sastavnih sedam cjelina koje tvore *Principles* naziva Russell jednostavno ‘dijelovi’ (*parts*), no radi jasnoće, tj. kako bismo izbjegli moguću zabunu oko toga govorimo li u tekstu o samim tim sastavnim cjelinama ili o proizvoljnim dijelovima *Principles*, nazivamo ih ovdje ‘glave’.

³¹ Detaljniji opis povijesti i konteksta nastanka *Principles*, kao i Russellovih ranijih stajališta i utjecaja na njih, može se naći u [18]. Kronološka tablica nastanka pojedinih dijelova knjige nalazi se na str. 294 i u njoj je dovršetak obaju dodataka datiran u početak studenoga 1902. Rani rukopis iz 1899.–1900., koji prethodi Peanovu utjecaju, i rekonstruirana inačica prve glave *Principles* iz svibnja 1901., u kojoj Russell prvi put piše o otkrivenom protuslovju, objavljene su u [33], s informativnim uvodom i pratećim tekstovima urednika G. H. Moorea. Kronološka tablica s podacima o Russellovu životu i filozofijskim i matematičkim radovima u razdoblju između kraja kolovoza 1899. i svibnja 1903. nalazi se na str. liii–lviii, no bez podatka o tome kada su napisani dodatci. Međutim, piše da je izdavač 15. studenog 1902. zaprimio dodatak A (usp. [33, str. lvi]), a Moore u uvodnom tekstu kaže da je dodatak B napisan vjerojatno u prosincu (usp. [33, str. xxxix]). Promjene koje je Russell u korekturi naknadno unio u predgovor i tekst prve glave sabrao je Blackwell u [1], a za drugu je glavu isto učinio Byrd u [2]. Blackwell u [1, str. 284] napominje da je moguće da su i oba dodatka naknadno revidirana.

gim mjestima u *Principles* međusobno znatno razlikuju – ponegdje čak izravno protusloveći jedna drugima – čuvajući tragove različitih “rješenja koja su [Russella] neko vrijeme zadovoljavala” [30, str. xix]. Slijedeći te tragove, pokušat ćemo rekonstruirati kako je Russell, u svjetlu svojih metafizičkih pogleda na narav razredā, sam razumio paradoks koji je otkrio i rješenje koje je na nj predložio. Kao i često do sada, određena pojednostavljivanja bit će pritom neizbježna.

Temeljna distinkcija u Russellovoj metafiziци razredā u *Principles* ona je između razreda kao mnoštva (*class as many*) i razreda kao jednoga (*class as one*). Razred kao mnoštvo jednostavno je mnoštvenost ili množina (*plurality*) danih oznaka ili pojedinačnosti – preciznije, svih onih oznaka koje zadovoljavaju neku stavačnu funkciju – no množina oznaka shvaćena upravo *qua* množina, tj. ne kao *jedan* entitet, ne kao zasebna *čjelina* konstituirana tim oznakama. Razred kao mnoštvo sâm nije oznaka, nije pojedinačnost, nije zaseban entitet, nego samo mnoštvenost oznaka, množina pojedinačnosti, mnoštvo danih entiteta. Za razred kao mnoštvo rabi Russell također izraz ‘kolekcija’.³² Iako kolekcije, strogo gledano, nemaju bitak – u smislu da nisu *zasebni* entiteti, nisu oznake – Russell i o njima govori kao o predmetima (*objects*), što, suprotno tekstu koji smo citirali na početku drugoga odsječka, upravo ‘predmet’ čini najširom riječi u njegovu filozofijskom rječniku.³³

Razred kao jedno, s druge strane, jest *čjelina* sastavljena od danih pojedinačnosti, jedinstven složeni entitet koji tvore pojedinačnosti u pitanju. Razred kao jedno i sam je, dakle, neko *jedno*, tj. zasebna oznaka ili pojedinačnost. Od stavaka, koji su, kako smo vidjeli, također cjeline ili složeni entiteti, razred kao jedno razlikuje se time što je “određen čim su njegove sastavnice znane” [30, str. 139]. Primjerice, imamo li oznake *A* i *B* i relaciju *R*, oni zajedno tvore jedan jedini razred kao jedno, iako su sastavnice dvaju različitih stavaka – naime, stavka da je *A* u relaciji *R*

³² Upozorava pritom da je, iako je riječ ‘kolekcija’ gramatički u jednini, ono na što se ona odnosi “bitno množina [*essentially plural*]” [30, str. 70]. Očito, sličan je problem s našim prijevodom ‘razred kao mnoštvo’. Precizniji bi prijevod bio ‘razred kao mnoga’, no činio nam se jezično ekstravagantnim.

³³ U podrubnici naknadno dodanoj u korekturi (usp. [1, str. 285]) Russell napominje da riječ ‘predmet’ rabi “u širem smislu od *oznaka*, tako da pokriva i jedninu i množinu”, no upozorava da mogućnost riječi sa širim značenjem od ‘oznaka’ vodi “teškim logičkim problemima” [30, str. 55].

prema *B* i stavka da je *B* u relaciji *R* prema *A*. Razred kao jedno, drugim riječima, ono je što Russell naziva ‘agregat’.³⁴

Razlika između razreda kao mnoštva i razreda kao jednoga, kaže Russell, često se povlači u jeziku. Uzevši neke od njegovih primjera, govorimo li o prostornim točkama, trenutcima, vojnicima i Francuzima, govorimo o razredima kao mnoštvima; govorimo li o prostoru, vremenu, vojsci i francuskoj naciji, riječ je o razredima kao jednima (usp. [30, str. 68, 524]). Stavci poput “Sokrat je jedan od ljudi” izražavaju relaciju između entiteta i množine entiteta, između oznake i razreda kao mnoštva kojemu ta oznaka pripada. S druge strane, stavci poput “Sokrat pripada ljudskoj vrsti” izražavaju relaciju između dvaju entiteta, između oznake i razreda kao jednoga – u ovom slučaju, “relaciju između Sokrata i ljudske vrste, tj. između oznake i njezina razreda shvaćenoga kao cjeline” [30, str. 78].

Razred kao mnoštvo i razred kao jedno mogu, dakako, sadržavati posve iste oznake, ali oni – barem u slučaju kada se sastoje od više od jedne pojedinačnosti – nisu time *isti* predmet. Odnosno, nije tako da je jedan te isti razred “jedno u jednom smislu, a mnoštvo u drugom” [30, str. 76]. Naime, kaže Russell, “kad god se razred sastoji od više od jedne oznake, može se dokazati da takvo poistovjećivanje nije dopustivo” [30, str. 76].³⁵ No i razred kao mnoštvo i razred kao jedno logički su

³⁴ Razlikovanje između kolekcija i agregata prisutno je već u Russellovu prepeanovskom rukopisu iz 1899.–1900., iako u njem još ne rabi izraze ‘razred kao mnoštvo’ i ‘razred kao jedno’. Dijelovi poglavlja “Meaning of Whole and Part” toga rukopisa uključeni su u poglavlje “Whole and Part” završnoga teksta *Principles*. Knjiga stoga sadržava malu terminološku neujednačenost glede izraza ‘cjelina’ i ‘agregat’. Naime, u spomenutom se poglavlju rukopisa ti izrazi rabe samo za *složene* (*complex*) oznake (usp. [33 str. 35]). To je značenje preneseno i u odgovarajuće poglavlje u *Principles* (usp. [30, str. 137]) pa se u njem ‘cjelina’ ne odnosi na *jedinični* razred kao jedno, tj. na razred kao jedno koji sadržava samo jednu oznaku. Stoga na kraju poglavlja Russell piše da “ono što smo zvali razredima kao jednima može se uvijek, osim ako sadržavaju jednu ili nijednu oznaku, [...] protumačiti kao agregat” [30, str. 141]. Međutim, u (kasnije pisanom) poglavlju “Classes”, Russell rabi ‘cjelina’ u širem smislu tako da uključuje i cjeline sastavljene od svega jedne oznake (usp. [30., str. 77]).

³⁵ Razlog koji navodi taj je što različiti pojmovi označavaju (*denote*) razred kao mnoštvo i razred kao jedno, no ako bi razred kao mnoštvo bio istovjetan odgovarajućemu razredu kao jednom, “slijedilo bi da što god denotira jedan, mora denotirati i drugi” [30, str. 76]. Kazati više o tome tražilo bi ekskurs u Russellovu teoriju označavajućih pojmova.

neophodni. S jedne strane, razred kao mnoštvo ono je čemu se pririču brojevi veći od jedan.³⁶ Drugim riječima, samo se u slučaju razreda kao mnoštva, sadržava li više od jedne oznake, radi o “predmetu” koji *su* oznake, a ne koji *je* oznaka (usp. [30, str. 68]).³⁷ Primjerice, stavak “Sokrat i Platon su dvojica” ne može biti o cjelini sastavljenoj od Sokrata i Platona jer cjelina je, kao zaseban entitet, “samo i bitno jedno” [30, str. 68]. No Russell upozorava da se razred kao mnoštvo, kako on *nije* zaseban entitet, ne smije shvatiti kao jedinstven logički podmet takvih stavaka. *Samo* oznake mogu biti logički podmeti: sve što je podmet nekoga stavka, kako je samim time pojedinačan predmet misli, ujedno je i oznaka. Stoga, kada se brojevi pririču razredu kao mnoštvu, stavak u pitanju nema jedan podmet, nego množinu podmetâ:

U takvim stavcima kao “*A i B su dvoje*” ne postoji logički podmet: ta tvrdnja nije ni o *A* ni o *B* ni o cjelini sastavljenoj od oboje, nego strogo i samo o *A i B*. Stoga se čini da tvrdnje nisu nužno o jedinstvenim [*single*] podmetima, nego mogu biti o mnogim [*many*] podmetima [30, str. 76–77].

S druge strane, “nužno je da možemo brojati svaki razred kao jedan” [30, str. 76] i uobičajeno “govorimo o jednoj kolekciji, jednoj mnoštvenosti [*manifold*] itd., a čini se da u svakom od tih slučajeva doista postoji nešto što je jedna oznaka” [30, str. 141]. Stoga su nam, uz razrede kao mnoštva, potrebni i agregati svih oznaka danoga razreda.

Isprva Russell upravo razred kao jedno uzima kao onaj smisao ‘razreda’ koji je relevantan za simboličku logiku i, samim time, za matematiku. Relacija pripadnosti razredu, koju označavamo simbolom ‘ $\in$ ’, relacija je između dviju oznaka – dakle, relacija oznake spram razreda kao jednoga.³⁸ Naime, piše Russell, razred kao mnoštvo “ne može biti *kao takav* predstavljen jednim slovom: stoga u svakoj mogućoj simboličkoj logici slova koja stoje za razrede ne mogu predstavljati razrede *kao*

³⁶ Preciznije govoreći, brojevi se *ne* pririču jer, prema Russellu, brojevi su određena vrsta razredâ (razredi ekvipotentnih razreda) i stoga nisu pojmovi nego stvari. Ispravno bi bilo napisati da su razredi kao mnoštva ono o čemu se može tvrditi da su nekoga broja, ali, kako nam to ovdje nije presudno, dopuštamo radi jednostavnosti jezika tu nepreciznost.

³⁷ Prisjetimo se, ‘predmet’ se ne odnosi samo na oznake, tj. entitete, nego i na množine oznaka (usp. bilj. 33).

³⁸ Malo preciznije, Russell ne rabi danas standardni simbol ‘ $\in$ ’, nego Peanov ‘ ϵ ’.

mnoštva” [30, str. 78].³⁹ Iz toga razloga ‘ $\in$ ’, povezujući u matematičkoj notaciji dva simbola, “ne može predstavljati relaciju oznake prema njezinoj razredu kao mnoštvu” [30, str. 78].⁴⁰ Matematički stavci o razredima, dakle, stavci su o agregatima. Ta se napomena može učiniti kao stvar puke filozofske sitničavosti jer moglo bi djelovati očitim – i svakako je isprva tako djelovalo Russellu – da za svaku kolekciju entiteta postoji ujedno agregat tih istih entiteta.⁴¹

Uočimo da razredi, bilo kao kolekcije bilo kao agregati, nisu za Russella apstraktni predmeti poput skupova. Razred kao mnoštvo nije nešto povrh ili pored množine pojedinačnosti koje ga tvore, već upravo ta sama množina pojedinačnosti. Kolekcija svih slika na zidu, primjerice, upravo je množina slika, te konkretne slike same. Slično, agregat svih slika na zidu cjelina je tih istih slika, konkretan složeni entitet kojega su slike u pitanju konkretni dijelovi. Stoga za Russella – za razliku od Peana – ne postoji razlika između razreda, bilo kao kolekcije bilo kao agregata, koji sadrži jednu jedinu oznaku i te njegove jedine oznake: mnoštvo od jedne jedine pojedinačnosti jednostavno je ta jedna pojedinačnost, a cjelina sastavljena od svega jednoga entiteta nije drugo do taj entitet sam. Peanova je distinkcija, kaže Russell, “nedopustiva u našem pogledu

³⁹ Russell je neko vrijeme neodlučan oko toga bi li kao središnje značenje ‘razreda’ uzeo razred kao mnoštvo ili razred kao jedno. U odsječku iz kojega citiramo rabi izraz ‘razred’ (bez kvalifikacije) za razrede kao mnoštva i stoga za relaciju $\in$ ne kaže da je relacija oznake i razreda kojega je član nego samo da je s tom relacijom povezana (*allied*). No drugdje (u odsječku koji se čini ranije napisanim) za $\in$ kaže da je relacija “pojedinačnosti prema razredu kojemu pripada, tj. relacija Sokrata prema ljudskoj vrsti” [30, str. 19]. Slično tomu, u naknadno umetnutom dijelu završnoga odsječka poglavlja o razredima (usp. [1, str. 286]) piše da stavak “Sokrat pripada ljudskoj vrsti” “izražava relaciju pojedinačnosti prema svojem razredu i, kako mogućnost relacije zahtijeva, uzima razred kao jedno, ne kao mnoštvo” [30, str. 81]. No u svim tim trima slučajevima – iako u prvom ‘razred’ rabi za razred kao mnoštvo, a u drugim dvama za razred kao jedno – $\in$ je relacija između oznake i cjeline, ne oznake i mnoštva.

⁴⁰ Nije potpuno jasno zašto je Russell smatrao da se razred kao mnoštvo ne može u matematičkoj notaciji prikazati jednim slovom. Moguće je da je jedan od razloga jednostavno taj što bi predstavljanje razreda kao mnoštva jednim slovom prirodno sugeriralo da razred kao mnoštvo može biti podmet stavka, tj. da je oznaka.

⁴¹ Primjerice, Russell u rukopisu iz 1899.–1900. izričito tvrdi da “svaka je cjelina jedno, a svaka kolekcija oznaka *A, B, C* itd. sastavlja [*composes*] cjelinu” [33, str. 16], da “[b]ilo koja kolekcija, iako je kao takva mnoštvo [*many*], ipak sastavlja cjelinu” [33, str. 37] i da “bilo koja kolekcija oznaka oblikuje cjelinu” [33, str. 39].

na razrede” [30, str. 73]. Jednako tako, u Russellovu pogledu na razrede ne postoji prazan razred: nema mnoštva bez ijedne pojedinačnosti ni cjeline bez ijednoga dijela.⁴²

Uočimo također da razred kao mnoštvo – već samim time što je kolekcija bitno *množina*, što kolekcija nije pojedinačnost – ne može sam biti član razreda, tj. ne može biti *jedna* od oznaka u razredu. Općenito, nema ograničenja što može biti član kolekcije “osim što nijedno ne smije biti mnoštvo [*many*]” [30, str. 71]. Razred kao jedno – samim time što je i sam pojedinačnost – može, dakako, biti jedna od oznaka u razredu. Međutim, primjećuje Russell, razred kao jedno, sastoji li se od više od jedne oznake, ne može kao jednu od svojih oznaka sadržavati *sam sebe* – “složena cjelina ne može nikada biti jedna od svojih sastavnica” [30, str. 69]. S druge strane, čini mu se jasnim da razred kao jedno može biti član “toga istoga” razreda kao mnoštva – preciznije govoreći, razred kao jedno može biti i sam jedna od mnoštva oznaka koje tvore taj isti razred kao jedno. Primjerice, razred svih razreda, uzet kao cjelina, i sam je jedan od razreda (usp. [30, str. 69, 102, 134]).

No pogledajmo sada one razrede kao mnoštva koji, uzeti svaki kao jedno, *nisu* jedan od članova “njih samih”. Malo preciznije, kolekcije takve da agregat svih njihovih oznaka sam nije jedna od oznaka u toj kolekciji. Kolekcija svih filozofa, recimo, takva je kolekcija jer agregat svih filozofa nije jedan od filozofa. Pretpostavimo li da svakomu mnoštvu odgovara neka cjelina sastavljena upravo od oznaka toga mnoštva – kako Russell isprva pretpostavlja – za svaku takvu kolekciju postoji njoj odgovarajući agregat.⁴³ Uzmimo sada kolekciju svih *tih* agregata, tj. razred kao mnoštvo svih razreda kao jednih koji nisu jedna od mnoštva svojih oznaka. Opet, prema pretpostavci, i za taj razred kao mnoštvo postoji razred kao jedno svih njegovih oznaka – drugim riječima, postoji agregat svih agregata koji nisu jedna od svojih oznaka. Uzmimo,

⁴² Iako Russell smatra da je prazan razred “fikcija” [30, str. 80], slaže se da nas simbolička notacija i ustaljena matematička praksa obvezuju “uvesti neki takav pojam” [30, str. 73] pa kao supstitut za ulogu praznoga razreda predlaže razred svih stavačnih funkcija koje nisu istinite ni za jednu oznaku. Iz sličnih je razloga potrebno i razlikovanje jediničnoga razreda i njegova jedinoga člana i u tom je smislu “nesumnjivo da je Peanova distinkcija ispravna” [30, str. 131], no svoj prijedlog kako to učiniti Russell kasnije uklanja u korekturi (usp. [2, str. 68]).

⁴³ Usp. bilj. 41.

dakle, taj razred kao jedno i zapitajmo se je li on sam jedna od oznaka iz mnoštva svih njegovih oznaka. Iz svakoga odgovora slijedi suprotno. To je Russellov paradoks, kako ga je opisao u pismu Fregeu.

Paradoks pokazuje, smatra Russell, da je pogriješna pretpostavka da za svaku kolekciju vrijedi da postoji agregat svih pojedinačnosti te kolekcije:

Smatrali smo aksiomatskim da se razred kao jedno može naći gdje god postoji razred kao mnoštvo; no taj aksiom ne mora biti općenito prihvaćen i izgleda da je on izvor protuslovlja [30, str. 104].

Ponekad za razred kao mnoštvo ne postoji odgovarajući razred kao jedno. Svaka nam stavačna funkcija daje razred kao mnoštvo, ali nema jamca da nam je time ujedno dan razred kao jedno. Sada napokon postaje jasna Russellova dijagnoza paradoksa iz pisma Fregeu: u nekim slučajevima definirljiva kolekcija – razred kao mnoštvo svih oznaka koje zadovoljavaju stavačnu funkciju – ne tvori sveukupnost, tj. nema odgovarajuće cjeline, razreda kao jednog svih oznaka dane kolekcije.⁴⁴

Paradoks vodi Russella k prihvaćanju kao središnjega i matematički jedinoga relevantnog značenja ‘razreda’ ne više razred kao jedno, već razred kao mnoštvo, “jedini predmet koji je uvijek definiran stavačnom funkcijom” [30, str. 491]. Prema revidiranomu pogledu na razrede, sada su stavci poput “Sokrat je jedan od ljudi” oni koji izražavaju pripadnost razredu u pravom smislu, ne više stavci poput “Sokrat pripada ljudskoj vrsti”. Matematički relevantna relacija članstva u razredu, relacija koju označavamo s ‘ $\in$ ’, za Russella nije više relacija između dviju oznaka, već relacija predmeta spram razreda kao mnoštva predmetā. Očito, znači to da sada dopušta da se u logičkoj notaciji razred kao mnoštvo prikazuje jednim simbolom.

Revidirajući u prvom dodatku *Principles* svoj nauk o razredima, Russell razmatra i odbacuje različite alternativne koncepcije razreda, uključujući onu prema kojoj se razred može poistovjetiti s cjelinom sastavljenom od svih oznaka toga razreda. Razlog odbacivanja upravo je to što,

⁴⁴ U rukopisu iz svibnja 1901., u kojem prvi put razmatra otkriveni paradoks, izrazi ‘razred kao mnoštvo’ i ‘razred kao jedno’ još se ne pojavljuju i Russell ‘razred’, bez dodatnih kvalifikacija, rabi samo za agregate. U rukopisu stoga jednostavno kaže da paradoks pokazuje “da ne tvori svaka definirljiva kolekcija razred” [33, str. 195].

po cijenu protuslovlja, odgovarajuća cjelina “ne može uvijek postojati kao oznaka” [30, str. 515], tj. što ne postoji za svako mnoštvo oznaka razred kao jedno svih oznaka u pitanju. S druge strane, problematična je i koncepcija razreda prema kojoj bi razred bio puka množina oznaka. Za razliku od razreda kao jednoga, postojanje razreda kao mnoštva uvijek je, doduše, logički neosporno, ali nevolja je u tome što razred kao mnoštvo sam nije zaseban entitet. Kako nije pojedinačnost, vidjeli smo gore, razred kao mnoštvo ne može biti *jedna* od oznaka u razredu. Općenito, kako razred kao mnoštvo nije entitet, razred kao mnoštvo ne može biti logički podmet stavka i zato uopće ne postoji nijedan stavak da je neki razred *kao mnoštvo* član nekoga razreda. Stoga, ako su razredi kao mnoštva jedino što imamo na raspolaganju, slijedi neprihvatljiva posljedica da više ne možemo govoriti o razredima razredā – razred razredā, naime, u tom slučaju ne bi bio mnoštvo oznaka, već “mnoštvo mnoštava [*many many's*]” [30, str. 516].

Odgovor na koji je Russell prisiljen svojevrсно je “postvarenje” razredā kao mnoštava:

Iako je razred mnoštvo, a ne jedno, ipak postoji istovjetnost [*identity*] i različitost [*diversity*] među razredima i stoga razredi mogu biti brojani kao da je svaki pravo jedinstvo [*genuine unity*]; i u tom smislu možemo govoriti o *jednom* razredu i o razredima koji su članovi razreda razredā [30, str. 516–517].

Dvije su važne posljedice toga postvarenja. Prva je da Russellovi razredi postaju vrlo tajnovita vrsta predmetā – otprilike ono što danas nazivamo ‘skupovi’. Dopuštanjem da je razred kao mnoštvo i sam u nekom smislu *jedno* – čime razred kao mnoštvo može biti *jedan* od članova nekoga daljnjeg razreda – omogućeni su razredi razredā (kao mnoštava) kao legitimna vrsta predmetā. No time se nameće peanovsko razlikovanje između jediničnoga razreda i njegova jedinoga člana, barem kada je član o kojem je riječ razred s više od jednoga člana. Argument je jednostavan. Neka je *S* razred s više od jednoga člana, a *S'* razred kojega je jedini član *S*. Tada, ako bi *S'* bio istovjetan *S*-u, *S'* bi, suprotno pretpostavci, imao više od jednoga člana. Russell stoga sada jednostavno prihvaća da je u svim slučajevima jedinični razred različit predmet od njegova jedinoga člana. No ako smo već prihvatili tu distinkciju, “možemo dakako također dopustiti [razred] u slučaju prazne stavačne funkcije” [30, str. 517], tj. dopustiti postojanje praznoga razreda. Problem je sada kako

točno shvatiti što su zapravo Russellovi razredi. U slučaju kada razred ima više od jednoga člana, razred možemo poistovjetiti s mnoštvom njegovih članova, kaže Russell (usp. [30, str. 517]), no prisiljeni smo ujedno dopustiti postojanje razreda koji su različiti od svojega jedinog člana, kao i razred bez ijednoga člana uopće.⁴⁵ Kolekcije su, kao i agregati, djelovale metafizički nevinima, ali to nije slučaj s praznim mnoštvom i mnoštvom različitim od jedinoga predmeta u tome mnoštvu.

Druga je posljedica postvarenja razreda hijerarhija logičkih tipova. Razred kao mnoštvo, smatra sada Russell, *može* i sam biti logički podmet, ali, kako razredi kao mnoštva nisu oznake, “u stavicima različite vrste od onih u kojima su njegove oznake podmeti” [30, str. 105]. Razredi kao mnoštva predmeti su logički drukčijega tipa nego oznake. Priroci i relacije koje se odnose na razrede “različiti su (uz neke iznimke) od onih koji se mogu pojaviti u stavicima koji imaju pojedinačne [*single*] oznake kao podmete” [30, str. 516]. Drugim riječima, stavačne funkcije kojima argument može biti bilo koja *oznaka* obično ne će kao svoje argumente imati razrede oznaka, razrede razreda itd. To je razlikovanje logičkih tipova u osnovi Russellove teorije tipova: “Temeljni nauk na kojem sve počiva nauk je da podmet stavka može biti množina [*plural*]” [30, str. 517].

Zaključak

“Filozofija je”, piše Bertrand Russell u *The Principles of Mathematics*, “uglavnom pitanje uvida i percepcije” [30, str. 129]. Svako matematičko istraživanje polazi od određenih nedokažljivih stavaka i nedefinirljivih entiteta, a uloga je filozofije u utemeljenju matematike, između ostaloga, raspoznavanje i razlučivanje tih polaznih matematičkih entiteta. Jedan od ciljeva knjige, kaže Russell u predgovoru, “objašnjenje je temeljnih pojmova koje matematika prihvaća kao nedefinirljive” [30, str. xv]. Taj

⁴⁵ Prihvativši u prvom dodatku *Principles* prazan razred i distinkciju između razreda i jedinoga člana toga razreda, Russell razrede (kao mnoštva), prevevši odgovarajući Fregeov termin (*Werthverläufe*), počinje nazivati ‘područja’ (*ranges*). Taj izraz dosljedno rabi u drugom dodatku, ograničivši izraz ‘razred’ na područja kojih su svi članovi međusobno istoga tipa.

je filozofijski pothvat “nastojanje vidjeti jasno, i učiniti da i drugi vide jasno, dotične entitete, kako bi um mogao imati onu vrstu upoznatosti s njima kakvu ima s crvenošću ili okusom ananasa” [30, str. xv]. Međutim, piše da nije uspio “percipirati nijedan pojam koji zadovoljava uvjete tražene za pojam *razreda*” [30, str. xv–xvi]. Razlog je, dakako, paradoks koji je otkrio tijekom pisanja knjige.

Svojim prijedlogom rješenja paradoksa Russell nije bio zadovoljan – između ostaloga, upravo zbog problematične metafizike razredā u pozadini toga prijedloga.⁴⁶ Množinama predmetā koje, iako nisu zasebni entiteti, jesu logički podmeti stavaka, mnoštvima različitima od jedinoga predmeta u mnoštvu i kolekcijama bez ijednoga predmeta u kolekciji nije jednostavno dati ontologijsku domovnicu. Također, reificiranjem razredā kao mnoštava, tj. činjenjem njih logičkim podmetima, razlika između razreda kao mnoštva i razreda kao jednoga ne djeluje više nimalo očitom. Razredi kao mnoštva i sami se čine punokrvnim entitetima prema svim Russellovim kriterijima: predmeti su mišljenja, sastavnice su stavaka, mogu se brojati (usp. [30, str. 43]). No tada teorijom tipova iznesenom u drugom dodatku *Principles* Russell doista napušta nauk o neograničenoj varijabli: unatoč tomu što ‘oznakama’ i ‘pojedinačnostima’ naziva samo predmete na najnižoj razini tipske hijerarhije, ti predmeti *ne* uključuju sve entitete.⁴⁷ Vidjeli smo, međutim, da je nauk o neograničenoj varijabli sastavni dio Russellova shvaćanja logike kao potpuno univerzalne znanosti. U godinama koje su uslijedile nakon *Principles* problemi tamo iznesene teorije tipova nagnali su Russella rješenje paradoksa potražiti u pokušajima potpune eliminacije razredā i drugih problematičnih entiteta iz pozadinske ontologije, kulminirajući u izgradnji razgranate teorije tipova u *Principia Mathematica*.

⁴⁶ Uz nevolje s razredima, Russellov je prijedlog i stavke učinio problematičnom vrstom predmetā. Kao prvo, kako za svaki predmet, uključujući razrede kao mnoštva, postoji stavak da je taj predmet istovjetan samomu sebi – i kako pritom razreda ima strogo više od pojedinačnosti – čini se da stavci ne mogu biti pojedinačnosti iako su, kako smo vidjeli, za Russella stavci cjeline, složeni entiteti. Kao drugo, Russell pokazuje da je, pod plauzibilnom pretpostavkom da stavci tvore jedinstveni tip, izvedljivo protuslovlje. Formalna analiza Russellova paradoksa stavaka može se naći u [12] i [3].

⁴⁷ Russell stoga, opisujući teoriju tipova, za oznaku ili pojedinačnost samo kaže da je “bilo koji predmet koji nije područje” [30, str. 523], tj. bilo što što nije razred kao mnoštvo (usp. bilj. 45).

Literatura

- [1] Blackwell, K. (1984), "Part I of *The Principles of Mathematics*", *Russell*, 4 (2), 271–288.
- [2] Byrd, M. (1987), "Part II of *The Principles of Mathematics*", *Russell*, 7 (1), 60–70.
- [3] Cantini, A. (2004), "On a Russellian Paradox about Propositions and Truth". U: *One Hundred Years of Russell's Paradox*, ur. G. Link, Berlin – New York: Walter de Gruyter, 259– 284.
- [4] Cantini, A. (2009), "Paradoxes, Self-Reference and Truth in the 20th Century". U: *Handbook of the History of Logic, Volume 5. Logic from Russell to Church*, ur. D. M. Gabbay i J. Woods, Amsterdam: Elsevier, 875–1013.
- [5] Carnap, R. (1929), *Abriss der Logistik, mit besondere Berücksichtigung der Relationstheorie und ihre Anwendungen*, Wien: Springer.
- [6] Cartwright, R. L. (2005), "Remarks on Propositional Functions", *Mind*, 114 (456), 915–927.
- [7] Copi, I. (1971), *The Theory of Logical Types*, London: Routledge & Kegan Paul.
- [8] Devidé, V. (1991), *Matematička čitanka*, Zagreb: Školska knjiga.
- [9] Frege, G. (1884), *Die Grundlagen der Arithmetik: eine logisch mathematische Untersuchung über den Begriff der Zahl*, Breslau: W. Koebner.
- [10] Frege, G. (1893), *Grundgesetze der Arithmetik*, Jena: Verlag Hermann Pohl.
- [11] Frege, G. (1980), *Philosophical and Mathematical Correspondence*, ur. G. Gabriel, Oxford: Blackwell.
- [12] Fuhrmann, A. (2001), "Russell's Early Type Theory and the Paradox of Propositions", *Principia*, 5 (1–2), 19–41.
- [13] Garciadiego, A. R. (1992), *Bertrand Russell and the Origins of the Set-Theoretic 'Paradoxes'*, Basel – Boston – Berlin: Birkhäuser Verlag.
- [14] Geach, P. (1972), *Logic Matters*, Oxford: Blackwell.
- [15] Gödel, K. (1990) [1944], "Russell's Mathematical Logic". U: Gödel, K. (1990), *Collected Works, Vol. II: Publications 1938–1974*, ur. S. Feferman et al., New York – Oxford: Oxford University Press, 119–141.
- [16] Godwyn, M. i Irvine, A. D. (2003), "Bertrand Russell's Logicism". U: *The Cambridge Companion to Bertrand Russell*, ur. N. Griffin, New York: Cambridge University Press, 171–201.
- [17] Grattan-Guinness, I. (2003), "Mathematics in and behind Russell's Logicism, and Its Reception". U: *The Cambridge Companion to Bertrand Russell*, ur. N. Griffin, New York: Cambridge University Press, 51–83.

- [18] Grattan-Guinness, I. (2003), "Russell's Way In: From Certainty to Paradoxes, 1895–1903". U: Grattan-Guinness, I. (2003), *The Search for Mathematical Roots, 1870–1940*, Princeton – Oxford: Princeton University Press, 268–332.
- [19] Griffin, N. (2004), "The Prehistory of Russell's Paradox". U: *One Hundred Years of Russell's Paradox*, ur. G. Link, Berlin – NewYork: Walter de Gruyter, 349–371.
- [20] Heijenoort, J. van (1967) (ur.) *From Frege to Gödel. A Source Book in Mathematical Logic, 1879–1931*, Cambridge, Massachusetts: Harvard University Press.
- [21] Heijenoort, J. van (1967), "Logic as Calculus and Logic as Language", *Synthese*, 17(3), 324–330.
- [22] Klement, K. C. (2010), "Russell, His Paradoxes, and Cantor's Theorem", *Philosophy Compass*, 5 (1), 16–41.
- [23] Klement, K. C. (2010), "The Functions of Russell's No Class Theory", *The Review of Symbolic Logic*, 3 (4), 633–664.
- [24] Landini, G. (1998), *Russell's Hidden Substitutional Theory*, New York – Oxford: Oxford University Press.
- [25] Landini, G. (2004), "Russell's 'Separation' of the Logical and Semantic Paradoxes", *Revue Internationale de Philosophie*, 58 (229), 257–294.
- [26] Peano, I. (1889), *Arithmetices principia, nova methodo exposita*, Augustae Taurinorum: Fratres Bocca.
- [27] Peano, G. (1967) [engleski prijevod [26] (dio)], "The Principles of Arithmetic, Presented by a New Method". U: *From Frege to Gödel. A Source Book in Mathematical Logic, 1879–1931*, ur. J. van Heijenoort, Cambridge, Massachusetts: Harvard University Press, 85–97.
- [28] Quine, W. V. (1955), "On Frege's Way Out", *Mind*, New Series, 64 (254), 145–159.
- [29] Russell, B. (1908), "Mathematical Logic as Based on the Theory of Types", *American Journal of Mathematics*, 30 (3), 222–262.
- [30] Russell, B. (1938) [1.izdanje1903.], *The Principles of Mathematics* (2. izdanje), New York: W. W. Norton & Company.
- [31] Russell, B. (1959), *My Philosophical Development*, New York: Simon and Schuster.
- [32] Russell, B. (1973) [rukopis iz 1906.], "On the Substitutional Theory of Classes and Relations". U: Russell, B. (1973), *Essays in Analysis*, ur. D. Lackey, London: Allen & Unwin, 165–189.
- [33] Russell, B. (1993), *The Collected Papers of Bertrand Russell. Volume 3: Toward the "Principles of Mathematics" 1900–02*, ur. G. H. Moore, London – New York: Routledge.

- [34] Schröder, E. (1890), *Vorlesungen über die Algebra der Logik*, sv. 1, Leipzig: Teubner.
- [35] Sobociński, B. (1949), “L’analyse de l’antinomie russellienne par Leśniewski”, *Methodos*, 1, 94–107, 220–228, 308–316.
- [36] Whitehead, A. N. (1898), *A Treatise on Universal Algebra*, Cambridge: Cambridge University Press.
- [37] Whitehead, A. N. i Russell, B. (1925–1927) [1. izdanje 1910–1913], *Principia Mathematica*, sv. 1–3, 2. izdanje, Cambridge: Cambridge University Press.
- [38] Zermelo, E. (1908), “Untersuchungen über die Grundlagen der Mengenlehre. I”, *Mathematische Annalen*, 65, 261–281.

Može li se naučiti kritički misliti?

Nenad Smokrović

Uvod

Glavni je cilj ovog teksta, pored razmatranja pojma 'kritičko mišljenje' u njegovoj suvremenoj upotrebi, ispitati mogućnosti stjecanja znanja ili naučivosti sposobnosti kritičkog mišljenja. Nije sporno da se kritičko mišljenje u nekim svojim segmentima može poboljšati (kao što je to slučaj sa svim vještinama). Pitanje je može li se ono naučiti (kao što je to slučaj sa stranim jezikom)?

Danas je govor o argumentiranoj raspravi, školskim kurikulumima, kao i rješavanju različitih vrsta dijaloških sporova postao nezamisliv bez spominjanja kritičkog mišljenja. Problematičnim, međutim, ostaje da je unatoč tome razumijevanje pojma kritičkog mišljenja danas u svom optimizmu nedovoljno elaborirano kako s obzirom na svoje temeljno određenje, tako i s obzirom na svoju ulogu i primjenu. Ono izražava nepokolebljivo i, čini se, nedovoljno osviješteno prosvjetiteljsko povjerenje u razum, u njegovu kritičku snagu koja kada je na djelu, doprinosi boljitku u svim segmentima ljudskog života. U ovom svjetlu, oslanjanje na razum kao moć rasuđivanja s jedne strane vodi do istinitijih vjerovanja i ispravnijih odluka, a s druge omogućuje odmicanje od osobnih interesa, svojevrstan *stepping back* te uvažavanje tuđih stavova i uviđanje vlastitih pogrešnih koraka. Mnogi će također optimistički povezati moralnost i optimalnost mišljenja. Tako, primjerice, Wittgenstein (*Filozofska istraživanja*): "Biti bolja osoba i biti pametniji jedno je te isto".

Koju god korist od kritičkog mišljenja (u daljnjem tekstu: KM) očekivali, ključna je ideja da se ono može naučiti. Od koga se, pak, očekuje sposobnost za KM? Suvremena i popularna upotreba termina, pogotovo u domeni pedagogije, pretpostavlja da je to način mišljenja u načelu dostupan svim racionalnim ljudskim bićima (gdje 'racionalno' uzima-

mo u najširem i najneodređenijem smislu), što uključuje (logički) neobrazovane ljude i školsku djecu, ali isto tako i sudionike znanstvenog diskursa, sudsku raspravu itd. Prešutno se uzima da je KM nešto što je u načelu poznato (stručnjacima), samo što *načela KM treba izraziti u terminima korisnih alata*, čemu onda ljude treba podučiti. Kao što se podučava ispravno skijanje, dobra vožnja, uspješno poslovno pregovaranje i slično. U pogledu alata, u knjizi *Arguing well* John Shand (2000) kaže: “Ova knjiga želi dati alate, logičke i psihološke, koji ljude sprečavaju da zanemare razum ili ga loše i pogrešno primijene, ..., a primjenjivati ih dobro ono jest što trebaju činiti” (str. 3). Naučimo li ljude ispravno primjenjivati alate, razum će se ispravno koristiti i mnogi problemi bit će riješeni.

Radi se o tome da ovakva pretpostavljena samorazumljivost KM prikriva potrebu za njegovim određenjem. No, u tome je upravo problem. Naivno shvaćanje KM u svojem izbjegavanju dubljeg promišljanja ovog pitanja prijeti postati radikalno nekritičkim. To je vidljivo u dijelu teorijskog diskursa koji se ovime bavi, poglavito u njegovu edukacijsko-pedagoškom aspektu. Ilustrirat ćemo to na primjeru eksplicitnog odbijanja jasnog određenja. Tako Davis i Barnett naglašavaju (2015): “Definicije kritičkog mišljenja nisu centralne za područja kao što su kritička pedagogija ili kritički feminizam. Učenje o takvim definicijama ne pomaže nam razviti kritički stav o društvu u kojem živimo.” (str. 3).

Određena doza samorazumljivosti u shvaćanju KM prisutna je i u elaboriranom filozofskom (epistemološkom) diskursu. Ona je vidljiva s obzirom na pitanje o kakvim se točno sposobnostima kada govorimo o KM radi, ali isto tako i s obzirom na pitanje o mogućnosti stjecanja ili naučivosti tih vještina. Ovdje nam se od posebne važnosti čini recentni članak autora Kotzee, Carter i Siegela (u daljnjem tekstu: K, C, S) (2019) pa ćemo se na njemu kratko zadržati. Autori članka žele pokazati da razvoj kritičkog mišljenja, s jedne strane i razvoj intelektualnih epitemičkih vrlina, s druge strane predstavljaju dva sasvim odvojena i nezavisna pedagoška cilja. Štoviše, ovo implicira da intelektualne vrline *nisu* dio kritičkog mišljenja, a s obzirom na pedagogiju “primarni cilj obrazovanja promocija je kritičkog mišljenja” (ibid, ...), za što pedagogija ima dostatnih resursa, što nije slučaj s epitemičkim vrlinama.” U ovom članku želimo pokazati upravo suprotno, naime da je određenje kritičkog mišljenje nepotpuno ukoliko ne uključuje epitemičke vrline

shvaćene na određeni način. Tvrdit ćemo da KM nije jednostavna, monolitna sposobnost. KM *složeni je teorijski konstrukt*¹ koji se sastoji od niza elemenata od kojih se neki lakše a neki teže mogu naučiti. Na ovu tvrdnju i daljnju raspravu s tezama Kotzee, Carter i Siegela vratit ćemo se u poglavlju o strukturi KM.

Izdvojimo na kraju ovog uvoda dvije tvrdnje koje su pretpostavljene u suvremenoj literaturi o KM i na koje će se fokusirati naša kritika. Prvo, to je ideja da je KM odjelita sposobnost zaključivanja koja se jasno može razlikovati od drugih i koja ne uključuje druga epistemička svojstva, prije svega epistemičke vrline i drugo, da je ovo sposobnost koju može steći svaki normalni, prosječni subjekt, odnosno da je to sposobnost koja se edukacijom može (i treba) u takvom subjektu razviti. Implicitno je također pretpostavljeno da je subjekt osposobljen za KM sposoban kritički misliti u svakoj domeni realnosti. Kritički mislitelj bit će jednako kritičan u domeni koja mu je dobro poznata, ali i u svakoj drugoj, primjerice, u prosuđivanju znanstvenog od neznanstvenog, u prosuđivanju religijskog mišljenja, u domeni problema koji se tiču socijalne jednakosti, pravde, u domeni politike, morala itd.

Nastanak pojma KM i kako se ono danas razumije?

Sam pojam sve je prije nego jednoznačan. Započnimo stoga s analizom suvremenog razumijevanja pojma KM i vrlo kratkim osvrtom na značenje ovog pojma kroz posljednja dva stoljeća. Termin KM u svom suvremenom određenju postao je općekorištenim relativno odnedavno (u posljednjoj trećini prošlog stoljeća. Prvu polovicu stoljeća, pak, karakterizira određenje KM kao 'kritike postojećeg'²). Isto tako, pozivanje na starije, filozofsko (kantovsko) određenje kritičkog pristupa shvaćenog kao proučavanje epistemičkih uvjeta svakog mogućeg mišljenja neće nam pomoću u razumijevanju suvremenog značenja ovog

¹ Pod konstruktom mislimo na standardno određenje gdje se pod tim misli na teorijski neopažljivi entitet čije se značenje određuje operacionalnom definicijom. Vidi Kerlinger, F. N., 1986.

² Ovo je određenje na djelu u radovima filozofa *Frankfurtske škole* te kasnije u radovima filozofa okupljenih oko časopisa *Praxis*.

pojma. Činjenica je da je u svojoj suvremenoj upotrebi značenje ovog termina s obzirom na ranije upotrebe doživjelo temeljit zaokret od prijašnjeg epistemičkog konteksta ili pak konteksta 'prevratničke kritike postojećeg' k užem, kognitivno-psihološkom određenju *ispravnog* mišljenja i mogućnosti *procjene ispravnosti personalnih ili interpersonalnih kognitivnih epizoda*.

Gledano iz povijesne perspektive, vrlo blizu suvremenom većinskom gledanju na određenje KM došao je John Dewey (on, doduše, ne spominje *kritičko mišljenje* nego *reflective thinking*) formuliravši ga kao: "Aktivno, ustrajno i pažljivo razmatranje vjerovanja ili pretpostavljenog oblika znanja u svjetlu razloga koji ga podržavaju i daljnjih zaključaka kojima ono teži." (1909). Ono što je za puno razumijevanje ove definicije potrebno odrediti jest relacija koja određuje kako razlozi (u klasičnom argumentu, premise) podržavaju (support) vjerovanje (u klasičnom argumentu, konkluziju). Standardni stav još od Kanta naovamo, ali posebno od formiranja formalne logike (Frege) do polovice 20. stoljeća bio je da premise podržavaju argument tek relacijom klasične, formalne valjanosti. Upravo ovakvo određenje formalne logike kao standarda za mjerenje sposobnosti za KM motivirao je suvremene teoretičare KM na potragu za adekvatnijim standardom. Ovi teoretičari, u težnji da odrede standarde i kanone mišljenja *stvarnih* ljudi u *aktualnom* svijetu, ne prihvaćaju ovaj stav, već relaciju podržavanja temelje na pravilima *neformalne logike*.

Pioniri u povezivanju pravila neformalne logike i KM jesu S. Toulmin, zasnivajući ono što on zove *working logic* (Toulmin, 1958) i C. Perelman, s prijedlogom nove retorike (Perelman i Olbrechts-Tyteca, 1958). Suvremeni autori, prije svega R. Johnson (2000) i D. Hitchcock (2007) preuzimaju ideju bliske povezanosti neformalne logike i KM. Naglasimo i to da je ideja same neformalne logike nedovoljno jasno artikulirana.

Zanimljivo je vidjeti kako jedan od najznačajnijih aktera suvremene teorije neformalne logike i KM Ralph Johnson (2012) vidi razloge za nastanak, najprije neformalne logike a onda i s njom povezanog KM. Neformalna logika, svjedoči Johnson, nastala je iz nezadovoljstva nastavnika i studenata mogućnošću primjene nastave formalne logike na ovladavanje aktualnim inferencijalnim problemima. Sedamdesetih i

osamdesetih godina prošlog stoljeća nastaje niz različitih artikulacija ideje neformalne logike, a na to se u isto vrijeme nadovezuje “pokret kritičkog mišljenja”. Unutar ovog pokreta javljaju se različite definicije KM: “Problem s ovim shvaćanjem nije naprosto to da postoje mnoge definicije kritičkog mišljenja, već da su neke inkompatibilne s drugima (ibid, 6)”. Zanimljivo je također vidjeti kako Johnson vidi odnos između neformalne logike i KM On kaže: “Neformalna logika označava tip logike dok kritičko mišljenje označava vrstu intelektualne prakse, kao i edukacijski ideal” (ibid, 8), te nekoliko pasusa dalje nastavlja: “Ako misliti kritički traži sposobnost procjene argumenata, za što mislim da bise mnogi složili, tada neformalna logika treba dati značajan doprinos ostvarenju cilja kritičkog mišljenja.” (ibid, 9). Ipak, sumnja u uspješnost ovog nastojanja izražava se na kraju članka: “Najvažnije pitanje nije to uče li studenti na našim tečajevima KM kritički misliti, kako je to određeno specifičnim testovima kritičkog mišljenja. Značajno pitanje jest traje li iskustvo s učenjem kritičkog mišljenja nakon samog tečaja i igra li ulogu u njihovim daljnjim kognitivnim aktivnostima.” (ibid, 13).

Struktura kritičkog mišljenja

Dakle, koncept KM sadrži ideju o tome da postoje pravila (neformalne logike) koja određuju strukturu “ispravne upotrebe razuma” te da se ta pravila regularnim edukacijskim procesom (i odgovarajućim kurikulumom) mogu implementirati u kognitivno ponašanje studenta. Sama priroda KM pa stoga i priroda toga što se treba podučavati nije dovoljno jasna. Je li to skup (neformalnih logičkih) pravila i njihovapraksa, kao što je tou slučaju učenja stranog jezika skup gramatičkih pravila i njihova praksa, koju većina ljudi može naučiti. Ili je to latentna sposobnost koju svatko u nekoj mjeri posjeduje a izrazitije se aktivira u određenim okolnostima, *posebno u argumentativnom dijalogu* gdje je svatko naročito oprezan ili kritičan u uočavanju oponentovih pogrešaka (Mercier i Sperber (2011) to zovu *vigilizam*). Je li to generalna sposobnost (kao što je to sposobnost za matematičko mišljenje) koju neki ljudi iskazuju u većoj a neki u manjoj mjeri, ali nije moguće značajnije ju poboljšati? Čini se da su suvremeni teoretičari kritičkog mišljenja skloniji ovoj prvoj mogućnosti. Mi ćemo kao plauzibilnu uzeti osnovnu ideju druge solucije i pokušati je detaljnije elaborirati.

Da bismo započeli s razradom ideje KM, vratimo se na sintagmu 'kognitivna epizoda mišljenja' koju smo upotrijebili u uvodu ovog rada. Ona pretpostavlja da se mnogoznačni i neprecizni pojam mišljenja može izraziti u nešto određenijim terminima *zaključivanja* i *argumentiranja*. Sasvim sumarno rečeno, pod *zaključivanjem* razumije se proces izvođenja ili derivacije kojim mislitelj od skupa iskaza ili propozicija dolazi do konkluzije. Proces zaključivanja može imati oblik argumenta, gdje se pod argumentom razumije formulacija iskaza (konkluzije) podržanog razlozima. U osnovi, to je tradicionalno shvaćanje. Tako već Kvintilijan (*Institutio oratoria*) kaže da je argument 'razlog koji putem stvari koje su izvjesne čini prihvatljivim ono što je dvojbeno' (*ratio per ea, quae certa sunt, fidei dubiis adheren*). Suvremeno standardno određenje zaključivanja u kognitivnoj znanosti kaže slično: "Zaključivanje je promjena (tranzicija) u mišljenju u kojoj neka vjerovanja (ili misli) nude osnovicu (ili razlog) za dolaženje do novog vjerovanja" (Rips i Adler, 2008, *Introduction*). Kognitivna epizoda utoliko pojedinačna je instanca mišljenja shvaćenog kao zaključivanje.

Argument je, pak, najčešća forma u kojoj se proces zaključivanje provodi. Da bismo razlikovali argument kao logičku formu nezavisnu od procesa zaključivanja od argumenta u kojem se provodi *proces zaključivanja*, nazovimo ovo posljednje argumentacija. U sažetom obliku izrazimo argumentaciju kao tromjesnu strukturu (Γ, R, α) . R označava proces zaključivanja kojim zaključitelj želi demonstrirati/dokazati da je njegov proces dolaženja do zaključka ispravan, odnosno da su premise u skupu Γ dobar razlog ili da *podržavanju* zaključak α . Sasvim je izvjesno da se KM u najopćenitijem obliku može opisati kao zaključivanje koje ima oblik argumentacije. Određenije, kao prvu aproksimaciju KM kažimo da je ono *zaključivanje podržano razlozima, izraženo u formi argumentacije*.

Prva polazna aproksimacija traži daljnje i detaljnije određenje prirode KM. Prije svega, koji su to elementi ili aspekti kojim se KM kao složeni konstrukt može izraziti te koji se dadu naučiti? Da bismo odgovorili na ovo pitanje, predložiti ćemo po našem sudu plauzibilnu strukturu KM KM, kao i druge više kognitivne sposobnosti, sadrži svoju fenomenološku, manifestnu razinu, ali također i razinu kognitivnih sposobnosti koje se dadu opisati kao mehanizmi, ponekad kao dispozicije³, koje su

³ U ovom tekstu nećemo ulaziti u razlikovanje između mehanizama i dispozicija

inherentne manifestnoj razini te koje stoje u njihovoj osnovi i omogućuju ih. Predlažemo podjelu na dvije manifestne razine na kojima se aktivnosti odvijaju u domeni svjesne kontrole subjekta, od kojih svaka ima sebi odgovarajuću razinu temeljnih sposobnosti (i/ili dispozicija) koje ih utemeljuju i omogućuju. Razlikovat ćemo, dakle, ukupno četiri aspekta KM. To su:

- a) sposobnost uočavanja ilatičkih odnosa (odnos razloga i tvrdnje),
- b) kognitivne sposobnosti koje omogućuju ilatičke sposobnosti,
- c) epistemičke vrline,
- d) dispozicije na kojima se zasnivaju epistemičke vrline.

Svaki od ovih aspekata detaljnije ćemo prikazati.

- Prva manifestna sposobnost jest *sposobnost uočavanje podržavajućih odnosa između razloga i tvrdnje*, odnosno u klasičnom argumentu, premisa i konkluzije. Uočavanje podržavajućih odnosa između razloga i tvrdnje uključuje kako sposobnost tvorbe vlastitih argumenata koje sam subjekt prosuđuje kao 'ispravne' (takve da su premise razlozi za konkluziju), tako i sposobnost prosuđivanja 'ispravnosti' tuđih argumenta. Nazovimo ovo *sposobnošću prosuđivanja ilatičkih odnosa* (koristeći termin koji je uveo Ralph Johnson). Pod ilatičkim odnosima razumijemo one koji utemeljuju tranziciju od skupa vjerovanja koji nude razloge za tvrdnju na samu tvrdnju. Malo formalnije, ako propozicije od γ_1 do γ_n predstavljaju dobre razloge za α (konkluziju), tada je opravdano zaključiti α . Ovaj odnos zasnovanosti tvrdnje na razlozima, rekli smo, može biti vrlo različit. Razlika je, naravno, radi li se o deduktivnom, induktivnom ili abduktivnom zaključivanju. Razlika se javlja i unutar svakog od ovih tipova zaključivanja. Svaki od ovih oblika zaključivanja može biti 'ispravan' ili 'neispravan'. 'Ispravnost' je, naravno, relativna s obzirom na kriterij prosuđivanja ispravnosti. Vidjeli smo već da se kod deduktivnog zaključivanja 'ispravnim' ili podržavajućim odnosom premisa i zaključka može smatrati formalno logička valjanost, ali isto tako i neka od svojstava utemeljenih u neformalnoj logici (primjerice, uvjerljivost

ni u razgranatu i složenu raspravu koje se o ovom problemu vodi. Povodit ćemo se za zatečenim terminima u literaturi i govoriti o mehanizmima tamo gdje se taj termin spominje, odnosno o dispozicijama kada se one spominju.

argumenta, prihvatljivost⁴ i slično). Ne ulazeći u raspravu o tome na čemu sve mogu biti zasnovani razlozi za tvrdnju, sposobnost o kojoj je riječ jest sposobnost prepoznavanja ilatičkih relacija.

- Dio ilatičkih sposobnosti odvija se na refleksivnoj, svjesnoj razini dok se u svom značajnom dijelu odvija na razini implicitnih, ponekad automatskih reakcija na zadatak s kojim se subjekt suočava. Radilo se o refleksivnoj ili spontanoj razini reakcije, ilatičko prepoznavanje značajno je određeno *kognitivnim sposobnostima*⁵, aktivnostima čije se procesiranje odvija na subpersonalnoj, manje-više nesvjesnoj razina. One se u (Stanovich, West i Toplak, 2016, str. 208) određuju ovako: “Kognitivne sposobnosti tipovi su kognitivnih procesa koje proučavaju istraživači procesiranja informacijama u potrazi za utemeljujućom kognitivnom osnovicom u temelju performansi na IQ testovima. *Brzina percipiranja, točnost razlikovanja, kapacitet radne memorije te efikasnost prizivanja informacija uskladištenih u dugoročnoj memoriji* primjeri su kognitivnih sposobnosti koje stoje u osnovi tradicionalne psihometrijske inteligencije”. Premda se u citiranom odjeljku spominje opća inteligencija (kao značajan segment racionalnosti), jasno je da su kognitivne sposobnosti pozitivno korelirane s provođenjem epizoda KM, posebno na razini ilatičkih sposobnosti. Što su kognitivne sposobnosti bolje, to je veća sposobnost prepoznavanja ispravnih ilatičkih odnosa. S obzirom na problem naučivosti KM, moramo pretpostaviti da je učenje u domeni KM povezano s poboljšanjem kognitivnih sposobnosti, što se opet mora moći postići edukacijom. Pokazat ćemo da to nije slučaj. Za sada dostaje analogija s inteligencijom. Na stupnju opće inteligencije zasnivaju se intelektualne razlike, međutim nije tako da se nizak stupanj IQ edukacijom može učiniti višim.

c) Drugi centralni oblik (pored ilatičkog) provođenja epizoda KM temelji se na epistemičkim vrlinama. Osnovno značenje pojma preuzimamo iz epistemološkog pravca nazvanog epistemologija vrlina (Zagzebski 1996, Battaly 2015, Baehr 2011) koji ova disciplina dijeli s etikom vrli-

⁴ U suvremenoj neformalnoj logici najčešće se za ispravnu ili podržavajuću relaciju koristi engleski termin “cogency”.

⁵ Termin preuzimamo od K. Stanovicha (2016). Vidi poglavlja 2 i 11.

na. U ovom tekstu značenje pojma *epistemička vrlina* želimo, što je više moguće, osloboditi svake etičke konotacije i maksimalno približiti ga kognitivnom kontekstu. Za razliku od sposobnosti prepoznavanja ilatičkih veza, za koje možemo reći da se manifestiraju na *objektnoj razini*, djelatnosti vezane za epistemičke vrline odvijaju se na *metarazini*. Epistemičke vrline odredit ćemo kao metainferencijalni aspekt kognitivnog djelovanja za koji djelatnik može preuzeti odgovornost.

Nabrajajući primjere kognitivnih vrlina J. Baehr, istaknuti predstavnik epistemologije vrlina, spominje (pored ostalog) sljedeće: *znatiželja (inquisitiveness), otvorenost uma, intelektualna hrabrost te intelektualno poštenje*. S druge strane, u području kognitivne znansti, barata se vrlo sličnim atributima. Tako Stanovich, West i Toplak (u daljnjem tekst: S, W, T) (2016), koji su kognitivni znanstvenici, umjesto o kognitivnim vrlinama radije govore o *dispozicijama racionalnog mišljenja* i ilustriraju ih primjerima koji su posve slični Beahrovim. Očito da je razlika među pojmovima epistemičke vrline nasuprot. dispozicija mišljenja samo terminološka. Evo stava Stanovicha, Westa i Toplaka: "Dispozicije racionalnog mišljenja one su koje se posebno odnose na adekvatnost formiranja vjerovanja i donošenja odluka. Primjeri dispozicija racionalnog mišljenja uključuju dispoziciju za *procjenu (odvagivanje) novih evidencija naspram čvrsto ukorijenjenih vjerovanja...*, *dispoziciju da se dobar dio vremena (ili kratko vrijeme) posveti problemu prije nego se odustane ili dispoziciju za procjenu mišljenja drugih prilikom formiranja vlastitih.*" (208). Da se u osnovi radi o vrlo sličnim stavovima, potvrđuje i Beahrovo temeljenje vrlina na dispozicijama. Prema njegovu mišljenju, svaka intelektualna vrlina pretpostavlja posjedovanje *dispozicije* za provođenje određene aktivnosti. Tako Baehr (2013, 249) u nešto formalnijem izričaju tvrdi: "Za svaku intelektualnu vrlinu V, subjekt S posjeduje V samo ako je S:

- (a) disponiran iskazati određenu aktivnost ili psihološku karakteristiku V
- (b) (i to čini) iz ljubavi prema epistemičkim dobrima".

Ovim prikazom kompletirali smo strukturu KM u kojoj smo istaknuli dva aspekta manifestnog djelovanja (ilatički i epistemičke vrline) te dvije utemeljujuće razine (kognitivne sposobnosti i dispozicije). Vratimo se sada na inicijalno pitanje: može li se KM naučiti? Ponuđena struktura

omogućuje da se naučivost razmotri s obzirom na svaki od aspekata ili elemenata koji čine KM. Pođimo od ilatičkih vještina i razmotrimo mogu li se one naučiti? Moj odgovor je negativan, što ću u nastavku obrazložiti.

To što nazivam ilatičkim sposobnostima u osnovi su inferencijalne sposobnosti i predmet su daleko najekstenzivnijeg proučavanja kroz više od četiri decenije. Opsežna psihološka i kognitivistička literatura nastala u ovom razdoblju eksperimentalno potvrđuje izrazito skeptičan stav kako s obzirom na sposobnosti zaključivanja aktualnog mislitelja, tako i s obzirom na mogućnost poboljšanja (naučivost) ovih sposobnosti. Spomenut ću samo najpoznatija istraživanja čiji su rezultati replicirani nebrojeno puta. To je prije svega Wasonov eksperiment, nazvan *problem izbora*, te istraživanja Kahnemana i Tverskog. Wasonov eksperiment⁶ pokazuje da ljudi izrazito loše i teško izlaze na kraj s elementarnim deduktivnim zadacima (preko 95% ispitanika pogrešno je odgovorilo na postavljeni zadatak) dok istraživanja Kahnemana i Tverskog dokumentiraju jednako loše rezultate u domeni indukcije (spomenimo samo tzv. *Linda problem*) i bejzovskog zaključivanja.

Značajno je napomenuti da spomenuta istraživanja prosuđuju ljudsku sposobnost zaključivanja s obzirom na standarde određene pravilima formalne logike (u slučaju deduktivnog zaključivanja) i formalnim pravilima (aksiomima) teorije vjerojatnosti (u slučaju induktivnog i bejzovskog zaključivanja). S obzirom na ove standarde, pokazuje se da su ljudi ako ne neracionalni, a ono barem suboptimalno racionalni. Ovi rezultati ni u kojem slučaju nisu neočekivani. Naime, načela klasične formalne logike suviše su zahtjevna da bismo njima mjerili ljudsku racionalnost. Dva njezina temeljna načela, *deduktivna zatvorenost* i *monotonost*, potpuno su neplauzibilna kao načela (ljudskog) zaključivanja⁷. Uzevši u obzir ove činjenice, možemo se pitati kako je ljude moguće naučiti onome što nisu u stanju provesti?

⁶ Wasonov zadatak izbora ili eksperiment četiri karte dizajniran je ovako: ispitanicima su prezentirane četiri karte prikazane tako da na sebi imaju na jednoj strani slovo, a na drugoj strani broj. Poredak karata bio je ovaj: A, D, 4, 7. Ispitanici su upoznati s pravilom: "Ako karta ima samoglasnik na jednoj strani, tada ima parni broj na drugoj". Zadatak s kojim su bili suočeni bio je: "Trebate reći koje od karata treba preokrenuti da biste utvrdili je li pravilo istinito ili lažno".

⁷ Detaljnije o tome u Smokrović, 2018.

Naravno, ljudsku je inferencijalnu sposobnost moguće mjeriti i drugačijim standardima. Naime, već u vrijeme formuliranja Wasonova zadatka izbora mnogi su kritičari primijetili da bi problem mogao ležati prije u formulaciji zadatka nego u ljudskoj sposobnosti. Originalni Wasonov zadatak formuliran je u apstraktnim terminima brojki i slova. Eksperimenti koji su slijedili zamijenili su u zadatku apstraktne brojeke i slova konkretnim pojmovima koji su prosječnom ispitaniku poznatiji i relevantniji za donošenje zaključka. Kod ovako formuliranog zadatka rezultati su se drastično popravili⁸. Znači li to da su ljudske ilatičke sposobnosti ipak pouzdane i da se one mogu naučiti? Prije nego odgovorimo na ovo pitanje, pogledajmo na čemu se za teoretičare neformalne logike zasniva odnos razloga (premise) i tvrdnje (konkluzije). U kontekstu eksperimenata koje smo prikazali, možemo modificirati pitanje i postaviti ga ovako: na osnovu čega u scenariju s pivom i kolom (vidi fusnotu 8) ispitanici daju ispravne odgovore (kad to već nije na osnovu formalno-logičkih svojstava)?

Teoretičari neformalne logike kao kriterije za procjenu argumenta najčešće predlažu sljedeće: relevantnost, prihvatljivost i dostatnost. Jedan od istaknutih teoretičara ovog pravca Blair kriterij za prepoznavanje argumenta opisuje ovako:

(...) argument je dobar ako su njegovi razlozi ili premise pojedinačno ili u kombinaciji relevantni kao podrška postavljene tvrdnje te zajedno (ako su relevantni i prihvatljivi) dostatni za podršku tvrdnji u prilog koje su bili ponuđeni. (Blair, 2011, 87).

Čini se da su sva tri kriterija (koja su zastupljena i u Blairovom argumentu) epistemološka, ne logička i u znatnoj mjeri relativni s obzirom na znanja i vjerovanja prosuditelja. Naročito je prihvatljivost izrazito subjektivna kategorija. Evaluator će prihvatiti argument s obzirom na kontekst izricanja. Primjerice, velika većina ispitanika prihvatit će kao ispravan ovaj argument:

⁸ Jedna od mnogih reformuliranih varijanti zadatka izbora uključuje sljedeći scenarij: zamisli da poslužuješ za šankom na kojem se nudi pivo i kola. Pravilo je da samo punoljetne osobe smiju piti pivo. Četiri karte sadrže informacije o ljudima za šankom i piću. Na jednoj strani karte je piće, a na drugoj godine starosti. Karte su sljedeće: pivo, kola, 25, 16. Pitanje je koju kartu treba okrenuti da biste utvrdili krši li se pravilo? U ovoj formulaciji zadatka oko 75% ispitanika ispravno je riješilo zadatak.

Ako Ana treba napisati seminarski rad, radit će u biblioteci.

Ana treba napisati seminarski rad,

dakle, radit će u biblioteci.

Ako, međutim, ovom argumentu dodamo još jednu premisu, što u skladu s *pravilom monotonosti* formalne logike ne može promijeniti njegovu valjanost, više od polovice ispitanika ga neće procijeniti valjanim. Evo primjera:

Ako Ana treba napisati seminarski rad, radit će u biblioteci.

Ana treba napisati seminarski rad,

radit će u biblioteci ako biblioteka bude otvorena,

dakle, radit će u biblioteci.

Vidimo da je prihvatljivost argumenta, a i relevantnost relacije premisa i konkluzije ovisna o kontekstu, o sadržaju premisa i prosuditeljevim postojećim vjerovanjima. Nadalje, eksperimenti su pokazali da ispitanici bolje zaključuju u kontekstu koji im je poznat, ali ostaju izrazito skloni pogreškama u kontekstima koji su im udaljeniji i nepoznatiji.

Vratimo se na problem naučivosti. Domena primjene kritičkog mišljenja po pretpostavci ne može biti ograničena na usko područje blisko mislitelju. Kritički mislitelj, odnosno ispravni prosuditelj argumenta, trebao bi iskazivati sposobnost kritičkog mišljenja u različitim kontekstima, ne samo u onima koju su mu poznati. Ako je napredak u ovom pogledu i moguć, on je ostvariv samo u kontekstima koji su prosuditelju poznati. Ne ulazeći dublje u ovu zanimljivu problematiku, kao ni i u brojne pokušaje osporavanja rezultata istraživanja, važno je reći da je dominantno mišljenje unutar kognitivne teorije to da je nerealno očekivati poboljšanja ili meliorativne pomake unutar inferencijalne domene bilo u deduktivnom, bilo u induktivnom području. Pogotovo je to nerealno očekivati kroz učenje 'ispravnih' pravila zaključivanja i ukazivanja na tzv. logičke pogreške, kako se to često naivno pretpostavlja u teorijama kritičkog mišljenja

Kognitivne sposobnosti, rekli smo, stoje u osnovi ilatičkih sposobnosti. Kod S, T, W (2013) pobrojene su one najvažnije: "Brzina percipiranja, točnost razlikovanja, kapacitet radne memorije i efikasnost prizivanja informacija uskladištenih u dugoročnoj memoriji". Uočimo da je jedna od dominantnih paradigmi u kognitivnoj znanosti teorija dualnih

procesa koja odvaja tzv. subpersonalne (ili *sistem 1*) od personalnih (ili *sistem 2*) kognitivnih procesa, pri čemu su oni prvi automatski, asocijativni i manje-više nesvjesni, dok su oni drugi refleksivni i utemeljeni na pravilima zaključivanja. Problem je u tome što kognitivne sposobnosti mahom spadaju u subpersonalne, nesvjesne i automatske procese. Teško je pretpostaviti da ono što se nalazi ispod praga svjesnog i refleksivnog može biti podvrgnuto učenju i poboljšanju kroz učenje. Što je još gore, kako je Reber (1993) pokazao: "Predsvjesni procesi imaju nizaki varijabilitet i ne pokazuju značajan odnos spram analitičke inteligencije". Zaključimo da je i u ovoj domeni nerealno očekivati poboljšanja kroz učenje.

Preostaje nam analiza posljednjeg elementa KM, epistemičkih vrlina. Mogu li se ove vrline steći učenjem? Za ovaj dio analize potrebno je vratiti se na članak Kotzee, Carter, Siegel, (u daljnjem tekstu: K, C, S) 2019, u kojem se tvrdi a) da kritičko mišljenje nema veze s epistemičkim vrlinama i b) da se kritičko mišljenje, za razliku od epistemičkih vrlina može naučiti. Pokušavajući pokazati što je to konkretno u KM što se može naučiti, autori nabrajaju upravo svojstva koja nedvojbeno pripadaju epistemičkim vrlinama. Evo citata: "Ono (kritičko mišljenje) uči studente što kritičko mišljenje jest, zašto je ono dobro (i zašto je nekritičko mišljenje loše) te kako misliti kritički. Pod 'dobrim' ovdje mislimo na dobro s epistemičke točke gledanja, to jest mišljenje koje zadovoljava epistemičke kriterije time što nudi razloge koji su usmjereni na istinu i dokazno snažne evidencije, koje izbjegava pogreške u zaključivanju i koje je poduprto širokim razumijevanjem epistemičkih kvaliteta i uvažavanjem kriterija u terminima kojih su te kvalitete određene...". Uočimo najprije da su sva navedena svojstva metainferencijalni predikati i da svaki od njih izražava neki aspekt epistemičkih vrlina. Sva se ona odnose na ono što S, W, T (2013, 249) određuju kao "adekvatnost formiranja vjerovanja i donošenja odluka", što je upravo određenje dispozicija mišljenja. Konkretno, svojstvo 'usmjerenosti na istinu' izravno odgovara Beahrovim vrlinama epistemičke znatiželje i otvorenosti uma. Ovo posljednje jest vrлина za koju S, W, T tvrde da je najuže povezana sa sposobnostima mišljenja. Također, 'dokazno (*probative*) snažne evidencije' naprosto su 'dispozicija za *procjenu (odvagivanje)* novih evidencija *naspram čvrsto ukorijenjenih vjerovanja*' kod S, W, T.

Vjerujemo da smo pokazali da teza a) kod K, C, S ne stoji. S obzirom na tezu b), čini se da je sami autori, pod pretpostavkom da je naša interpretacija ispravna, demantiraju. Naime, ako se kritičko mišljenje (odvojeno od epistemičkih vrlina), kako tvrde, može naučiti, a pokazuje se da je to što se može naučiti upravo u domeni epistemičkih vrlina, onda su epistemičke vrline onaj element u strukturi kritičkog mišljenja koji se može naučiti.

Vidjeli smo da K, C, S smatraju da je studente potrebno podučiti tome zašto je KM “dobro”, a to specificiraju: “Pod ‘dobrim’ ovdje mislimo na dobro s epistemičke točke gledanja”. Ovo određenje upravo odgovara Beahrovom spominjanju epistemičkih dobara čije stavove autori žele osporiti. Beahr, naime, kaže da netko posjeduje vrlinu samo ako je:

(a) disponiran iskazati određenu aktivnost ili psihološku karakteristiku V

(b) (i to čini) iz ljubavi prema epistemičkim dobrima.

Naše je pitanje, bez obzira radi li se o nekom monolitnom KM ili o epistemičkim vrlinama kao segmentu KM: kako se ljubav prema epistemičkim dobrima ili u najmanju ruku korist od epistemičkih dobara može edukacijom prenijeti na studente? Već smo pokazali da je nerealno očekivati da se to ostvari kroz učenje ‘ispravnih’ pravila zaključivanja i ukazivanjem na tzv. logičke i druge pogreške. Smatramo da korist od epistemičkih dobara množe postati očigledna tek kroz praksu i upotrebu segmenata mišljenja u kojima epistemička dobra igraju značajnu ulogu. Praksa na koju mislimo jest argumentativna praksa koja ima za cilj pokazati (demonstrirati) da je argument ispravan, odnosno da je oponentov argument neispravan. Svatko tko sudjeluje u argumentativnoj praksi želi pokazati ispravnost, a time i nadmoć svojeg stava zasnovanog na argumentu. Drugim riječima, sudionik argumentativnog procesa želi dokazati adekvatnost svog argumenta s obzirom na cilj, bila to uvjerljivost, valjanost ili istinitost. Korist od “epistemičkih dobara” može se prepoznati tek kroz praksu koja jasno razdvaja ona svojstva koju su efikasna za postizanje argumentacijskog cilja od onih koja to nisu.

Dakle, da rezimiramo naš stav: nerealno je očekivati da mišljenje može postati “kritičko” podučavanjem formalno-logičkih ili neformalnih pravila. Ako uopće postoji mogućnost učenja u domeni KM, tada se ona može ostvariti kroz usvajanje epistemičkih vrlina. Segment kroz

koji se može prepoznati, pa onda i prihvatiti, efikasnost epistemičkih dobara (epistemičkih vrlina) nalazi se na metainferencijalnoj razini na kojoj mislitelj prosuđuje uspješnost svojeg argumentiranja⁹. One strategije koje se pokazuju uspješne usvajaju se dok se one neuspješne izbjegavaju. Naravno, ne vodi svaka argumentativna praksa ovakvoj soluciji. Najčešće agumentacija završava u polarizaciji mišljenja i neslaganju. Međutim, ako postoji mogućnost napretka, onda ona leži u spremnosti da se napuste strategije koje vode neslaganju i prihvate one koje mogu polučiti razrješenje sukoba¹⁰.

Literatura

- Baehr, J., 2013, The Cognitive Demands of Intellectual Virtue. In T. Henning and D. P. Schweikard, eds., *Knowledge, Virtue, and Action: Putting Epistemic Virtues to Work*. London: Routledge, str. 99–117.
- Blair, J., 2011, *Groundwork in the theory of argumentation*. New York: Springer.
- Davies, M., Barnett, R., 2015, *The Palgrave Handbook of Critical Thinking in Higher Education*, Palgrave Macmillan.
- Johnson, R., 2000, *Manifest Rationality*, Mahwah, NY: Lawrence Erlbaum.
- Johnson, R., 2012, When Informal Logic Met Critical Thinking, *Inquiry: Critical Thinking Across the Disciplines*, 27, 3: 5–14.
- Kahneman, D., Slovic, P., Tversky, A., 1992, *Judgment under uncertainty: Heuristics and biases*, Cambridge University Press.
- Kerlinger, F. N., 1986, *Foundations of Behavioral Research*, Harcourt Brace.
- Kotzee, B., Carter, J. A., Siegel, H., 2019, Education for Intellectual Virtue: a critique from action guidance, *Episteme*: 1–23.
- Kuhn, D., 2016. A Role for Reasoning in a Dialogical Approach to Critical Thinking, *Topoi*, DOI 10.1007/s11245-016-9373-4.
- Mercier, H., Sperber, D., 2011, Why do humans reason? Argument for an argumentative theory, *Behavioral and Brain Sciences* 34: 57–111.
- Perelman, C., Olbrecht-Tyteca, L., 1958, *Traité De L'argumentation: La Nouvelle Rhétorique*, Paris: Perss Universitaires de France.
- Rips, L. J., Adler, J. E., 2008, *Reasoning*, Studies for Human Inference and its Foundations, Cambridge University Press.

⁹ Detaljnije o tipu argumentacije koji može dovesti do proširenja znanja u: Smokrović, 2015.

¹⁰ Sličan stav našem zastupa Deanna Kuhn (Kuhn 2016)

- Shand, J., 2000, *Arguing Well*, Routledge.
- Smokrović, N., 2015, Argumentation as a Means for Extending Knowledge, *Croatian Journal of Philosophy*, 44.
- Smokrović, N., 2018, Informal Reasoning and Formal Logic, *Croatian Journal of Philosophy*, no. 54: 455–471.
- Stanovich, K., West, R., Toplak, M., 2016, *The Rationality Quotient*, MIT Press.
- Toulmin, S., 1958, *The Use of Argument*, Cambridge: Cambridge UP.
- Wason, P. C. 1968, Reasoning about a rule, *Quarterly Journal of Experimental Psychology*, 20 (3): 273–281.

Naturalizam “quattro stagioni”

Luca Malatesti, Predrag Šustar

1. Uvod

Metafilozofska pitanja o odgovarajućim ciljevima, predmetima i metodama filozofije uvijek su bila konstitutivni dio filozofskog istraživanja i sada ih se istražuje u metafilozofiji, koja je poprimila oblike specifične filozofske discipline (Overgaard, Gilbert, i Burwood 2013; Williamson 2010; 2018). Usporedba filozofije s drugim disciplinama i granama znanja te istraživanje njezinih odnosa s njima često su u središtu metafilozofskih istraživanja. Konkretno, odnos filozofije i znanosti važan je problem za metafilozofsko istraživanje. Jedna od temeljnih dimenzija metafilozofske rasprave odnosi se na pitanje treba li filozofija biti u tematskom i metodološkom kontinuitetu sa znanostima.

Neki filozofi prihvaćaju gledište *diskontinuiteta*. Prema toj praksi ili teoriji, filozofija je metaznanost koja bi, iz specifične filozofske perspektive, trebala istraživati, kritizirati i preporučiti znanstvene metode i tako potvrditi ili odbaciti znanstvene pretenzije na znanje. Takvo je stajalište ukorijenjeno u kantovskoj tradiciji, s posebnim filozofskim fokusom na transcendentalne preduvjete bilo kojeg znanja. Slično tome, logički pozitivisti, sa svojim posebnim filozofskim metodološkim fokusom na logičke i semantičke značajke jezika, prihvaćali su takvo gledište. U ovom slučaju, postoji diskontinuitet između predmeta i metoda znanosti te filozofije. Međutim, prema drugačijem pristupu, može se smatrati da su filozofija i znanost u *kontinuitetu*. Važan pobornik ovog je pristupa Willard Van Orman Quine. Ovo stajalište, između ostalog, preporučuje metodološki naturalizam, prema kojemu bi filozofija u razmatranju pripadnih problema trebala koristiti metode empirijskih i/ili formalnih znanosti (Quine 2013).

U ovom poglavlju, bez ulaska u raspravu o krajnjim zaslugama spomenutih stajališta, cilj nam je unaprijediti i potkrijepiti primjerima istraživanja koja trenutno provodimo, a koja zastupaju naturalistički pristup između ovih dviju krajnosti. Glavna načela našeg stava jesu sljedeća: *prvo*, u suprotnosti s gledištem diskontinuiteta, u pragmatičnome duhu; filozofija može reagirati na specifične znanstvene probleme s fokusom na pružanju odgovarajućih rješenja, a ne na osiguravanju ultimativnog temelja određenog područja istraživanja i povezane praktične primjene. Stoga kao *drugo*, istraživanje treba biti znanstveno informirano na odgovarajući način. *Treće*, filozofsko istraživanje moglo bi biti metodološki autonomno spram znanosti zbog skupa specifičnih pitanja i, u određenoj mjeri, konceptualne analize, koja bi međutim ipak trebala biti informirana, ograničena i opravdana pitanjima koja su relevantna za područje istraživanja.

Prvi primjer naturalističkog pristupa koji zastupamo u ovom poglavlju odnosi se na doprinos koji filozofija biologije može ponuditi u recentnoj “debati o informaciji” u genetici i općenito biomedicinskim znanostima. Ovo je metodološki primjer koji preporučujemo u odnosu na filozofski rad koji treba provoditi ‘unutar’ znanstvene discipline. Naime, filozofi mogu doprinijeti znanosti pojašnjavajući pojmove koji su relevantni u specifičnim znanstvenim disciplinama ili čak specifičnim istraživačkim projektima. Kao što ćemo nastojati pokazati, to ne mora biti ‘izvanjski’ doprinos znanostima, kao što je to najčešće slučaj s gore spomenutim stajalištima. Umjesto toga, očekuje se *prvo*, da filozofi ekspliciraju znanstvene prakse i ciljeve koji se pozivaju na ključne pojmove. *Drugo*, filozofi mogu ponuditi pojašnjenja tih ključnih pojmova koji usklađuju same znanstvene prakse i ciljeve. I u konačnici *treće*, takva vrsta pojmovne analize na određeni način ažurira poveznicu između znanstvene prakse i tradicionalnih filozofskih problema, kao što to korištenje pojma informacije u molekularnoj biologiji ponajviše čini s nekim aspektima problema uzročnosti, ali i drugih problema tradicionalno pripadnih filozofiji.

Drugi primjer odnosi se na doprinos koji filozofija može ponuditi u primjeni znanstvenih spoznaja u praksama pripisivanja pravne i kriminalne odgovornosti. U ovom slučaju, osim etičkih ili normativnih problema koji bi neposredno očito trebali privući filozofsku pozornost, postoje i ozbiljni teorijski izazovi pri čijem nam rješavanju filozofija može

pomoći. Što se tiče ove teoretske dimenzije, čini se da se nekoliko pitanja odnosi na tradicionalne filozofske probleme kao što su to problem odnosa uma i tijela ili problem slobodne volje. Međutim, u nastavku argumentiramo da pobliže fokusiranje na pravnu praksu i ograničenja trenutnih znanstvenih spoznaja otvara konkretnija pitanja koja imaju i konceptualne dimenzije kojima se bave i filozofi. U tom kontekstu smatramo kako ne isključivo pragmatični razlozi nalažu kako se istraživanje ne mora nužno proširiti i na standardne filozofsko-ontološke probleme.

U sljedeća dva odsječka predstavljamo dvije studije slučaja koje se odnose na interakciju filozofije sa znanošću. U svakom odsječku opisujemo neke od temeljnih pretpostavki koje karakteriziraju te probleme i dajemo nacrt kako bi trebalo 'prakticirati' odnos između znanosti, pripadnih praksi i filozofije.

2. Rasprava o genetičkoj informaciji: konceptualna analiza u (znanstvenoj) praksi

Sustavna filozofska i znanstvena rasprava o pojmu biološke, odnosno ponajviše genetičke informacije sada već broji više od dva desetljeća (usp., primjerice, Okasha 2019). Isto vrijedi za znanstvene pojmove u biologiji koji su srodni *biološkoj informaciji* kao što su to *genetički kod*, unutar i/ili međustanično *signaliziranje* i drugi pojmovi u uporabi u sve sofisticiranijoj eksperimentalnoj praksi recentne biologije. Kao što ćemo ukratko vidjeti u nastavku, ova rasprava ima više krakova ili aspekata. No mi ćemo se zadržati upravo na izvjesnoj *diskrepanciji* između spomenute eksperimentalne sofisticiranosti, posebice molekularne biologije i tome nasuprot, relativno relaksiranoj uporabi pojmova kao što je to biološka, ovdje specifično genetička, informacija.

Krenut ćemo od tvrdnji koje su standardno u uporabi u molekularnoj biologiji, kako u samom istraživanju, tako i u udžbeničkim prikazima. Tako se uporaba pojma informacije unutar molekularne biologije može sažeto prikazati pomoću sljedećih iskaza:

- (i) makromolekularna sekvenca tipa S prenosi informaciju za sintezu makromolekularne sekvence tipa S^* ;
- (ii) gen G je o G^* (ali obrnuto ne vrijedi).

Kao standardnu ilustraciju iskaza (i) možemo uzeti DNK ili RNK sekvencu nekog individualnog organizma koja sintetizira druge makromolekularne sekvence, primjerice različite vrste RNK molekula te polipeptidnu sekvencu budućeg proteina. U konačnici, ta inicijalna makromolekularna nukleinska sekvenca u određenoj mjeri dovodi do fenotipskog svojstva pripadnog organizma. Vidjet ćemo u nastavku kako "centralna dogma" molekularne biologije u osnovi sažimlje iskaz (i) i njegove varijante. Ovdje ćemo više istaknuti način na koji se pojam genetičke informacije pojavljuje i koristi u (i) i srodnom iskazu (ii). Naime, (i) implicira postojanje dviju domena barem minimalno razdvojenih: domenu *uzročnih odnosa*, prije svega, izravnih i posrednih između glavnih skupina biomakromolekula, kao i brojnih pripadnih pozadinskih uvjeta za pravilno odvijanje tih uzročnih procesa, primjerice kod prethodno spomenute sinteze proteina; te *informacijsku* domenu, koja posljedično uključuje prijenos genetičke informacije kroz različite vrste i razine biološke organizacije, osim same makromolekularne.

Uvođenjem potonje domene nameće se više pitanja. U kakvom su odnosu dvije domene? Nadalje, ukoliko zadržimo pojam informacije u molekularnoj biologiji, je li opravdano tvrditi kako postoje 'privilegirani' nositelji informacije za sintezu biomakromolekularnih sekvenci i odgovarajućih fenotipskih svojstava organizama? Potonje pitanje tiče se prije svega DNK, tj. specifično bioloških uloga gena kao primarnih repozitorija i vektora informacije u odnosu na ostale sastavnice živoga svijeta (usp. Waters 2007). Tako iskaz (ii) eksplicitno ilustrira drugo pitanje. Osim toga, čini se kako (ii) dodatno ističe prije spomenuto razdvajanje domena u biologiji, naznačujući kako geni nisu samo uzročni čimbenici, već posjeduju semantičke karakteristike. Drugim riječima, primarna karakteristika gena bila bi to da su 'o' nekom svojstvu, skupini svojstava ili biološkom entitetu na način kao što je to, primjerice, recept za kolač 'o' samoj pripravi kolača. Dakle, iskaz (ii) razlikuje se od iskaza (i) upravo zbog isticanja gledišta kako puko uzročno razumijevanje aktivnosti gena ili općenito DNK nije dovoljno (za više o dva glavna shvaćanja uloge gena u prijenosu informacije vidi Šustar 2007). Pogledajmo klasifikaciju odgovora na prethodno spomenuta dva pitanja (usp. Downes 2006, str. 67):

(1) Informacija je prisutna u DNK-u i drugim nukleotidskim sekvencama (primjerice kod RNK virusa). Ostali stanični mehanizmi ne sadrže informaciju;

(2) Informacija je prisutna u DNK-u, isto tako kod drugih nukleotidskih sekvenci te kod ostalih staničnih mehanizama (primjerice mehanizama koji se tiču enzima ili općenito proteina. Oni također mogu prenositi biološku informaciju drugim biomolekulama s kojima su u odgovarajućem kontaktu. Dakle, prijenos informacije nije isključivi privilegij DNK i RNK, odnosno, gena). Konačno, po (2) slijedi kako i okoliš nekog organizma ili živog sustava prenosi informaciju u funkcioniranju organizma i/ili njegovih dijelova;

(3) DNK i druge nukleotidske sekvence ne sadrže informaciju niti je sadrže ostali stanični mehanizmi.

(1) tvrdi da postoji informacijska domena. No, samo za nukleinske kiseline, točnije gene.¹ Odgovori u (2) maksimalno proširuju uporabu informacijske terminologije i na ostale biološke mehanizme – sve do relevantnih uvjeta u okolišu. Posljedično, DNK i RNK gube privilegiranu ulogu u prijenosu biološke informacije. Konačno, odgovori u (3) u potpunosti eliminiraju samostojnost informacijske domene u biologiji, uključujući same nukleinske kiseline, tj. gene. (3) stoga jasno zagovara pristup po kojem u biologiji treba razgraničiti različite vrste uzročnih struktura; tako uzročno djelovanje RNK, posebice DNK, odlikuje se izrazito preciznim određivanjem drugih biomakromolekula kao što su to same nukleinske kiseline tijekom njihove replikacije te polipeptida, odnosno proteina tijekom genske ekspresije. Drugim riječima, jedan od razloga za odgovore u točki (1) bila bi upravo navedena preciznost, tj. specifičnost uzročnog djelovanja DNK i RNK u proizvođenju njihovih učinaka, analogno srodnim karakteristikama u prijenosu informacija u ljudskim komunikacijskim sustavima i tehnologijama (usp. primjerice, Stegmann 2014). Pogledajmo sada neke poznatije znanstvene odgovore i njihov razmještaj s obzirom na prethodne točke (1)–(3).

Najprije metafizički smjeliji odgovori. To su znanstveni odgovori koji otvoreno postuliraju opstojnost zasebne informacijske domene u kojoj

¹ Dodatno pitanje koje se tiče statusa pojma genetičke informacije, također u velikoj mjeri prisutno u predmetnoj raspravi (usp., primjerice, Godfrey-Smith i Sterelny 2016), dobilo je nekoliko razrađenih odgovora: od toga kako je ovaj pojam tek *metafora* ili određena *analogija* s pojmovima iz područja informacijskih i komunikacijskih tehnologija do odgovora kako je pojam genetičke informacije, uvjetno biološke informacije nešto više od metafore; naime, omogućuje *heuristiku* u stjecanju znanja o temeljnim biološkim mehanizmima (za potonji odgovor, usp. Levy 2011). Više o ovom pitanju i drugačijem odgovoru usp. završni dio Odsječka 2.

DNK igra glavnu, ako ne isključivu ulogu. Tako George C. Williams (1992, str. 11) ističe kako molekularna biologija vodi k specifičnoj dualističkoj ontologiji po kojoj genetička informacija i njezini pojmovni srodnici, kao što je to primarno kodiranje proteina te općenito fenotipskih svojstva formiraju autonomnu biološku domenu. Drugim riječima, čine tzv. "kodirajuću domenu" (engl. *codical domain*) koja u molekularnoj biologiji postoji paralelno uz onu materijalnu. Williams ilustrira takvu ontologiju sljedećom tvrdnjom: "Gen nije DNK-molekula; on je informacija koja se može transkribirati, a kodirana je DNK-molekulom." (Williams 1992, str. 11). Sličan dualizam prisutan u opisima bioloških pojava na molekularnoj razini možemo pronaći i kod Johna Maynarda Smitha i Eörsa Szathmáryja (1999, str. 11) u gledištu kako fenomen života na svim razinama biološke organizacije pokazuje dualnost "metaboličke" i "informacijske" domene. Kao ilustraciju zamislimo, s jedne strane, transkripciju DNK na glasničku RNK (mRNK), a s druge, metaboličke strane, utrošak ATP kao izvora energije u funkcioniranju tog istog biološkog mehanizma; ukratko, prijenosa genetičke informacije i samih metaboličkih aktivnosti tzv. 'biokemijske mašinerije' unutar stanice. Konačno, Sydney Brenner (1999, str. 1964) zaključuje kako uz temeljni tijek materije i energije u živim sustavima postoji i karakterističan paralelni tijek genetičke informacije.

No, uz ove odgovore uglednih znanstvenika koji se u osnovi svrstavaju u prethodno spomenutu skupinu (1), povijest molekularne biologije dala je i metafizički umjerenija gledišta. Vjerojatno najeksplicitnija po pitanju samog pojma genetičke informacije sadržana je u samostalnim radovima Francis H. C. Cricka, jednog od ključnih utemeljitelja suvremene biologije (usp. više o povijesti molekularne biologije Divan i Royds 2016). Smatramo kako upravo njegove definicije pojma informacije predstavljaju prikladno polazište za filozofski tip konceptualnog razmatranja znanstvenih pojmova u molekularnoj biologiji. Stoga ćemo najprije dati prikaz Crickovih definicija, potom nacrt naše pozicije – posebice imajući u vidu metafilozofsku naturalističku težnju.

Crickove definicije genetičke informacije. Crickov poznati članak *On Protein Synthesis* (1958, str. 144) donosi prvu definiciju:

"Pod informacijom podrazumijevam specifikaciju aminokiselinskog slijeda (*sequence*) u proteinu."

Druga definicija pojma genetičke informacije u povijesti molekularne biologije isto tako pripada Cricku. Točnije, dana je unutar njegove formulacije najopćenitije generalizacije na ovom području, a to je tzv. centralna dogma molekularne biologije:

"*Centralna dogma*. Ona tvrdi da kad jednom 'informacija' prijeđe u protein *ne može više izići*. Detaljnije, prijenos informacije s nukleinske kiseline na nukleinsku kiselinu ili s nukleinske kiseline na protein jest moguć, ali prijenos s proteina na protein ili s proteina na nukleinsku kiselinu nije moguć. Informacija ovdje znači *precizno* određivanje sljedova (*sequence*) ili baza u nukleinskoj kiselini, ili sljedova aminokiselinskih ostataka u proteinu." (Crick 1958, str. 153)

Potonja Crickova definicija vjernije ocrtava prevladavajuće shvaćanje pojma genetičke informacije u znanstvenoj praksi suvremene molekularne biologije. Razlozi su sljedeći: (i) definicija koja se zasniva na terminu 'određivanja' (engl. *determination*) obuhvaća sve temeljne makromolekularne odnose u biologiji, ne isključivo sintezu proteina, već i replikaciju DNK i RNK te transkripciju informacije s DNK na mRNK, pored ostalih procesa; (ii) sam Crick znatno češće koristi relaciju 'određivanja' između nukleinskih kiselina i proteina nego što koristi termin 'specifikacija' iz prve definicije. To vrijedi neovisno od činjenice da je druga definicija genetičke informacije neposredno povezana s formulacijom poznate centralne dogme; konačno, (iii) glavne empirijske generalizacije u ovom znanstvenom području pozivaju se na relaciju uzročnog 'određivanja' svojstvenu prije svega za DNK, a kada se i pozivaju na relaciju 'specifikacije', u krajnjoj liniji svode je na manje prijeponu i češće korištenu vrstu relacije. Međutim, prije eksplikacije uzročnog određenja u prijenosu informacije s DNK (i RNK) na proteine, barem što se tiče Crickovih definicija, ukratko o središnjoj poteškoći s kojom se treba suočiti prethodno naznačena uzročna interpretacija pojma genetičke informacije.

Poteškoća ima dva dijela: *prvi* se odnosi na opravdanost pripisivanja prijenosa genetičke informacije samo u jednom smjeru. Naime, od DNK i/ili RNK na polipeptide, tj., proteine. Drugim riječima, ako i interpretiramo Crickove definicije na temelju odgovarajućih uzročnih odnosa, zašto bi nukleinske kiseline bile *privilegirani* uzročni čimbenik? *Drugi* dio poteškoće tiče se razlikovanja, ukoliko prihvatimo uzročnu interpretaciju Crickova korištenja informacijskih analogija u molekularnoj biologiji, samih *uzroka* naspram tzv. 'pozadinskih uvjeta' (engl. *backgro-*

und conditions). Primjerice, u standardnoj situaciji kad upalimo šibicu, što je uzrok, a što su tek pozadinski uvjeti (količina kisika u prostoriji, vlažnost šibice i sl.)? U našem slučaju, kako razlikovati uzročni primat nukleinskih kiselina u proizvođenju odgovarajućih učinaka – polipeptida – naspram pozadinskih uvjeta, koji se obično u znanstvenoj praksi molekularne biologije pojednostavljeno karakteriziraju kao ‘biokemijska mašinerija’, sastavljena uglavnom od enzima, pripadnih kofaktora i ostalih uvjeta koji moraju biti ispunjeni za ispravan prijenos genetičke informacije koji spominje glavna Crickova definicija?

Slijedom druge Crickove definicije, standardno korištena formulacija kako se genetička informacija prenosi sa skupine biomakromolekularnih sekvenci *S* na skupinu biomakromolekularnih sekvenci *S** ekvivalentna je tvrdnji kako *S* određuje sintezu *S** na specifičan uzročni način. Eksperimentalna praksa molekularne biologije pokazuje kako se informacijska domena, po Crickovu shvaćanju, može ustvari izjednačiti s odgovarajućom uzročnom domenom, odnosno uzročnom strukturom.² Tako, primjerice, u procesu sinteze primarne strukture proteina ili tzv. ‘genske ekspresije’ biomakromolekularna sekvenca baza poliadenin na jednom od DNK lanaca određuje sintezu biomakromolekularne sekvence polifenilalanina kod novonastalog proteina. Vjerojatnost *preciznog određivanja* putem odgovarajuće uzročne strukture još je izraženija u procesu DNK replikacije te u procesu transkripcije iako na drugačiji način, prije svega zbog kompleksnosti organizacije genoma kod eukariota (vidi Sarkar 2000). Stoga prethodno istaknuta vrsta i razina preciznosti u određivanju pripadnih učinaka jest to što izdvaja uzročnu ulogu nukleinskih kiselina u temeljnim biološkim procesima u odnosu na uzročne uloge enzima u pratećoj ‘biokemijskoj mašineriji’ unutar stanice. U osnovi, proteini u svojoj funkcionalnoj formi nemaju *linearnost*, odnosno *sekvencijsku uzročnu strukturu* te rezultirajući način djelovanja kao što vrijedi za nukleinske kiseline. Smatramo kako su upravo ti razlozi motivirali Crickovo uvođenje informacijske domene u molekularnu genetiku. Dodatno, za izdvajanje uzročnog djelovanja

² Za recentni trend u filozofiji biologije kojim se diverzificiraju uzročne strukture za znanstveno objašnjenje bioloških pojava, uključujući tijek (*flow*) genetičke informacije u procesu sinteze proteina, usp. Ross 2021. Smatramo kako se Crickove definicije, posebice definicija pojma informacije koja je sadržana u centralnoj dogmi, uklapaju u taj trend.

nukleinskih kiselina, posebice DNK u odnosu na ostale skupine biomakromolekula.

Naše zaključno gledište po ovom pitanju stoga nije *eliminacijsko*, tj. isključivanje pojma genetičke informacije i srodnih pojmova iz rječnika molekularne biologije, već svojevrsni *prosvijećeni pluralizam* po kojem možemo koristiti informacijsku terminologiju, kao što to primjerice čini Crickova formulacija centralne dogme. No, to također vrijedi i za eksplicirano značenje pojma u terminima uzročnih struktura i načina njihovog djelovanja.³

Dakle, filozofi mogu doprinijeti znanosti pojašnjavajući pojmove koji su relevantni u specifičnim znanstvenim disciplinama ili čak specifičnim istraživačkim projektima. Kao što se može vidjeti u prethodnoj raspravi, to nije "izvanjski" doprinos znanostima. Umjesto toga očekuje se *prvo*, da filozofi ekspliciraju znanstvene prakse i ciljeve koji se pozivaju na ključne pojmove. *Drugo*, filozofi mogu ponuditi pojašnjenja tih ključnih pojmova koji usklađuju same znanstvene prakse i ciljeve. I u konačnici *treće*, takva vrsta pojmovne analize na određeni način ažurira poveznicu između sastavnica znanstvene prakse i tradicionalnih filozofskih problema, kao što to prethodno razmatrano korištenje pojma informacije u molekularnoj biologiji čini s nekim aspektima problema uzročnosti. U sljedećem odsječku krećemo u razmatranje druge instance filozofskog konceptualnog rada u *kontinuitetu* s drugim disciplinama koji osim znanstvenih istraživanja, razmatra i pravnu praksu.

3. Znanost u sudnici: moguća uloga za filozofiju

Relevantan dio neuroetike u području neuroprava (engl. *neurolaw*) istražuje normativni značaj napredaka u znanosti o mozgu, umu i ponašanju za administraciju zakona (Clausen i Levy 2015). Istražuju se pravni aspekti neuroznanstvenih tehnika detekcije laži (Danaher 2018), pravni izazovi predikcije antisocijalnog ponašanja na temelju biomarke-

³ Korišteni naziv 'prosvijećeni pluralizam' nalazi primjenu ne samo u filozofiji biologije, već i u samoj znanstvenoj praksi molekularne biologije, primjerice kod Cricka. Tako da u raspravi o pojmu genetičke informacije i njegovu statusu postoji značajna srodnost gledišta, koju smo ovdje naposljetku nastojali istaknuti.

ra (Singh, Sinnott-Armstrong i Savulescu 2014), pravni ili etički aspekti različitih vrsta kognitivnog ili moralnih biopoboljšanja (engl. *bioenhancement*) (Dubljević 2019). Dio rasprave bavi se utjecajima koje bi trenutno dostupne neuroznanstvene spoznaje, kao i srodne discipline, trebale imati na prakse u kojima važnu ulogu igra pojam pravne odgovornosti (Vincent 2013).

Neke neuroetičke rasprave odvojene su od sadašnje znanosti i pravnih načela, doktrina i praksi. Raspravlja se o *budućim* ili *moćnim* posljedicama neuroznanosti i kognitivnih znanosti za primjenu kriminalnog zakona. To istraživanje često se odnosi na tradicionalnu raspravu o slobodi volje. Na primjer, dok neki tvrde da će nas neuroznanost, koja navodno podržava deterministički pogled ljudskog djelovanja, prisiliti da napustimo ili u velikoj mjeri revidiramo trenutnu teoriju i praksu povezanu s pravnom odgovornošću (De Ridder, Verplaetse i Vanneste 2013), drugi se ne slažu (Mele 2014). Oni koji sudjeluju u tim raspravama često promiču neuroznanstvenu “ushičenost” koja graniči s pseudoznаноšću. Na primjer, većina etičkih rasprava o biološkom poboljšanju naših kognitivnih ili moralnih sposobnosti temelji se više na hipotetičkim znanstvenim spoznajama nego na trenutno dostupnim rezultatima (Wiseman 2016).

Međutim, neki su autori uvjerljivo argumentirali da je potreban konceptualni okvir za povezivanje napretka u znanosti i prava (Pardo i Patterson 2013; Hirstein, Sifferd i Fagan 2018; Buckholtz, Reyna i Slobogin 2016). Filozofija može u tom kontekstu, usporedno sa znanošću i pravom, pomoći u razvoju tog konceptualnog okvira imajući u vidu dva glavna cilja. Prvo, taj okvir bi omogućio utvrđivanje značaja stvarnog znanstvenog napretka za trenutnu pravnu praksu. Drugo, ponudio bi pojmovne temelje za program promicanja znanstvenog istraživanja i pravne prakse s ciljem razvoja plodne interakcije između tih domena. Možemo to ilustrirati na slučaju obrane na suđenjima za krivična djela koje se baziraju na nebrojivosti zbog duševnih smetnja.

Obično ta vrsta obrane ima dvije komponente (Grozđanić, Škorić, Martinović 2013; Meynen 2016). Osoba se ne može smatrati odgovornom za kazneno djelo ako kada je to počinila: (1) nije bila sposobna shvatiti značenje svojeg postupanja ili vladati svojom voljom; (2) i to zbog mentalnog poremećaja. Na primjer, osoba koja ubije drugog pojedinca jer zbog psihoze misli da je žrtva opasan demon smatrat će se

neubrojivom. Zapravo, ta osoba zbog svoje patološke zablude nije sposobna razumjeti prirodu svog postupka. Slično tomu mentalni poremećaj mogao bi natjerati nekoga na počinjenje kaznenog djela čak i ako je osoba svjesna da je djelo nezakonito.

U slučaju obrane neubrojivosti pojmovni bi okvir omogućio prevladavanje poteškoća koje onemogućavaju povezivanje nesposobnosti koje su potrebne za neubrojivost i relevantnih eksperimentalnih paradigmi te neuroznanstvenih rezultata i hipoteza. Filozofi bi mogli pomoći u izgradnji tog okvira u suradnji sa stručnjacima iz drugih relevantnih disciplina. Okvir treba uključiti jasne i zakonski opravdane formulacije psiholoških nesposobnosti koje su relevantne za neubrojivost. Važno je istaknuti da o tome ne postoji konsenzus o formulacijama dviju komponenti obrane neubrojivosti. Na primjer, u odnosu na drugu komponentu i dalje se raspravlja treba li na koji način pojam mentalnog poremećaja biti uključen u pravne formulacije pojma krivične neubrojivosti (Mala-testi, Jurjako, i Meynen 2020).

Štoviše, postoje različita tumačenja nesposobnosti koja su prema prvoj komponenti obrane nužne za neubrojivost. Na primjer, nije kontroverzno da ubrojiva osoba kako bi razumjela prirodu svog postupka, treba imati neki bazični dodir sa stvarnošću. Dakle, potpuno zabludjena osoba nema tu sposobnost te se ne bi se trebala smatrati ubrojivom. Međutim, srž uvjeta prema kojem ubrojiva osoba razumije značenje svog kriminalnog postupka jest da ona razumije da to nije dopušteno. Debatira se kako to tumačiti (Yannoulidis 2012). U nekim pravnim tradicijama to znači da osoba zna da je postupak bio nezakonit jer je svjesna da postoje zakoni koji to zabranjuju. Razmotrimo na primjer osobu koja ubija drugu osobu jer misli da je to nezakonito, ali je morala to učiniti jer joj je Bog to naredio. Prema ovom tumačenju, ta osoba bila bi ubrojiva. Međutim, prema drugoj tradiciji osoba treba imati prikladno moralno razumijevanje koje na neki način opravdava postojanje zakona prema kojemu je određeni postupak zabranjen. U ovom slučaju optuženik u gore navedenom primjeru mogao bi se smatrati neubrojivim jer izgleda da nije sposoban razumjeti moralnu motivaciju zakona protiv ubojstva. Slično tome, postoje različita tumačenja toga što znači da netko nije sposoban vladati svojom voljom ili da ima deficite samokontrole. Nedavna povijest razvoja pojma krivične neubrojivosti u SAD-u odražava postojanje ovih različitih tumačenja (Meynen 2016, str. 22–32).

Filozofi mogu sudjelovati pri izradi pravnih dimenzija relevantnog okvira. Povijesni i suvremeni filozofski pokušaji karakterizacije moralne ili pravne odgovornosti mogu obogatiti pravne rasprave koje se odnose na ove pojmove (Hart 2008). Međutim, kako bi filozofski doprinos bio koristan, filozofsko pojašnjenje neubrojivosti mora ostati u doticaju s pravnom praksom na nekoliko načina. Na primjer, filozofska analiza treba se oblikovati na temelju pučko-psihološkog pojma djelovanja, treba izbjegavati sporne metafizičke pretpostavke o osnovama slobodnog djelovanja i ne uključivati psihološke sposobnosti koje nadmašuju kapacitete većine statistički prosječnih ljudi (Malatesti i Jurjako 2016, str. 89–90). Filozofski pojam praktične racionalnosti, na primjer, bio je koristan za usavršavanje karakterizacije psiholoških nesposobnosti potrebnih za određivanje pojma moralne ili krivične neubrojivosti. Filozofi i pravni stručnjaci tvrde da su racionalne psihološke sposobnosti barem nužne za kognitivnu i voljnu sposobnost potrebnu za kaznenu odgovornost (Glannon 2011; Morse 2008). Stephen Morse izražava to gledište kako slijedi:

“Pravno odgovorni djelatnici jesu stoga ljudi koji imaju općenitu sposobnost shvatiti i voditi se dobrim razlogom u određenom pravnom kontekstu. Moraju biti sposobni za racionalno praktično zaključivanje.” (Morse 2000, str. 253)⁴

Međutim, takvo pojašnjenje nije dovoljno za okvir koji bi omogućio povezivanje prava i znanosti. Nesposobnosti koje se nalaze u podlozi neubrojivosti trebaju se karakterizirati detaljno, čak i kada su jednoznačno karakterizirane unutar određene pravne tradicije ili se trebaju razjasniti u okviru nerevizionističke ili zakonski prihvatljive filozofske analize. Zapravo, zakon razmatra sposobnost odgovornosti u smislu razumijevanja prirode djela ili upravljanja ponašanjem. Čak i kada su ti pojmovi pročišćeni, na primjer kao racionalne sposobnosti, nije jasno kako se neuroznanstvena dokazna građa može koristiti za određivanje jesu li te sposobnosti prisutne u određenim slučajevima kada se dogodi neki zločin ili nisu.

Primamljiv filozofski potez bio bi oslanjanje na uvid koji dolazi iz rasprave o problemu uma i tijela. Napokon, pojmovi poput razumijevanja

⁴ “Legally responsible agents are therefore people who have the general capacity to grasp and be guided by good reason in particular legal contexts. They must be capable of rational practical reasoning.” (Morse 2000, str. 253).

relevantnih razloga, djelovanje koje se temelji na tom razumijevanju i tako dalje pripadaju psihološkom rječniku koji se odnosi na mentalna stanja ili procese. S druge strane, neuroznanost se odnosi na stanja i procese u mozgu. Dakle, moglo bi se pretpostaviti da bi preliminarno rješenje problema uma i tijela, tj. neko gledište na opći odnos između mentalnih sposobnosti i neuronskih procesa moglo utemeljiti prijenos podataka iz neuroznosti u područje primjene obrane neubrojivosti. U ovom bi slučaju filozofska zadaća bila odlučiti koju bi poziciju u pogledu odnosa uma i tijela, poput glavnih stajališta kao što su redukcionizam, dualizam, eliminativizam itd. trebalo prihvatiti (Pečnjak i Janović 2016). Ovom iskušenju treba se oduprijeti u onoj mjeri u kojoj takvo gledište doprinosi promicanju diskontinuiranog pristupa u odnosu s trenutnom pravnom i znanstvenom praksom. Metafizička rasprava o odnosu uma i tijela u filozofiji uma vrlo je komplicirana i njezino konačno razrješenje nije na vidiku. Štoviše, razrada i obrana detalja takvih metafizičkih gledišta u konačnici nije relevantna za rješavanje praktičkog problema odnosa zakona i neuroznosti.

Potrebni filozofski doprinos odnosi se na analizu prirode sposobnosti koje su relevantne za atribuciju odgovornosti i podspobnosti koje su nužne za njihovo posjedovanje. To bi omogućilo određivanje različitih tipova eksperimentalnih paradigmi, a time i podataka relevantnih za utvrđivanje prisutnosti ili odsutnosti tih sposobnosti. Stoga bi poseban fokus trebao biti na operacionalizaciji središnjih sposobnosti koje su relevantne za pripisivanje ubrojivosti. Drugo, ovo bi se uglavnom trebalo povezati s eksperimentalnim paradigmama koje se odnose na istraživanje ponašanja i relevantnih kognitivnih ili neuronskih procesa (Buckholtz, Reyna i Slobogin 2016). Na taj način, ti podaci mogli bi predstavljati dokaznu građu da neki nužni uvjeti za postojanje relevantnih sposobnosti nisu ispunjeni. Jasno, ova se posljednja pretpostavka oslanja na minimalno materijalističkoj pretpostavci da su svi ljudski psihološki kapaciteti koji se nalaze u podlozi pravne odgovornosti, a time i nesposobnosti koje se odnose na obranu neubrojivosti zbog duševne bolesti, ukorijenjeni u procesima u mozgu i ovise o njima. Da bismo pojasnili kako se ovaj okvir može razviti, koristit ćemo pojam samokontrole koja je relevantna za obranu neubrojivosti zbog mentalnog poremećaja.

Pojmovni okvir trebao bi omogućiti istraživanje načina na koje mentalni poremećaji utječu na sposobnost samokontrole optuženika. To se može ostvariti na različite načine. Jedan se odnosi na pitanje nedostaju li pojedincu ove nesposobnosti ako pripada klasi pojedinaca koji pate od određene vrste poremećaja. Alternativno, istraživanje se može odnositi na određene pojedince da bi se uspostavilo, neovisno o njihovoj pripadnosti klasi pojedinaca s određenom neurokognitivnim profilom, da nema sposobnosti za samokontrolu. U svakom slučaju, relevantno je da je pojam kontrole ekspliciran u terminima koji se mogu povezati s neuroznanstvenim spoznajama o sposobnostima koje se nalaze u podlozi samokontrole (Amaya 2020). U nedavnoj raspravi o ovim stvarima, na primjer Buckholtz i kolege (2016) upozoravaju da trenutni napredak u proučavanju uma, mozga i ponašanja, a pod specifikacijom pravno relevantnog pojma samokontrole, potiče istraživanje odnosa pravnog pojma samokontrole s domenama fleksibilnosti ponašanja, izbora i djelovanja koje istražuje suvremena kognitivna znanost. Ove veze zatim omogućavaju pomoću specifičnih eksperimentalnih zadataka koji mjere određene procese određivanje toga jesu li relevantne sposobnosti prisutne u tim domenama. Na primjer, fleksibilnost ponašanja djelatnika mogla bi se riješiti izučavanjem procesa koji se mogu mjeriti psihološkim zadacima kao što su Flankerov zadatak, Stroopov zadatak, zadatak obrnutog učenja itd.

Slično tome, Hirstein i kolege (Hirstein, Sifferd i Fagan 2018) tvrde da bi sposobnosti ili nesposobnosti relevantne u sadašnjim zakonskim formulacijama za izuzimanje od krivnje trebale biti povezane s izvršnim funkcijama. Općenito govoreći, izvršne funkcije predstavljaju skup kognitivnih sposobnosti kao što su radna memorija, pažnja, inhibicija djelovanja itd. koje suvremena kognitivna znanost smatra nužnim za kontrolu ponašanja i koje na više načina mogu biti oštećene različitim vrstama mentalnih poremećaja (Wood, Allen i Pantelis 2010).

Ne ulazeći u raspravu o detaljima ovih prijedloga, oni dijele neke metodološke aspekte koji su relevantni za našu raspravu. Povezivanje pravnog pojma samokontrole s domenama, zadacima i procesima koje proučava suvremena neuroznanost također zahtijeva otkrivanje ili uspostavljanje konceptualnih veza između pravne i znanstvene domene. Jasno, u odnosu na ove domene filozofski pokušaj definiranja sposobnosti i pod sposobnosti potrebnih za samokontrolu može igrati važnu

ulogu (Bermúdez 2018). Međutim, osim što je značajno povezano sa znanstvenim istraživanjima usmjerenim na utvrđivanje prisutnosti ili odsutnosti tih sposobnosti, njihovu važnost za pravo, kako sugeriraju Buckholtz i kolege (2016, str. 27), trebaju vrednovati i pravni stručnjaci.

4. Zaključak

Kako bismo rekapitulirali, ponudili smo dva primjera čija je svrha pokazati kako se znanost i druge ljudske prakse poput prava mogu suočiti s akutnim konceptualnim pitanjima. Štoviše, pokazali smo da filozofi mogu doprinijeti jasnoj formulaciji tih pitanja, ali i biti uključeni u njihovo rješenje. Neke metafilozofske pouke mogu se izvući iz ova dva slučaja. *Prvo*, u određenim kontekstima filozofi bi trebali pokušati "ostati blizu" problema, fokusirajući se na razvoj empirijskog znanja i poboljšanju društvenih praksi. *Drugo*, trebali bi pažljivo njegovati osjećaj relevantnosti svojih doprinosa, cijeneći nekoliko pragmatičnih ograničenja. Neka proizlaze iz problema o kojima je riječ dok neka iz spremnosti sudjelovanja u plodonosnim interdisciplinarnim istraživanjima, kako to ilustriraju primjeri koje smo detaljnije razmotrili u ovom poglavlju.

Zahvale

Želimo zahvaliti Martini Blečić, Zdenki Brzović, Marku Jurjaku i Neli Malatesti na čitanju i komentiranju rukopisa.

Istraživanje za ovaj rad omogućeno je financijskom potporom projekata Hrvatske zaklade za znanost (HRZZ): (LM) *Odgovori na anti-socijalne osobnosti u demokratskom društvu* (RAD) (šifra Istraživačkog projekta: IP-2018-01-3518) i (PŠ) *Teorijske pretpostavke molekularne biologije* (ThUMB) (šifra Istraživačkog projekta: IP-2018-01-3378) te se ovim putem zahvaljujemo na potpori. (PŠ) Također zahvala sveučilišnoj potpori *Klasifikacije u biologiji i medicini* (KUBIM) (šifra Potpore: uniri-human-18-265, Sveučilište u Rijeci).

Literatura

- Amaya, Santiago. 2020. "The Science of Self-Control". *John Templeton Foundation – Research report*. URL: <https://www.templeton.org/wp-content/uploads/2020/08/JTF-Self-Control-Final.pdf> (preuzeto: 28.01.2021.)
- Bermúdez, José Luis. 2018. *Self-Control, Decision Theory, and Rationality: New Essays*. <https://doi.org/10.1017/9781108329170>.
- Brenner, Sydney. 1999. "Theoretical Biology in the Third Millennium". *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences* 354 (1392): 1963–65. <https://doi.org/10.1098/rstb.1999.0535>.
- Buckholtz, Joshua W, Valerie Reyna, i Christopher Slobogin. 2016. "A Neuro-Legal Lingua Franca: Bridging Law and Neuroscience on the Issue of Self-Control". *Mental Health Law & Policy* 32 (16): 1–24.
- Clausen, Jens, i Neil Levy, ur. 2015. *Handbook of Neuroethics*. Dordrecht Heidelberg New York: Springer Reference.
- Crick, F. H. 1958. "On Protein Synthesis". *Symposia of the Society for Experimental Biology* 12: 138–63.
- Danaher, John. 2018. "Brain-Based Lie Detection and the Mereological Fallacy: Reasons for Optimism". U *Neurolaw and Responsibility for Action*, uredio Bebhinn Donnelly-Lazarov, 160–81. Cambridge University Press. <https://doi.org/10.1017/9781108553339.008>.
- De Ridder, Dirk, Jan Verplaetse, i Sven Vanneste. 2013. "The Predictive Brain and the 'Free Will' Illusion". *Frontiers in Psychology* 4. <https://doi.org/10.3389/fpsyg.2013.00131>.
- Divan, Aysha, and Janice A. Royds. 2016. *Molecular Biology: A Very Short Introduction*. First edition. Oxford: Oxford University Press.
- Downes, Stephen. 2006. "Biological Information". U *The Philosophy of Science: An Encyclopedia*. Vol. 1, A-M. New York: Routledge/Taylor and Francis Group: 64–8.
- Dubljević, Veljko. 2019. *Neuroethics, Justice and Autonomy: Public Reason in the Cognitive Enhancement Debate*. Cham: Springer International Publishing. <https://doi.org/10.1007/978-3-030-13643-7>.
- Ginther, Matthew R., Richard J. Bonnie, Morris B. Hoffman, Francis X. Shen, Kenneth W. Simons, Owen D. Jones, i René Marois. 2016. "Parsing the Behavioral and Brain Mechanisms of Third-Party Punishment". *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience* 36 (36): 9420–34. <https://doi.org/10.1523/JNEUROSCI.4499-15.2016>.
- Glannon, Walter. 2011. *Brain, Body, and Mind: Neuroethics with a Human Face*. Oxford: Oxford Univ. Press.

- Godfrey-Smith, Peter and Kim Sterelny. 2016. "Biological Information". U *The Stanford Encyclopedia of Philosophy*. Summer 2016 Edition. Edward N. Zalta (ed.). URL <https://plato.stanford.edu/archives/sum2016/entries/information-biological/>
- Grozđanić, Velinka, Marissabell Škorić, i Igor Martinović. 2013. *Kazneno Pravo: Opći Dio*. Rijeka: Pravni fakultet Sveučilišta u Rijeci. <https://www.bib.irb.hr/664008>.
- Hart, H. L. A. 2008. *Punishment and Responsibility: Essays in the Philosophy of Law*. 2nd Revised edition. Oxford i New York: Oxford University Press.
- Hirstein, William, Katrina L. Sifferd, i Tyler K. Fagan. 2018. *Responsible Brains: Neuroscience, Law, and Human Culpability*. Cambridge, Massachusetts London, England: The MIT Press.
- Levy, Arnon. 2011. "Information in Biology: A Fictionalist Account". *Nous* 45 (4): 640–657.
- Malatesti, Luca, Marko Jurjako, i Gerben Meynen. 2020. "The Insanity Defence without Mental Illness? Some Considerations". *International Journal of Law and Psychiatry* 71: 101571. <https://doi.org/10.1016/j.ijlp.2020.101571>.
- Malatesti, Luca, i Marko Jurjako. 2016. "Metaphilosophy in Practice: The Responsibility of Psychopathic Offenders as a Case Study". *Anthropology & Philosophy* 12: 85–100.
- Maynard Smith, John, i Eörs Szathmáry 1999. *The Major Transitions in Evolution*. Reprinted. Oxford: Oxford University Press.
- Mele, Alfred R. 2014. *Free: Why Science Hasn't Disproved Free Will*. New York: Oxford University Press.
- Meynen, Gerben. 2016. *Legal Insanity: Explorations in Psychiatry, Law, and Ethics*. Cham: Springer International Publishing. <https://doi.org/10.1007/978-3-319-44721-6>.
- Morse, Stephen J. 2008. "Psychopathy and Criminal Responsibility". *Neuroethics* 1 (3): 205–12. <https://doi.org/10.1007/s12152-008-9021-9>.
- Okasha. 2019. *Philosophy of Biology: A Very Short Introduction*. New York: Oxford University Press.
- Overgaard, Søren, Paul Gilbert, i Stephen Burwood. 2013. *An Introduction to Metaphilosophy*. Cambridge: Cambridge University Press. <https://doi.org/10.1017/CBO9781139018043>.
- Pardo, Michael S., i Dennis M. Patterson. 2013. *Minds, Brains, and Law: The Conceptual Foundations of Law and Neuroscience*. New York, NY: Oxford University Press.
- Pećnjak, Davor, i Tomislav Janović. 2016. *Prema dualizmu: Ogledi iz filozofije uma*. Zagreb: Ibis grafika.

- Quine, W. V. O. 2013. *Word and Object*. New edition, First edition 1960. Cambridge, Mass.: MIT Press.
- Ross, Lauren N. 2021. "Causal Concepts in Biology: How Pathways Differ from Mechanisms and Why It Matters". *The British Journal for the Philosophy of Science*. 72 (1): 131–58. <https://doi.org/10.1093/bjps/axy078>.
- Sarkar, Sahotra. 2000. "Information in Genetics and Developmental Biology: Comments on Maynard Smith". *Philosophy of Science* 67 (2): 208–13. <https://doi.org/10.1086/392771>.
- Singh, Ilina, Walter P. Sinnott-Armstrong, i Julian Savulescu, ur. 2014. *Bioprediction, Biomarkers, and Bad Behavior*. New York: Oxford University Press.
- Šustar, Predrag. 2007. "Crick's Notion of Genetic Information and the 'Central Dogma' of Molecular Biology". *The British Journal for the Philosophy of Science* 58 (1): 13–24. <https://doi.org/10.1093/bjps/axl018>.
- Vincent, Nicole A., ur. 2013. *Neuroscience and Legal Responsibility*. Oxford Series in Neuroscience, Law, and Philosophy. New York, NY: Oxford University Press.
- Waters, C. Kenneth. 2007. "Causes That Make a Difference". *Journal of Philosophy* 104 (11): 551–79. <https://doi.org/10.5840/jphil2007104111>.
- Williams, George Christopher. 1992. *Natural Selection: Domains, Levels, and Challenges*. Oxford Series in Ecology and Evolution 4. New York, NY: Oxford University Press.
- Williamson, Timothy. 2010. *The Philosophy of Philosophy*. The Blackwell/Brown Lectures in Philosophy. Malden, MA: Blackwell.
- Williamson, Timothy. 2018. *Doing Philosophy: From Common Curiosity to Logical Reasoning*. Oxford: Oxford University Press.
- Wiseman, Harris. 2016. *The Myth of the Moral Brain: The Limits of Moral Enhancement*. Basic Bioethics. Cambridge, Massachusetts London, England: MIT Press.
- Wood, Stephen J, Nicholas B. Allen, i Christos Pantelis. 2010. *The Neuropsychology of Mental Illness*. Cambridge: Cambridge University Press.
- Yannoulidis, Steven. 2012. *Mental State Defences in Criminal Law*. Farnham: Ashgate.

On the concept of design as a very general idea in philosophy of science

Tonći Kokić

I. SCHEME

An article of this size can study only a thin layer of the topic of scientific legitimacy of design in general, and concerning OoL (Origin of Life, the shorthand 'OoL' will be used throughout the paper) in particular. This article is an abstract (meta)theoretical study which does not deal with the details of the design problem nor with the current research in the area of OoL. Instead, it is concerned with the concept of design and with general metaphysical (and epistemological) beliefs about the nature of the world, as well as about OoL, which are not in the foreground. These metaphysical beliefs are then transformed into the methodological commitments for any scientific inquiry, complementary with classical requests for scientific theories or hypotheses – explanation and theoretical entrenchment of some kind and degree of predictability and testability – all of which is achieved through empirical observation and experiment. To clarify this, the second section of the essay argues that design is loosely defined, if at all, and it presents a few attempts at its definition, connected with observation of processes that run counter to natural ones, as well as possible ways of recognizing design activity in nature. Because the concept of design is banned, or at least under suspicion as being scientifically illegitimate, the third section discusses two analogies, thereby hoping to comprehend an obvious or camouflaged metaphysical stratum of thought which is frequently presented as a pure scientific matter of fact. These analogies attempt to clarify the scientific legitimacy of other logically (and nomically) possible explanations. The fourth and the fifth sections investigate the logical permissibility of metaphysical assumptions (in the fourth) and epistemological assumptions

of designer activity (in the fifth). Based on this investigation, the sixth section conjectures that design cannot be excluded as a source of first life on both metaphysical and epistemological grounds and it proposes that *the concept of design could have scientific legitimacy as a very general idea inside a conceptual hierarchy of science*, similar to many other metaphysical concepts, such as uniformity, predictability, scientific laws and the intelligibility of the universe. Finally, this makes it reasonable to conclude that, under some interpretations of philosophy of science, the scientific legitimacy of design could be reconsidered.

II. ON THE CONCEPT OF DESIGN

The concept of design generally, and in the case of OoL research, is not a serious option in scientific explanations; more precisely, the concept of supernatural design is entirely prohibited. To put it mildly, non-intentional explanations are to be preferred before all others. This default metaphysical position seriously emerged in the appearance of Darwin's theory of evolution by which the concept of design was removed as needless, because there is ... "no need for the God hypothesis".¹ This assumption was later transformed in the widespread and loose definition of naturalism according to which reality is exhausted by nature and does not contain anything 'supernatural'. Naturalism becomes a fundamental metaphysical (ontological) and epistemological position which serves as a signpost intended to be a methodological obligation for all further scientific reasoning. Recently, design has been viewed exclusively from the perspective of the Intelligent Design (ID) movement. Philosophical dictionaries and encyclopaedias do not have a heading on the concept of design but direct towards Argument from Design,² Teleological Arguments for God's Existence³ and Design Arguments for the Existence

¹ Francisco Ayala and Robert Arp, "Is There a Place for Intelligent Design in the Philosophy of Biology?", in: Francisco Ayala and Robert Arp (ed.), *Contemporary Debates in Philosophy of Biology*, Blackwell Publishing Ltd., Malden, MA, 2009, p. 338 [337–340].

² Robert Audi (ed.), *The Cambridge Dictionary of Philosophy*, Cambridge University Press, Cambridge 1999, p. 227.

³ Edward Zalta (ed.), *Stanford Encyclopedia of Philosophy*, Stanford, The Metaphysics

of God and Intelligent Design.⁴ Elsewhere in the literature it is possible to find more design definitions focusing on the logical relation of contraries: Dembski's definition relies on the negation of regularity and chance,⁵ Behe defines (intelligent) design by irreducible complexity at the molecular level,⁶ and Davies sees design as a product of some kind of 'cosmic computer'.⁷ On the contrary, there is an opinion that design (in biology) means "first, that organisms evince to have been designed; ...second, that only God could account for the design".⁸ Ayala also thinks that Intelligent Design is more of a political movement than a 'theory' or hypothesis which can be subjected "to critical examination and empirical testing" and which avoids references to God "so that the 'theory' of ID can be taught in the public schools as an alternative to the theory of evolution".⁹ So, Ayala rejects design as a legitimate part of science. A trivial definition of design determines design as a noticeable and easily understandable action of a supernatural designer by intervention which opposes natural processes. As well, there is a definition of design as an intentionally produced order or set of patterns, which are mind correlative and which must include some kind of counterflow.¹⁰ Counterflow means that observed things are "running contrary to what, in the relevant sense, *would* (or *might*) have resulted or occurred *had* nature operated freely".¹¹ In this case, the designer does not have to be

Research Lab Center for the Study of Language and Information Stanford University, 2017, <https://plato.stanford.edu/> (24.3.2020.).

⁴ Dowden Bradley (ed.), *Internet Encyclopedia of Philosophy*, ISSN 2161-0002, <https://www.iep.utm.edu/> (24.3.2020).

⁵ William Dembski, *The Design Inference: Eliminating Chance through Small Probabilities*, Cambridge University Press, New York, 1998, p. 36.

⁶ Michael Behe, *Darwin's Black Box*, Free Press, New York, 1996, p. 257.

⁷ Paul Davies, *The Mind of God. The Scientific Basis for a Rational World*, Simon & Shuster Paperbacks, New York, 1996, p. 97, 123, 124.

⁸ Francisco Ayala, "There is No Place for Intelligent Design in the Philosophy of Biology: Intelligent Design Is Not Science.", in Francisco Ayala & Robert Arp (ed.), *Contemporary Debates in Philosophy of Biology*, Blackwell Publishing Ltd., Malden, MA, 2009, p. 364 [364–391].

⁹ Ayala, "There is No Place for Intelligent Design in the Philosophy of Biology"...p. 375, 365.

¹⁰ Del Ratzsch, *Nature, Design and Science. The status of Design in Natural Science*. SUNY Series in Philosophy of Biology, New York, 2001, p. 4, 5.

¹¹ Ratzsch, *Nature, Design and Science*...p. 5.

supernatural – although that option is not excluded, it does not have to act opposite to nature but in ways nature is seldom working, its actions can be part of natural laws or in acts that preceded natural laws or is a fuzzy quantum processes which cannot be explained by known laws. Also, its action could be epistemologically inaccessible to humanity. All these different definitions of design would probably agree with one fact: proof of design or an intentional agent should indicate that at least some marks of observed phenomena are not natural: repetitious macroscopic patterns,¹² symmetry, geometric characteristics and other types of patterns in mesocosmos¹³ and/or the presence of some type of counterflow. There are also ‘secondary’ marks of agent activity and counterflow which are “typically appealed to in traditional design arguments – complicated development, complex structures, coordination of components, adjustment of means to ends, interlocking functions, extreme improbability, purpose like behaviours etc.”¹⁴ Also, in order for a scientific theory or hypothesis to be considered scientifically legitimate, and to be one that provides evidence of intentional doing and/or counterflow, it is necessary to complete standard requirements of “(a) testability and falsifiability through empirical observation and experiment, as well as (b) results, predictions, and explanations”¹⁵

III. TWO ANALOGIES

There is a strong aversion to considering design as legitimate scientific concept, based on the belief that design could not satisfy the standard requests for scientific theories or hypotheses. The two analogies discussed below, representing two broad metaphysical positions (one which allow only mind-independent world, and one which does not restrict world on that way), may be helpful in revealing the metaphysical (or methodological) stratum of thought instead of a pure scientific matter of fact.

¹² Jacques Monod, *Chance and Necessity: An Essay on the Natural Philosophy of Modern Biology*, Vintage Books, New York, 1972, p. 11, 81, 120.

¹³ Ratzsch, *Nature, Design and Science*...p. 12.

¹⁴ Ratzsch, *Nature, Design and Science*...p. 12.

¹⁵ Ayala and Arp, “Is There a Place for Intelligent Design in the Philosophy of Biology?”...p. 339.

Also, these analogies could elucidate the scientific legitimacy of other logically (and nomically) possible explanations and compare their status with regard to the abovementioned requests for scientific validity.

1.1 Space Odyssey Monoliths

There are different definitions and interpretations of black monoliths that appear in the book and movie *2001: A Space Odyssey* and other novels and movies from the same series.¹⁶ Putting aside complexity of meaning and numerous interpretations which are not significant here, it is possible to say that extremely flat solid black monoliths observed in a three dimensional Euclidean space have the shape of a rectangular cuboid whose sides extend in a ratio of 1x4x9 (and probably the ratio of proportions continued in a similar ratio in additional dimensions up to an unknown size to humans). Monoliths are as large as necessary. The human observer has limited knowledge of the composition, internal structure and functions of monoliths because they are almost indestructible and impenetrable,¹⁷ resisting deeper analyses. But, at least for the monolith TMA-2,¹⁸ there is evidence for mass and density (net greater than of air), with the possibility to teleport and replicate itself, as well as the possibility of its copies to assemble themselves into a bigger structure capable of performing tasks not achievable by a single unit (emergent properties). The monoliths are directed by some kind of internal program and the consciousness of human and artificial intelligence could be incorporated as computer programs into its software. Monoliths communicate within some sort of hierarchical organization of lower and higher monoliths in accordance to an authorization for decision making. This interaction has some characteristics of 'sociability' and work division. Among some monoliths it is possible to notice magnetic fields (for example, TMA-1), the broadcasting of radio waves

¹⁶ 2010: *Odyssey Two* (novel and movie), 2061: *Odyssey Three* (novel), 3001: *The Final Odyssey* (novel).

¹⁷ Except the structurally destructible TMA-4, and the functionally destructible monolith TMA-2.

¹⁸ The abbreviation TMA-2 stands for Tycho Magnetic Anomaly and the serial number of observed monolith.

and an autonomous supply of electric power from a huge energy-storage system of cosmic superconductors. At least the function of one monolith (TMA-0) appears to be conjoined with the increase in speed of human evolution, but without any proved causal relation. In the series *A Space Odyssey* human observers conclude:

- (i) Monoliths are artificially produced by operation of intentional (finite) rational agent or agency.

Explanation of (i): although the composition, structure and operation of monoliths are known only superficially – flat surface, geometric shape and ratio of the sides (1x4x9) which is the squares of the first three positive integers, they indicate intentionally produced patterns which correlate to rationality, as well as a functional bias to purpose. This solution includes some kind of counterflow. Counterflow should be part of the explanation of the origin of monoliths because nature does not make cuboids (especially not with that side-ratio) and the monolith's functions make it difficult not to believe that we are dealing with some kind of intentional agency.

Still, some other logically (and nomically) possible inferences and explanations are available:

- (ii) Monoliths are naturally produced by chance process.

Explanation of (ii): The degree of probability of the chance origin of the monoliths with their observed structure and functions seems to be too low to be treated as a serious option¹⁹ – it is nearly impossible – virtually zero, yet, because of this, this option cannot be logically (and nomically) excluded. In fact, in OoL research such highly improbable events could be commonplace.²⁰ Counterflow is not part of the explanation. The structure and functions of monoliths look purposeful, but that is only an epiphenomenon of underlying physical components, processes and natural laws.

¹⁹ As Hoyle's well-known chance argument which challenges the probability of abiogenesis by comparing the chance that a tornado sweeping through a junkyard might assemble a Boeing 747 plane.

²⁰ Christian De Duve, *Vital Dust*, Basic Books, New York, 1995., p. 112; Robert White, "Does Origins of Life Research Rest on a Mistake", *Noûs*, 2007, 4, [453–477].

- (iii) Monoliths are naturally produced by necessary and inevitable determined processes.

Explanation of (iii): This option assumes that monoliths appear as an inevitable or at least highly expected result of certain prior physical and/or some other initial conditions and laws somewhere in the 'past'. So to speak, the monoliths are the outcomes of some brute physical constraints and conditions; purposeful intentional activity or happy accident are excluded.

- (iv) Monoliths are naturally produced by a long-lasting process of millions of small intermediate steps governed by some kind of evolutionary principle, but for sure they are the natural consequence of a specific physical (and chemical) organization of matter and natural laws.

Explanation of (iv): This explanation considers the low probability of the sudden appearance of the monoliths' structure and functions and excludes design as explanation, and considers the occurrence of long-term natural processes consisting of a large number of small steps in which available material was moulded by existing natural laws and principles as a more probable explanation. All of this has happened as the natural consequence of a specific physical self-organization of matter – for instance crystals, by some kind of evolution of physical components, and laws which have finally shaped matter into the monolith form. As in the two above explanations, counterflow or purposeful activity are not observed and the final product is the pure consequence of natural forces.

All explanations (i) – (iv) acknowledge some semblance of purposeful activity in the monoliths' functions and presumed structure. But, while hypothesis (i) claims obvious counterflow (nature was somehow disturbed or projected) and purposeful activity as a mechanism that enables the monoliths' machines to maintain their systems and perform certain functions and tasks, as something consciously planned by an intentional (finite) agent, so far hypotheses (ii) – (iv) do not accept counterflow and reject purposeful activity as the deliberate anticipation of some rational agent – a purely natural mechanism must be responsible for the semblance of 'goal-directed' processes in the monoliths. All explanations (i) – (iv) equally fail to fulfil previously mentioned re-

quests for testability and falsifiability through empirical observation and experiment,²¹ so, we cannot decide based on that.

1.2 The Astronauts

Let's imagine for a moment that the monoliths are a lure intended for the investigation of desirable objects. In this way Dr. Heywood R. Floyd and his associates in Moon base Clavius are investigated: data shows they are 'machines' made of just a few basic chemical components (CHON),²² but their energetic mechanisms differ from the energy flow in the rest of universe, specifically related to heat work and energy, meaning that the system of these machines is not in thermal equilibrium, which allows for mechanical work. This means that the asymmetric process of heat transfer, which cannot be converted entirely into work, degrades toward an equilibrium state in which, finally, no energy is available to do 'useful' work. This situation is named entropy or the degree of disorder of the system. These machines, similarly to monoliths, take negative entropy²³ or free energy from the environment for their work and order, while increasing disorder in the universe. Furthermore, in the chemical sense these are highly integrated cybernetic systems with the complex 'command' management and control of chemical activities, with the ability to create stable, diverse, large and complex molecules. It appears that these machines, in the structural sense, have a hierarchical order in which every higher level is based on all the intact lower levels and each higher level provides some new properties not found on any lower level.²⁴ Physical constants, components and laws of nature are necessary for the origin of life, maybe even fine-tuned for it,²⁵ but

²¹ Ayala and Arp, "Is There a Place for Intelligent Design in the Philosophy of Biology?"...p. 339.

²² 99.9% of the biochemical content of living organisms includes just four elements: carbon, hydrogen, oxygen and nitrogen (C, H, O i N), and in total just 25 chemical elements.

²³ Erwin Schrödinger, *What is Life. The Physical Aspect of the Living Cell*, Cambridge University Press, Cambridge, 2013. The book was first time published in 1944.

²⁴ The following levels are noted: subatomic, atomic, molecule, macromolecule, organelle, cell, tissue, organ, organ system and organism, but some other higher and unknown levels of organization are possible.

²⁵ The list of fine-tuned parameters is extensive: from cosmic constants, initial con-

are still not sufficient to explain it. A large number of these machine parts are in an enormously complex relationship between substance and energy modification (metabolic and catabolic reactions), which is not spontaneous and involves things like *homochirality*, whose origin is mysterious, or at least unknown.²⁶ Considering replication, investigated machines contain something like 100 trillion cells per unit, with a total of three billion bases in (DNA) molecules. Micro machines found inside them (*E. coli*) can replicate their entire genome of roughly 5 million bp in time approximating to 40 minutes. The requested clue here was to find some fundamental physical law or principle in the foundations of gene replication or gene expressions, as well as the origin of coded information for the machinery of protein production. The knowledge of the underlying physical principle of gene replication and the origin of this coded system remains unavailable. These machines have a 'control unit' made of one hundred billion²⁷ components (neurons), where each neuron can make contact with thousands or even tens of thousands of other neurons, via synapses, and they make a million new connections during every second of its function. And finally, these machines possess subjective consciousness – *qualia*, language, they can think in general concepts and grasp meaning. Other numerous important properties could not be listed here.

Hypothesis and explanation of the origin of these 'machines', as made by the monoliths, are as follows:

- (v) These machines are produced by deliberate rational agent activity or a designer.

Explanation of (v): Conditions favourable to the (life) origin of these machines on Earth are registered, they are necessary but not sufficient for 'building' it, as well. No laws of nature that directed or instructed matter to produce these machines are observed, therefore it is reasonable to conclude that nature was disturbed at least to a degree of nomic

ditions and local planetary conditions to their fundamental effects for the origin of life as the polarity of the molecule of water.

²⁶ David Deamer, *First Life. Discovering the Connections Between Stars, Cells and How Life Began*, University of California Press, Berkeley, 2011, p. 79.

²⁷ One followed by 11 zeros.

discontinuity – or even less – to a level where the origin of this machine just circumvented the laws of nature. So, counterflow could be implied. The decisive element in this inference is the extraordinary complexity of parts and their interconnections, as well as adapting the organization of means to an end unrecorded outside of this kind of machines. All of this correlates to rationality.

However, some other inferences and explanations are logically (and nomically) available:

(vi) Machines are naturally produced by chance process.

Explanation of (vi): The degree of probability of the chance origin of the amazing complexity of the arrangements which are needed to produce these machines is close to zero – virtually zero. The proofs show that these machines originated very quickly, in some kind of burst mode, immediately upon the creation of the conditions for it on Earth.²⁸ It is too short a time to do the job of assembling functionally and structurally complex, as well as mutually dependent and interlocked, parts. This obstacle could be hardly resolved without the introduction of some kind of counterflow as a trigger which puts ‘system at work’.

(vii) Machines are produced naturally by a necessary and inevitably determined process.

Explanation of (vii): According to this hypothesis the appearance of machines is inevitable and highly expected as the result of certain prior physical conditions, some other initial-primordial conditions and deep underlying laws and constants. Machines are the outcomes of some brute physical and/or chemical constraints. The appearance of machines is expected, or at least not surprising. Counterflow could be part of the explanation and purposeful intentional activity cannot be excluded, albeit both are not directly accessible to the observer.

²⁸ Iris Fry, *The Emergence of Life on Earth: a Historical and Scientific Overview*, Rutgers University Press, New Brunswick, 2000, p. 125.

James Griesemer, “Origins of Life Studies”, in: Michael Ruse (ed.), *The Oxford Handbook of Philosophy of Biology*, Oxford University Press, New York, 2008, p. 271 [263–290].

- (viii) Machines are the result of a long-lasting process of millions of small intermediate steps governed by some kind of self-assembling principle, they are the natural consequence of a specific physical organization of matter and natural laws. Anyway, agent activity could not be excluded because of the impossibility of direct access to evidence.

Explanation of (viii): This hypothesis presupposes that the origin of these machines more likely was a long-term process consisting of a large number of small steps in which available material was moulded by natural laws and principles – or a mechanism which still has to be functional now.²⁹ Counterflow or purposeful activity is not obvious, but the impact of some kind of agency could not be excluded from the area of initial conditions and processes – nomic discontinuity is at least possible, there exists a process on a deeper level which does not include any kind of even mild nomic violations.

Only hypothesis and explanation (v) explicitly includes counterflow and/or purposeful activity as part of its explanation. On the other hand, hypotheses (vi)–(viii) do not include but neither do they reject counterflow and they do not exclude the possibility of the purposeful activity of an intentional agent as a designer of these machines. Either way, similarly to the previous analogy, explanations (v)–(viii) equally fail to fulfil the requirement of testability and falsifiability through empirical observation and experiment, which prevents us from deciding based on this kind of postulate.

IV. ON METAPHYSICS

Human observers conclude that the monoliths were built by an unseen and unknown extra-terrestrial species, but still, their conclusion features the credentials of scientific legitimacy: it does not matter that information about the structure and functions of the monoliths are strictly limited and a designer is unobserved. An unobserved designer is not a cause for a proclamation of the scientific illegitimacy of the design

²⁹ David Penny, “An Interpretative Review of the Origin of Life Research”, *Biology and Philosophy*, 2005, 20, p. 640 [633–671].

hypothesis, even scientific realism does not reject both unobservable and unobservable *per se* phenomena as unscientific. The problem could be found in some interpretation of scientific realism as “true or approximately true descriptions of observable and unobservable aspects of a mind-independent world”³⁰ because there is no proof of the monoliths’ nature – mind correlate. Likewise, it is not problematic if a conclusion is not based on empirical data from the origin of monoliths. Further, it could be mentioned that the scientific legitimacy of theories is not just derived from empirical data: Popper mentioned the example of astrology or psychoanalysis which relies on a staggering amount of empirical evidence based on observation – horoscopes and biographies. Also, this type of theories needs additional assumptions to explain its phenomena:

... requires revisions and added ad hoc assumptions to be able to explain ... The ad hoc theory (in psychoanalysis, AN): is based on a mixture of theory and private language seasoned with a sprinkling of body wisdom, some vague memories of clinical moments, and perhaps a mixture of other sensations aroused by the ambience of the consulting room ³¹

In our case, an observer’s inference is reached on the basis of a mix of metaphysical assumptions, ad hoc hypotheses and speculations derived from deficient empirical evidence. In fact, it seems the conclusion is reached more on the conviction that hypotheses (ii)–(iv) are unacceptable than on any positive evidence for hypothesis (i). But, those who argue against the scientific legitimacy of the concept of design in philosophy of biology claim, as Ayala does, that “It is not sufficient for a theory to be accepted because some alternative theory has failed”.³² Moreover, there is no security in the failure of hypotheses (ii)–(iv) nor does hypothesis (i) fulfil the necessary conditions of scientific legitimacy: it is not testable, it is not falsifiable by means of empirical observation and experiment, nor does it offer clear predictions. Hypothesis (i) and its explanation are the result of a metaphysical preconception of what might

³⁰ Anjan Chakravartty, “Scientific Realism”, in: Edward N. Zalta (ed.), *The Stanford Encyclopedia of Philosophy* (Summer 2017 Edition), <https://plato.stanford.edu/archives/sum2017/entries/scientific-realism> (24.3.2020.).

³¹ Aviezer Tucker, “Unique events: The Underdetermination of Explanation”, *Erkenntnis*, 1998, 48, p. 67 [59–80].

³² Ayala, “There is No Place for Intelligent Design in the Philosophy of Biology”...p. 375.

be rationally warranted in this case: an observed regularity of patterns and geometry, a uniformity of material, geometric shape (plan surface), a ratio of sides which indicates a correlation to the mind and functions similarly to purposeful activity that does not look like a habitual flow in nature.³³ But, a logically (and nomically) possible origin of such complex entities as monoliths, based on hypotheses (ii)–(iv) which are without counterflow and intentional agent activity, is rejected as an explanation without any (counter) evidence. So, neither hypothesis (i) nor hypotheses (ii)–(iv) can be proved or denied, and acceptance of hypothesis (i) is based on deep metaphysical assumptions about the nature of the universe (including monoliths) and the methodological commitments (methodological materialism or naturalism) we find in science. The flow of the argument is almost the same as the one we find in the so-called argument from design: the observer concludes from the assumption *what nature would not do* to the conclusion of *what the mind would do*. According to this, hypothesis (i) itself has a modest exploratory power (low degree of testability and falsifiability), and the connection between metaphysical and empirical levels is rather weak. Still, human observers undoubtedly believe in the scientific legitimacy of hypothesis (i) and scientific illegitimacy of hypotheses (ii)–(iv).

On the other hand, monoliths receive data showing that machine-humanity has extraordinary structure and functions, and only one of its parts (the cell) has an intricate network of functions which goes beyond any artificial object, and it is possible to say that its brain represents the most complex type of structural-functional entity in the universe. First life was for sure simpler, but nonetheless “the kind of system required for a process of natural selection to get going would also have to be extraordinary intricate and complex”.³⁴ Also, the question is “How can a universe of mindless matter produce beings with intrinsic ends, self-replication capabilities, and ‘coded chemistry’”.³⁵ Just because we still do not know how, we cannot laconically and without any proof credit blind forces with this. In what way does the validity of the logical structure of

³³ Comprehensive list of properties suitable for counterflow identification can be found in Ratzsch (2010).

³⁴ White, “Does Origins of Life Research Rest on a Mistake”...p. 470.

³⁵ Anthony Flew, *There is a God: How the World's Most Notorious Atheist Changed his Mind*, Harper Collins Publishers, New York, 2007, p. 124.

the argument according to which it is possible to reasonably assume that monoliths are a product of design, and that the most complex entities in the universe – humanity and the human brain – stand unshakable in the belief they must be products of unconscious physical and mechanical unidirectional forces, remain unclear? The main reason for accepting this kind of solution is metaphysical assumption: hypothesis (i) deals with natural design and hypothesis (v) with supernatural. That is, it testifies to the scientific legitimacy of the hypothesis of (neo)panspermia according to which life was sent to Earth by an automatically controlled spacecraft, as Francis Crick proposed.³⁶ This hypothesis does not explain the OoL, but it moves it to another remote location, and it hardly can be scientifically legitimate.

The key reason for attributing the scientific legitimacy to hypothesis (i) and scientific illegitimacy to hypothesis (v), both of which assume intentional activity, is an assumption of metaphysical materialism or naturalism as part of *background knowledge* – later transformed into a commitment of methodological materialism or naturalism – regarding the existence of solely natural entities in a mind-independent world. Of course, it can be said that natural science deals with its own realm, the realm of natural objects and forces, and that even if not excluded, supernatural entities must be separated into other area, as NOMA (*nonoverlapping magisteria*) posits.³⁷ NOMA postulates two worlds and two areas of epistemological competence: first, the magisterium of the material, empirical world which recognizes the scientific method as suitable and second, the transcendent ontological magisterium which is not accessible by epistemology of natural science. The NOMA, in principle, does not deny the possible existence of supernatural entities but posits them as epistemologically and methodologically needless in natural science.

There is also possibility of POMA (*partially overlapping magisteria*), a position which assumes the possibility of the permeation of two areas, as proposed by former Head of the Human Genome Project (1994 – 2008) Francis Collins.³⁸ According to this, the concept of design would

³⁶ Francis Crick, *What Mad Pursuit*, Penguin Books, New York, 1988, p. 148.

³⁷ Stephen J. Gould, “Nonoverlapping magisterial”, *Natural History*, 1997, 106, 2, p. 4 [16–22].

³⁸ Francis Collins, *The Language of God: A Scientist Presents Evidence for Belief*, Simon & Schuster UK Ltd., London, 2007.

be acceptable as a metaphysical assumption which is a basis for some concrete scientific theory: field theory was derived from Faraday's belief in God as creator and sustainer of the universe, Maxwell field equations was based on his understanding of interactions among the Trinity.³⁹ Flew holds similar views and lists a few other famous names who believe the laws of nature mirror the Mind of God: Planck, Heisenberg, Schrödinger, Dirac and even Einstein. So to speak, metaphysical assumptions are at least (i)ndirectly a matter of science: quite different models can give very similar results. The well-known, although fairly crude example, compares Ptolemy's and Copernicus' models of the universe: the first is geocentric and the second heliocentric, and while they have several other differences, both models were capable of predicting planetary positions. Although imprecisely,⁴⁰ the two theories did not differ in their predictions regarding the astronomical data. This situation was recognized as the problem of the underdetermination of scientific theory by evidence, claiming that the available evidence may be insufficient to determine what beliefs we should hold appropriate due to that evidence. I think this puts forward the concept of design as at least permissible on the deep metaphysical ground. The next step is checking the situation on the epistemological side of the issue.

V. EPISTEMOLOGY OF THE (SUPER)NATURAL

The epistemological position of hypothesis (i) is straightforward: it is hard to avoid an explanation which includes a touch of counterflow somewhere along the causal chain of events which preceded the origin of the monoliths. Explanatory reasoning in generating hypothesis (i) rests on abduction – inference to the best explanation, as well as justifying it. Although abduction is a kind of inductive argument which is not logically valid, strictly speaking, it seems quite reasonable in this case (as in many other cases we deal with every day). Other explanations

³⁹ Ratzsch, *Nature, Design and Science...*p. 128.

⁴⁰ Copernicus' predictions agreed better with observations of the superior planets and solar eclipses, while Ptolemy's predictions were more accurate for the inferior planets and lunar eclipses. At the end and in total, the Copernicus predictions were superior.

are possible but they do not look rational according to our previous knowledge about the world or according to our best theories about the world: nature does not make cuboids which reproduce themselves and have functions that were described earlier. In either case, hypothesis (i) and hypotheses (ii) – (iv), as the final or ultimate cause of investigated phenomena, belong to the realm of purely natural, because even the presupposed finite designer of the monoliths is a part of nature.

The explanation of hypothesis (v) also included a touch of counterflow somewhere in the causal past of the investigated objects. But it would be hasty to put up hypothesis (v) as a model of an easily observed and knowable direct creation of things by supernatural designer through intervention, contrary to natural and logical laws. This is a pre-theoretical (or naive) assumption because it uncritically magnifies human epistemological capability and minimizes the power of a possible supernatural designer.

Considering the first, the human epistemological subject has limited sensory capacities even when the phenotype is extended through different highly developed technical devices and conceptual inventions or theories. Quantum phenomena are an example of this kind: we cannot prove on the quantum level whether nature produced something on its own or did a designer intervene somewhere inside fuzzy quantum processes or initial conditions prior to them. Also, any knowledge of quantum phenomena necessarily includes interactions between quantum phenomena and the observer. In this way, a possible designer intervention on the quantum level could be scientifically inaccessible in principle. The second example is the random nature of mutation processes as one of two basic principles of evolution, but the concept of randomness just means our lack of knowledge about (possible) real causality in that processes. Darwin wrote about variations due to chance as “wholly incorrect expression ... serves to acknowledge plainly our ignorance of the cause of each particular variation”.⁴¹ Moreover, at least a few cases proved that mutations are not-random or spontaneous.⁴² These cases might belong to the class of so-called *God of the Gaps* cases which do not

⁴¹ Charles Darwin, *The Origins of Species by Means of Natural Selection*. Random House Value Publishing, New York, 1979.

⁴² Barbara E. Wright, “A Biochemical Mechanism for Nonrandom Mutations and Evolution”, *Journal of Bacteriology*, 2000, 182, 2993–3001.

have a good reputation in regards to their scientific legitimacy, which does not mean that it is possible to ignore them nor that they would be problematic for the acceptance of supernatural design as an explanation. An assumption about the natural inability of producing life is not *per se* scientifically illegitimate: this is not just a rejection of the metaphysical commitment in natural explanations for whatever phenomena but is also a fact recognized by science itself: it is not unscientific to claim *in principle* the inability of things to go faster than light or the absurdity of *perpetuum mobile* constructions. The long-term resistance of OoL to the scientific explanations indicate that it could in principle be epistemologically unattainable, situated above the human ability to understand and explain. For instance, unique events could be epistemologically inaccessible or at least problematic: "It seems that at least in the case of singularities ... there are events beyond the purview of science".⁴³ So, a statement about the impossibility of a naturalistic explanation of OoL is not scientifically illegitimate but could be metaphysically unacceptable for someone.

Considering the second, this assumption overlooks the potencies of a supernatural designer and that possible design activity does not have to include the violation of natural laws: it could just set up initial conditions (before laws existed or set up these laws at a primordial stage) or it could use natural processes we do not know or understand in a nomic way (it could operate in a probabilistic mode or without causation). Also, the question of why (and how) the universe originated, together with uncertainty about some other possible results of that process (as different universes, with different laws and conditions or even without them – as a disordered universe) remain as one of things which are addressed mostly by the metaphysical realm. In this way, a supernatural designer could set up a universe in such a way that creation could be consistent with natural laws through initial primordial conditions, processes and results, or could set up nomically forbidden conditions out of which life would then purely naturally arise. So, counterflow could be undiscovered on a (pre)historical time scale because designer interven-

⁴³ Tucker, "Unique events: The Underdetermination of Explanation"... p. 61. In the same publications Tucker offers a comprehensive philosophical debate on problem of unique events and relates issues of definition, underdetermination of explanation, types of unique events etc.

tion stays deeply hidden somewhere in prehistorical or protohistorical time 'when' nature was not completely shaped by the full constants and laws we recognize as active today, as well as time itself. It is very naive to imagine a supernatural designer being forced to leave traces of counter-flow or who designs directly against the lawful natural flow. Due to this, there is a possibility that "design activity operates within the stream of natural laws and process, but laws are neither suspended nor broken".⁴⁴ A supernatural designer could have restricted his/her/its intervention on the determination of the primordial conditions in the beginning of the universe without any interference into historical time and processes: by this approach supernatural intervention could be qualified as natural at the level of processes (and result) that led to the origin of first life within time. Due to this, design cannot be excluded as a source of first life at some stage of the (pre)development of universe.

VI. DESIGN CAN BE A LEGITIMATE PART OF SCIENCE

Mainstream science is quite optimistic about the prospect of a scientific explanation of the origin of first life, claiming that the explanation of the origin of life is, albeit difficult, in principle solvable. In addition, there is a widespread belief regarding a giant advance in science reached in this area of research and rapid growth in understanding the key steps in processes that lead to life, which could give a full scientific explanation of the OoL. On the contrary, it seems that the current stage of 'business' is not so promising and convenient results are not available, both in the theoretical and experimental domains. It is possible to speak about weakness of the scientific hypotheses about the OoL because:

There are significant impediment to the scientific explanation of OoL... (i) there is no scientific hypothesis that successfully explains OoL, and (ii) there is an evident experimental failure of the attempts to (re)create life from inorganic components... the scientific study of OoL rest on a mistake or science must reduce the scope of phenomena that can be scientifically explained.⁴⁵

⁴⁴ Ratzsch, *Nature, Design and Science*...p. 30.

⁴⁵ Tonći Kokić, "Weakness of the Scientific Hypotheses about the Origin of Life", *Filozofia Nauki*, 2018, 1, p. 9 [9–23].

Instead of a strong scientific theory supported by successful experiments, there are a number of barren hypotheses followed by failed attempts in recreating life. In fact, from the theoretical view of the problem, OoL is still in a pre-paradigmatic phase or an immature science, and from the empirical side of view, the confirmation of these hypotheses indicates only "...tiny steps on the long road to life".⁴⁶ So, the devastating fact is that both hypotheses and experiments guided by them lack some (and the most important) components required for full description of the origin of life.⁴⁷ Also, it remains unclear whether supernatural design as an explanation in principle is banned in science or if such theories / hypotheses are banned because they are incorrect or wrong, meaning that they do not follow scientific methodology (predictability, experimental verifiability/falsification). It seems reasonable that supernatural design could not be legitimate as a particular scientific theory or hypothesis, but it also seems quite reasonable that it could be legitimate as a metaphysical assumption. The assumption of design can be contained in the general metaphysical layer of assumptions which exist in all specific scientific theories: design is placed in the deep metaphysical zone which is the basis for particular scientific theories and which directs the explanation of natural phenomena. Building on the previously mentioned Faraday, Maxwell has a similar notion about scientists who believe in the ontological reality of a supernatural designer (and sustainer) which could be ascribed to Herschel, Newton, Copernicus, Mendel, Kelvin, Planck, etc. – a list of many others is long and cannot be presented here. At worst for the illegitimacy position, this list clearly suggests that the metaphysical assumption of a supernatural design(er) should not be an obstacle in creating scientific theories and explanation of phenomena by natural causes. What is not excluded here is the assumption that the naturally produced design of natural things could have been built 'during' some prehistorical, protohistoric events or activity deeply hidden in principally epistemologically inaccessible 'time', processes, primordial conditions, places or otherwise outside of the space and time now in operation. Whatever anybody thinks about the *God of the Gap* category of argument, the claim that nature *per se*

⁴⁶ Robert Hazen, *Genesis. The Scientific Quest for Life Origin*, Joseph Henry Press, Washington, DC, 2005, p. 93.

⁴⁷ Penny, "An Interpretative Review of the Origin of Life Research"...p. 638.

does not have the ability to behave in some way or produce some class of things, for example moving things faster than light or the origin of life, is in principle not unscientific. Moreover, as a matter of fact, the laws of nature *per definitionem* do not allow some kind of processes, behaviour or result. The logical structure and validity of assumptions of what nature can and cannot do are equivalent. According to this, the assumption of supernatural design is a legitimate layer of scientific thought: the analogies of monoliths and astronauts demonstrated this scientific judgment and the judgement of its legitimacy is in large scope dependent on metaphysical beliefs and not on the empirical aspects of science. Because of all of this, it is reasonable to give a chance to the process of reconsidering the scientific legitimacy of design *as a very general idea inside a conceptual hierarchy of science*. This is accompanied by a request for the right to take a neutral metaphysical position about the underlying metaphysical nature of physical and mechanical forces responsible for the origin of life, as *terminus a quo*, in OoL research.

REFERENCES

- AUDI Robert (ed.), *The Cambridge Dictionary of Philosophy*, second edition, Cambridge University Press, Cambridge, 1999.
- AYALA Francisco and ARP Robert, "Is There a Place for Intelligent Design in the Philosophy of Biology?", in: AYALA Francisco and ARP Robert (ed.), *Contemporary Debates in Philosophy of Biology*, Blackwell Publishing Ltd., Malden, MA, 2009, 337–340.
- AYALA Francisco, "There is No Place for Intelligent Design in the Philosophy of Biology: Intelligent Design Is Not Science", in: AYALA Francisco and ARP Robert (ed.), *Contemporary Debates in Philosophy of Biology*, Blackwell Publishing Ltd., Malden, MA, 2009, 364–391.
- BEHE Michael, *Darwin's Black Box*, Free Press, New York, 1996.
- CHAKRAVARTTY Anjan, "Scientific Realism", in: Edward N. Zalta (ed.), *The Stanford Encyclopedia of Philosophy*, (Summer 2017 Edition), <https://plato.stanford.edu/archives/sum2017/entries/scientific-realism> (24.3.2020.).
- COLLINS Francis, *The Language of God: A Scientist Presents Evidence for Belief*, Simon & Schuster UK Ltd., London, 2007.
- CRICK Francis, *What Mad Pursuit*, Penguin Books, New York, 1998.
- DARWIN Charles, *The Origins of Species by Means of Natural Selection*, Random House Value Publishing, New York, 1979.

- DAVIES Paul, *The Mind of God. The Scientific Basis for a Rational World*, Simon & Shuster Paperbacks, New York, 1992.
- DEAMER David, *First Life. Discovering the Connections Between Stars, Cells and How Life Began*, University of California Press, Berkeley, 2011.
- DE DUVE Christian, *Vital Dust*, Basic Books, New York, 1995.
- DEMBSKI William, *The Design Inference: Eliminating Chance through Small Probabilities*, Cambridge University Press, New York, 1998.
- FLEW Anthony, *There is a God: How the World's Most Notorious Atheist Changed his Mind*, Harper Collins Publishers, New York, 2007.
- FRY Iris, *The Emergence of Life on Earth: a Historical and Scientific Overview*, Rutgers University Press, New Brunswick, 2000.
- GOULD Stephen Jay, "Nonoverlapping magisterial", *Natural History*, 1997, 106, 16–22.
- GRIESEMER James, "Origins of Life Studies", in: Ruse Michael (ed.), *The Oxford Handbook of Philosophy of Biology*, Oxford University Press, New York, 2008, 263–290.
- HAZEN Robert, *Genesis. The Scientific Quest for Life Origin*, Joseph Henry Press, Washington, DC, 2005.
- KOKIĆ Tonći, "Weakness of the Scientific Hypotheses about the Origin of Life", *Filozofia Nauki*, 2018, 1, 9–23.
- MONOD Jacques, *Chance and Necessity: An Essay on the Natural Philosophy of Modern Biology*, Vintage Books, New York, 1972.
- PENNY David, "An Interpretative Review of the Origin of Life Research", *Biology and Philosophy*, 2005, 20, 633–671.
- RATZSCH Delatzsch, *Nature, Design and Science. The status of Design in Natural Science*, SUNY Series in Philosophy of Biology, New York, 2001.
- SCHRÖDINGER Erwin, *What is Life. The Physical Aspect of the Living Cell*, Cambridge University Press, Cambridge, 2013.
- TORRANCE, Thomas, *The Christian Frame of Mind*, Handsel Press, Edinburg, 1985.
- TUCKER Aviezer, "Unique events: The Underdetermination of Explanation", *Erkenntnis*, 1998, 48, 59–80.
- WHITE Robert, "Does Origins of Life Research Rest on a Mistake", *Noûs*, 2007, 4, 453–477.
- WRIGHT Barbara, "A Biochemical Mechanism for Nonrandom Mutations and Evolution", *Journal of Bacteriology*, 2000, 182, 2993–3001.
- ZALTA Edward (ed.), *Stanford Encyclopedia of Philosophy*, Stanford, The Metaphysics Research Lab Center for the Study of Language and Information Stanford University, 2017, <https://plato.stanford.edu/> (24.3.2020.).

Novi relativizam^{1,2}: može li relativist reći da Hitler nije trebao narediti istrjebljenje Židova?³

Boran Berčić

Od Platonova *Protagore* do Kurosawina *Rašomona*, ideja relativizma zaokuplja refleksivne umove (Berčić 1995; Berčić 2006). Prihvaćali je ili je odbacivali, mnogi su o njoj raspravljali. Platon je smatrao da se ideja relativizma ne može konzistentno formulirati jer je u samoj svojoj osnovi samopobijajuća pozicija. Antički skeptici smatrali su da je relativnost ljudske spoznaje očita činjenica. Enesidem je smatrao da je spoznaja trostruko relativna: u odnosu na (1) predmet spoznavanja, u odnosu na (2) medij spoznavanja i u odnosu na (3) organ spoznavanja. Pascal je smatrao da je relativizam apsurdan. Pitao se kako nešto može biti istinito s ove strane Pirineja, a neistinito s druge.

¹ Ovaj rad financiralo je Sveučilište u Rijeci projektom *Metametafizika* broj uniri-human-18-239.

² Zahvaljujem na pomoći Maji Brali s Odsjeka za anglistiku u Rijeci i Dušanu Dožudiću s Instituta za filozofiju u Zagrebu, zatim svim sudionicima radionice *Kontekstualizam i relativizam* koja se u veljači i ožujku 2020. održavala na Filozofskom fakultetu u Rijeci, na koncu organizatorima i sudionicima konferencije *Contextualism and Relativism* koja se još 2012. održala na IUC-u Dubrovniku.

³ Nisam baš sasvim siguran da je primjereno koristiti ovaj primjer kao ilustraciju u filozofskoj raspravi. Možda bi to u nekom smislu moglo biti uvredljivo za Židove pa i za one koji smatraju da je pogrešno ljude dijeliti po akcidentalnim svojstvima kao što su nacija, konfesija itd. No primjer koriste Harman, Kölbel i ostali u raspravi pa ga koristim i ja. Na koncu, ovo je primjer nečega što je očito pogrešno, a upravo se takvi ekstremni primjeri i koriste u raspravi o relativizmu.

Novi relativizam ili analitički relativizam

Krajem 20. i početkom 21. stoljeća ideja relativizma dobila je novi *impetus* razvojem u filozofiji jezika. John MacFarlane, Max Kölbel, Peter Lasersohn i drugi pokušali su ideju relativizma formulirati na konzistentan, precizan i prihvatljiv način. Zbog novog pristupa njihovi stavovi dobili su naziv *novi relativizam* ili *analitički relativizam* (Baghramian i Carter 2020). To je pozicija koja svoje uporište prvenstveno ima u radovima Davida Kaplana o indeksikalima i demonstrativima. To su radovi koji predstavljaju suvremeni klasik filozofije jezika. Pozicija dijelom ima uporište i u metaetičkim radovima Gilberta Harmana i Jamesa Dreiera, u kojima oni nastoje precizno formulirati tezu moralnog relativizma. Novi relativizam prvenstveno je semantička pozicija. MacFarlane, Kölbel, Lasersohn i drugi nastoje razviti semantički okvir za relativizam. Oni nisu radikalni relativisti, ne zastupaju opći relativizam u pogledu istine ili znanstvene spoznaje. No očekuju da bi relativistički okvir mogao imati primjenu u području etike, estetike, sudova osobnog ukusa i eventualno drugih područja u kojima bi relativizam mogao predstavljati plauzibilnu poziciju. Oni smatraju da u tim područjima može postojati takozvani *faultless disagreement*. (Kölbel 2004). To bismo mogli prevesti kao *bezgrešno neslaganje*.⁴ To je situacija u kojoj netko tvrdi neki sud *p*, netko drugi tvrdi negaciju tog suda *ne-p*, a da pri tome nitko nije u krivu. Moguće je da su obojica u pravu. Ako iskreno tvrde to što tvrde, njihovi iskazi mogu biti istiniti iako predstavljaju kontradikciju. Novi relativisti najčešće analiziraju upotrebu predikata kao što su *ukusan*, *zabavan* i slični. Tipični su primjeri sljedeći: vlak smrti jest zabavan. (Lasersohn 2005: 643), Picasso je bolji od Matissea. (Kölbel 2007), riblji štapići jesu ukusni. (MacFarlane 2014: 14) itd.⁵

⁴ Jasno, apsolutist smatra da *faultless disagreement* nije moguć. Ako je *disagreement*, ne može biti *faultless*. Ako je *faultless*, ne može biti *disagreement*. Ako obje strane mogu biti u pravu, onda tu ne može biti nikakvog neslaganja. Ako se ne slažu, onda mora biti da netko griješi, ako nitko ne griješi, onda nema neslaganja. (Iacona 2008)

⁵ Ako u sudovima ukusa doista postoji bezgrešno neslaganje, onda nije baš sasvim jasno kako je moguće birati najbolje vino, najbolji sir itd. Nije jasno ni čemu služe somelijeri. U gostioni *Stara pošta* u Permania kod Rijeke svake godine biraju se najbolje kobasice. Žiri obrazlaže svoju odluku to jest *navodi karakteristike* koje pobjedničku kobasicu čine boljom od drugih. Teško da može postojati bezgrešno neslaganje ako se navode karakteristike.

Ovdje treba napomenuti da novi relativisti ne raspravljaju o tome da netko voli riblje štapiće, a netko ne ili o tome da ih ista osoba voli dok je mala, a ne voli kad odraste. To nije upitno i to nije problem, u tome nema kontradikcije. Oni čak ne raspravljaju ni o značenju predikata *ukusno*, ono je po pretpostavci fiksno i univerzalno.⁶ Ono o čemu oni raspravljaju jest *istinosna vrijednost suda* da su riblji štapići ukusni. Oni nastoje pokazati da je moguće da je jedan te isti sud za jednog govornika istinit, a za drugog neistinit. To je predmet rasprave i o tome se radi. Peter Lasersohn kaže:

“Istinosna vrijednost rečenica koje sadrže predikate “osobnog ukusa” kao što su *zabavno* i *ukusno* moraju biti relativizirane u odnosu na individue. Ova relativizacija odnosi se samo na istinosnu vrijednost i ne uključuje relativizaciju semantičkog sadržaja: ako vi kažete da su vlakovi smrti (*roller coasters*) zabavni, a ja kažem da nisu, ja negiram isti sadržaj koji vi tvrdite i direktno vam proturječim. Ipak, oba iskaza mogu biti istinita (relativno u odnosu na svoje kontekste). (Lasersohn 2005)

Berit Brogaard upozorava na istu pretpostavku:

“Prema relativističkoj semantici, semantički sadržaj relativnih izraza invarijantan je u odnosu na kontekst. Ali istina rečenice relativizirana je u odnosu na kontekst upotrebe i kontekst vrednovanja.” (Brogaard 2008: 386)

U ovom radu nastojat ću prikazati osnovne crte novog relativizma, objasniti motivaciju za uvođenje novih tehničkih rješenja, analizirati relevantne primjere te pokušati vrednovati ukupnu primjenu i doseg te nove varijante relativizma.

⁶ U ovakvim situacijama prirodno nam pada na pamet opcija da se tu radi o *semantičkom* neslaganju. Neslaganje je semantičko ako ljudi koji se ne slažu oko toga što je *zabavno* ili *ukusno* pod *zabavno* ili *ukusno* misle na različite stvari, dakle ako tim izrazima pripisuju *različita značenja*. Na primjer, mogli bismo reći da se Zagorac i Dalmatinac ne slažu koje je vino dobro zato što izraz *dobro vino* za njih ima *različita značenja*. Kada tvrdi da je neko vino dobro, Zagorac misli na kiselo vino dok Dalmatinac misli na suho. Moguće je da oni imaju različite kriterije vrednovanja i da ti kriteriji sadrže različite pH vrijednosti. No to ne znači da izraz *dobro vino* za njih ima različito značenje. Izraz ima isto značenje za obojicu. Kriterij ne ulazi u značenje. Ili u punoj formulaciji, kriterij vrednovanja dobrog vina ne ulazi u značenje izraza *dobro vino*. Možda ovo i nije točno, ali u svakom slučaju predstavlja pretpostavku novog relativizma.

Ispovijest njemačkog vojnika

Osamdesetih godina 20. stoljeća susreo sam starijeg gospodina, Nijemca. Bio je vrlo ugodan i zanimljiv sugovornik. Izrazito pristojan, s klasičnim manirama, a opet jednostavan i otvoren.⁷ Pričao je kako je u Drugom svjetskom ratu bio njemački vojnik i kako se borio na Istočnom frontu. Bez ikakvog uvijanja, direktno i eksplicitno, priznao je kako je i sam sudjelovao u stravičnim zločinima koje su njemačke snage počinile nad civilnim stanovništvom. Između ostaloga, njegovo svjedočenje bilo je filozofski vrlo zanimljivo. Imalo je vrlo jaku i zanimljivu relativističku poantu. Govoreći o zločinima koje je počinio, rekao je:

- Onda smo vjerovali da je to normalno.
- Danas vjerujemo da je to monstruozno.

Mislio sam da relativističku poantu koristi da bi pokazao koliko je bio jak utjecaj nacističke propagande. Pitao sam ga da li to znači da su s vremenom uvidjeli da je to bilo monstruozno *već onda* kada su to učinili. Na moje iznenađenje, reagirao je momentalno te jasno i glasno rekao: “Ne, ne! Nisam to rekao! Rekao sam samo to da smo 40-ih vjerovali da je to normalno, a da 80-ih vjerujemo da je to monstruozno. Nisam rekao ništa više od toga.”⁸ Ovaj odgovor leđi krv u žilama, no filozofski je vrlo zanimljiv. Naime, ovim odgovorom sugovornik čini dvije stvari:

(1) On odbija postupak *sagledati apsolutistički* i priznati da su zločini koje su počinili 40-ih moralno monstruozni uvijek, svuda i za svakoga. Ideja apsolutizma jest da ako su neki postupci monstruozni, onda su monstruozni *po sebi*, bez obzira na perspektivu iz koje ih se sagledava i bez obzira na to koliko kome treba da to shvati.

(2) On odbija čak i bitno slabiji pristup. On odbija zločine koje su počinili 40-ih *sagledati iz perspektive 80-ih*. Odbija standarde iz 80-ih

⁷ Upravo je zato priča koja slijedi još više zastrašujuća. Da je iste stavove iznio neki divlji i primitivni stari nacist, oni ne bi bili toliko zastrašujući.

⁸ Taj odgovor bio je toliko brz i toliko dobro pripremljen da sam imao dojam da je *naučen*. Ne mogu to tvrditi, ali možda su psiholozi ili socijalni radnici radili na tome s njemačkim vojnicima koji su se vraćali iz sovjetskog zarobljeništva. Ova verzija relativizma doista bi mogla biti dobra psihološka brana osjećaju krivnje nekome tko je počinio ratne zločine.

bez zadrške primijeniti na postupke iz 40-ih. To znači da odbija priznati da su 40-ih bili u krivu, odbija priznati da su ti zločini bili monstruozi već onda i odbija pokajati se za ono što su učinili.

Iako još nismo objasnili distinkciju između *kontekstualizma* i *relativizma*, možemo odmah reći da stav njemačkog veterana odgovara poziciji koja se u suvremenoj semantičkoj raspravi naziva *kontekstualizam* dok stav izražen u opciji (2) odgovara poziciji koja se naziva *relativizam*. Da ne bi bilo zabune, i kontekstualizam i relativizam relativističke su pozicije ako se o relativizmu govori u općem i uobičajenom smislu. Kontekstualizam i relativizam (u užem smislu) samo su dvije varijante relativizma (u širem smislu). Kontekstualizam i relativizam (u užem smislu) jesu pozicije koje su se diferencirale unutar semantičke rasprave o tome kako točno formulirati tezu relativizma (u širem smislu). Distinkcija će biti izložena i detaljno raspravljena u nastavku teksta.⁹

O čemu ovisi istinosna vrijednost iskaza?

Prema tradicionalnom gledištu istinosna vrijednost iskaza ovisi o dva faktora: ovisi o (1) značenju iskaza i o (2) stanju stvari u svijetu. Ako su stvari takve kao što iskaz kaže da jesu, onda je on istinit, a ako nisu, onda je neistinit. Dakle, da bismo znali je li neki iskaz istinit ili nije, moramo znati (1) što on znači i (2) kakav je svijet. Ovu ideju može se prikazati formulom:

$$T(s,w)$$

U ovoj formuli T stoji za istinosnu vrijednost. s stoji za sadržaj iskaza tj. za značenje iskaza. Drugim riječima, s stoji za propoziciju ili za sud koji se iskazom izriče. w stoji za svijet tj. za činjenice ili za stanja stvari na koja se iskaz odnosi.

Tradicionalno gledište na prvi pogled može izgledati sasvim uvjerljivo i plauzibilno. Međutim, ono *ne uzima u obzir kontekst u kojem*

⁹ Jasno, sama činjenica da postupke iz 40-ih vrednuje standardima iz 80-ih već predstavlja odstupanje od kontekstualizma i djelomično približavanje relativizmu. Naime, striktni kontekstualist trebao bi tvrditi da postupke iz 40-ih nije moguće vrednovati standardima iz 80-ih.

se iskaz izriče. A istinosna vrijednost iskaza može ovisiti o kontekstu izricanja. Zbog toga ova formula nije potpuna i treba biti nadopunjena. To je posebno očito kod indeksikala i demonstrativa.¹⁰ Ako Boran Berčić izrekne rečenicu “Ja sam gladan.”, onda je time izrečena propozicija *Boran Berčić je gladan*. Njezina istinosna vrijednosti ovisi o tome je li Boran Berčić gladan ili nije. Ako Marin Biondić izrekne istu rečenicu “Ja sam gladan.”, onda je time izrečena druga propozicija *Marin Biondić je gladan*. Njezina istinosna vrijednosti ovisi o tome je li Marin Biondić gladan ili nije. Dakle, ista rečenica, ovisno o tome tko je izgovara, može izražavati različit sud i može imati različitu istinosnu vrijednost. Ista stvar vidi se i kod vremenski indeksiranih iskaza. Ako u nedjelju 20. 9. 2020. kažem “Jučer je bila subota.”, onda taj iskaz znači da je 19. 9. 2020. bila subota i taj je iskaz istinit. Ako pak u ponedjeljak 21. 9. 2020. izrekнем istu rečenicu “Jučer je bila subota.”, onda sam time rekao da je u nedjelju 20. 9. 2020. bila subota, što je netočno i iskaz je neistinit. Niz je zanimljivih i duhovitih primjera. Najdraži mi je “Ja sam kraljica Engleske.” Jasno, istinosna vrijednost ovog iskaza ovisi o tome tko ga izriče. Zbog ovakvih i sličnih primjera postalo je jasno da istinosna vrijednost iskaza ovisi i o kontekstu u kojem se iskaz izriče. Zbog toga u gore navedenu formulu treba uvrstiti i kontekst izricanja iskaza.

$$T(s,w,c)$$

U ovoj formuli znakovi imaju isto značenje kao i u prethodnoj. Jasno, s time da c stoji za kontekst u kojem se iskaz izriče.

Kontekstualizam

Iz očitih razloga gore izložena pozicija naziva se *kontekstualizam*. Kao što smo rekli, ideja je da istinosna vrijednost iskaza, pored toga što ovisi o (1) značenju iskaza i o (2) činjenicama u svijetu, ovisi i o (3) kontekstu u kojem se iskaz izriče. Ideja kontekstualista jest da su ovo *svi relevantni faktori* i da istinosna vrijednost iskaza *ne ovisi ni o čemu drugome*.

¹⁰ Indeksikali su izrazi čije značenje ovisi o tome tko ih, gdje i kada izgovara. To su izrazi kao što su: *ja, sada, ovdje, desno, ispred, jučer, lani*, itd. Demonstrativi su izrazi koji sadrže pokazni element: *ovaj, onaj*, itd.

Relativizam i perspektiva

Međutim, autori kao što su John MacFarlane, Max Kölbel, Peter Lerssohn i drugi smatraju da ova formula ne obuhvaća sve relevantne faktore i da je treba nadopuniti. Oni smatraju da pored navedenih faktora postoji i *perspektiva* iz koje se procjenjuje istinosna vrijednost suda. Budući da je ideja da postoji perspektiva iz koje se procjenjuje istinosna vrijednost suda bitno relativistička, oni svoju poziciju nazivaju relativizam. Tako Max Kölbel kontekstualizam naziva *indeksikalni relativizam* dok svoju poziciju koja uključuje i perspektivu vrednovanja naziva pravi/autentični/stvarni *relativizam*. (Kölbel 2004). Razlika je u tome što u kontekstualizmu kontekst u kojem je iskaz izrečen, da tako kažemo, *fiksira* istinosnu vrijednost i ona zauvijek ostaje takva kakva je bila u trenutku izricanja iskaza. S druge strane, u relativizmu istinosna vrijednost suda ovisi i o perspektivi iz koje se procjenjuje, a ta perspektiva može varirati. Istinosna vrijednost jednog te istog suda može se procjenjivati iz niza perspektiva, u principu iz beskonačnog broja različitih perspektiva. Zbog toga u relativizmu istinosna vrijednost suda može varirati ovisno o perspektivi iz koje se vrednuje. Jedan te isti sud može imati različite istinosne vrijednosti ovisno o perspektivi iz koje se vrednuje. Dakle, prema relativistima istinosna vrijednost T ovisi o značenju iskaza s , stanju stvari u svijetu w , kontekstu izricanja c i perspektivi vrednovanja p .

$$T(s,w,c,p)$$

Prema ovoj slici, istinosna vrijednost suda bila bi funkcija navedenih četiriju faktora. Upravo je u tome razlika između kontekstualizma i relativizma.

Sada treba detaljnije objasniti ideju perspektive, to jest razliku između konteksta izricanja i konteksta vrednovanja.

Kontekst izricanja i kontekst vrednovanja

Za razliku od kontekstualista, koji smatraju da kontekst izricanja suda ujedno jest kontekst koji određuje njegovu istinosnu vrijednost, relativisti smatraju da ova dva konteksta treba razdvojiti. Oni smatraju da su

to dvije različite stvari i da se kontekst u kojem se sud izriče ne mora podudarati s kontekstom u kojem se procjenjuje je li istinit ili nije. Tako Berit Brogaard kaže:

“Max Kölbel, John MacFarlane i drugi predlažu radikalnu reviziju standardne teorije. Dok standardna teorija istinu rečenice relativizira u odnosu na kontekst upotrebe, relativisti predlažu da se istina rečenice relativizira i u odnosu na kontekst upotrebe, i u odnosu na kontekst vrednovanja.” (Brogaard 2008: 391)

U nekom smislu i kontekstualizam može sadržavati parametre vrednovanja, no oni moraju biti uključeni u kontekst izricanja i ne mogu biti od njega nezavisni. Maria Baghramian i Adam Carter kažu:

“Kontekstualist može, isto kao i relativist, prihvatiti parametar ‘standarda’ ili ‘sudenja’. No za kontekstualista on će biti sadržan u kontekstu upotrebe dok relativist smatra da je sadržan u kontekstu vrednovanja, potpuno nezavisno od konteksta upotrebe.” (Baghramian i Carter 2015)

John MacFarlane radio je na *problemu budućih kontingencija* kada je uvidio potrebu da razdvoji ta dva konteksta (*context of use* i *context of assessment*). (MacFarlane 2003) Naime, izgleda da iskazi o budućim kontingentnim događajima nemaju istinosnu vrijednost u trenutku izricanja, nego da je stječu tek kasnije, kada se dogode ili ne dogode događaji o kojima govore. (Berčić 2019). Vremenska prognoza (izrečena sredinom ljeta) da će jesen biti kišovita u trenutku izricanja *još nije ni istinita ni neistinita* jer jesen još nije nastupila. Prognoza *postaje istinita ili neistinita* tek kada završi jesen. U ovom primjeru *kontekst izricanja* jest sredina ljeta, a *kontekst vrednovanja* jest kraj jeseni. Razlika između ova dva konteksta pogotovo je uočljiva u slučaju budućih kontingencija jer one u trenutku izricanja još nemaju istinosnu vrijednost nego je dobivaju tek kasnije. Jasno, relativisti generaliziraju ovaj uvid i smatraju da on u principu vrijedi za sve sudove ili barem za velike domene sudova. I drugi relativisti nezavisno su došli do istog uvida. Max Kölbel razlikuje *context of use* i *circumstance of evaluation*. (Kölbel 2004) dok Peter Lasersohn razlikuje *autocentric perspective* i *exocentric perspective* (Lasersohn 2005).

Nedostaci kontekstualizma

Ovdje ćemo izložiti dva problema na koja nailazi kontekstualizam (Brogaard 2008).¹¹ Prvi je da kontekstualizam ne može stvarno objasniti neslaganja u sudovima etike, ukusa, itd. Drugi je da ne može objasniti povlačenje prethodno danih sudova.¹²

Kontekstualizam je nastao kao detaljna semantičko-logička formulacija opće relativističke ideje. Na prvi pogled može se učiniti da kontekstualizam sa svojom formulom $T(s,w,c)$ dobro zahvaća općerelativističke intuicije i da može dobro objasniti slučajeve vrijednosnog neslaganja. Budući da moralni kod govornika ulazi u relevantni kontekst c , istinska vrijednost iskaza relativna je u odnosu na moralni kod govornika. Drugim riječima, istinska vrijednost moralnog suda ovisi o tome kojem moralnom kodu pripada govornik. Što je istinito za jednoga, ne mora biti istinito i za drugoga! Izgleda da je ovime dobro zahvaćena općerelativistička ideja. Međutim, problem je u tome što prema ovoj rekonstrukciji *neslaganje zapravo nije moguće*. Da bi neslaganje bilo neslaganje, *moramo smatrati da je onaj drugi u krivu*. Ali ako prihvatimo formulu $T(s,w,c)$, onda ne možemo onome tko se s nama ne slaže pripisati grešku. Ako je on svoj sud donio na osnovi svojeg koda, nemamo na osnovi čega tvrditi da griješi. Mi možemo tvrditi suprotno, ali ne možemo tvrditi da on griješi. Mi ponekad doista zauzimamo ovakav stav prema nekim ljudima i stavovima koje zastupaju. No to su oni koje smatramo beznadnim slučajevima. Takav stav možemo zauzeti prema izrazito primitivnim ljudima ili prema ljudima koji dolaze iz udaljenih krajeva svijeta. Nećemo se ljutiti na njih i nećemo osjetiti potrebu da ispravimo njihove stavove. Naprosto ćemo smatrati da su oni van dosega moralne argumentacije. Radi se o tome da takav stav ne zauzimamo prema *svima* koji se s nama ne slažu. Za neke smatramo da su u krivu i da griješe. Problem s kontekstualizmom jest u tome što

¹¹ Inače, Brogaard smatra da uz neke izmjene kontekstualizam može izaći na kraj s ovim problemima i da ukupno gledajući predstavlja bolju teoriju. Varijantu koju predlaže naziva *perspectivalism*.

¹² Relativisti navode i druge karakteristike jezika i govora koje prema njihovom mišljenju kontekstualizam ne može objasniti, a relativizam može. U svojoj knjizi iz 2008. MacFarlane na dva mjesta navodi pet takvih karakteristika, str. 21 i 147. No ovdje ćemo izložiti samo dvije najočitije.

nam ne omogućava da zauzmemo takav stav. Prema kontekstualizmu neslaganje zapravo predstavlja *semantičku pogrešku*: onaj tko se s nekim ne slaže, taj zapravo ne razumije jezik. A to nije tako. Berit Brogaard jasno ilustrira ovaj uvid. Ovdje navodim primjere koje je izložila u radu od 2008. Usporedimo sljedeća dva kratka dijaloga:

John: Žensko genitalno sakaćenje jest pogrešno

Mary: Ne slažem se. Uopće nije pogrešno.

John: Gladan sam.

Mary: Ne slažem se. Ja uopće nisam gladna.

Očito je da u prvome dijalogu Mary pokazuje jezičnu kompetenciju, a u drugome ne. U prvome dijalogu neslaganje ima smisla dok u drugome nema. Međutim, kada bi kontekstualizam bio istinit, onda bi Mary jednako pokazivala jezičnu manjak jezične kompetencije u oba dijaloga. Prema kontekstualizmu, neslaganje u prvom dijalogu jednako je besmisleno kao i neslaganje u drugome. A to očito nije tako. Uvjereni smo da je u prvom slučaju neslaganje smisljeno, a da u drugome nije. Dakle, kontekstualizam ne predstavlja dobar okvir za razumijevanje neslaganja.

Drugi problem na koji nailazi kontekstualizam jest da ne može objasniti povlačenje prethodno danih sudova (*retraction*). Ako promijenimo mišljenje o nečemu, prirodno je smatrati da je naše prethodno mišljenje bilo pogrešno. Ako smo prije vjerovali da *p*, a sada smo uvidjeli da ne-*p*, onda to znači da smo prije bili u krivu i da smo pogrešno vjerovali da *p*. Zamislmo da je John vjerovao da je žensko genitalno sakaćenje pogrešno, a da je onda, nakon što je proveo nekoliko mjeseci u zemlji u kojoj je to uobičajena praksa, došao do zaključka da nije pogrešno. Dijalog iz Brogaard 2008:

John: Žensko genitalno sakaćenje uopće nije pogrešno.

Mary: Ali prije nekoliko mjeseci tvrdio si da je pogrešno.

John: Znam. Bio sam u krivu.

U ovakvoj situaciji izgleda prirodno da John povuče svoje prethodno mišljenje i da kaže da je bio u krivu. Međutim, kontekstualizam to ne dopušta. Boravkom u zemlji u kojoj je žensko genitalno sakaćenje uo-

bičajena praksa promijenio se relevantni kontekst c. Kontekst iz kojega John sada govori različit je od onoga iz kojega je govorio prije. Zbog toga se prema kontekstualizmu njegovi iskazi zapravo *uopće ne mogu uspoređivati*. Jedino što se može reći jest da je prije za njega bila istina da je žensko genitalno sakaćenje pogrešno, a da je sada za njega istina da nije. U okviru kontekstualizma ne može se reći *da je prije bio u krivu*. Budući da je povlačenje prethodnog mišljenja prirodna posljedica promjene mišljenja, kontekstualizam ne predstavlja dobru rekonstrukciju načina na koji mislimo i govorimo. Dakako, relativisti smatraju da se taj problem može lako i prirodno riješiti uvođenjem dodatnog parametra perspektive, to jest razlikovanjem konteksta izricanja i konteksta vrednovanja.¹³

Istina, mi ponekad kažemo *Onda je bilo tako!* Pri tome ne želimo vrednovati to kako je bilo. Ne želimo reći da je bilo bolje ili gore nego što je sada. Samo želimo reći da je bilo *drugačije* nego što je sada. *Onda je bilo onako, a sada je ovako!* Prije se služio vojni rok, a sada više ne. Prije se išlo u Trst po odjeću, a sada više ne. Tko je imao *Cortinu*, bio je glavni frajer. ... Međutim, ovakav stav zauzimamo kada smatramo da su okolnosti bile toliko različite da nema smisla uspoređivati ono što je bilo onda s onime što je sada. Ovakav stav *ne* zauzimamo kada se radi o sudovima i istini. Ako je John prije tri mjeseca vjerovao da je nešto pogrešno, a sada vjeruje da nije, onda mora vjerovati i to da je prije bio u krivu. Zato je kontekstualizam pogrešan i zato njemački veteran griješi kada svoje postupke iz 40-ih odbija sagledati iz perspektive 80-ih. U slučaju moralnih vjerovanja ne možemo reći *Onda smo vjerovali ono a sada vjerujemo ovo!* i ostati na tome. Ako smo promijenili vjerovanje, onda to automatski implicira da smo prije bili u krivu.

¹³ Treba napomenuti da kontekstualizam i relativizam imaju empirijske posljedice. Zbog toga je u principu moguće lingvističkom anketom utvrditi koja je pozicija točna, a to nije čest slučaj kod filozofskih rasprava. Markus Kneer smatra da empirijska evidencija presuđuje u korist kontekstualizma (Kneer 2017). No pitanje je možemo li u ovakvim slučajevima dobiti direktnu i čistu evidenciju koja ne bi bila kontaminirana interpretacijom i sugestijama ispitivača. Dobra ilustracija jest MacFarlaneov primjer "Riblji su štapići ukusni." Prva reakcija publike (uključujući lingviste, logičare i filozofe) uvijek je "Pa nema tu nikakve kontradikcije. Kad je bio mali, volio ih je, a sada ih više ne voli. U čemu je problem?" No, kada im se objasni da se radi o *sudu*, a ne o gustativnoj preferenciji, počinju naginjati relativističkom stavu da treba povući prethodni sud.

Kontekstualizam i relativizam u etici

Nakon što smo izložili razliku između kontekstualizma i relativizma te vidjeli na koje probleme nailazi kontekstualizam, možemo vidjeti kako se ta rasprava odražava na pokušaje točne formulacije etičkog relativizma (u širem smislu). Etički relativizam u pravilu se formulira kao *semantička teza* o tome kako treba shvatiti normativne sudove.¹⁴ Ideja je da su normativni sudovi uvijek relativni u odnosu na moralni kod onoga tko ih izriče. Dakle, iako su naizgled apsolutni, oni su zapravo relativni. Dakle, teza je ozbiljno *revizionistička* u pogledu razumijevanja normativnog diskursa. Iako je iskaz "A treba učiniti *x*." po svom logičkom i gramatičkom obliku apsolutan (nije relativan u odnosu na bilo što), njime se zapravo tvrdi da A treba učiniti *x* prema moralnom kodu grupe kojoj pripada ili nešto slično tome. Kao što sam naveo na početku teksta, novi *impetus* za etički relativizam došao je iz filozofije jezike, i to iz analize indeksikala kao što su *ja*, *sada*, *ovdje* itd. Očito je da istinsosna vrijednost iskaza koji sadrže indeksikale ovisi o kontekstu, dakle da ovisi o tome tko ih, gdje i kada izgovara. Etički relativisti smatraju da isto tako i istinsosna vrijednost normativnih iskaza ovisi o kontekstu, dakle da ovisi o tome tko ih, gdje i kada izgovara. Oni vjeruju u relevantnu analogiju između indeksikala i normativnih iskaza. Najjednostavnije rečeno, etički relativisti vjeruju da su normativni iskazi po svojoj prirodi indeksikalni. Kontekstualisti Björnsson i Finlay kažu:

"Mi smatramo da su normativni "treba" iskazi dvostruko relativni u odnosu na kontekst. Relativni su i u odnosu na (i) dostupne informacije i u odnosu na (ii) standarde ili ciljeve. Prema ovom gledištu, svako smisleno izricanje rečenice "A treba učiniti Φ " izražava propoziciju da A treba-relativno-u-odnosu-na-informacije-*i*-i-relativno-u-odnosu-na-standard-*s* učiniti Φ , gdje su *i* i *s* određeni kontekstom iskaza." (Björnsson i Finlay 2010: 8)

Tezu je učinio popularnom Gilbert Harman svojim radovima sedamdesetih godina dvadesetog stoljeća. On moralni relativizam izlaže kao logičku tezu:

"Formulirajući ovo kao logičku tezu, želim tretirati moralno "treba" kao četveromjesni predikat (ili "operator") "Treba (*A*, *D*, *C*, *M*)," koji dovodi u vezu

¹⁴ Neki autori, pogotovo domaći, inzistiraju na razlici između normativnih (treba) i vrijednosnih (dobro) sudova. To razlikovanje nije relevantno u kontekstu ove rasprave i zato ću ga ovdje zanemariti.

djelatnika *A*, vrstu postupka *D*, razloge *C*, i motivacijski stav *M*. ... Da budem nešto precizniji, dakle, moja teza je ovo. “Treba (*A*, *D*, *C*, *M*)” otprilike znači da pod pretpostavkom da *A* ima motivacijske stavove *M* i razloge *C*, *D* je postupak za koji *A* ima najbolje razloge da ga učini. U sudovima u kojima se “treba” javlja u ovom smislu, *C* i *M* često nisu eksplicite navedeni, ali kontekst iskaza upućuje na njih. U takvim slučajevima *C* će obično biti “uzevši sve u obzir”, a *M* će biti stavovi koje govornik dijeli s publikom. (Harman 1975: 38)

Harmanova ideja da normativno *treba* zapravo predstavlja četveromjensni predikat izražen formulom *Treba* (*A*, *D*, *C*, *M*) u svojoj je osnovi kontekstualistička jer *ne razlikuje kontekst izricanja od konteksta vrednovanja*. To jest *ne sadrži izvanjsku perspektivu iz koje bismo mogli vrednovati normativne iskaze*. To znači da nečije iskaze i postupke možemo vrednovati samo iz njegove perspektive i ni iz koje druge. Dakle, ako je nečiji sud istinit u odnosu na njegove razloge i u odnosu na motivacijske stavove koje dijeli s publikom, onda mi ne možemo učiniti ništa drugo nego priznati da je njegov sud istinit. Izgleda da nam kontekstualizam ne ostavlja nikakvu osnovu da kritiziramo njegove sudove i da tvrdimo da nisu istiniti. Ako Hitlerov sud da treba istrijebiti sve Židove zadovoljava Harmanovu formulu *Treba* (*A*, *D*, *C*, *M*), ako slijedi iz njegovih razmatranja *C* i ako ga izriče pred publikom koja dijeli njegove vrijednosti i motivaciju *M*, onda ne možemo učiniti ništa drugo nego moramo priznati da je taj sud istinit.¹⁵ Harman je svjestan ove posljedice svoje teorije, raspravlja o njoj i nastoji je prikazati na prihvatljiv način. Harman smatra da postoji smisao u kojem ne možemo reći da Hitler nije smio narediti istrjebljenje Židova. To je takozvani *unutrašnji sud* o Hitleru:

¹⁵ Rasprava o moralnom relativizmu u pravilu je zasnovana na implicitnoj ili eksplicitnoj pretpostavci da postoje različiti, striktno odvojeni i nesumjerljivi moralni kodovi. Zato Harman i kaže da je Hitler van dosega moralne argumentacije kakvu mi poznajemo. Uvjeran sam da je ova pretpostavka *potpuno pogrešna* te da navodi na niz krivih zaključaka. Hitler nije imao neki svoj posebni nacistički moralni kod. On je pripadao istom moralnom kodu kojem i mi pripadamo, zapadnoevropskom ili srednjeevropskom, ako je uopće primjereno govoriti o moralnim kodovima. Vjerojatno je smatrao da je pošteno da svatko živi od svog rada, da za jednaki posao ljudi trebaju biti jednako plaćeni itd. Može biti zastrašujuće to što je svoje monstruozne stavove formirao *unutar istog moralnog koda* koji i mi imamo, no to je činjenica s kojom se moramo suočiti. Ne možemo nastanak nacizma objasniti moralnim relativizmom i tvrditi da je Hitler van dosega moralne argumentacije kakvu poznajemo. Na koncu, njemački i ruski vojnici na Staljingradu nisu se borili zbog suprotstavljenih moralnih kodova. Oni su imali istu etiku. I jedni i drugi smatrali su da trebaju odraditi dug prema domovini, da trebaju pomoći suborcima, da dezertere treba kažnjavati itd.

“Vrijedi napomenuti da unutrašnji sudovi zvuče preslabo, ne zbog stravičnih razmjera onoga što je Hitler učinio, nego zato što smatramo da je svojim postupcima pokazao da ne podliježe moralnim razmatranjima na osnovi kojih mi donosimo naš sud. On je u relevantnom smislu van dosega i zato ne možemo donositi unutarnje sudove o njemu.” (Harman 1975: 37)

Internalizam

Ovdje treba objasniti što je to *unutrašnji sud*. Harman prihvaća *internalizam* u etici.¹⁶ To je gledište prema kojemu možemo za nekoga reći da nešto treba učiniti samo ako u njegovim motivacijskim stanjima možemo pronaći želju ili dispoziciju koja bi ga navela da to učini. Dakle, *A* treba učiniti *x* samo ako među *A*-ovim motivacijskim stanjima postoji neko stanje koje bi ga navelo da učini *x*. Ako *A* nema takvog motivacijskog stanja, onda nema smisla tvrditi da *A* treba učiniti *x*. Shodno ovom uvidu, *unutrašnji treba-sud* jest sud da netko nešto treba učiniti obzirom na motivacijska stanja koja ima, a *vanjski treba-sud* jest sud da netko nešto treba učiniti bez obzira na motivacijska stanja koja ima. Harman prihvaća internalizam, a prema internalizmu vanjski treba-sudovi naprosto nemaju smisla. Budući da ne relativiziraju istinosnu vrijednost treba-suda u odnosu na motivacijski sklop onoga tko ih izriče, oni ne zadovoljavaju Harmanovu *Treba* (*A*, *D*, *C*, *M*) formulu. Zbog toga se u okviru Harmanove varijante relativizma ne može reći da Hitler nije trebao ili da nije smio narediti istrjebljenje Židova. I to iz dva razloga. Prvi je *semantički kontekstualizam*, koji istinosnu vrijednost treba-suda vezuje uz kontekst u kojem ga govornik izriče. Drugi je *etički internalizam*, koji istinosnu vrijednost treba-suda vezuje uz motivacijska stanja govornika koji izriče sud.

Jasno, za relativiste *istinit* uvijek znači *istinit za njega*, a nikada *istinit po sebi*. No, priznati da sud da treba istrijebiti jednu cijelu konfesionalnu zajednicu jest istinit na bilo koji način, u bilo kojem smislu i za bilo koga izgleda potpuno neprihvatljivo. Uvjereni smo da mora postojati način da kažemo da taj sud nije istinit. I to da nije istinit ni za onoga tko se nalazi u odgovarajućem kontekstu i tko ima odgovarajuća motivacijska

¹⁶ Rasprava između internalizma i eksternalizma u etici vrlo je razvijena. Prikaz osnovnih uvida i argumenata može se naći u Berčić 2012: 277–281.

stanja. Želimo tvrditi da sud da treba istrijebiti Židove nije istinit ni kada ga izgovara sam Hitler. Stoga od semantičke teorije očekujemo da nam omogućí da kažemo da taj sud nije istinit. Apsolutist nema nikakvih problema s time jer ne relativizira treba-sudove.¹⁷ Za absolutiste istinsna vrijednost treba-sudova ne ovisi o kontekstu izricanja ni o motivaciji govornika. Oni su istiniti ili neistiniti *po sebi*, a ne u odnosu na ovo ili ono. Zbog toga absolutist može bez ikakvih problema reći da Hitlerov sud da treba istrijebiti Židove nije istinit. No, pitanje o kojem raspravljamo jest može li i relativist to reći? Može li, s jedne strane, tvrditi da su treba-sudovi relativni u odnosu na ovo ili ono, a s druge strane, tvrditi da sud da Židove treba istrijebiti nije istinit čak ni kada ga izriče sam Hitler?

Novi relativizam

Max Kölbel, John MacFarlane i ostali novi relativisti smatraju da uvođenjem parametra perspektive i razlikovanjem između konteksta izricanja i konteksta vrednovanja mogu to postići. Prema novom relativizmu jedan te isti sud može se vrednovati iz više perspektiva, u principu beskonačno mnogo perspektiva. Jedan te isti sud može biti istinit iz jedne perspektive, a neistinit iz neke druge. Tako Kölbel kaže:

“Parametar može biti moralni kod, skup principa, perspektiva itd. Jednostavnosti radi, recimo da se radi o relativnosti u odnosu na perspektivu. Genuini relativizam, dakle, jest gledište da istinsna vrijednost sadržaja koji su izraženi iskazima varira od perspektive do perspektive.” (Kölbel 2004: 306)

Ova pozicija, uvodeći parametar perspektive, omogućava nam da procjenjujemo istinsnu vrijednost bilo kojeg iskaza, ma tko ga i ma u kojim okolnostima izrekao. Tako nam relativistička formula $T(s,w,c,p)$ omogućava da kažemo da je Hitlerov sud da Židove treba istrijebiti bio neistinit čak i kada ga je on sam izgovarao:

¹⁷ Zanimljivo je da MacFarlane i apsolutističku ideju izražava na relativistički način. Apsolutizam u pogledu estetskog standarda definira kao tezu da *estetski sud ima istinsnu vrijednost u odnosu na Božji estetski standard*. (MacFarlane 2014: 67) Dakle, opet je ideja da sud ne može biti istinit po sebi ili naprosto nego da mora biti istinit u odnosu na neki standard.

“Čak ni Hitler ne može izbjeći fokus mojih treba-sudova: moj sud da nije trebao narediti istrjebljenje Židova može ne biti istina u odnosu na Hitlerovu perspektivu, ali to me ne sprječava da donesem taj sud, kao ni to da on bude točan u odnosu na moju perspektivu.” (Kölbel 2004: 308)

Može nam se učiniti da relativizam doista predstavlja napredak u odnosu na kontekstualizam. Pod pretpostavkom relativizma, prihvaćanjem parametra perspektive možemo reći da Hitlerov sud nije bio istinit, što pod pretpostavkom kontekstualizma navodno nismo mogli reći. No je li tu doista postignut napredak? I ako jest, koliki? Pod pretpostavkom relativizma možemo reći da Hitlerov sud nije istinit iz naše perspektive, ali ujedno moramo priznati i da je istinit iz njegove. Je li to napredak u odnosu na kontekstualizam? Naime, i kontekstualizam nam omogućava da tvrdimo da Hitler nije bio u pravu. I pod pretpostavkom kontekstualizma naš je sud istinit u odnosu na kontekst kojem mi pripadamo, a njegov je istinit u odnosu na kontekst kojem je on pripadao. Istina, relativizam nam omogućava da kažemo da Hitlerov sud nije bio istinit ni u kontekstu kojem je pripadao, što kontekstualizam ipak ne može dozvoliti. Dakle, iako novi ili analitički relativizam predstavlja daljnje rafiniranje opće relativističke teze, pitanje je koliki je njegov doseg i kolika je njegova plauzibilnost. Pravo pitanje jest može li opća ideja relativizma ikada biti formulirana na konzistentan i prihvatljiv način.

Literatura

- Baghrmian, M. i Carter, A. (2020): “Relativism”, *The Stanford Encyclopedia of Philosophy*, (ur.) E. N. Zalta. <<https://plato.stanford.edu/archives/fall2020/entries/relativism/>>
- Berčić, B. (1995): *Realizam, Relativizam, Tolerancija*, HKD Rijeka.
- Berčić, B. (2006): “Relativism”, *Bienal of Quadrilateral* (ed. Baccarini, E.) Museum of Modern and Contemporary Arts, Rijeka, 9–18.
- Berčić, B. (2012): *Filozofija 1*, Ibis-grafika, Zagreb.
- Berčić, B. (2019): “I Was Right!”, *Prolegomena* 18 (2), 209–225.
- Björnsson, G. i Finlay, S. (2010): “Metaethical Contextualism Defended”, *Ethics* Vol. 121, No. 1, 7–36.
- Brogaard, B. (2008): “Moral Contextualism and Moral Relativism”, *The Philosophical Quarterly*, Vol. 58, No. 232, 385–409.
- Garcia-Carpintero, M. i Kölbel, M., (ur.) (2008): *Relative Truth*, Oxford University Press.

- Harman, G. (1975): "Moral Relativism Defended", *Philosophical Review* 84 (1), 3–22, pretiskano u Shafer-Landau (ur.) 2013.
- Iacona, A. (2008): "Faultless or Disagreement", u Garcia-Carpintero i Kölbel (ur.) 2008.
- Kneer, M. (2017): "Predicates of Personal Taste: Empirical Data", rukopis, https://www.academia.edu/34561974/Predicates_of_Personal_Taste_Empirical_Data
- Kölbel, M. (2004): "Faultless Disagreement", *Proceedings of the Aristotelian Society* Vol. 104, 53–73.
- Kölbel, M. (2004): "Indexical Relativism versus Genuine Relativism", *International Journal of Philosophical Studies* Vol. 12(3), 297–313.
- Lasersohn, P. (2005): "Context Dependence, Disagreement, and Predicates of Personal Taste", *Linguistics and Philosophy* Vol. 28, 643–686.
- MacFarlane, J. (2003): "Future contingents and relative truth", *Philosophical Quarterly* 5.
- MacFarlane, J. (2014): *Assessment Sensitivity – Relative Truth and its Applications*, Oxford University Press.
- Shafer-Landau, R. (ur.) (2013): *Ethical Theory: An Anthology*, John Wiley & Sons.

Metafizika vremena i šteta smrti¹

Marin Biondić

Metafizika i priroda stvari

Suvremena filozofska rasprava o vrijednosti smrti – je li smrt dobra, loša ili vrijednosno neutralna za osobu koja umre – otvara čitav niz metafizičkih problema. Zapravo, čini se da bez određenih metafizičkih rješenja nema napretka u raspravi o vrijednosti smrti. Pitanje prirode zla², pitanje osobnog identiteta tj. trajanja u vremenu i prostoru i pitanje same prirode vremena, jesu ključna metafizička pitanja u raspravi o vrijednosti smrti. Odgovori na ta pitanja otkrivaju što je dobro/loše za određene kategorije bića, kada točno prestajemo postojati, a kada nastavljamo postojati kroz određene promjene, tj. kada samo mijenjamo (bitna) svojstva, a kada doživljavamo supstancijalnu promjenu i naravno prestajemo postojati. Od svih ovih pitanja, čini se da je pitanje prirode vremena po samoj svojoj prirodi najsloženije. Kada bismo znali stvarnu prirodu vremena, imali bismo fundamentalni metafizički okvir koji bi u konačnici mogao odrediti naš vrijednosni stav o smrti.

Općenito govoreći, otkrivanje stvarne prirode nečega može odrediti naše vrijednosne stavove prema tome. Npr. kada mala djeca otkriju da je ukusno, mirisno jelo koje jedu zapravo zec ili janje, mogu smatrati da je to nedopustivo ili odvratno te prestati jesti. Kada dijete otkrije prirodu stvari, to utječe na njegovu vrijednosnu prosudbu, naime, ona se mijenja

¹ Članak je nastao kao rezultat rada i suradnje na dva znanstvena projekta čiji sam član, redom to su: a) *Šteta, namjera i odgovornost* – Institut za filozofiju, Zagreb (Hrvatska zaklada za znanost), b) *Metametafizika* (Sveučilište u Rijeci).

² Engleski izraz jest "harm", hrvatski prijevod "šteta" zvuči pomalo nezgrapno pa u istom značenju ponekad koristim i izraz "zlo". "Šteta" ima specifično teoretsko značenje koje uključuje neiskustvo i neznanje subjekta štete o samoj šteti koju "trpi", a može uključivati i nepostojanje subjekta.

i u konačnici otkriće prirode stvari utječe na ponašanje. Otprilike dijete promišlja na sljedeći način: "Ukusna hrana jest *zapravo* ubijeni zec, to je odvratno i nedopustivo, to više neću jesti". Odrasli već znaju pravu prirodu stvari u tanjuru i čini se da im to nimalo ne smeta, ali djeca često ne znaju i to im, kada saznaju, često smeta. Dakle, otkriće *prirode stvari* ima potencijal promjene vrijednosnih sudova i djelovanja.

Pitanje je, dakle, kako nove informacije koje steknemo o temeljnoj strukturi stvarnosti mogu utjecati na naše vrijednosne stavove i u konačnici na samo djelovanje? Ako je cilj metafizike istražiti krajnju, fundamentalnu stvarnost, tj. ono što je u temelju svega i što "leži" iza pojavnosti, onda naše svakodnevno, zdravorazumsko shvaćanje svijeta može biti u opreci s metafizičkim teorijama. Način na koji doživljavamo svijet, razlikovat će se od načina na koji svijet stvarno jest, tj. svakodnevne činjenice bit će u opreci sa stvarnim stanjem stvari. Npr. predreflektivno vjerujemo u svakodnevne činjenice o postojanju crvenih jabuka koje jedemo, čvrstih stolova za kojima sjedimo i sl. i mislimo da dajemo opis stvarnosti, a zapravo, tvrdi se iz određene metafizičke pozicije, konačan opis stvarnosti uključuje samo atomske ili subatomske čestice i prazninu.³ Pitanje je koja je onda istinostna vrijednost rečenica kao što su "Postoje crvene jabuke" i "Postoje čvrsti stolovi"? Ako se konačni, fundamentalni, opis materijalnog svijeta može dati u terminima nedjeljivih čestica i praznine, jesu li onda te rečenice neistinite? Ako su te rečenice neistinite, onda se čini da bismo trebali prestati vjerovati da uistinu postoje crvene jabuke i čvrsti stolovi iako nam u svakodnevnom funkcioniranju takve tvrdnje mogu biti od koristi.⁴ Analogno tome, dakle pitanje je: mogu li metafizičke teorije vremena na sličan način utjecati na naše vrijednosne stavove o smrti? Mogu li metafizičke teorije vremena imati presudnu ulogu u tome može li smrt biti loša za osobu koja umre, ma što mi o tome mislili na svakodnevnom zdravorazumskom nivou?

³ Vidi van Inwagen 1990.

⁴ Za konciznu raspravu o problemu metafizičkih teorija i svakodnevnih činjenica vidi Sider 2013.

Metafizika vremena – prezentizam i eternalizam

Dvije su osnovne metafizičke teorije vremena, prezentizam i eternalizam.⁵ Prema prezentizmu, jedino što postoji, tj. što je stvarno sadašnje je vrijeme i njegov sadržaj. Možemo reći da osim 3D (trodimenzionalnoga) prostora i vremena ono što postoji jesu samo *sadašnji* objekti. Sadašnji objekti naprosto postoje, bez obzira na njihovu prostornu lokaciju i udaljenost. Stol za kojim sjedim, Barack Obama, Koloseum u Rimu, planet Jupiter i galaksija Andromeda naprosto postoje. Stol je u Rijeci, Barack Obama u SAD-u, Koloseum u Italiji, Jupiter u Sunčevom sustavu oko 870 milijuna kilometara od Zemlje, Andromeda u svemiru, udaljena od nas oko 2,538 milijuna svjetlosnih godina. Svi ti objekti postoje. Nije važno gdje su. Gdje god da nešto jest, ako je *sada*, ono postoji. Prema prezentizmu, prostornost i sadašnjost jesu kriteriji koji presuđuju postoji li nešto ili ne postoji.

Što ne postoji prema prezentizmu? Prema prezentizmu ne postoje prošli i budući objekti. Tako ne postoje dinosauri, ne postoji Sokrat, ne postoji Pangea, ne postoji ljudska postaja na Marsu. Nešto od ovoga nabrojenog jest postojalo, nešto će vjerojatno postojati, nešto je moglo postojati, ali sada ne postoji.⁶ Dakle, temeljna ideja prezentizma jest da sve ovo *sada* ne postoji. Drugim riječima, kada prezentist kaže da stvari prestaju postojati, on misli da stvari prestaju postojati *naprosto*, tj. da *naprosto ne postoje*. To znači da ono što je istinito u jednom vremenu, može biti neistinito u drugom vremenu. Tako je rečenica "Sokrat postoji" istinita u jednom vremenu (469. pr. Kr. – 399. pr. Kr.), ali nije istinita danas. Istinosna vrijednost rečenica varira ovisno kada ih izričemo.

S druge strane, prema eternalizmu jednako postoji prošlo, sadašnje i buduće vrijeme i njihov sadržaj. Bez obzira na vremensko određenje, prošlo, sadašnje ili buduće, jednako postoje Sokrat, Barack Obama ili budući unuk tek rođenog djeteta. Što se tiče postojanja, nije važno je li

⁵ Prezentizam ili još A-teorija ili 3D+vrijeme. Vidi npr. Markosian 2004. Eternalizam ili još B-teorija ili 4-dimenzionalizam. Vidi npr. Silverstein 2010.

⁶ Što je s postojanjem fiktivnih entiteta prema prezentizmu? Neki prezentisti kao Yourgrau 1987. negiraju *mogućnost* postojanja fiktivnih entiteta kao što su zlatne planine, Pegaz, magarac koji govori i sl., ali prezentist može smatrati da fiksijski i mogući entiteti iako ne postoje u vremenu, mogu postojati kao npr. apstraktni entiteti.

nešto u prošlosti, sadašnjosti ili budućnosti. Kao što objekti imaju prostorne dijelove, tvrde eternalisti, tako objekti imaju i temporalne dijelove. Kao što kažemo da *postoji* Jupiter, ali nije *ovdje*, tako prema eternalistima trebamo tvrditi da *postoji* Sokrat, ali nije *sada*. Eternalisti naime, u ontološkom smislu izjednačavaju prostornu udaljenost s vremenskom udaljenosti. Ako nešto nije prisutno u nekom sadašnjem vremenskom odsječku, to ne znači da naprosto ne postoji. Kako bismo jasno razumjeli eternalističku poziciju, valja razlikovati lokacijski i ontološki pojam postojanja. *Lokacijski* pojam postojanja možemo definirati kao “biti prisutan u nekom prostornovremenskom odsječku (t_1, t_2, t_3)” dok *ontološki* pojam postojanja možemo definirati kao “postojanje naprosto”.⁷ Tako svi mi, prema eternalizmu, koji smo sada prisutni imamo isti *ontološki* status kao Sokrat i Pangea, ali nemamo isti *vremensko-lokacijski* status. Mi, Sokrat i Pangea naprosto postojimo, jedina je razlika u tome da je naš prostorno-vremenski odsječak lociran u sadašnjem prostornovremenskom odsječku, a prostorno-vremenski dio Sokrata i Pangee nije. Dakle, kada eternalist kaže da stvari prestaju postojati, on misli da stvari prestaju biti u sadašnjem trenutku, ali i dalje u strogom ontološkom smislu postoje.

Vidjeli smo da za prezentiste istinitost rečenica varira s obzirom na vrijeme kada su rečenice izrečene. Tako su neke rečenice kojima tvrdimo postojanje određenih entiteta u jednom vremenu istinite dok u nekom drugom nisu. Za eternaliste istinitost je rečenica fiksna. Ako je nešto istinito, onda je uvijek istinito. Pogledajmo sljedeći par rečenica i manevar kojim se služe eternalisti:

“Barack Obama postoji.”

“George Washington postoji.”

Za prvu rečenicu, o Baracku Obami, možemo tvrditi da je ona istinita svugdje iako je on samo negdje, u SAD-u. Analogno tome, druga rečenica, o Georgeu Washingtonu, prema eternalistima je uvijek istinita iako je on trajao od t_1 do t_2 (1732. – 1799.) Gornje rečenice možemo zapisati na sljedeći način:

⁷ Za detaljan prikaz eternalističke i prezentističke pozicije vidi u Gilmore 2013. U ovom kontekstu “lokacijsko postojanje” ne znači puko prostorno postojanje, već se misli na “vremensko postojanje”.

“Barack Obama postoji iako ne postoji *ovdje*.”

“George Washington postoji iako ne postoji *sada*.”⁸

Dakle, eternalist tvrdi ako su rečenice istinite *svugdje* iako su njihovi istinosni uvjeti samo *negdje*, onda su i rečenice koje su istinite istinite *uvijek* iako su njihovi istinosni uvjeti samo u nekom *vremenskom odsječku*. Drugim riječima, ako je rečenica istinita, onda je ona naprosto uvijek istinita.

Možemo sumirati glavne odrednice prezentizma i eternalizma. Ako smo prezentisti, trebamo tvrditi da postoji 3D prostor i sadašnje vrijeme te njegov sadržaj, da ono što nije sada naprosto ne postoji i da rečenice koje su istinite u nekom vremenu mogu biti neistinite u nekom drugom vremenu ili uopće nemati istinosnu vrijednost. Ako smo eternalisti, onda trebamo tvrditi da postoji 4D prostor-vrijeme, tj. postoji jedan ogromni prostornovremenski entitet, tzv. blok svemir i njegov sadržaj, a rečenice koje su istinite uvijek su istinite.⁹

Metafizički okvir prezentizma i eternalizma, kakav je gore izložen, implicitno ili eksplicitno prisutan je u metafizičkim raspravama o vrijednosti smrti. Međutim, prije same primjene metafizičkih teorija vremena na raspravu o vrijednosti smrti ukratko ću izložiti glavne teoretske pozicije rasprave vrijednosti smrti te se zatim vratiti na primjenu metafizičkih teorija vremena u samoj raspravi.

Vrijednost smrti – šteta smrti

U metafizičko-vrijednosnim raspravama o smrti u suvremenoj analitičkoj filozofiji kristaliziralo se sljedeće pitanje:

“Je li smrt loša za osobu koja umre?”¹⁰

⁸ Vidi Silverstein 2013. u Taylor 2013. Čini se da bi za svijet bilo bolje da Donald Trump nije uključen ni u lokacijski ni u ontološki pojam postojanja.

⁹ Isto. Možemo reći da 4D implicira postojanje prošlog, sadašnjeg i budućeg vremena, ali samo kao psiholoških pojmova, objektivno postoji 4D blok-svemir u kojemu te ontološke razlike nema.

¹⁰ Ovim problemom opširno se bavim u knjizi *Smrt - priroda i vrijednost prenatalnog i postmortalnog nepostojanja* 2015.

Ovako formulirano vrijednosno pitanje u sebi sadrži nekoliko pretpostavki koje valja navesti. Prvo, pitanje se odnosi na perspektivu prvog lica, tj. je li smrt loša za samu osobu koja je umrla, a ne za druge osobe koje su povezane s njom ili za svijet u cjelini i sl. Drugo, kada se kaže "smrt", misli se na "stanje" bivanja mrtvim koje razumijevamo kao iskustvenu prazninu, a ne misli se na npr. proces umiranja koji sa sobom može nositi psihofizičku patnju. Treće, pretpostavlja se da nakon događaja smrti ne postoji nikakav oblik nastavka postojanja osobe, ni u materijalnom ni u nematerijalnom obliku. Drugim riječima, smrt predstavlja trajno i ireverzibilno uništenje bića. Ova pretpostavka formulira se kao teza terminacije:

TT: Ako živuća stvar (osoba) umre u nekom trenutku, tada ta živuća stvar *naprosto prestaje postojati* u tom trenutku.¹¹

Valja uočiti da je teza terminacije formulirana na način da se tvrdi da neki *x* *naprosto prestaje postojati*, za razliku od tvrdnje da *x* prestaje postojati kao nešto, a nastavlja postojati kao nešto drugo. U tom slučaju radilo bi se o nastavku postojanja jednog te istog entiteta, samo u drugom obliku dok teza terminacije govori o prestanku postojanja naprosto. Imajući u vidu navedene pretpostavke, filozofi analitičke tradicije, počevši od Thomasa Nagela¹², odgovaraju na pitanje:

"Je li smrt loša i ako jest, na koji način, za osobu koja umre pod pretpostavkom nepostojanja?"

Dvije su suprotstavljene teoretske struje koje problematiziraju gornje pitanje čiji su korijeni u antici. Jedna je epikurejska pozicija koja će braniti Epikurov stav da smrt nije ni dobra ni loša za osobu koja umre jer "dok postojimo mi nema smrti, a kada nastupi smrt više nema nas".¹³ Druga je aristotelijanska, koju nastavlja Thomas Nagel¹⁴, koja će smrt tretirati kao zlo "...a smrt je najstrašnija, ona je kraj..."¹⁵ Suvremeni branitelji Epikurove pozicije kao što su Rosenbaum, Suits i Taylor¹⁶ tako će u ko-

¹¹ Vidi Feldman 2000.

¹² Vidi Nagel 1970.

¹³ Epikur "Pismo Menoikeju" u Bošnjak 1983. (144–153).

¹⁴ Vidi Nagel 1970.

¹⁵ Aristotel Nikomahova etika (1115a).

¹⁶ Rosenbaum 1986, Suits 2001, Taylor 2013.

načnici tvrditi da smrt (bivanje mrtvim) nije *nikada* loša za osobu koja umre dok će suvremeni branitelji tvrdnje da je smrt loša za osobu koja umre tvrditi ili da je smrt *uvijek* loša za osobu koja umre¹⁷, ili u nešto blažem obliku, da smrt može biti loša za osobu koja umre, ali da može biti i dobra¹⁸, tj. da je smrt *ponekad* loša za osobu koja umre.

Temeljno opravdanje tvrdnje da smrt može biti loša za osobu koja umre poziva se na ideju deprivacije, tj. lišavanja. Tvrdi se da je smrt loša za osobu koja je umrla zato što je deprivira životnih iskustava koje je osoba mogla imati da nije umrla tada kada je umrla.¹⁹ Dakle, ovakvo kontrafaktično deprivacionističko opravdanje pripisuje subjektu svojstvo "biti depriviran dobara života". Naravno, deprivacija u slučaju smrti jest neiskustvena, to nije deprivacija kao npr. "biti depriviran odsijecanjem noge ili ruke". Međutim, prema deprivacionistima iskustvo nije važno, ako je netko depriviran tada je on depriviran, znao za to ili ne, iskusio to ili ne.

Ovo je problematična tvrdnja i epikurejci se s njom neće složiti. Dva su uvjeta koje epikurejci navode kao nužna i/ili dovoljna za pripisivanje vrijednosti. Prvi je uvjet iskustva, a drugi je uvjet postojanja. Prema prvom uvjetu pripisivanje vrijednosti subjektu moguće je akko (ako i samo ako) je moguće iskustvo; budući da se uništenjem subjekta uništava i *mogućnost* iskustva, tada nije moguće niti pripisivanje vrijednosti smrti za samoga subjekta. Dakle, nema mogućnosti iskustva, nema vrijednosti. Ovaj epikurejski uvjet deprivacionisti osporavaju pozivanjem na neiskustvena zla kao što su izdaja i prijevara za koja ne znamo, ali su i dalje loši za nas.²⁰ Drugim riječima, deprivacionisti tvrde da postoje neiskustvena zla i da je smrt jedno od tih zala.²¹ Međutim, čini se da je drugi uvjet uvjet s kojim će deprivacionisti puno teže izaći na kraj, tj. to je uvjet koji će nas uvući u metafiziku vremena. Iako možemo pripisivati osobama neiskustvenu vrijednost, tj. iako su stvari koje ne možemo *ni u*

¹⁷ Nagel 1970.

¹⁸ Feldman 1991, Williams 1973, Luper-Foy 1987.

¹⁹ Nagel 1970, Feldman 1991.

²⁰ Nagel 1970.

²¹ Epikurejci će tvrditi da neiskustvena zla mogu biti zla samo dok postojimo jer postoji *mogućnost* njihova iskustva, a kada umremo te mogućnosti više nema te nema ni zla.

principu iskusiti uistinu loše za nas, pitanje je kome uopće pripisujemo takvu vrijednost u slučaju smrti? Ako nema subjekta zla, tada se čini da ne može biti ni logički konzistentne kontrafaktičke komparacije kakvu iznose deprivacionisti. Prema uvjetu postojanja, pripisivanje vrijednosti moguće je akko postoji subjekt kojemu pripisujemo određenu vrijednost, u slučaju smrti trebao bi postojati postmortalni subjekt kojemu možemo pripisati vrijednost, ali to nije slučaj budući da se u raspravi prihvaća teza terminacije kao istinita teza, tj. osoba kada umre, *naprosto* prestaje postojati.

Čini se da centralni problem, problem subjekta kako ga je postavio Epikur, nema zadovoljavajuće rješenje prema deprivacijskoj teoriji. Ono što ne postoji ne može imati nikakva svojstva. Mrtvi ne postoje, dakle ne mogu imati svojstva pa tako ni svojstvo “biti depriviran dobara života”.²² Deprivacijska teorija *pretpostavlja* da postoji subjekt kojemu pripisujemo svojstvo “biti depriviran dobara života”. Međutim, to je moguće samo u određenom metafizičkom okviru. Može li se pripisati svojstvo “biti depriviran dobara života” ovisi upravo o tome koju metafiziku vremena zastupamo, prezentističku ili eternalističku. Pogledajmo sada na koji način metafizičke teorije vremena pokušavaju riješiti problem nepostojećeg subjekta i time opravdati tvrdnju o šteti smrti za osobu koja umre.

Eternalizam – nepostojeći subjekt i šteta smrti

Eternalist tvrdi da objekti koji su u vremenu *naprosto* postoje iako nisu sada. Tako i mrtve osobe postoje iako nisu sada. Mrtvi imaju isti ontološki status kao i prostorno udaljeni planeti. Jupiter postoji iako nije ovdje, mrtvi postoje iako nisu sada. Dakle, imamo subjekta kome možemo pripisivati svojstva pa tako i svojstvo “biti depriviran dobara života”. Ako možemo pripisati svojstvo “biti depriviran dobara života”, onda i smrt može biti loša za osobe. Drugim riječima, eternalizam nam vraća subjekta u raspravu i omogućava pripisivanje štete tom subjektu. Kada to ne bi bio slučaj, završili bismo u ekstremnoj formi epikureizma, prema kojoj ne samo da smrt nije loša za osobu koja umre, već bi i sami

²² Vidi Sider 2013, (161).

iskazi o takvim "osobama" bili besmisleni jer bez referenta ne možemo uopće vršiti vrednovanje.²³

Problem s eternalističkim rješenjem nepostojećeg subjekta u raspravi o vrijednosti smrti prvenstveno je u samom prihvaćanju eternalističke teorije. Mnogi filozofi koji smatraju da je smrt loša za osobu koja umre ne prihvaćaju 4D metafiziku prostor-vremena, npr. Yourgrau kaže da "...za objekte u vremenu što ne postoji sada, ne postoji uopće" i smatra da specijalizacija vremena nije uopće provediva.²⁴ Jednako tako prominentni autori obrane teze štete smrti, kao Nagel i Feldman, ne zastupaju eternalističku poziciju kako bi branili tvrdnju da je smrt loša. Čini se da većina filozofa i osoba koje smatraju da je smrt loša na predreflektivnom nivou to smatraju bez obzira na eternalističku metafiziku. Uvođenje eternalizma *ad hoc* u raspravu o vrijednosti smrti, samo kako bismo se "izvukli" iz problema nepostojećeg subjekta, ne može biti opravdano. Možda je eternalizam istinit, ali ako je istinit, onda je istinit neovisno od rasprave o vrijednosti smrti. Pretpostavljati njegovu istinitost samo kako bi se mogla obraniti teza štete smrti znači uzeti za istinito ono što treba dokazati. Nadalje, čak i kada bismo prihvatili eternalizam, epikurejci se mogu povući na rezervnu poziciju uvjeta iskustva i nastaviti raspravu (što i čine u takvim slučajevima) bez obzira na eventualno postojanje subjekta kojemu možemo pripisivati svojstva. Taj subjekt nema mogućnost iskustva deprivacije te u konačnici ni smrt ne može biti loša za njega.²⁵

Međutim, pitanje je može li se deprivacionistička pozicija konzistentnije braniti u prezentističkom metafizičkom okviru? Nije li slučaj da Epikurov argument svoju snagu ima upravo u okviru prezentizma? Tvrdnja "Dok smo mi nema smrti, kada dođe smrt nema nas" pretpostavlja upravo prezentističku metafiziku vremena.

²³ Silverstein smatra da se problem nepostojećeg subjekta može riješiti isključivo u eternalističkom okviru, u suprotnom bismo trebali prihvatiti neki oblik epikurizma u pogledu smrti ili da smrt nije loša za osobu koja umre, ili da ne može uopće biti koherentnog vrednovanja. Vidi Silverstein 1980. i 2000.

²⁴ Yourgrau 1987. u Fischer 1993. (141).

²⁵ Vidi Rosenbaum 1986. Iz eternalističke perspektive moglo bi se tvrditi da je rasprava o šteti smrti psihološke prirode, a ne metafizički problem (recenzent ovoga rada). Međutim, to se odnosi na metafizički problem postojanja subjekta, ali i dalje opstoji navedeni problem vrijednosti tj. štete i nemogućnosti iskustva.

Prezentizam – nepostojeći subjekt i šteta smrti

Prisjetimo se, prema prezentizmu mrtvi naprosto ne postoje. Mrtvi imaju isti ontološki status kao i bilo koji drugi nepostojeći entiteti (prošli, budući, mogući). Svi prošli entiteti, svi budući entiteti i svi mogući entiteti koji nisu aktualni *naprosto* ne postoje. Mi možemo govoriti o “njima”, ali naše rečenice koje se pozivaju na takve entitete ne mogu biti istinite. Ne postoje referenti koji korespondiraju izrazima u takvim rečenicama. Dakle, nema subjekta kojemu možemo pripisati ikakvo svojstvo pa tako ni svojstvo “biti depriviran dobara života”. Samim time, smrt ne može biti loša za mrtve osobe zato što *ne postoje* mrtve osobe. I ne samo to, već se čini da nas prezentistička metafizika dovodi do kontradiktornosti u samim iskazima. Pogledajmo sljedeće rečenice koje svakodnevno možemo čuti i izgovarati:

“Mojsije je mrtav, tj. on ne postoji.”²⁶

“Mrtvi jesu oštećeni zato što ne postoje.”

U prvoj rečenici tvrdimo nešto o Mojsiju – to da on ne postoji. Ali samim time pripisujemo svojstvo nepostojanja *nekome*. Impliciramo da Mojsije na neki način postoji kako bismo mu pripisali svojstvo nepostojanja.²⁷ Kako je to moguće? Čini se da je ovo kontradiktorno. Ne može nešto u isto vrijeme postojati i ne postojati. Kao da kažemo “Mojsije postoji i ne postoji”.

Isti je slučaj i u drugoj rečenici. Tvrdimo da su mrtvi oštećeni zato što ne postoje. Ali samim time pripisujemo svojstvo nepostojanja *nekome*. Čini se da taj “netko” na neki način postoji pa mu onda pripisujemo svojstvo nepostojanja. Kao da kažemo “Mrtvi postoje i ne postoje”. Problem prošlih singularnih iskaza notorni je problem prezentizma u koji neću ovdje dublje ulaziti, ali ono što je relevantno za raspravu o vrijednosti smrti jest da se čini da prezentizam čini ovakve iskaze besmislenima.²⁸ S druge strane, čini nam se da su rečenice kao što je “Mojsije je mrtav” istinite rečenice, odnosno da činjenica ili propozicija koja je izre-

²⁶ Primjer s Mojsijem koristi Yourgrau 2013. u Taylor 2013.

²⁷ Isto.

²⁸ Za moguća rješenja kontradiktornih iskaza i formuliranje prošlih singularnih iskaza u okviru prezentizma vidi Markosian 2004.

čena tom rečenicom postoji, tj. da je istinita. Trebamo li tvrditi da “ako činjenica o Mojsiju postoji, onda i Mojsije mora postojati” ili to da “ako Mojsije ne postoji, onda i činjenica o Mojsiju ne postoji”?²⁹ Za Pallea Yourgraua očito je da činjenica o Mojsiju postoji, tj. da je propozicija o Mojsiju istinita, a samim time i Mojsije mora imati određeni ontološki status.³⁰ Međutim, Yourgrau je prezentist i smatra da je eternalizam naprosto pogrešan te Mojsije ne može postojati jer postoji samo ono što je sada. Ono što je relevantno u ovome kontekstu to je da Yourgrau smatra da je smrt loša za osobu koja umre te da je problem subjekta rješiv upravo u prezentističkom okviru. Jednako kao što Mojsije ima ontološki status, tako ga imaju i mrtvi, ali nikako u eternalističkom metafizičkom okviru. Kako je to moguće?

Yourgrau zastupa modalnu metafiziku bivanja čineći distinkciju između *bivanja* i *postojanja*.³¹ “Biti” i “postojati” dvije su odvojene ontološke kategorije. Prema Yourgrauu, jedna je stvar “biti objekt”, a druga je stvar “biti postojeći objekt”. U ovakvom metafizičkom okviru postojanje nije ništa doli svojstvo realno bivstvjućih entiteta koji mogu *biti*, a da *ne postoje*. Postojanje je predikat bivanja. Na taj način realno bivajući entiteti mogu imati svojstvo postojanja ili nepostojanja. Za Yourgraua očito je da ne postoje prošli i budući entiteti zato što sve što postoji, postoji sada, ali to ne znači da oni uopće nisu. Oni *jesu* samo što *ne postoje*. Prošla, sadašnja i buduća bića, jesu jednako stvarna, ali neka od njih ne postoje, neka postoje, neka će postojati, neka nikada neće postojati, a neka će postojati i ponovo neće postojati. Svi oni jesu, a samo neki imaju svojstvo postojanja.

Jednako tako mrtvi ne postoje, ali jesu. Mrtvi, živi i nerođeni jesu u istoj ontološkoj kategoriji bivanja, ali samo živi imaju svojstvo postojanja. Sve su to *moгуća* bića, a samo neka postoje. I mi koji sada živimo jesmo *moгуća* bića jer da nismo *moгуća*, ne bi nas ni bilo. Mogućnost je realitet koji je preduvjet postojanja, tj. modalni princip kojega zastupa Yourgrau glasi:

²⁹ Vidi Yourgrau 2013. u Taylor 2013.

³⁰ Isto.

³¹ Eng. “being” i “existence”. U logičkom zapisu $(\exists x) \leftrightarrow (E!x)$ “There is an object, and that object is nonexistent”, Yourgrau u Fischer 1993. (142).

MP: Mogućnost nužno prethodi aktualnosti.³²

Dakle, sama mogućnost daje realitet bićima.³³ Moguća bića prema Yourgrauu možemo, dakle, razvrstati u četiri kategorije:

- moguća bića koja nikada neće postojati,
- moguća bića koja će postojati,
- moguća bića koja postoje,
- moguća bića koja su mrtva, tj. koja su postojala, ali više ne postoje.

Sve rečenice koje govore o ovim kategorijama bića u ovom ontološkom okviru jesu smislene rečenice jer su moguća bića realna bića. Sve takve rečenice imaju svoje referente. “Mojsije” ima svog referenta, “Mrtvi” imaju svoje referente. Sve propozicije o tim referentima mogu biti istinite ili neistinite. Dakle, imamo subjekta kojemu možemo pripisivati svojstva pa tako i svojstvo “biti depriviran dobara života”.

Iz ovoga ne slijedi samo to da možemo smisleno tvrditi da su mrtvi “deprivirani dobara života”, već to da smisleno možemo tvrditi da su svi oni koji ne postoje “lišeni postojanja” uopće. Ne samo da su mrtvi lišeni postojanja, već su i nerođeni lišeni postojanja, a tu spadaju i oni koji nikada neće biti rođeni. Yourgrau smatra da upravo oni “trpe” najveću štetu jer nikada neće postojati. Što manje postojanja za realna bića, to više štete. Maksimum nepostojanja, tj. nepostojanje uopće, maksimalna šteta. Prema ovoj slici, čini se kao da beskrajan broj bića čeka u redu na svoje postojanje, a bića ulaze i izlaze iz postojanja. Ontološka realnost bivanja vezana je uz modalnu dimenziju, a ne uz prostornu i vremensku. Sve ovo može zvučati čudno. I zvuči čudno. Ali to su posljedice ovakve prezentističko-modalističke ontologije u kojoj “imamo” subjekta štete smrti.

Fundamentalna metafizika vremena i vrijednosti

Vidjeli smo da prema eternalizmu, postoji 4D prostorno-vremenski blok svemir i njegov sadržaj, a samim time mrtvi ulaze u ontološku kategoriju postojanja. Ono što nije sada postoji iako trenutno nije prisutno

³² Yourgrau 1987. u Fischer 1993.

³³ Valja naglasiti da fiktivni likovi prem Yourgrauu nisu realna bića.

pa tako i mrtvi postoje iako nisu sada. Rečenice o mrtvim osobama jesu smislene i imaju svoje referente. To nisu samo prazne riječi, to su rečenice koje imaju istinosne uvjete. Tezu terminacije u eternalističkom metafizičkom okviru treba, dakle, formulirati na sljedeći način:

TTE: Ako živuća stvar (osoba) umre u nekom trenutku, tada ta živuća stvar *prestaje biti prisutna* u sadašnjem trenutku (ali i dalje postoji).³⁴

S druge strane, vidjeli smo da prema prezentizmu postoji jedino sadašnje vrijeme i njegov sadržaj, a što nije sada, naprosto ne postoji, a samim time i mrtvi *naprosto ne postoje*. Rečenice kojima tvrdimo nešto o mrtvim (nepostojećim) osobama nisu istinite rečenice zato što nemaju svog referenta i samim time nemaju istinosnih uvjeta. Tezu terminacije u prezentističkom metafizičkom okviru, dakle, formuliramo u njenom ishodišnom značenju:

TTP: Ako živuća stvar (osoba) umre u nekom trenutku, tada ta živuća stvar *naprosto prestaje postojati* u tom trenutku.

Budući da u metafizičkoj raspravi o prirodi vremena nema konkluzivnih argumenata koji bi presudili u korist eternalizma i da postoji niz autora koji plauzibilno brane prezentizam, onda je nadogradnja prezentizma modalnom ontologijom, kakvu iznosi Yourgrau, plauzibilan način na koji teza terminacije neće "uništiti" subjekta te ćemo moći konzistentno tvrditi da je smrt loša i za one koji ne postoje.³⁵

Dakle, ako smatramo da bez rješenja problema nepostojećeg subjekta ne možemo braniti tvrdnju da je smrt loša za osobu koja umre, onda se čini da smo prisiljeni braniti eternalizam ili prezentizam koji je nadopunjen modalnom ontologijom. Ukoliko odbacimo eternalizam ili modalnu ontologiju u okviru prezentizma, čini se da gubimo subjekta i da je epikurejska pozicija u konačnici istinita, tj. da je problem subjekta nepremostiv problem.

Međutim, mnogi nisu spremni prihvatiti "ontološku prtljagu" eternalizma i Yourgrauovog tipa prezentizma. Čini se da je cijena koju je

³⁴ Iz eternalističke perspektive preciznije je reći "vremenski dio živuće stvari" umjesto "živuća stvar".

³⁵ Za obranu prezentizma vidi npr. Markosian 2004.

potrebno platiti kako bi se riješio problem subjekta prevelika. Ovo nas vraća na početno pitanje o odnosu metafizičkih teorija i svakodnevnih činjenica. Može li se teza štete smrti opravdati *isključivo* pod pretpostavkom istinitosti eternalizma ili Yourgrauovog tipa prezentizma ili njeno opravdanje može ostati na nivou svakodnevnih zdravorazumskih činjenica?

Naime, kao što sam već ranije spomenuo, čini se da je ono što veliki broj deprivacionističkih filozofa misli kada brane tezu da je smrt loša za osobu koja umre, to da je loša *bez obzira* na određenu metafizičku teoriju vremena. Pogledajmo Feldmanov kontrafaktički princip štete kojim se brani teza štete smrti:

KPŠ: Svaki događaj koji čini život osobe *x* sveukupno gorim nego što bi on bio da se taj događaj nije dogodio jest šteta za osobu *x*.³⁶

Ako je taj događaj smrt (ili bivanje mrtvim), onda smrt može učiniti život sveukupno gorim nego da se nije dogodila. Drugim riječima, da osoba *x* nije umrla tada kada je umrla, život bi joj bio sveukupno bolji nego što aktualno jest.³⁷ Ovakva formulacija, prema Siderovu mišljenju, ne pretpostavlja nikakav vremensko-metafizički okvir, već je neutralna prema fundamentalnoj metafizici.³⁸ Tvrdnja “da osoba *x* nije umrla tada kada je umrla život bi joj bio sveukupno bolji nego što aktualno jest” svakodnevna je činjenica, ako je činjenica. Naime, tvrdi Sider, svakodnevne činjenice ostaju svakodnevne činjenice bez obzira na fundamentalnu realnost koja im stoji u podlozi. To da postoje crvene jabuke i čvrsti stolovi jest svakodnevna činjenica bez obzira na to što je npr. fundamentalna fizička realnost svodiva na subatomske čestice i prazninu. Prema ovom gledištu, fundamentalna realnost stoji u podlozi svakodnevnih činjenica, ali ona ne “poništava” svakodnevne činjenice i ne treba imati utjecaj na vrijednost koju pripisujemo tim činjenicama. Npr. ako se konačni opis ljudskih tijela može dati u terminima subatomske čestice, to ne znači da ne postoji svakodnevna zdravorazumska činjenica o postojanju ljudi i da nam je zbog toga manje stalo do naših bližnjih ili pak da se može tvrditi da nema smrti, već postoji samo reorganizacija subatom-

³⁶ Vidi Feldman 1991.

³⁷ Smrt lišava mogućeg dobra, koje god to bilo, da se nije dogodila kada se dogodila.

³⁸ Vidi Sider 2013. u Bradley, Feldman, Johansson 2013.

skih čestica. Ono što Sider zapravo tvrdi jest da moramo paziti na kojim razinama iznosimo tvrdnje. Tvrdnja o šteti smrti za osobu koja umre tako je tvrdnja o svakodnevnom činjenicama, a ne tvrdnja o fundamentalnoj realnosti, a samim time čini se da pozivanje na fundamentalnu realnost ne dovodi u pitanje tu činjenicu i njenu vrijednost: "...moramo imati na umu razliku između svakodnevnih činjenica i fundamentalne stvarnosti kada evaluiramo etičke implikacije metafizičkih teza".³⁹ Dakle, pitanje je zapravo koji je odnos fundamentalnih metafizičkih teorija i vrijednosti svakodnevnih činjenica?

S jedne strane, autori kao što su Silverstein i Yourgrau, kao što smo vidjeli, smatraju da se opravdanje štete smrti za osobu koja umre ne može dati bez pozivanja na fundamentalnu metafiziku vremena (eter-
nalistička u Silversteinovu slučaju ili prezentističko-modalna u Yourgrauovu slučaju). S druge strane, Feldman prilikom opravdanja teze štete smrti nigdje ne pretpostavlja istinitost ni jedne određene metafizike vremena dok Sider eksplicitno tvrdi da to uopće nije potrebno jer vrijednost svakodnevnih činjenica ne ovisi o fundamentalnoj metafizici. Prema tome, određuje li fundamentalna metafizika vrijednost svakodnevnih činjenica, kao što je činjenica o zlu smrti – ako je činjenica?

Čini mi se da u slučaju rasprave o vrijednosti smrti za osobu koja umre ne možemo metafizičke teorije postaviti u zagrade i naprosto tvrditi da vrijednost svakodnevnih činjenica ne ovisi o njima. Čitava je rasprava o vrijednosti smrti od 1970. pa nadalje strukturirana kao odgovor na Epikurov stav da smrt nije ništa za osobu koja umre. Taj Epikurov vrijednosni stav upravo proizlazi, među ostalim, iz pretpostavljanja istinitosti prezentizma, tj. kod Epikura fundamentalna metafizika određuje vrijednost svakodnevnih činjenica. Vrijednost je smrti uništena jer je uništen subjekt u prezentističkom metafizičkom okviru. To da mi naprosto prestajemo postojati kada umremo, dakle postoji samo ono što je sada, prema Epikuru povlači za sobom to da smrt nije ništa za osobu koja umre. Kada bi fundamentalna metafizika bila drukčija, i vrijednosti bi bile drukčije. Dakle, iako veza između fundamentalnih metafizičkih teorija i vrijednosti nije nekakve automatske, mehaničke prirode, mislim da promjene na metafizičkoj razini imaju direktan utjecaj na vrijednosti, barem u slučaju rasprave o vrijednosti smrti.

³⁹ Isto (163).

Zaključak

Smatram da nema u potpunosti koherentnog opravdanja teze štete smrti bez određene fundamentalne metafizike vremena, kao što to i smatraju Silverstein i Yourgrau, međutim za razliku od njih, ne vidim rješenje u eternalizmu ili prezentizmu Yourgrauova tipa upravo zbog "ontološke prtljage" ovih teorija. Pitanje je može li se opravdanje teze štete smrti dati u okvirima prezentizma ili pak trebamo prihvatiti stav da problem subjekta nema zadovoljavajuće rješenje na metafizičkoj razini i da je Epikur u konačnici bio u pravu? Drugim riječima, ako nema rješenja problema nepostojećeg subjekta na fundamentalnoj metafizičkoj razini, onda nema ni opravdanja vrijednosti svakodnevnih činjeničnih iskaza o šteti smrti za osobu koja umre.

Literatura

- Aristotel, 1988. *Nikomahova etika*, Zagreb: Globus.
- Biondić, M., 2015. *Smrt-priroda i vrijednost prenatalnog i postmortalnog nepostojanja*, Zagreb: Ibis-grafika.
- Epikur, 1983. Pismo Menoikeju, u Bošnjak, ur., *Od Aristotela do renesanse*, 144–153.
- Feldman, F., 1991. "Some Puzzles About the Evil of Death," *The Philosophical Review*, 100: 205–27; u Fischer 1993, 307–326.
- Feldman, F., 2000. "The Termination Thesis," *Midwest Studies in Philosophy*, 24: 98–115.
- Fischer, J. M., ed., 1993. *The Metaphysics of Death*, Stanford: Stanford University Press.
- Gilmore, C., 2013. "When Do Things Die?" u Bradley, Feldman, Johansson 2013, 5–60.
- Luper-Foy, S., 1987. "Annihilation," *The Philosophical Quarterly*, 37 (148): 233–52; u Fischer 1993, 267–290.
- Markosian, N., 2004. "A Defence of Presentism" *Oxford Studies in Metaphysics*, 1:47–82.
- Nagel, T., 1970. "Death," *Noûs*, Vol. 4, 1: 73–80; u Fischer 1993, 61–69.
- Rosenbaum, S., 1986. "How to Be Dead and Not Care: A Defense of Epicurus," *American Philosophical Quarterly*, 23 (2): 217–25; u Fischer 1993, 119–134.
- Silverstein, H., 1980. "The Evil of Death," *Journal of Philosophy*, 77 (7): 401–424; u Fischer 1993, 95–116.

- Silverstein, H., 2000. "The Evil of Death Revisited," *Midwest Studies in Philosophy*, 24: 116–135.
- Silverstein, H., 2010. "The Time of the Evil of Death" u *Time and Identity*, ur. Rourke, Campell, Silverstein, Cambridge MA: MIT Press.
- Suits, D., 2001. "Why Death Is Not Bad for the One Who Died," *American Philosophical Quarterly*, 38 (1): 269–84; u Benatar 2010, 267–287.
- Sider, T., 2013. "The Evil of Death: What Can Metaphysic Contribute", u Bradley, Feldman, Johansson 2013, 155–166.
- Taylor, J.S., 2013. *The Metaphysics and Ethics of Death*, New York, Oxford University Press.
- van Inwagen, P., 1990. *Material Beings*, Ithaca, NY: Cornell University Press.
- Williams, B., 1973. "The Makropulos Case: Reflections on the Tedium of Immortality," u Williams ur., *Problems of the Self*, Cambridge: Cambridge University Press; u Fischer 1993, 71–92.
- Yourgrau, P., 1987. "The Dead," *Journal of Philosophy*, 86 (2): 84–101; u Fischer 1993, 137–156.
- Yourgrau, P., 2013. "Kripke's Moses", u Taylor 2013, (134–149).

Je li problem slobode volje pseudoproblem?¹

Filip Čeč

Što se misli reći kada se kaže da je problem slobode volje pseudoproblem? Što uopće jest pseudoproblem? Kada netko izriče takvu tvrdnju? Iz kojeg bi razloga netko uopće tvrdio takvo što? U radu ću prikazati stajališta pojedinih filozofa koji su iznijeli takve tvrdnje te ponuditi analizu takvog pristupa problemu slobode volje te ću nastojati ponuditi odgovor na nekoliko pitanja koja se nameću: tvrde li različiti filozofi različite stvari kada tvrde da je problem slobode volje pseudoproblem? Odnosno, javlja li se u raspravi samo jedna ili više shvaćanja slobode volje kao pseudoproblema? Je li moguće razvrstati te interpretacije u nekoliko srodnih, ali ipak različitih skupina? Govorimo li u tim slučajevima zaista o pseudoproblemu ili se u raspravi odvija nešto drugo? Naposljetku, pomoću koje argumentacijske metodologije možemo odbaciti optužbe da se u raspravi o slobodi volje bavimo pseudoproblemom?

U prvom dijelu rada oslikat ću narav problema, u drugom dijelu prikazat ću kako su pojedini klasični kompatibilisti argumentirali da sloboda volje jest pseudoproblem te ću ponuditi argumente kojima se takvo stajalište može dovesti u pitanje, u trećem dijelu prikazat ću Balaguerovo inovativno tretiranje problema slobode volje kao pseudoproblema s libertarijanističkog pijedestala te ću u četvrtom poglavlju ponuditi razloge zašto smatram da je takav tretman problema slobode volje neutemeljen. U konačnici ponuđene bi kritike trebale pokazati da problem slobode volje ne predstavlja pseudoproblem.

¹ Rad je napisan u sklopu znanstveno-istraživačkog projekta *Metametafizika* Sveučilišta u Rijeci (šifra projekta: uniri-human-18-239).

I

Prvi korak jest utvrditi što uopće jest problem slobode volje. Mark Balaguer svoj će napad na taj problem započeti tradicionalnom formulacijom problema slobode volje koja je dana u obliku inkonzistentne trijade:

- (1) Determinizam je istinit (svaki je događaj uzročno određen od strane ranijih događaja zajedno s uzročnim zakonima);
- (2) Ljudska bića posjeduju slobodu volje;
- (3) Sloboda volje inkompatibilna je s determinizmom. (Balaguer, 2010: 1–2)

Takav oblik formulacije problema slobode volje uobičajen je i nalazimo ga u brojnim uvodnim udžbenicima iz filozofije u kojima prezentacija rasprava započinje upravo na gore navedenom ili ponešto izmijenjenom primjeru. Tako primjerice u svojem uvodu u filozofiju *Philosophical Problems and Arguments: An Introduction* Cornman, Lehrer i Pappas tvrde:

Koji paradoks povezujemo s problemom slobode i determinizma? Sljedeći: determinizam je teza univerzalne uzročnosti, teza prema kojoj je sve uzrokovano. S druge strane, doktrina slobode smatra da su neki naši postupci slobodni. Obje se teze čine istinitima. Mi vjerujemo da je sve uzrokovano i da su neki naši postupci slobodni. No ova vjerovanja dovode nas do inkonzistentnog rezultata. Jer ako je sve uzrokovano, onda su i naši postupci, koje smatramo slobodnima, uzrokovani. ... U tome se sastoji paradoks naših svakodnevnih vjerovanja. (Cornman, Lehrer, Pappas 1992: 91)

Svaka od triju gore navedenih tvrdnji ima svoju inicijalnu plauzibilnost utemeljenu na određenim *prima facie* razlozima zbog kojih ćemo i biti skloniji pojedinu opciju inkonzistentne trijade izdvojiti kao pogrešnu.²

Vratimo se ranije navedenim pitanjima: što se misli reći kada se kaže da je problem slobode volje pseudoproblem? Kako tada postupiti? Zašto bi netko uopće tvrdio da se susrećemo s pseudoproblemom? Što uopće

² U okvirima tradicionalne debate razrješenje inkonzistentne trijade vrši se tako da se prihvaćaju dvije tvrdnje, a jedna odbacuje. Odbacivanje prve premise implicira prihvatanje libertarijanizma, odbacivanje druge premise vodi ka skepticizmu po pitanju postojanja slobode volje, a odbacivanje treće premise postavlja nas u kompatibilistički tabor. U sofisticiranim suvremenim raspravama moguća su i drugačija rješenja.

jest pseudoproblem? Nekoliko se odgovora može ponuditi na posljednje pitanje od kojih bih izdvojio tri moguća shvaćanja koja se mogu primijeniti kod klasificiranja statusa rasprave o problemu slobode volje³: pseudoproblem kao besmislica, pseudoproblem kao trivijalna semantička analiza pojmova te pseudoproblem kao pogrešno korištenje pojmova.

Filozofi skloni verifikacionističkom programu Bečkog kruga težili bi jednoj, nazovimo je oštroj, striktnoj interpretaciji pseudoproblema. Pseudoproblem jest problem koji je formuliran na temelju pseudoiskaza, pri čemu pseudoiskaz jest “niz riječi koji ima oblik iskaza, ali nema smisla, tj. značenja” (Berčić 2002: 41). Ne radi se o tome da su iskazi koji se koriste radi prikaza problema istiniti ili neistiniti, naprotiv – stav koji se zauzima znatno je oštrije, korišteni iskazi za opisivanje problema naprosto su – besmisleni. Lišeni značenja. Što napraviti s takvim problemom? Ukoliko taj problem uistinu jest pseudoproblem, tada ga valja eliminirati iz filozofske rasprave. Učiniti pravu stvar, spriječiti nepotrebno gubljenje vremena i truda te pročistiti filozofiju nepotrebnih besmislica. Klasičan primjer takve argumentacije predstavlja Carnapova kritika Heideggera.⁴

Neke se rasprave mogu okarakterizirati kao trivijalne, metafizički isprazne ukoliko te rasprave ne mogu doprinijeti spoznaji svijeta koji nas okružuje, odnosno ako se prezentiraju kao jedna vrsta analize koja želi konstatirati što se pod određenim pojmom misli i je li ponuđena analiza pojma koherentna. Za takve se rasprave može reći da predstavljaju pseudoproblem shvaćen kao trivijalna semantička analiza pojmova. Nerijetko pojedine rasprave mogu stvoriti privid da su metafizički relevantne, ali u pojedinim slučajevima tome nije tako. Primjerice ukoliko se utvrdi da se ne radi o raspravi koja može doprinijeti spoznaji svijeta koji nas okružuje ili ako se radi o nekoj vrsti tričavog nadmudri vanja oko značenja pojedinog pojma, onda takve rasprave valja odbaciti kao irelevantne.

Manje radikalna interpretacija što jest pseudoproblem mogla bi se formulirati na način da se tvrdi da je pseudoproblem problem koji je

³ Pri tome ne smatram da je ponuđena interpretacija pseudoproblema potpuna, no ona jest dovoljna ukoliko se promatra u sklopu rasprave o slobodi volje kao pseudoproblemu.

⁴ Radi detaljnijeg prikaza pristupa filozofa Bečkog kruga pseudopojmovima u metafizici vidi: Berčić, *Filozofija Bečkog Kruga*, 40–58.

nastao zbog pogrešnog shvaćanja ili korištenja određenog pojma ili pogrešno postavljene pretpostavke te se zbog te inherentne greške ne može niti ponuditi odgovarajući odgovor na postavljeno pitanje ili se zbog te greške generira problem kojeg zapravo nema. Takvo shvaćanje nije radikalno poput ranije navedenih shvaćanja jer se problem više ne tretira kao besmislica lišena značenja ili kao trivijalnost, nego se radi o pogrešnom razumijevanju ili interpretaciji ili korištenju smislenih pojmova.

Može li se rasprava o slobodi volje podvesti pod jedan od gore navedenih interpretacija što pseudoproblem jest?

II

U raspravi o slobodi volje stav da rasprava u svojoj biti predstavlja pseudoproblem višestruko je bio potenciran od strane klasičnih kompatibilista. Prema mišljenju klasičnih kompatibilista njihova teorija svoju plauzibilnost duguje činjenici da je povezana sa zdravorazumskim promišljanjem svijeta, običnim jezikom, pučkopsihološkim poimanjem, kao i uvažavanjem empirijskih činjenica koje suvremena znanost nudi. Stoga teret dokaza leži na suprotnim, kontraintuitivnim pozicijama koje su optužene da se pozivaju na *ad hoc* rješenja, općenito nerazumijevanje ili krivo baratanje osnovnim pojmovima u debati, odnosno optužuje ih se da se bave “opšurnom i paničnom metafizikom” (Strawson 1962: 36) koja brzopleto nastoji spasiti što se spasiti može.

Kada prikazuju kritike inkompatibilizma, odnosno libertarijanizma koji se optužuje za pseudoproblemsku argumentaciju, McKenna i Pereboom u svojem udžbeniku izdvajaju četiri klasična autora koji su pristupili takvom radikalnom odbacivanju argumenata koji suprotna strana nudi: Humea, Hobarta, Schlicka i Ayera (McKenna, Pereboom 2016: 52–54). Osim što ti autori iznose neke od najčešće upotrebljivanih kompatibilističkih argumenata, oni se izdvajaju i zbog svojih nemilosrdnih stavova prema oponentima u debati koja nerijetko završava u otvorenim optužbama za potpuno nerazumijevanje. Tako će Hume istaknuti da je cjelokupna rasprava “puka verbalna rasprava” (prema Bradley 1958: 34), Hobart će istu raspravu prozvati “ogromnim gubitkom u području samoga rasuđivanja” te jednim “lošim snom” (Hobart 1934: 27), na što će se kasnije nadovezati Moritz Schlick koji će pak izreći znamenite riječi kojima tvrdi:

... to je doista jedan od najvećih skandala filozofije kojemu se uvijek iznova toliko papira i pišaće tinte mora posvetiti, da ne govorimo o utrošku promišljanja koje bi se moglo primijeniti na važnije probleme [...]. Stoga bih se doista trebao sramiti što pišem poglavlje o slobodi. (Schlick 1930: 143)

Na kojim se argumentima temelje oštre optužbe gore navedenih autora? Hume u *Raspravi o ljudskoj prirodi* (Hume 1983: 348–352) iznosi optužbu koja se zasniva na problemu koji nastaje zbog neadekvatnog razumijevanja uzrokovanja. Kritika započinje definicijom uzrokovanja koje se shvaća kao konstantna konjunkcija između dvaju događaja te odbacivanjem postojanja nužne relacije između uzroka i učinka. Problem nastaje kada nužnost projiciramo i nastojimo je ukalupiti tamo gdje je nema, kao vezu između uzroka i učinka, tada kao posljedica takvog promišljanja nastaje inkompatibilističko viđenje uzrokovanja kao nužnoga slijeda događaja, što pak implicira interpretaciju uzrokovanja kao procesa koji prisiljava pojedinca da se ponaša na točno određen način. Može se reći da tako nastaje pseudoproblem zbog “viđenja nečeg čega nema”: nužnosti u relaciji uzrokovanja. Na takvu kritiku u kasnijoj se raspravi nadovezuju drugi autori poput Alfreda Ayera koji će istaknuti da samo neke vrste uzrokovanja predstavljaju prisilu poput slučajeva u kojoj nas netko tjera da činimo nešto protiv naše volje, da se, primjerice, bacimo s mosta, ali, Ayer naglašava, samo uzrokovanje nas ni na što ne tjera, a samim time ono nije instanca prisile: “uzrokovanje nije u suprotnosti sa slobodom, prisila jest.” (Ayer 1954: 114).

Kritiku koja govori o krivoj interpretaciji jastva koja uvelike nalikuje humeovskoj kritici krive interpretacije uzrokovanja donosi R. E. Hobart. Kritika je usmjerena na libertarijance koji su skloni interpretirati jastvo kao posebnu supstanciju, različitu od uzročno generiranog karaktera pojedine osobe. Ako se tako koncipira jastvo, odnosno sposobnost odlučivanja pojedinca koji može nadvladati svoje karakterne osobine koje su prouzrokovane vanjskim faktorima poput okoline, životnih događaja, genetike, onda govorimo o jastvu koje nadilazi karakterne osobine pojedinca i omogućuje generiranje novih kauzalnih lanaca. Stoga će pojedini autori poput Thomasa Reida ili Rodericka Chisholma standardnom, uobičajenom prirodnom uzrokovanju pridodati djelatnikovo uzrokovanje, vrstu uzrokovanja koju posjeduje djelatnik kao posebna vrsta supstan-

cije koja može stvarati nove kauzalne lance.⁵ Hobartova dijagnoza jest da se ovdje radi o nerazumijevanju prirode jastva. Uzročno formiranom karakteru osobe dodaje se neko posebno jastvo s posebnim uzročnim moćima kojeg u stvarnosti, naposljetku – nema. Hobart iznosi primjer iz devetnaestostoljetne ruralne Njemačke u koju je upravo pristigla željeznica: začuđeni seljaci promatraju lokomotivu, a lokalni pastor koji je upućen u funkcioniranje iste u tančine im objasni način na koji radi, dijelove od kojih je sastavljena te po završetku izlaganja dobiva pitanje: “Da, da, Herr Pastor, ali konj je u njoj, zar ne?” (Hobart 1934: 3). Primjer jasno oslikava ljudsku težnju pridodavanja posebnih moći onome što si ne može objasniti: konja lokomotivi, odnosno specifičnih kauzalnih moći jastvu.

Iz prikazane argumentacije evidentno je da Hume i Hobart karakteriziraju problem slobode volje kao pseudoproblem zato što vjeruju da su inkompatibilisti krivo interpretirali uzročnost, odnosno jastvo. Striktno govoreći, inkompatibilisti ne iznose besmislice. Njihovi iskazi imaju značenje, ali prema iskazanoj kritici u potpunosti krivo shvaćaju narav određenih pojmova i generiraju problem tamo gdje ga ne bi ni trebalo biti! Možemo se priupitati govorimo li uopće o pseudoproblemu? Prije nego li se odgovori na to pitanje, valja preispitati i Schlickovo viđenje slobode volje kao pseudoproblema koje predstavlja opće mjesto te rasprave i koje je stvorilo jedan grozd odgovora na tu temu, među kojima se ističu Campbellovi i Bradleyevi radovi⁶.

U knjizi iz 1930. *Problemi etike* u poglavlju naslovljenom “Kada je čovjek odgovoran?” Moritz Schlick iznosi svoje viđenje načina na koji je rasprava o slobodi volje kontaminirana raznim pogrešnim interpretacijama te time pretvorena u pseudoproblem. Prema njegovu mišljenju, sve što je potrebno jest analiza značenja koje svakodnevno pridajemo iskazima “odgovornost” i “sloboda”. Razjašnjenje načina korištenja tih pojmova dovest će do nestanka pseudoargumenta koji se nalazi u korijenu pseudoproblema. Prema Schlicku, valja naglasiti da nesporazum ima svoj korijen u pojmu “odgovornost” te je upravo na taj pojam usmjerena analiza dok se o pojmu “sloboda volje” govori sporadično. Zanimljivo

⁵ Radi detaljnijeg prikaza Reidove ili Chisholmove teorije djelatnikovog uzrokovanja vidi: (Kane 2005: 44–46) i (McKenna, Pereboom 2016: 64–66).

⁶ Vidi: (Campbell 1951) i (Bradley 1958)

je što Schlick raspravu prvenstveno vidi kao etičku raspravu, kao raspravu o odgovornosti, a ne kao metafizičku raspravu o slobodi volje, što sugerira da metafizičku raspravu i ne smatra filozofski relevantnom, odnosno tretira je kao pseudoproblem.

Schlick, poput brojnih drugih kompatibilista,⁷ napominje da mi ne znamo je li princip uzročnosti, odnosno determinizam univerzalno valjana teza, ali istovremeno znamo da je nemoguće razriješiti spor između determinizma i indeterminizma promišljanjem i vaganjem razloga za i protiv određene pozicije. A rasprava o problemu slobode volje započinje upravo na taj način: jedna skupina koja ističe taj problem nastoji odbaciti valjanost uzročnog principa dok ga druga brani (Schlick 1930: 144–145).

Inkompatibilistički argument koji Schlick odbacuje kao pseudoargument može se formalizirati na sljedeći način:

- (1) Ako je determinizam istinit, onda je i moja volja determinirana zakonima.
- (2) Ako je moja volja determinirana zakonima, onda su i moje odluke nužno determinirane.
- (3) Da bih bio moralno odgovoran za svoje postupke, ja moram moći utjecati na njih.
- (4) Ako je determinizam istinit, ja ne mogu utjecati na odluke koje donosim jer su nužno determinirane.
- (5) Stoga, ja ne mogu biti odgovoran za svoje postupke.
- (6) Dakle, moralna odgovornost pretpostavlja slobodu, odnosno izuzeće iz uzročnosti.

Prema Schlicku, argument koji nudi inkompatibilist obiluje nizom nesuvislosti, a osnovni problem sastoji se u nesposobnosti razlikovanja dviju vrsti zakona: preskriptivnih i deskriptivnih zakona (Schlick 1930: 146–148). Preskriptivni zakoni oni su koje nam nameću određeni au-

⁷ P. F. Strawson u svojem antologijskom radu "Sloboda i zamjeranje" znamenito izjavljuje: "Neki filozofi kažu kako ne znaju što je teza determinizma. Drugi kažu, ili podrazumijevaju, da znaju što ona jest. (...) Ako me se pita kojoj strani pripadam, moram reći da je to ona koja prethodi svima, strana onih koji ne znaju što je teza determinizma." (Strawson, 1962: 19)

toriteti iz naše društvene zajednice te mogu predstavljati jedan oblik prinude kada se pojedinca tjera da im se pokori. No, za razliku od preskriptivnih zakona, pri opisu prirode koristimo se deskriptivnim zakonima koji samo opisuju činjenice koje se događaju, stoga zakonomjerne uzročnosti u prirodi ne djeluju na nas na način na koji to čine društvene zakonitosti: ne prisiljavaju nas. Takvi zakoni ništa nam ne nameću, ne tjeraju ni na što, oni naprosto jesu. Keplerovi zakoni ne tjeraju planete da se kreću na određeni način. Kada kažemo "čovjek se pokorava psihološkim zakonima", ne mislimo da mu se određuje koje želje mora imati. Oni opisuju na jednak način na koji to rade astronomski zakoni. Prisila se događa kada nas se sprječava da činimo ono što želimo činiti. Prisila nije prisutna u prirodnim zakonima.

Tom krivom, antropomorfiziranom shvaćanju zakona prirode inkompatibilist pridodaje pogrešnu interpretaciju prisile i na taj način čini drugu grešku u argumentaciji. Schlick nas, u duhu humeovske tradicije, upozorava da kada govorimo o prirodnim zakonima, onda ne valja govoriti o nužnosti nego o univerzalnoj valjanosti zato što ćemo o prirodnom zakonu govoriti tak kada pronađemo neko pravilo za koje smo ustanovili da vrijedi bez iznimaka, da je univerzalno valjano (Schlick 1930: 148–149). Dakle, prirodni zakoni nisu nužni, oni su univerzalno valjani. Što onda s prisilom? Pojam prisile toliko je pojmovno udaljen od pojma univerzalne valjanosti da "jednom kada steknemo uvid u grešku, više nećemo moći ni pojmiti mogućnost zabune." (Schlick 1930:149). Pojam suprotan univerzalnoj valjanosti jest indeterminizam, odnosno "akauzalnost", a pojam suprotan prisili jest sloboda. Stoga, ukoliko netko interpretira univerzalnu uzročnost kao nužnost, onda taj griješi, a ukoliko pretpostavi da je ona u suprotnosti sa slobodom, onda čini dodatnu grešku. Schlick, kao i drugi brojni kompatibilisti, ističe da iz takve argumentacije proizlazi često korištena besmislica prema kojoj sloboda znači "izuzetost iz kauzalnih principa" ili "nepodlijevanje zakonima prirode".

Stoga ranije prikazano tradicionalno shvaćanje problema slobode volje predstavlja pseudoproblem koji se s lakoćom može razriješiti. Valja samo razjasniti što se misli pod pojmom zakon i koji se pojmovi (ne) nalaze u suprotnosti te ćemo tada uvidjeti da problema i nema. Pseudoproblem nastaje zbog neprimjerene pretpostavke da uzročni zakoni uključuju prinudu.

Jesu li klasični kompatibilisti u pravu kada problem slobode volje tretiraju kao pseudoproblem? Ako je tome tako, iznose li inkompatibilisti besmislice? Implicira li to da oni ni ne mogu započeti smislenu raspravu? Kako pokazati da je takav stav neopravdan, odnosno kako odbaciti optužbe da se u raspravi o slobodi volje radi o pseudoproblemu? Smatram da se te optužbe mogu odbaciti ako se pokaže da:

- (1) neosnovano je pozivanje na evidentnost na koju se pozivaju kompatibilisti,
- (2) rasprava koja se vodi ne predstavlja pseudoproblem shvaćen kao iznošenje besmislica,
- (3) rasprava koja se vodi jest uobičajena, metafizička rasprava.

Neosnovanost pozivanja na evidentnost lako je uvidjeti ako se promotre načini pomoću kojih inkompatibilisti započinju raspravu. Pozivanje na introspeksijsko iskustvo često će predstavljati prvi korak, a u nekim slučajevima i posljednji. Tako će, primjerice, Samuel Johnson izjaviti “Mi *znamo* da je volja slobodna i tu je kraj.”⁸ Pri čemu on smatra da je evidentno ne samo da tu rasprava počinje, nego i završava: u evidentnosti naše slobode koju spoznajemo introspekcijom. Nije potrebno ulaziti u evaluaciju takvog argumenta, ono što je bitno jest da se na taj način libertarijanski tabor poziva na evidentnost vlastite pozicije, istu onu evidentnost na koju se poziva i kompatibilistički tabor. Kako razriješiti spor? Ako je evidentnost upitna, onda se spor valja riješiti argumentima.⁹

No, ukoliko rasprava zaista jest jedan pseudoproblem kako to tvrde kompatibilisti, onda bi inkompatibilistička argumentacija zaista bila besmislena te bi rasprava stala na ovoj točki i ne bi se dalje razvijala. No, kao što je ranije prikazano, prema kompatibilističkoj kritici ne radi se o besmislenim argumentima. Optužbe ne idu u tom smjeru. Ono što Schlick, Hobart, Hume i drugi nastoje pokazati jest da inkompatibili-

⁸ Citat preuzet iz (Bradley 1958: 38)

⁹ Jedna metoda pomoću koje bi se moglo utvrditi koja pozicija jest evidentna, a koja nije bilo bi testiranje ljudskih intuicija po tom pitanju. Takvo što provedeno je u okviru eksperimentalne filozofije, ali rezultati do kojih se došlo inkonkluzivni su. Za više detalja o intuicijama o tradicionalnom problemu slobode volje vidi (Nahmias 2011), o doprinosu eksperimentalne filozofije u okviru rasprave o slobodi volje vidi (Čeč 2013).

sti u potpunosti krivo shvaćaju narav određenih pojmova i generiraju problem tamo gdje ga ne bi ni trebalo biti. Ali ako je tome tako, onda raspravu o tradicionalnom problemu slobode volje ne možemo okarakterizirati kao pseudoproblem shvaćen kao iznošenje besmislica, nego se radi o drugoj vrsti pseudoproblema, pseudoproblemu koji nastaje zbog pogrešnog korištenja pojmova.

No, je li opravdanu takvu raspravu tretirati kao pseudoproblem, kao pseudodiskusiju? Kako se inkompatibilisti ne slažu s kompatibilističkom dijagnozom, onda oni nastoje istim oruđem odgovoriti kompatibilistima: argumentima. Time započinje jedna klasična filozofska rasprava u kojoj svaka pozicija iznosi argumente za svoju poziciju i pobija tuđe. Zašto onda govoriti o pseudoproblemu? Ne bi li onda trebali reći da kada se govori o kategoriji pseudoproblema koji nastaju zbog pogrešnog korištenja pojmova, da onda uopće ne govorimo o pseudoproblemima, nego da se radi o uobičajenoj argumentaciji kojom se nastoji pokazati neosnovanost suprotne pozicije. Dakle, kategorija pseudoproblema koja nastaje zbog pogrešnog korištenja pojmova uopće ne postoji. To je uobičajena filozofska rasprava. Stoga Schlickove ratoborne riječi kojima se poziva na gubljenje vremena zbog prisutnosti pseudoproblema čine se poput retoričkog zahvata kojim se želi obezvrijediti suparnička pozicija. Optužbe kompatibilista možda jesu osnovane, možda suprotan tabor zaista krivo koristi pojmove iz rasprave, suprotstavlja determinizam slobodi umjesto da joj suprotstavlja prisilu, ali ako je tomu tako, to ne znači da se oni koriste pseudoargumentacijom, nego da su u krivu. A da su u krivu, nastoji se dokazati argumentima. Dakle, vodi se jedna punokrvna, uobičajena metafizička rasprava.

No, ostaje još jedna mogućnost, a to je da je rasprava trivijalna, da se suprotstavljene pozicije ne bave metafizikom nego nekom vrstom semantičkog nadmudrivanja. Na tu mogućnost upozorava nas Mark Balaguer.

III

Pristup za koji se zalaže Mark Balaguer u svojim radovima, poglavito u svojoj knjizi *Free Will as an Open Scientific Problem* (Balaguer 2010), može se karakterizirati kao inovativan i radikalan, barem kada govorimo o raspravi o slobodi volje. Naime, Balaguer je inovativan u tome što je

njegov libertarijanizam postavljen na tvrdo naturalističkim osnovama, što ga uvelike razlikuje od dominantne slike libertarijanizma iz prošlih stoljeća, koja je često bila postavljena na upravo suprotnim, antinaturalističkim premisama. Po tom se pitanju Balaguerova argumentacija preklapa s argumentacijom klasičnih kompatibilista kada nastoje odbaciti sliku svijeta koju nudi antinaturalistički libertarijanizam. Pozicija koju Balaguer brani jest tvrdnja da je sloboda volje, kao što naslov njegovog rada i sugerira, otvoreno znanstveno pitanje.

Što to znači? Jednostavno, jedino metafizički relevantno pitanje jest ono postoje li odluke koje su u svojoj biti indeterminističke, i to na takav način da omogućuju djelatniku slobodu izbora. Takvo je pitanje prema Balagueru otvoreno pitanje jer sadašnji stupanj znanstvenog razvitka onih relevantnih empirijskih disciplina koje bi mogle ponuditi odgovor na ovo pitanje još uvijek nije dovoljan da takav odgovor mogu konkluzivno ponuditi. Dakle, to je posve empirijsko pitanje i zato je metafizički zanimljivo. Sve ostalo u raspravi o slobodi volje nema taj metafizički karakter, naprotiv, ostale rasprave poput pitanja "Što je sloboda volje?" ili pitanja "Je li ona kompatibilna s determinizmom ili nije?" u svojoj su biti semantičke rasprave, lišene ikakvog pravog metafizičkog naboja. Koristeći se ratobornom terminologijom klasičnih kompatibilista, mogli bismo reći da su takve rasprave pseudorasprave. Takav stav uvelike se protivi uobičajenom, tradicionalnom metafizičkom pristupu problemu slobode volje za koji je rasprava o kompatibilnosti slobode volje i determinizma itekako metafizički relevantna i zasigurno neće reći da je priroda te rasprave isključivo semantička.

Mark Balaguer svoju argumentaciju započinje ranije navedenom tradicionalnom formulacijom problema slobode volje koja je dana u obliku inkonzistentne trijade. Uvodna rasprava nastavlja se analizom pojmova slobode koje posjedujemo. Zapочиње zdravorazumskom definicijom slobode za koju se zalažu kompatibilisti i koja je utemeljena na idejama koje je David Hume iznio u svojem *Istraživanju o ljudskom razumu*. Definicija tako shvaćene slobode, *humeovske slobode*, glasi: "Osoba S ima humeovsku slobodu ako je osoba S sposobna djelovati u skladu sa svojim izborima i izabirati u skladu sa svojim željama." (Balaguer 2010: 4) Ali, to nije jedina sloboda koju nude kompatibilisti, Balaguer navodi još nekoliko kompatibilističkih pristupa slobodi te zatim iznosi jednu neobičnu tvrdnju:

Važno je naglasiti da su, poput humeovske slobode, sva ta druga kompatibilistička shvaćanja slobode ne samo kompatibilna s determinizmom, nego je tome *evidentno* tako. (...) Frankfurtovska sloboda može se definirati (ukratko) kao sposobnost kontrole, pomoću naših stavova drugog reda, koja će od naših želja prvog reda utjecati na naše ponašanje, no čini se da je frankfurtovska sloboda jasno kompatibilna s determinizmom. (Balaguer 2010: 5)

Te nastavlja tvrdeći:

Do iste točke gledišta može se doći u slučaju svih ostalih kompatibilističkih poimanja slobode u literaturi, ali o tome neću sada argumentirati jer za svaki slučaj rezultat je jednako evidentan i nekontroverzan kao što je u slučaju humeovske ili frankfurtovske slobode. Drugim riječima, u slučaju svakog od tih drugačijih poimanja slobode kontroverzno pitanje nije je li ono kompatibilno s determinizmom, nego daje li ono točnu analizu poimanja slobode volje – točnije, je li ono sloboda volje? (Balaguer 2010: 5)

Za Balaguera je evidentno da posjedujemo humeovsku slobodu, da možemo djelovati na način na koji želimo djelovati, također je evidentno da smo slobodni na frankfurtofski način te isto više ili manje vrijedi i za svaku drugu kompatibilističko poimanje slobode. Sve su one evidentno kompatibilne s determinizmom te ih sve više ili manje posjedujemo. Kako Balaguer dolazi do takvog zaključka? On smatra da je tome tako jer tradicionalna briga o postojanju ili nepostojanju slobode volje ne potječe iz problematike kompatibilizma, nego je u biti ona jedna “briga t o postojanju indeterminističke slobode volje ili točnije, libertarijanističke vrste slobode volje.” (Balaguer 2010: 7) Dakle, do problema s kojima se susrećemo u ranije navedenoj trijadi dolazi zbog libertarijanističkog poimanja slobode volje. Po tom se pitanju Balaguer slaže s klasičnim kompatibilistima jer i prema njihovom shvaćanju problem se generira zbog libertarijanističkog shvaćanja slobode volje, ali za razliku od njih, Balaguer neće to pitanje odbaciti kao pseudoproblem, nego će na tom temelju graditi vlastitu argumentaciju.

No, za libertarijanističku slobodu nije dovoljna prisutnost indeterminizma, kako nam jasno sugerira problem inteligibilnosti,¹⁰ nego se zahtijeva i prisutnost, kako je Balaguer naziva “prikladne neslučajno-

¹⁰ Prema problemu inteligibilnosti, libertarijanističkim pozicijama nije dovoljno pretpostaviti postojanje indeterminizma, nego im treba i smisleno objašnjenje kako indeterminizam može otvoriti vrata slobodi volje. Dakle, radi se o problemu kompatibilnosti slobode volje i indeterminizma.

sti” (Balaguer 2010: 7). odnosno, ideja da indeterminizam sam po sebi nije dovoljan, odnosno da ne smije biti slučajan i mora biti prikladan na način da dovodi do slobode volje. *Prikladna neslučajnost* implicira prisutnost djelatnika koji sam mora biti uključen u proces odlučivanja na način da autorizira i kontrolira odluku, čime ta odluka biva njegova (Balaguer 2010: 7). Dakle, ljudska sloboda nije samo ugrožena determinizmom nego i određenim vidovima slučajnosti ili manjka prikladne neslučajnosti. Stoga Balaguer smatra da bi bolje bilo reći da nije determinizam onaj koji je u ranije navedenoj tradicionalnoj trijadi problematičan, nego je to ono što Balaguer naziva *Determinizam-ili-slučaj-izam*, odnosno *D-ili-S-izam*¹¹. Ukoliko je *D-ili-S-izam* istinit, onda su naše odluke ili determinirane, ili su indeterminirane, ali na način da nisu i prikladno neslučajne. Dakle, *D-ili-S-izam* jest ono što, ako postoji, čini našu slobodu nemogućom (Balaguer 2010: 9–10).

Sada je vrijeme da se pitamo o kojoj slobodi govorimo te ovim putem dolazimo do definicije one punokrvne libertarijaniističke slobode za koju se Balaguer zalaže, *L-slobode*. Slobode prema kojoj stoji da:

Osoba je L-slobodna ako i samo ako njene su odluke takve da (a) istovremeno su indeterminirane i prikladno neslučajne te (b) indeterminizam koji je tu prisutan relevantan je za prikladnu neslučajnost na način da generira neslučajnost ili je osigurava, ili je obogaćuje, ili povećava, ili čini nešto u tom duhu. (Balaguer 2010: 26)

No, osim što smo došli do definicije L-slobode, došli smo i do definicije onoga što onemogućuje ljudskim bićima da imaju tu istu slobodu: *FE-determinizma*, odnosno indeterminizma koji ne omogućava slobodu, koji možemo zvati i *Determinizam**.¹² Konačni rezultat jest da smo sada

¹¹ U originalu: “*Determined-or-Randomism (D-or-R-ism)*” (Balaguer 2010: 8).

¹² U knjizi Balaguer se koristi pojmom “FE-determinism” da bi govorio o onome što naziva i “No Freedom-Enhancing Indeterminism”, odnosno indeterminizam koji ne omogućava slobodu. Pojam “FE-determinism” čini mi se nespretno formuliranim jer njegov doslovni prijevod glasi “determinizam koji omogućuje slobodu”, što je u suprotnosti s onime što Balaguer želi reći. Stoga ću se koristiti sintagmom “determinizam*” koju koristi Michael McKenna (McKenna 2012), kojom se označava ono što smo ranije nazvali indeterminizmom koji ne omogućava slobodu ili FE-determinizmom. Ta sintagma obuhvaća sve determinizme i indeterminizme koji su takvi da onemogućuju L-slobodu. Tome je tako jer se do ovog pojma došlo preko ranije navedenog “determinizma-ili-slučajnost-izma”, čija je domena dodatno sužena L-slobodom kakvom je koncipira Balaguer. Dakle, pojmovi “determinizam*”, “FE-determinizam” i “indeterminizam koji

došli do rafinirane, preciznije formulirane inkonzistentne trijade od one tradicionalne koja je bila ranije ponuđena:

- (1) Determinizam* je istinit (libertarijanizam je neistinit).
- (2) Ljudska bića posjeduju slobodu volje;
- (3) Sloboda volje inkompatibilna je s determinizmom*. (Balaguer 2010: 13)

Sljedeći korak sastoji se u tome da valja odrediti koje su tvrdnje iz trijade točne, a koje valja odbaciti. Zato Balaguer smatra da ako želimo riješiti problem slobode volje, moramo se pitati sljedeće:

- (1) Determinističko* pitanje: je li determinizam* istinit?
- (2) Pitanje o tome imamo li slobodu volje: posjeduju li ljudska bića slobodu volje?
- (3) Kompatibilističko pitanje: je li sloboda volje kompatibilna s determinizmom*? (Balaguer 2010: 26)

Gore navedena pitanja mogu se reducirati na dva osnova pitanja:

- (a) Pitanje o tome što je sloboda volje?
- (b) Pitanje o vrstama slobode volje koje mi posjedujemo: Imaju li ljudska bića L-slobodu, ili humeovsku, ili frankfurtsku, ili neku drugu? (Balaguer 2010: 27)

Ta redukcija može se izvršiti na sljedeći način:

- (i) kompatibilističko pitanje (3) svodivo je na pitanje o tome što je sloboda volje (pitanje a),
- (ii) determinističko* pitanje (1) svodivo je na pitanje o vrstama slobode koje posjedujemo (b) te
- (iii) pitanje o tome imamo li slobodu volje (2) svodivo je na oba pitanja (a) i (b). (Balaguer 2010: 27)

ne omogućava slobodu” istoznačni su pojmovi. U ovom radu gdje god je Balaguer koristio termin “FE- determinism”, ja koristim onaj koji nudi McKenna: “determinizam*”, bez ikakve suštinske promjene značenja.

Stoga Balaguer smatra da odgovaranjem na pitanje (a) i na pitanje (b) dajemo odgovore na sve što nas zanima u vezi problema slobode volje. Tezu (ii) Balaguer smatra evidentnom jer kako je determinizam* negacija libertarijanizma, iz toga slijedi da je pitanje o postojanju determinizma* (pitanje 1) u biti pitanje o postojanju libertarijanističke slobode, a ono pripada pitanju koje slobode posjeduju ljudska bića (pitanje b) (Balaguer, 2010: 27). Na jednak način evidentna je i teza (iii): ako možemo odgovoriti na pitanje (a): što je sloboda volje? i na pitanje (b): koje vrste slobode posjedujemo?, tada ćemo znati i odgovor na pitanje (2): imamo li slobodu volje (Balaguer, 2010: 27).

Posebno zanimljiva jest teza:

(iv) pitanje (a) što je sloboda volje? nije na relevantan, metafizički, netrivialan način bitno pitanje za pitanje (2) imamo li slobodu volje?. (Balaguer 2010: 27–28)

Štoviše, to pitanje, smatra Balaguer, “nije relevantno ni za jedno bitno pitanje o prirodi ljudskog procesa odlučivanja, osim na trivialan način” (Balaguer, 2010: 27–28). Ako se sada kombiniraju teza (i) i teza (iv), onda se dolazi do:

(v) kompatibilističko pitanje jest irelevantno pitanje u pronalazanju odgovora na pitanje o prirodi ljudskog procesa odlučivanja. (Balaguer 2010: 28)

Zaključak je ovo koji svakako predstavlja oštro odbacivanje polaznih točaka tradicionalne debate, koja se onda tretira kao metafizički irelevantna, odnosno kao jedna vrsta pseudodebate. Tomu je tako zbog izuzetno naturalističkog karaktera Balaguerove metafizike. Kada se on koristi terminom ‘metafizičko’, onda se značenje tog termina veže uz nešto o svijetu, kada se govori o “ljudskom procesu odlučivanja i ljudskoj slobodi, tada se ‘metafizički’ koristi na način da znači nešto kao *o ljudskim bićima ili o ljudskom procesu odlučivanja*.” (Balaguer, 2010: 28).

Sam argument u prilog teze (iv) te samim time i teze (v) počiva na sljedećem misaonom eksperimentu: zamislimo dva lingvisti koji nastoje odrediti kako bi se trebao koristiti termin ‘planet’ te istovremeno zamislimo da su astronomi otkrili neki objekt koji kruži oko Sunca iza

orbite Plutona. Prema teoriji prvog lingvista, taj će objekt biti planet, prema drugom neće. Balaguer smatra da bi nam se moglo učiniti da smatramo da se lingvisti bave astronomijom, ali on smatra da tome nije tako – oni se bave empirijskom semantikom (Balaguer, 2010: 31–32). Zašto je tomu tako? Ako situaciju iz primjera malo preoblikujemo i učinimo je takvom da je analogna problemu slobode volje, onda ćemo doći do boljeg razumijevanja. Pretpostavimo da ne znamo postoji li planet, odnosno sloboda volje, nego da smo u posjedu nekih naznaka postojanja tih entiteta: perturbacija u orbiti ili introspekcijskih stanja. Zatim postavljamo pitanje “Postoji li planet?” odnosno “Postoji li sloboda volje?” Tada rasprava dvaju lingvista postaje evidentno trivijalna jer ona ne govori ništa o samoj prirodi Sunčevog sustava, odnosno ljudskog procesa odlučivanja. Ona samo govori kako se riječ “planet”, odnosno “sloboda volje” koristi. Zato

“Kada postavljamo pitanje imamo li slobodu volje, mi navodno nastojimo spoznati nešto o prirodi ljudskog procesa odlučivanja, ali pitanje što je sloboda volje za taj je cilj relevantno u najbolju ruku na trivijalan način jer to kako se izraz ‘sloboda volje’ koristi u svakodnevnom diskursu ne govori nam ništa o prirodi ljudskog procesa odlučivanja.” (Balaguer 2010: 33)

Ovomu treba nadodati da Balaguer smatra da pri određivanju kakvi su ljudski procesi odlučivanja možemo koristiti razne tehničke definicije slobode volje, poput one koju on sam nudi, L-sloboda, ili nekom drugom definicijom, primjerice *frankfurtovska sloboda*. Tada ne moramo brinuti o tome kako odrediti što obični ljudi misle pod pojmom sloboda volje (Balaguer, 2010: 34). Dovoljan nam je neki radni pojam i možemo započeti s istraživanjem. Zato odrediti što sloboda volje jest nije metafizičko pitanje dok istraživanje imamo li, primjerice, L-slobodu jest. Stvari su još gore za tradicionalni pristup metafizici kada se gore navedenim tezama pridoda njegov stav da pojmovna analiza ima dva aspekta: “Prvi je artikulacija pojmova, a drugi je iznalaženje argumenata za ispravnost različitih analiza koje su bile ponuđene ili protiv nje” (Balaguer, 2010: 35), pri čemu prvi aspekt jest metafizički relevantan jer daje nove prijedloge o tome što sloboda volje jest i samim time valja istražiti imaju li ljudi to poimanje slobode. Drugi je aspekt irelevantan ili kako Balaguer kaže:

“Ono što ja tvrdim da nije relevantno za metafizička pitanja o ljudskim bićima jest pitanje koje od brojnih analiza slobode volje koje su ljudi artikulirali jest ispravno. A to je upravo ono o čemu se radi u pitanju što jest sloboda volje” (Balaguer 2010: 35)

Tim putem dolazi se i do kritike samog kompatibilističkog pitanja jer ni ono nije relevantno za prirodu ljudskog procesa odlučivanja. I ono je trivijalno, semantičko pitanje o tome što mislimo pod određenim pojmom i je li ta analiza pojma koherentna. Zašto? Pa jednostavno, kao što je ranije navedeno, kompatibilističke definicije slobode volje očito su kompatibilne s determinizmom i očito ih imamo, zato daljnja rasprava o njima jest puka analiza značenja bez ikakve veze sa stvarnošću. To i nije toliko čudan zaključak za nekoga tko se zalaže za jaki naturalistički pristup, pristup koji analizu pojmova stavlja po strani i favorizira metode tvrdih znanosti. U takvom pristupu onda nema mjesta za kompatibilizam, tezu koja po mnogima jest “visoko sofisticirana filozofska teza koja zahtijeva pojmovnu analizu da bi se mogla izgraditi te se samim time ne može braniti isključivo na znanstvenim temeljima.” (De Caro 2007: 258) Iz svega toga slijedi da će bilo kakva rasprava o kompatibilizmu, prema Balagueru, biti trivijalna i zato je jedino relevantno metafizičko pitanje koje nam preostaje pitanje imamo li libertarijaniističku slobodu, točnije onu ranije definiranu L-slobodu? No, to pitanje nema konačni odgovor jer kako u četvrtom poglavlju svoje knjige argumentira (Balaguer, 2010: 129–169), ne postoje dovoljno dobri empirijski razlozi ni za odbacivanje, ni za potvrđivanje L-slobode, odnosno sadašnji stupanj znanstvenog razvitka ne može nam dati konačni sud o statusu L-slobode te je stoga njeno postojanje jedino otvoreno znanstveno pitanje koje preostaje u raspravi o slobodi volje.

IV

Za gore opisani Balaguerov pristup može se reći da počiva na četiri problematična temelja. Ako se ta četiri temelja dobro potkopaju, onda će se i Balaguerova argumentacija urušiti. Problematične su postavke Balaguerove teze sljedeće:

- (1) Pretpostavka da je svaka kompatibilistička definicija slobode evidentno kompatibilna s determinizmom te da takvu slobodu očigledno posjedujemo;

- (2) L-sloboda kao 'istinsko' poimanje slobode;
- (3) Pretpostavka da se do određenog broja zaključaka došlo argumentacijom koja nije metafizička;
- (4) (Pre)strogi naturalizam.

Započnimo s prvom točkom i pretpostavkom da su definicije slobode volje dane od strane kompatibilista evidentno kompatibilne s determinizmom te da je očigledno da takve slobode posjedujemo. Čini se da stvar nije tako evidentna kao što Balaguer pretpostavlja. Uzmimo za primjer frankfurtovsku definiciju slobode volje. Ona se često povezivala s teorijama djelatnikova uzrokovanja zbog ideje identifikacije koju Frankfurt nudi:

"Velleman nam je pokazao u jednoj izuzetno pronicavoj diskusiji Frankfurtova prijedloga da identifikacija, kako je početno figurirala u Frankfurtovoj teoriji slobode, nije ništa više od jedne etikete za djelatnikovo uzrokovanje koje se najprirodnije shvaća u libertarijaniističkim terminima." (McKenna 2012: 7–8)

Ako je tomu tako, onda kompatibilist koji se zalaže za frankfurtovsku slobodu mora pokazati zašto je ta teorija kompatibilna s determinizmom, zašto nije samo još jedna etiketa libertarijaniističkih teorija djelatnikovog uzrokovanja te onda samim time više nije evidentno da će sve kompatibilističke teorije slobode volje biti evidentno kompatibilne s determinizmom. Naprotiv, takvo što treba pokazati argumentima. Onda pitanje više nije semantičko, prestaje biti trivijalno i postaje metafizički interesantno i relevantno.

Isto vrijedi i za sve rasprave koje su nastale oko argumenta posljedice, argumenta koji se bavi kompatibilizmom naše mogućnosti da postupimo drugačije i determinizma. U potpunosti se slažem s McKennom kada kaže: "Iz moje perspektive [filozofskog] krajolika, tamo gdje se čini da Balaguer ne vidi ništa metafizički zanimljivo, meni je sve što vidim [metafizički zanimljivo]." (McKenna 2012: 9). Što ako nam argument posljedice pokaže da su neka kompatibilistička poimanja neutemeljena? Onda više nije evidentno da ih imamo. Naprotiv, ne bismo ih ni trebali moći imati. Dakle, čini se da s kompatibilizmom stvari nisu tako jednostavne kako Balaguer pretpostavlja.

Tu ne prestaju problemi s kompatibilizmom jer se sada može dovesti u pitanje tezu (2), tezu da je L-sloboda 'istinsko' poimanje slo-

bode. Tomu je tako zbog toga što se možemo pitati što ako određena poimanja kompatibilističke slobode prežive pojmovnu analizu? Valja li onda i dalje preferirati L-slobodu nad tim poimanjima? No, ovaj se problem može i produbiti. Možemo se pitati zašto govoriti o jednom ili 'istinskom' problemu slobode volje? Brojni autori smatraju da postoji čitav niz pristupa i formulacija problema slobode volje pa se, primjerice, McKenna i Pereboom u svom udžbeniku propituju treba li nam za slobodu volje mogućnost drugačijeg postupanja ili je ključno da djelatnik bude krajnji izvor radnji (McKenna, Pereboom 2016: 37–43)? Čini se da Balaguer smatra da je izraz "sloboda volje" jedan od onih izraza čiji se semantički sadržaj može utvrditi ili izučavanjem svakodnevnog jezika, ili izučavanjem svakodnevnog jezika u konjunkciji s nekim pitanjem normativnosti (McKenna 2012: 5). No, je li tomu tako ili je situacija takva da se ne može izučavanjem svakodnevnog jezika utvrditi za što termin "sloboda volje" stoji? Čini se vjerojatnijim tvrditi da je to termin koji figurira kao jedan 'krovni' termin kojim govorimo o brojnim međusobno povezanim problemima poput ljudskog donošenja odluka, moralne odgovornosti, kontrole te ih u okviru rasprave koju nazivamo sloboda volje nastojimo razriješiti. Ili kako McKenna kaže:

"Ne smatram da bi čak ni dio onoga što mi filozofi činimo prilikom rješavanja pitanja što sloboda volje jest trebao biti izgrađen u terminima onoga što obični ljudi misle pod izrazom 'sloboda volje'. (...) Mislim da je najbolje razmišljati o izrazu 'sloboda volje' kakav se pojavljuje u filozofskom kontekstu, kao o jednom isključivo tehničkom terminu (...). Mi bismo ga trebali tretirati i tretirali smo ga kao tehnički termin i pokazali smo njegovu važnost kao zanimljiva filozofskoga problema vezujući ga za druga promišljanja koja *jesu* direktno povezana s našim svakodnevnim jezičnim iskazima." (McKenna 2012: 6)

U prijevodu, filozofski pojam "slobode volje" jest tehnički pojam koji svoju filozofsku relevantnost duguje tome što je povezan s brojnim drugim pojmovima koji se javljaju u svakodnevnom i u filozofskom diskursu. Samim time ne postoji jedno shvaćanje što sloboda volje jest, ne postoji 'istinsko' shvaćanje što sloboda volje jest nego više međusobno povezanih problema koji se nalaze pod istim krovom. Dakle, Balaguerova teza (2) da je L-sloboda 'istinsko' poimanje slobode, jedino poimanje koja ima metafizičku težinu i koju vrijedi istražiti jednostavno – ne stoji.

Iz gore navedenog lako se dolazi do toga da je velik dio rasprave metafizički itekako relevantan i samim time do negacije teze (3), pretpo-

stavke da se do određenog broja zaključaka došlo argumentacijom koja nije metafizička, a paralelno s time i do negacije teze (4), teze strogog naturalizma za koji se Balaguer zalaže.

Čini se da se moramo pitati je li podjela na materijalna, empirijska pitanja i terminološka, semantička pitanja baš toliko oštra? Brani li Balaguer samo nominalno takvo poimanje ili je posrijedi nešto drugo? Je li zaista tako da je pitanje što je sloboda volje semantičko pitanje i samim time trivijalno pitanje?

Zanimljive su, i po mojemu mišljenju, točne kritike koje, između ostalih, iznosi Neil Levy na prepisci koja se nalazi na blogu "The Garden of Forking Paths"¹³ i do koje je došlo neposredno pred izlazak Balaguerove knjige. Kritike se uglavnom fokusiraju na Balaguerov metafizički pristup koji je on prezentirao u nekim ranijim radovima (Balaguer 2004), kao i u samoj knjizi. Kritika započinje time da se pitamo:

"...koliko vrsta slonova postoji? Dugo su vremena znanstvenici smatrali da postoje samo dvije različite vrste slonova (afrički i azijski). No, sada neki znanstvenici argumentiraju da bismo trebali razlikovati dvije različite vrste afričkih slonova. (...) Jedna grupa znanstvenika smatra da nam ovi podaci daju pravo na identifikaciju dviju vrsta dok drugi smatraju da nam ovi podaci daju pravo identifikacije dviju podvrsta." (Levy 2009)

Zatim se Levy pita je li ovo samo terminološko pitanje kao što to Balaguer smatra kada iznosi svoj primjer s "planetom" ili se ipak radi o sadržajnom pitanju? I jedna i druga skupina znanstvenika zalaže se za očuvanje biološke raznolikosti, odnosno zalažu se za očuvanje čim većeg broja vrsta organizama. Stoga se čini da problem sa slonovima implicira da postoji neko sadržajno pitanje o kojemu valja raspravljati.

Problem slične vrste javlja se kada se govori o problemu proširenog uma: postoje određeni aspekti ljudske kognicije koji uključuju uzročne procese koji se odvijaju izvan lubanje. Hoćemo li te procese nazvati umskim procesima ili ne? Smatramo li da nam pripadaju ili predstavljaju nešto nama izvanjsko? (Levy 2009).

Iz ovih primjera proizlazi da govor o terminima, pitanje o tome kada i kako ih upotrebljavati uistinu ima neku metafizičku težinu; pitamo se koliko vrsti slonova postoji, što je um i gdje se nalazi. Isto vrijedi i za

¹³ The Garden of Forking Paths. Blog.http://gfp.typepad.com/the_garden_of_forking_pat/2009/06/the-garden-reading-group-comment-now-open.html

pojmove unutar rasprave o slobodi volje: kako odrediti što je kontrola, kako odrediti što je autorstvo nad radnjom? Levy smatra da ako ne odgovorimo na pitanje što je sloboda volje, onda nećemo moći ni opisati, ni razumjeti svojstva svijeta:

“...mi ne možemo znati jesu li svojstva poput ‘kontrola’, ‘autorstvo’ ili ‘prikladna neslučajnost’ realna, a da ne odradimo teške pojmovne napore, osim ako se ne reduciramo na tupo pokazivanje, na iskaze da *ovo* jest ljudsko odlučivanje.” (Levy 2009)

Dakle, u osnovi je ideja da moramo znati što tražimo. Ako ne znamo što jest ‘prikladna neslučajnost’ ili ‘kontrola’, nećemo ni znati postoji li ona, ako ne znamo postoji li ona, nećemo znati jesmo li L-slobodni ni znati kako to istražiti. Vezano uz to, pitanje što jest sloboda volje upravo se tiče kontrole, prikladne neslučajnosti te kada odredimo jedno, odredit ćemo i drugo. Stoga valja zaključiti da kada se bavimo pojmovnom analizom, mi se itekako bavimo metafizikom.

Balaguer smatra da je uloga pojmovne analize u slučaju bavljenja metafizikom trivijalna i taj stav u svom odgovoru Levyjevoj kritici obrazlaže tako što tvrdi da “Ako su naše odluke prikladno neslučajne, onda ih mi autoriziramo i kontroliramo, a da ne kažemo što autorstvo i kontrola jesu.” (Balaguer 2009). Ili slikovito rečeno, kada govorimo o osobama, onda: “Ja mogu znati da je Kobe Bryant osoba, a mjesec to nije (...), a da ne posjedujem pojam osobe” (Balaguer 2009) pa analogno možemo prepoznati da je jedna osoba, primjerice, L-slobodna, a da ne posjedujemo jasno poimanje recimo kontrole. Da budem precizniji, za Balaguera nije problem ako netko promišljanjem dođe do nove spoznaje i ponudi novo poimanje kontrole, novo poimanje slobode. Takvo razumijevanje pojmovne analize jest metafizički korisno, problematično je ono razumijevanje pojmovne analize koje iznalazi argumente za ponuđene analize pojmova ili protiv njih. Zbog toga ponuditi pojam po kojemu je Kobe Bryant osoba, a mjesec to nije nije problematično. Problematično je pronalaziti argumente protiv toga jer time ne otkrivamo ništa novog o tome što osoba jest. Isto vrijedi i za slonove i za prošireni um. Terminološka pitanja neće nam reći ništa novog o vrstama slonova, nego će nam samo govoriti o tome koja je najbolja definicija pojma vrste slona.

Odgovor koji se može ponuditi Balagueru tvrdit će da određeno poimanje što vrsta jest ili što um jest, ili što sloboda volje jest može manje ili

više, bolje ili lošije zahvatiti svijet. Odnosno, koristeći se tzv. mesarskom metaforom kada se brani realizam u metafizici, određeni će pojam bolje ili lošije rezati prirodu po spojevima. Određeni će nam pojmovi bolje ili lošije znati reći što je mišić, što hrskavica, a što kost. Odnosno, bolje će ili lošije zahvatiti prirodu onakvu kakva jest. Balaguer prilikom razmatranja te metafore smatra da moramo biti oprezni da ne postavimo semantiku ispred metafizike jer je redoslijed upravo suprotan: metafizika prethodi semantici. (Balaguer 2009 i komentari s bloga “The Garden of Forking Paths”)

Osobno se ne slažem s tako jednostavnim shvaćanjem programa realizma u metafizici te dijelim ideju Neila Levya da je taj odnos jedan odnos međuovisnosti kao što primjer slonovskih vrsta i pokazuje. Ignorirati svijet predstavlja bijeg od svijeta, ali bez jasne slike pojmova koje posjedujemo nećemo moći ni istraživati svijet. Odnosno:

“Pojmovna analiza bez izučavanja stvarne uzročne strukture svijeta samo je jedna vrsta sistematske pučke psihologije. Ali metafizika je bez pojmovne analize slijepa prema svijetu koji bi željela istraživati.” (Levy 2009)

Na gore navedenu kritiku Balaguer će reagirati tako da će opet ponoviti svoj stav i nanovo iznijeti gore navedene argumente, na što mu se može odgovoriti istom mjerom pa se čini kao da nema izlaska iz ovog vrtloga. Pesimistički rečeno, čini se kao da rasprava vodi u prazno. Ipak, smatram da tomu nije tako i da se može puno jednostavnije pristupiti njenom rješavanju, i to na način kako ga u neku ruku pomalo banalnim komentarom na blogu oslikava Eddy Nahmias: “Čini li se to samo meni ili se jako mnogo pojmovne analize odvija u ovoj raspravi?”

Ideja je vrlo jednostavna: čitava ova rasprava sadrži veliku količinu pojmovne analize, analizu koju Balaguer priznaje kao metafizički relevantnu kada se ona bavi artikulacijom pojmova, ali je ne priznaje kao relevantnu kada nastoji argumentima pokazati da je jedan pojam inkohherentan, bolji od drugog ili jednostavno zbog određenog argumenta neprihvatljiv. Čini mi se da je drugi aspekt pojmovne analize, iznalaženje argumenata za određenu analizu ili protiv nje, u funkciji prvog aspekta. Mi ne možemo započeti, a kamoli pak formirati i predložiti prvi aspekt bez drugog. Primjerice, Balaguer se u raspravi na blogu često koristi sintagmom “jasno poimanje slobode” ili “jasno poimanje indeterminiranosti” (Balaguerov komentar na blogu), ali do jasnoće dolazimo upravo pročišćavanjem zamagljenih, nedorečenih, emocijama

ili drugim shvaćanjima okaljanih pojmova i pronalaženjem jasne slike onoga o čemu govorimo – argumentima. To je pogotovo važno u raspravi o slobodi volje unutar koje se miješa čitav niz ne nužno koherentnih intuicija koje čine ovaj problem još kompleksnijim. Zbog svega ranije navedenog, čini mi se da Balaguer ne želi priznati da dio argumentacije koja se odvija, a tiče se samih pojmova, jest metafizički. Mislim da se tijekom rasprave Balaguer implicitno bavi metafizičkom argumentacijom *par excellance*, ali je eksplicitno kao metafizičku ne želi priznati. No, pri tome ide predaleko, tvrdeći da su debata o tome što sloboda volje jest i je li kompatibilna s determinizmom semantičke rasprave lišene metafizičke težine.

Možda Balaguerova želja da se izbjegne govor o pitanju što je sloboda volje i njegova želja da to pitanje klasificira kao nebitno proizlazi iz zbrke koju nudi taj pojam. Kao što sam već pokazao kroz primjer koji nudi McKenna, mi možemo govoriti o više vrsta kontrole pa samim time i o više vrsta slobode volje, ali kako i sam McKenna govori, pojam slobode volje svojevrstan je tehnički termin koji možda ne govori samo o jednoj, 'istinskoj' slobodi volje, nego je jedna vrsta zbirne imenice pod čiji krov svrstavamo niz srodnih problema koji se javljaju u debati. Odrediti koje je ono pravo poimanje izuzetno je teško jer vjerojatno ni jedno ne uspijeva zadovoljiti sve zahtjeve koje postavljamo pred takav pojam. Možda iz tog razloga Balaguer odbija govoriti u tim terminima, o pitanju što sloboda volje jest i stipulira L-slobodu kao nešto na što se može fokusirati suvremena znanost. U tome mu pristup nije pogrešan, ali griješi kada interpretira izuzetno kompleksno, čak možda nemoguće pitanje "što sloboda volje jest?" kao trivijalno. A čak i za tu tvrdnju, za nemogućnost, odnosno trivijalnost, treba nam argument, koji Balaguer i daje!

Zaključno, ono što se može reći na temelju kritika koje sam ponudio jest da je težnja k interpretaciji rasprave o slobodi volje kao besmislenom pseudoproblemu, odnosno kao trivijalnom problemu – neutemeljena. Takve težnje ne uspijevaju jer se i same, ulaskom u raspravu, bave metafizikom. Borbenost, pozivi na evidentnost vlastite pozicije, optužbe za pseudoraspravu onda predstavljaju samo retoričko oruđe, intuiciju koju tek treba dokazati argumentima. Time se pokazalo da se u raspravi o slobodi volje ne iznose besmisleni ili trivijalni iskazi nego naprotiv, da se vodi smisljena, punokrvna metafizička rasprava.

Popis literature

- Ayer, A. J., 1954. "Freedom and Necessity". U: *Philosophical Essays*. Greenwood Press. str. 3–20.
- Balaguer, M. 2004. "A Coherent, Naturalistic, and Plausible Formulation of Libertarian Free Will". *Noûs* 36 (3). str. 379–406.
- Balaguer, M. 2009. "Reply to Levy". http://gfp.typepad.com/the_garden_of_forking_pat/2009/06/the-garden-readinggroup-comment-now-open.html (stranica posjećena 9.9.2020.)
- Balaguer, M. 2010. *Free Will as an Open Scientific Problem*. MIT Press.
- Berčić, B. *Filozofija Bečkog Kruga*. KruZak. Zagreb. 2002.
- Bradley, R. D. 1958. "Free Will: Problem or Pseudo-problem?". *Australasian Journal of Philosophy*. 36:1. str. 33–45.
- Campbell, C. A. 1951. "Is Free Will a Pseudo-Problem?". *Mind* 60. str. 245–261.
- Čeč, F. 2013. "Utjecaj eksperimentalne filozofije na problem slobode volje". U: Boršić, L., Skuhala Karasman, I.(ur.). *Filozofija u dijalogu sa znanostima*. Institut za filozofiju. Zagreb. str. 291–329.
- Cornman, J. W., Lehrer, K., Pappas, G. 1992. *Philosophical Problems and Arguments: An Introduction*. Hackett Publishing.
- De Caro, M. 2007. "How To Deal with the Free Will Issue: The Roles of Conceptual Analysis and Empirical Science". U: Maraffa, M., De Caro, M., Feretti F. (ur.). *Cartographies of the Mind*. str. 255–268.
- Hobart, R. E. 1934. "Free Will as Involving Determinism and Inconceivable Without It". *Mind*. 43 (169). str. 1–27.
- Hume, D. 1983. *Rasprava o ljudskoj prirodi*. Veselin Masleša. Sarajevo.
- Kane, R. 2005. *A Contemporary Introduction to Free Will*. Oxford University Press.
- Levy, N. 2009. "Just What is that is Metaphysically Irrelevant? A Response to Balaguer". http://gfp.typepad.com/the_garden_of_forking_pat/2009/06/the-gardenreading-group-comment-now-open.html (stranica posjećena 9.9.2020.)
- McKenna, M. 2012. "The Metaphysical Importance of the Compatibility Question: Comments on Mark Balaguer's *Free Will as an Open Scientific Problem*". *Philosophical Studies Online*. str. 1–12.
- McKenna, M., Pereboom, D. 2016. *Free Will: A Contemporary Introduction*. Routledge.
- Nahmias, E. 2011. "Intuitions About Free Will, Determinism, and Bypassing". U: Kane, R. (ur.). *The Oxford Handbook of Free Will*. 2nd edition. Oxford University Press.

Schlick, M. 1939. *Problems of Ethics*. Prentice-Hall Inc.

Strawson, P. F. 1962. "Freedom and Resentment". U: McKenna, M., Russell, P. (ur.) 2008. *Free Will and Reactive Attitudes: Perspectives on P. F. Strawson's "Freedom and Resentment"*. Ashgate. str. 19–3.

Alternativne mogućnosti i odgovornost: diskusija o načelu alternativnih mogućnosti¹

Davor Pećnjak

Problem slobode volje i problem moralne odgovornosti dva su različita problema, općenito govoreći, no oni su u nekim pogledima blisko povezani i imaju važnu vezu na mjestu gdje je naše poimanje moralne odgovornosti snažno povezano uz način na koji konstruiramo naše sposobnosti ili nesposobnosti da odlučujemo, namjeravamo i izvodimo djelovanja. S druge strane, mi ne moramo imati krajnje rješenje problema slobode volje (i, u skladu s tim, našeg slobodnog ili neslobodnog odlučivanja, namjeravanja i djelovanja) da bismo bili u poziciji za diskutiranje moralne odgovornosti. Rješenje problema slobode volje ne treba biti dano prije rješavanja problema koji se tiču moralne odgovornosti. Mi možemo diskutirati i argumentirati o različitim mogućim rješenjima moralne odgovornosti čak i samo unutar pretpostavke što *bi bio* slučaj ako je ovo ili ono stajalište o slobodi volje, slobodi odlučivanja ili slobodi djelovanja, istinito. Mi možemo formulirati neke općenite načela moralne odgovornosti i ispitivati ih pod različitim mogućim pretpostavkama i scenarijima.

Jedan od najvažnijih, ili čak najvažnije načelo u diskusiji o odgovornosti sadržan je u frazi “mogao je učiniti drugačije”. Naime, i intuitivno, i više tehnički, odnosno filozofski, mi mislimo da je netko odgovoran za ono što je učinio samo ako je mogao učiniti i drugačije. Naravno, ovdje se podrazumijeva da je netko mogao učiniti i drugačije pod potpuno istim prethodnim uvjetima. Kada bi okolnosti bile čak minimalno drugačije i kada bi u takvim, minimalno drugačijim okolnostima osoba postupila drugačije, tada bi zapravo cijela situacija bila potpuno drugačija, pa to ne bi bilo važno za razmatranje moralne odgovornosti pod *potpuno istim* antecedentnim uvjetima za djelovanje u pitanju.

Samo jedna kratka opaska prije nego se upustimo u glavnu diskusiju. Kada bi se radilo o čvrstom determinizmu odnosno inkompatibilističkom determinizmu, diskusija o moralnoj odgovornosti ne bi bila potrebna. Takav tip determinizma drži da sve što se dogodi jest – bez obzira je li to protjecanje rijeke, razvoj svemira, igra biljara, padanje kamena, filozofovo pisanje članka, igranje tenisa od strane profesionalca ili amatera, jednostavno podizanje ruke ili bilo što drugo – u potpunosti određeno početnim uvjetima u svemiru i zakonima prirode. Ništa ne bi moglo biti drugačije nego što jest, nijedna stvar, nijedan proces, nijedna radnja ne bi mogla biti drugačije učinjena, uključujući sve ljudske radnje. U takvom svijetu ne bi bilo pravih djelovatelja, niti čovjekove slobode volje; sve ono što su ljudi činili ili će tek učiniti ne bi ovisilo o njihovom vlastitom djelovanju, nego samo o prethodnim uvjetima, stanjima i prirodnim zakonima, te bi bilo neizbježno. Tada ne bi bilo mjesta za bilo kakvu vrstu odgovornosti budući da su ti uvjeti izvan utjecaja i kontrole svakog pojedinca, te su sami za sebe dostatni da bi se nešto dogodilo. U takvom svijetu moralna odgovornost jednostavno ne bi postojala. Mislim da je to ispravno stajalište moralne odgovornosti ako se radi o čvrstom determinizmu.

No, budući da još uvijek u potpunosti ne znamo kakav je naš svijet, razmotrit ćemo mogućnosti moralne odgovornosti. Kao što sam već spomenuo, trebali bismo se usredotočiti na načelo izraženo u frazi “mogao je učiniti drugačije”. Ono se naziva “načelo alternativnih mogućnosti” (u daljnjem tekstu PAM) i formulirano je na sljedeći način:

(PAM) Osoba je moralno odgovorna za svoja djelovanja samo ako je mogla učiniti drugačije.

Neko vrijeme ovo se načelo smatralo toliko očitim da se uglavnom držalo kao neosporivo. Početak njegova osporavanja nalazimo u poznatom Frankfurtovom (1969) članku, kojim se trudio dokazati pogrešnost navedenog načela. Ovdje navodim Frankfurtov protuprimjer i njegove argumente. Budući da ga je najbolje objasniti njegovim vlastitim riječima, citirat ću najznačajniji dio Frankfurtova (1969, str. 835) članka kako to i ostali čine:

Pretpostavimo da netko – nazovimo ga Black – želi da druga osoba – Jones – izvrši određeno djelo. Black je spreman poduzeti određene mjere, ali se radije

ne bi nepotrebno izlagao. On će, dakle, pričekati da Jones donese odluku, i neće poduzeti ništa ukoliko mu ne postane jasno (Black može odlično prosuditi takve stvari) da je Jones odlučio nešto različito od onoga što Black želi da on učini. Pa ako, dakle, postane jasno da će Jones odlučiti nešto drugo, Black će poduzeti učinkovite mjere kako bi osigurao da Jones donese odluku kakva će odgovarati Blacku, i da to zatim i učini. Bez obzira na to kakvu će odluku donijeti Jones, na kraju će se dogoditi onako kako to Black želi...

...Pretpostavimo sada da Black nikada nije iskazao Jonesu svoje namjere, jer je Jones, iz vlastitih razloga, odlučio poduzeti radnju koju je Black želio da on poduzme.

Da bih dodatno razjasnio prethodni citat, poslužit ću se donekle izmijenjenom pričom iz Ginetovog (1996) članka. Pretpostavimo da Jones razmišlja o ubojstvu određene osobe – Smitha. Tu je i druga osoba, Black, koja također želi da Jones ubije Smitha. Ali Black se radije ne bi izlagao ako to nije nužno. On može prisiliti Jonesa da se odluči na ubojstvo i zatim ubije Smitha ako to Jones sam ne odluči učiniti; on postavlja mehanizam koji će dostatno uzročno utjecati na Jonesa da ubije Smitha u slučaju da Jones odluči da ga neće ubiti. Naravno, Jones nije svjestan postojanja takvog mehanizma i o njemu ništa ne zna.

Ovaj primjer je zamišljen tako da bude očito da osoba koja je u pitanju, Jones, nije mogla učiniti drugačije (ako odluči da neće učiniti ono što Black hoće da on učini, tj. da ubije Smitha, Black će poduzeti mjere koje će neizbježno rezultirati time da ga Jones ipak ubije). Bez obzira na svoju početnu odluku, Jones čini A; ne može to izbjeći. Ali, zapravo, Jones može ubiti Smitha na dva ponešto različita načina: prvi, da to učini na temelju vlastite odluke, bez utjecaja Blackovog mehanizma; i drugi, dostatno uzrokovano Blackovim mehanizmom. U slučaju kad na djelovanje utječe mehanizam, pa je Jones primoran ubiti Smitha, Jones neće snositi odgovornost. No, iako Jones ne može učiniti drugačije, Frankfurt smatra da će Jones biti odgovoran ako ubije Smitha iz vlastitih razloga, bez djelovanja mehanizma. Iako Jones ne može učiniti drugačije, u posljednjem slučaju bit će odgovoran (jer su ga na ubojstvo nagnali samo vlastiti razlozi). Dakle, načelo alternativnih mogućnosti nije valjan.

Naime, kada Jones ubije Smitha zbog vlastitih razloga, bez pritiska mehanizma, čini istu radnju koju bi učinio i da mehanizma nije ni bilo. Bit će odgovoran za ubojstvo Smitha i u slučaju kad ga ubija bez posredstva mehanizma koji bi ga na to mogao prisiliti, kao i u slučaju kad

postoji mehanizam, ali se ne aktivira jer Jones sam donosi odluku o ubojstvu Smitha.

Čini se da je tvrdnja da PAM nije generalno vrijedeći prilično uvjerljiva. Ali ne misle svi tako. Bilo je pokušaja opravdavanja PAM-a pokazivanjem kako Frankfurtov primjer nije dovoljno uvjerljiv. Volio bih, ispitujući Ginetove (1996) napade na Frankfurtov članak, razmotriti ovaj detaljniji slučaj i ponovo promisliti o primjeni PAM-a i njegovoj mogućoj općoj nevaljanosti.

Glavni se problem sastoji u sljedećem. Slučaj je prerađen na takav način da ne odudara puno od originalnog Frankfurtovog primjera. Kao što smo već vidjeli, postoje tri različita momenta u vremenu. Black postavlja mehanizam koji će uzrokovati da Jones ubije Smitha do t_3 ako Jones to već nije učinio do određenog trenutka u t_2 . Dakle, ako Jones nije ubio Smitha do trenutka t_2 , u t_2 se aktivira mehanizam i čini uzročno nužnim da Jones ubija Smitha do t_3 . To aktiviranje onemogućuje Jonesa da izbjegne ubijanje Smitha do t_3 . Mehanizam nije aktiviran zbog toga što Jones u t_1 ubije Smitha. Uzmimo sada nešto dulji citat iz Ginetovog (1996, str. 406–407.) članka koji bi trebao poslužiti da nam pokaže, prema ovom gledištu, kako je Frankfurt pogriješio te da obrani PAM:

A sada presudno pitanje. Je li Jones moralno odgovoran kada do t_3 učini B, una-toč činjenici da nije mogao izbjeći da nastane takvo stanje stvari? Frankfurtov je zaključak da je on za to odgovoran, budući da Jones čini B do t_3 na temelju svoje odluke, bez posredstva mehanizma, i da bi učinio B do t_3 na isti način zbog istih razloga čak i da mehanizam nije ni postojao i da je on mogao izbjeći činjenje B do t_3 . Ne slažem se s tom intuicijom. Ipak, imam intuiciju da za nešto Jones mora biti odgovoran, za nešto što za nužnu posljedicu ima njegovo činjenje B do t_3 , tj. njegovo činjenje B baš točno u ono vrijeme u kojem je to zaista učinio, t_1 . Kasnije stanje stvari je ono stanje čije je nastajanje Jones mogao izbjeći (iako, da ga je izbjegao, učinio bi B u neko kasnije vrijeme, između t_2 i t_3). Jonesovo činjenje B u t_1 razlikuje se od činjenja do t_3 samo prema tome što je temporalno specifičnije. Moja je intuicija da Jones nije odgovoran za nastajanje stanja stvari koje je manje temporalno specifično, jer, zbog postojanja Blackovog mehanizma, Jones to nije mogao izbjeći, ali može biti odgovoran za nastajanje temporalno specifičnijeg stanja stvari koje je mogao izbjeći. Ovaj fenomen o postojanju odgovornosti za nastajanje specifičnijeg stanja stvari dok istodobno ne postoji odgovornost za implicirano nastajanje određenog manje specifičnog stanja stvari nije uopće neuobičajen. Ja sam, na primjer, odgovoran za svoj trenutni boravak u nekoj sobi, ali nisam odgovoran za to što se nalazim na udaljenosti unutar milijun milja od centra zemaljske kugle. Sklon sam smatrati da privlačnost Frankfurtovih zaključaka o navedenom primjeru proizlazi iz neuspjeha u razlikovanju dva stanja stvari, Jonesovo činjenje B u t_1 i Jonesovo činjenje B do t_3 , koji se međusobno razlikuju samo po temporalnoj specifičnosti. (...)

Zamijetite da ne govorim da Jones može biti odgovoran za to što čini B na temelju vlastite odluke prije nego kao rezultat Blackovog mehanizma. Bilo bi pogrešno tvrditi da je tako. Budući da Jones nije bio ni najmanje svjestan postojanja Blackovog mehanizma, nije znao ni mogao znati za činjenicu koju izražavam govoreći "činim B na temelju vlastite odluke, prije nego kao rezultat Blackovog mehanizma", no on je mogao biti odgovoran za postizanje tog učinka samo ako je, u to vrijeme, znao ili morao znati da tako postupa. Smatram da Jones može biti odgovoran za to što je učinio B u t_1 . On je znao da će postići taj učinak.

Prvo, mi ne određujemo, u ovakvim primjerima, nečiju odgovornost prema tome što netko zna ili čega je svjestan. Odgovornost se određuje prema činjenicama u određenom slučaju ili prema stanju stvari koje jest ili nije postignuto, neovisno o tome što netko subjektivno zna ili čega je svjestan. Dakle, Jonesova odgovornost se ne prosuđuje prema onome čega je svjestan i što može reći. Treba je prosuditi prema onome što Ginet smatra pogrešnim, a to je je li to Jones učinio svojevolumno ili nije (potpuno je nevažno da li on zna ili je svjestan toga da je to učinio samovoljno ili je na njega utjecala neka sila). Stoga ovaj Ginetov potez ne uspijeva pokazati da je pogrešno procjenjivati Jonesovu odgovornost na temelju toga je li što učinio svojom voljom ili nije.

Primijetite da u Ginetovom članku postoji opravdavanje Jonesa isto kao i u takvim situacijama kod Frankfurta, i to samo u slučaju kada je Blackov mehanizam aktiviran. Ginet nije dao objašnjenje zašto bi temporalno specifičniji ili manje specifičan učinak bio relevantan za postojanje moralne odgovornosti. Ginet bi pobio Frankfurtovu teoriju samo ako bi definitivno dokazao da, ako Jones ubije Smitha do t_3 uz prisutnost Blackovog mehanizma, ali bez njegovoga aktiviranja (što znači da Jones ubija iz vlastitih razloga, pa bi do istog rezultata došlo i da nije postojao mehanizam) on neće biti odgovoran za učinjeno do t_3 . Ali Ginet nigdje ne kaže takvo što, i čak se i ne trudi to pokazati. No sama prisutnost mehanizma, iako on stvara situaciju u kojoj Jones ne bi mogao učiniti drugačije, u slučaju da on nije aktiviran, ne dovodi Jonesa do ubijanja Smitha u bilo kojem trenutku. Tako da se čini kako stanje stvari koje je temporalno ili više ili manje specifično ne igra znatnu ulogu u procjeni moralne odgovornosti.

Kako bismo razjasnili stvar, ponovo ćemo razmotriti primjenjivost PAM-a. U nastavku slijedi pomalo drugačija situacija, koja bi trebala pomoći da jasnije sagledamo problem.

Zamislimo sljedeću situaciju:

Jones u t_1 odlučuje ubiti Smitha, i priprema se na to, ali u t_2 , prije nego to učini, mehanizam se aktivira, neovisno o Jonesovoj ranijoj odluci da ubije Smitha. Mehanizam, naravno, dovodi do toga da u t_3 Jones ubije Smitha. Nakon aktivacije mehanizma, Jones sigurno ne bi mogao odlučiti da postupi drugačije, jer je potpuno onemogućen u promjeni odluke i ne može izbjeći ubijanje Smitha nakon t_2 . Do t_2 mogao je odlučiti da učini drugačije, ali nakon t_2 više ne. U t_3 on ubija Smitha. Da nije postojao mehanizam, Jones bi mogao odlučiti da postupi drugačije, i tako i postupiti u vremenu između t_1 i t_3 . Dakle, od t_2 onemogućeno mu je odlučiti da učini nešto drugo, što bi inače, da nije došlo do aktiviranja mehanizma, mogao. Njegovo vrijeme za odluku o drugačijem postupanju bilo bi dulje. Je li u takvoj situaciji on odgovoran za ubojstvo Smitha? Čini se da odgovor ovisi o razmatranju kako, odnosno pod kojim je okolnostima, Jones ubio Smitha, a ne kada, ili kada specifičnije ili manje specifično. Jaka intuicija dovela nas je do toga da uzmemo u obzir u kojoj mjeri je on prisiljen ili uzročno određen nekom silom izvan sebe to učiniti, ili to učiniti bez prisile, odnosno iz nekih vlastitih razloga.

Pa, kako se čini da frankfurtovski protuprimjer za PAM nije odbačen, odgovorio bih na pitanje pozitivno – on je odgovoran, ali možda tek djelomično. Kako to? U vremenu prije t_2 , prije aktiviranja mehanizma, namjeravao je zbog vlastitih razloga ubiti Smitha. Nakon t_2 , aktivacija mehanizma je dovoljan uzrok da natjera Jonesa da ubije Smitha. Od t_2 Jones nije mogao promijeniti odluku i nije mogao odustati od ubijanja Smitha. Ali to, napokon, ne oslobađa Jonesa od djelomične odgovornosti. Zašto? Jer je od t_1 do t_2 mogao učiniti drugačije i mogao je odlučiti da ne ubije Smitha. Nakon t_2 , Jonesovo ubijanje Smitha bilo je nužno uzrokovano mehanizmom, bez obzira na njegovu raniju odluku. U tom slučaju Jones ne bi bio odgovoran. Ali budući da je već odlučio ubiti Smitha (zbog vlastitih razloga), on će biti djelomično odgovoran za ubojstvo Smitha. Naravno, neće biti odgovoran za vremenski interval između t_2 i t_3 (u t_3 on ubija Smitha u ovom izmijenjenom slučaju) jer je bio potpuno onemogućen u promjeni svoje odluke, određen mehanizmom, odnosno okolnostima potpuno izvan njegove kontrole; bez mogućnosti donijeti drugačiju odluku i učiniti nešto drugo (naime, ne ubiti Smitha). Iz toga slijedi da će, iako nije mogao učiniti drugačije, biti

djelomično odgovoran u takvoj situaciji. To će se dogoditi samo zato što u fazi u kojoj je mogao odlučiti drugačije, nije tako odlučio i tako je sam sebi onemogućio da bude oslobođen odgovornosti kada se, u t_2 , aktivira mehanizam. U t_2 , dakle, postoje dvije mogućnosti: Jones ulazi u t_2 sa vlastitom odlukom da ubije Smitha, ili ulazi u t_2 bez nje, dakle, odlučuje da ga ne ubije. Situacija u kojoj se mehanizam aktivira kada Jones nije donio nikakvu odluku i nema namjeru ubiti Smitha, razlikovat će se od one kada je Jones donio takvu odluku i namjerava ga ubiti. Sigurno je da će biti odgovoran za odluku i svoju namjeru koju je donio iz vlastitih razloga.

Nikoga nećemo smatrati odgovornim zbog onoga što je učinio, ako je odjednom stroj potpuno preuzeo kontrolu nad njim, i ako ta osoba nije ranije donijela nikakvu odluku o djelovanju na koje ga je stroj prisilio; drugačije će biti ako je osoba prije aktiviranja mehanizma već donijela odluku da će učiniti isto na što je kasnije bila prisiljena djelovanjem stroja (sve to, dakako, bez obzira na to da li osoba zna ili ne zna za aktiviranje mehanizma).

Ulaskom u t_2 s već donjetom odlukom da ubije Smitha, na Jonesa pada barem djelomična odgovornost, čak i kada je, od t_2 nadalje, stroj bio uzročno dostatan izazvati Jonesovo ubijanje Smitha, bez obzira na njegovu prijašnju odluku.

Nadam se da ova razmatranja, iako kratka, pokazuju kako se još uvijek čini da je Frankfurt bio u pravu – odnosno, da PAM nije generalno primjenjiv. Međutim, smatram da primjenjivost navedenog načela ne treba u potpunosti isključiti, zbog važnosti pojmova kao što su “biti prisiljen učiniti nešto”, “prisila”, “uzročno prinuđen nešto učiniti” itd. Oni također imaju značajnu ulogu u određivanju odgovornosti. U njihovoj kombinaciji s PAM-om, možda možemo dati preciznije određenje PAM-a: kako i kada se može primjenjivati, iako samo načelo nije univerzalno primjenjivo. No, to je već tema za drugi članak.

Bilješka

1. Tekst je izvorno objavljen na engleskom jeziku pod naslovom “Alternate Possibilities and Responsibility: A Discussion of PAP” u *Studia hermeneutica*, Neue Folge, band 3, 2005: 65–73. Urednici: Hans Ineichen i Jure Zovko. Zahvaljujem se urednicima na dozvoli da se objavi prevedeni tekst na hrvatskom jeziku. Zahvaljujem se Lidiji na prijevodu.

Literatura

- Frankfurt, H. G. (1969): “Alternate Possibilities and Moral Responsibility”, *Journal of Philosophy*, vol. 66: 829–839.
- Ginet, C. (1996): “In Defense of the Principle of Alternate Possibilities: Why I Don’t Find Frankfurt’s Argument Convincing”, *Philosophical Perspectives*, vol. 10: 403–417.

Post scriptum 2019.

Mogli bismo reći i ovako: kada Jones čini ono što čini kada se aktivira Blackov mehanizam, dakle pod neizbježnim uzročnim utjecajem Blackovog mehanizma, tada on više nema svoju volju niti nadzor nad samim sobom, svojim odlučivanjem i svojim činjenjem – dakle, Jones to ne čini više kao osoba i djelovatelj, nego je potpuno izmanipuliran te postaje samo sredstvo tuđeg činjenja. Od trenutka aktiviranja Blackova mehanizma, naravno da Jones više ne može učiniti drugačije nego što “diktira” Blackov mehanizam. Ovo nam pokazuje samo ono što smo i prije znali – da ako je na djelu potpuni determinizam čimbenicima koji nisu ni u kakvoj moći djelovatelja – onda djelovatelj nije odgovoran za to što čini. No, ako je Jones odlučio ubiti Smitha te se Blackov mehanizam ne uključuje, i doista Jones ubije Smitha bez ikakvog uključivanja rečenog mehanizma, tada zapravo još ne znamo automatski da li je Jones odgovoran za to što je odlučio i učinio; jer nam ništa u primjeru nije rečeno u kakvom je to svijetu učinio. Ako je svijet deterministički, tada Jones nije, iako se Blackov mehanizam ne uključuje, odgovoran za odluku da ubije Smitha i za ubojstvo Smitha, jer su njegove odluke i ono što

čini (na osnovi odluka) također učinci čimbenika nad kojima on nema kontrolu – naime to su učinci prethodnih stanja svijeta i zakona prirode. Ako je svijet libertarijanski, onda bi Jones bio odgovoran, jer do samog ubojstva, *ex hypothesi* za libertarijanski svijet, može promijeniti odluku i postupiti u skladu s promijenjenom odlukom i ne ubiti Smitha. No, u libertarijanskom svijetu, tada nastupa Black, i nakon što Jones donese odluku da ne ubije Smitha, uključuje se mehanizam koji deprivira Jonesa od njegove vlastite volje, vlastitog odlučivanja i vlastitog činjenja i potpuno ga, čimbenicima potpuno izvan Jonesova nadzora, prisiljava da se s njime čini ono što Black želi, te Jones postaje samo Blackovo sredstvo (i nije djelovatelj u pravom smislu).

U stvari, smatram da u izvornoj situaciji opisanoj od stane Frankfurta postoji alternativna mogućnost za Jonesa samo ako je svijet libertarijanski – da li će *odlučiti* ubiti Smitha svojom voljom ili će *odlučiti* ne ubiti ga pa će ga ubiti zapravo Black, rabeći Smitha samo kao sredstvo. U slučaju frankfurtovske situacije u determinističkom svijetu, Jones nije odgovoran za ubojstvo Smitha čak i ako se ne aktivira Blackov mehanizam.

Jones, dakle, može izbjeći ubiti Smitha zbog svojih razloga (u libertarijanskom svijetu). U determinističkom svijetu ionako nema svojih razloga niti svoju volju – svi razlozi i volja koji se čine da su njegovi, zapravo to nisu jer su proizvod prethodnih stanja i zakona prirode nad kojima Jones nema nikakvog utjecaja pa nema utjecaja ni na misli, razloge i volju koje se njemu zapravo samo pojavljuju i izgledaju samo subjektivno kao da su njegov vlastiti proizvod.

U ovom svjetlu, možda sam trebao dati nešto drugačiji zaključak nego što sam ga iznio na kraju izvornog teksta.

01/2022.

ISBN 978-953-352-081-0



80,00 kn