

Društvo za analitično filozofijo in filozofijo znanosti

časopis za kritično misel

ANαliZA

3-4

2016 letnik 20.

ANαliZA, časopis za kritično misel, številka 3-4, leto XX., 2016

Izdajatelj in založnik: Društvo za analitično filozofijo in filozofijo znanosti,
Aškerčeva 2, 1000 Ljubljana
<http://www.drustvo-daf.si/analiza/>
e-naslov: analiza@drustvo-daf.si

Naročila sprejemamo na zgornji naslov.

Na leto izidejo trije zvezki Analize (dve enojni in ena dvojna številka).
Cena zvezka v prosti prodaji je 6,00 EUR, za naročnike 4,00 EUR oz. letna naročnina 12,00 EUR (z vključenim DDV).

Izdajateljski svet:

Bojan Borstner, David Owens, Matjaž Potrč, Marko Uršič, Matthias Varga von Kibéd, Elizabeth Valentine, Bojan Žalec

Glavna in odgovorna urednica: Maja Malec

Uredniški odbor:

Smiljana Gartner, Božidar Kante, Vojko Strahovnik, Toma Strle, Borut Trpin, Sebastjan Vörös

Lektor: Borut Cerkovnik

Oblikovanje: K.reg@

Tisk: Demat, d.o.o., Stegne 3, 1000 Ljubljana

Naklada: 200 izvodov

Izid te revije je sofinancirala

JAVNA AGENCIJA ZA KNJIGO REPUBLIKE SLOVENIJE.

Revija Analiza je indeksirana v THE PHILOSOPHER'S INDEX.

ISSN 1408-2969

Kazalo

Miselni eksperiment v Platonovi *Državi*

Nenad Miščević

Platonova *Država* kot politični miselni eksperiment..... 5

Filozofija znanosti

Sebastijan Vörös in Blaž Cotman

Znanost proti psevdoznanosti: problem ali psevdoproblem?..... 21

Andraž Ivančič

Popperjev falzifikacionizem: kritični pregled 43

Nika Škof in Tadej Pirc

Nedosegljivost ideala znanstvene objektivnosti 55

Filozofija narave

Peter Lukan

Temeljni pojmi Boškovičevega atomizma v perspektivi vznikanja moderne znanosti... 71

Praktična etika

Nika Češnovar

Etični vidiki prekinitve aktivnega zdravljenja 89

Sandra Vajda

Abolicionizem in moralna dopustnost uporabe živali v raziskovalne namene..... 105

Naturalistični pristop k svobodi in moralnosti

Šemsudin Džopa

Naturalistični vidiki svobodne volje in moralne odgovornosti 121

Utelešena umetna inteligenca

Tine Kolenik

Kaj imata skupnega Heidegger in Roomba: utelešena umetna
inteligenca kot peskovnik kontinentalne filozofije..... 147

Utelešena kognicija in reprezentacije

Martina Tomšič

Reprezentacije pri telesnem zavedanju 167

Nenad Mišćević

Prevedla Tina Oblak

Platonova *Država* kot politični miselni eksperiment*

Platonova *Država* je političen miselni eksperiment. Miselno eksperimentiranje je napovedano v zgodbi o Gigesovem prstanu, opravljeno pa je na temeljit in sistematičen način skozi vrsto političnih scenarijev: družbena lastnina blaga, žensk in otrok, izobraževalnega sistema in vladanje filozofa. Članek obravnava staro vprašanje verjetnosti, znotraj sedanjih razprav o miselnih poskusih, in pa vprašanje zamenljivosti: ali je mogoče določen politični miselni eksperiment zamenjati z argumentom, ki vsebuje samo norme in empirične podatke? Članek prav tako postavlja miselni eksperiment o *Državi* v širši zgodovinski kontekst, s tem ko ga predstavi kot točko izvora utopične literature na eni strani in miselno-eksperimentalne tradicije politične filozofije na drugi strani. Pri tem ga primerja s socialno-pogodbenim miselnim eksperimentom, ki je tudi nakazan v *Državi*, popolnoma razvit pa v sodobni misli.

Ključne besede: politika, miselni eksperimenti, Gigesov prstan, epistemski, motivacijski, utopija, Platon

1. Uvod

Miselni eksperimenti v antiki so v zadnjem času postali modna tema v profesiji, o čemer pričajo članki pa tudi kakšna knjiga.¹ Avtorji razpravljajo o (pro-)to)znanstvenih (kot je Arhitasov dokaz končnosti fizičnega sveta) in filozofskih miselnih eksperimentih (kot je Karneadov splav-deska), ne omenjajo pa kaj dosti miselnih eksperimentov s političnega področja. Naj omenim, da izgradnja idealne države v *Državi* ni splošno priznana kot miselni eksperiment. Santas (2006) v poglavju o metodologiji v *Državi* sploh ne omeni miselnega eksperimenta, nedavna avtoritativna knjiga Schofielda (2006) pa se prav tako izogiba temu izrazu. Omembo, da gre za miselni eksperiment, mi je uspelo najti v učbeniku za dodiplomske študente, ki jo je napisal Cohen (2005). Ta opiše zgradbo knjige kot »skrbno izdelan miselni eksperiment«, in tudi omeni, da njen miselno-

* Mišćević, N. (2012). »Plato's Republic as a Political Thought Experiment«. *Croatian Journal of Philosophy*, 12(2), str. 153–165.

¹ Glej na primer Ierodiakonou in Roux (2011).

eksperimentalni značaj, kot se blago izrazi, »ni tako pogosto prepoznan« (Cohen, 2005: 2). Nicholas Smith (1999 in 2007) je sicer prišel dokaj blizu ugotovitvi, da je Platonova *Država* miselni eksperiment, vendar se je bolj osredotočil na naravo podob in pa na funkcijo domišljije v izobraževanju.²

V tem prispevku želim obravnavati prav temo političnih miselnih eksperimentov, in počastiti spomin na Majo Hudoletnjak Grgić tako, da se osredotočim na najbolj klasični, prvi in pionirski primer političnega miselnega eksperimenta iz Platonove *Države*.³

Najobsežnejši del *Države*, ki se nanaša na vprašanja v politični filozofiji, je sestavljen iz niza manjših miselnih eksperimentov, organiziranih v en velik miselni eksperiment. Vsaj tako sam trdim. Naslednji razdelek na kratko predstavi pojem miselnega eksperimenta na primeru Gigesovega prstana. Tretji razdelek tvori osrednji del prispevka in obravnava *Državo*. Prvi podrazdelek se začne s kratko obravnavo strukture »mikro-miselnih eksperimentov«, ki so sestavni deli velike celote, in se vpraša o njihovi verjetnosti in nepogrešljivosti. Potem se osredotoči na zgradbo celote, spet s poudarkom na pravkar omenjenih dveh vprašanjih. Drugi podrazdelek obravnava status *Države* kot utopije in razmerje med političnim miselnim eksperimentom in utopijo. Tretji podrazdelek v grobem oriše osnovno idejo Aristotelove kritike in predlaga dokaj grobo sliko dveh tradicij – eno za in drugo proti miselnim eksperimentom v politični filozofiji –, pri čemer na kratko prikaže, da njuna debata zajema največji del zgodovine metodoloških vprašanj v politični filozofiji.

2. Gigesov prstan in narava miselnih eksperimentov

Naj začnem s samo idejo miselnih eksperimentov (v nadaljevanju »ME«). *Država* vsebuje očiten in neproblematičen primer – Gigesov prstan. Da bi pokazal, da tisti, ki ravna po pravičnosti, to počnejo proti svoji volji in zato, ker nimajo moči, da bi delali krivico, Glavkon predlaga, da si zamislimo osebo, ki je povsem v njeni moči, da dela krivico, jo nato spremljamo v domišljiji in opazujemo, kaj bo storila. Potem bi se lahko postavili v njene čevlje in tako spoznali, kako bi se večina ljudi obnašala v takem položaju. V zgodbi bo ta moč izhajala iz prstana, ki svojemu lastniku omogoči, da po želji postane neviden. Scenarij bo konkreten, z določeno osebo, Gigesom, kot protagonistom. Prvi rezultat se bo nanašal samo nanj, nato na »nas« in končno na ljudi nasploh. Poslušalec/bralec, pravi Protarh, bo prišel do trdnega prepričanja, da bo nevidni Giges res zagrešil veliko nepravilnih stvari, vključno s krajo, posilstvom in umorom. Protarh se lahko nato vpraša, če

² Tu se moram zahvaliti Nicholasu Smithu, da mi je posredoval svoje članke in pa za razpravo o teh vprašanjih na konferenci.

³ Za bolj splošni pregled miselnega eksperimentiranja v politični filozofiji glej Miščević (v tisku).

bi se v podobni situaciji bil sam sposoben upreti skušnjavam, in spozna, da se ne bi mogel. Iz tega in ker je Giges povprečen posameznik, sam Protarh pa vsekakor nič slabši od večine svojih sodržavljanov, lahko sklepa, da bi večina njegovih sodržavljanov storila enako. Nadalje, če bi večina Atenčanov ravnala tako, potem bi zagotovo tudi večina prebivalcev manj kultiviranih mest. Torej bi se ljudje na splošno slabo obnašali, če jih ne bi omejeval nek zunanji motiv; v tem primeru je to strah pred kaznijo.

Takšne domišljjske vaje, opravljene »v naslonjaču«, brez fizičnega eksperimentiranja in opazovanja, in namenjene odkrivanju resnice (ali nečesa, podobnega resnici) o nekem določenem vprašanju, so osrednji primeri ME. Običajno predstavljajo neresnične situacije, čeprav bi nekateri izmed scenarijev lahko bili resnični ali pa bi jih lahko uresničili. Posamezna sodba, denimo, »Giges bi gotovo storil vse te slabe stvari« ali pa »Tudi jaz bi jih naredil«, je pogosto opisana kot »intuicija«.

Namen ME o Gigesu je odkriti psihološka dejstva. V etiki in politični filozofiji se ME pogosteje uporablja za priklic moralnih intuicij kakor pa dejstvenih prepričanj. Običajno se sprašuje, ali bi bil tak in tak ukrep dopusten, pravičen ali obvezen. Takšne ME najdemo povsod v etiki in politični filozofiji – pomislite na Rawlsovo tančico nevednosti, problem vozička, ali Nozickov stroj izkušenj, če omenimo le nekaj znanih primerov. O njih so se porodila številna vprašanja in v tem prispevku obravnavam dve pomembni v povezavi s Platonovo strategijo. Prvo vprašanje je, kaj se v zamišljenem scenariju predlaga kot novi element in kaj ostane, večinoma implicitno, nespremenjeno. Ko Giges postane neviden, obdrži svoje telesne zmožnosti (odvzame lahko materialne predmete, suče meč in se zelo uspešno ljubi s kraljico) ter svojo motivacijo (najpomembneje, ne postane kar naenkrat boljši človek). Če uporabimo sliko iz filozofije znanosti – nekatere spremenljivke (vidnost) se spremenijo, druge pa se ohranijo (telesne zmožnosti, osebnost). Pa je nastali konstrukt verjeten in verodostojen? Ali nam pove kaj o nas, kakršni resnično smo? Julian Baggini (2007) je poimenoval to težavo problem »nadzora spremenljivk«, kmalu pa bomo videli, kako je bil (vsaj deloma) zapažen že v helenističnih časih. Alternativni način opisovanja te težave se sklicuje na vsebovani scenarij kot »protidejstveno stanje« (ali celo »svet«). Avtorji, kot je denimo Williamson (2007), se sprašujejo, kako se to stanje razlikuje od dejanskega in kako oddaljeno je od njega. No, bil naj bi najbližji svet dejanskemu s samo eno veliko spremembo – nevidnostjo – in nekaj manjšimi spremembami, ki jih ta prinaša (morda Giges lahko hodi okoli gol, nihče ga namreč ne more videti). Kako vemo kaj spremeniti in kaj pustiti tako, kot je v dejanskem svetu? Ta formulacija problema, za razliko od formulacije s pomočjo nadzora spremenljivk, poudarja modalno oddaljenost. Moje vprašanje bo sledeče: ali je Platon uspel nadzorovati svoje spremenljivke oziroma, alternativno, ali je uspel predlagati primeren, bližnji možni svet, ki ni neverjeten ali celo protisloven? Pri tem mi bo Aristotel za pričo, da so težave resne.

Drugi sklop vprašanj se nanaša na načelno uporabnost ME. Ali so ti nepogrešljivo in neizločljivo orodje ali samo hevristično koristna »slikovna« ali celo »metaforična« sredstva? Na primer, Rawls opiše svojo tančico nevednosti kot »sredstvo za predstavljanje«. Ali to pomeni »samo sredstvo« ali »ključno in nepogrešljivo sredstvo«?⁴ Nekateri filozofi, imenujmo jih »inferencialisti« (najbolj znan je John Norton), trdijo, da lahko ME načeloma nadomestimo z argumenti, ki ne vsebujejo premis o tem, kaj je zamisljivo. Torej, domišljjsko »sredstvo za predstavljanje« je zgolj pripomoček. Na primer, Norton opiše svojo teorijo, ki se osredotoča na ME v znanosti, kot »razlago miselnih eksperimentov kot navadnega sklepanja, ki se skriva izza žive slikovne ali pripovedne oblike« (Norton, 2004: 45). Naj poimenujem njihov izziv ME kot »inferencialistični ugovor«. V politični filozofiji bi ugovor bil sledeč: politični ME so zgolj domišljjska sredstva za predstavljanje, ki so načeloma in največkrat tudi v praksi zamenljiva z bolj abstraktnimi argumenti, ki se opirajo na normativna načela in empirične premise. Vrnimo se k Gigesovem prstanu, in sicer h Glavkonu, ki pravi takole:

Da pa ljudje, ki (pravico) uresničujejo, to uresničujejo neprostoovoljno, ker nimajo moči, da bi delali krivico, najlažje zaznamo, če si v mislih predstavimo (gr. *τῇ διανοίᾳ*) naslednje. Obema, pravičnemu in krivičnemu človeku, dajmo možnost, da delata, kar hočeta; nato jima sledimo in glejmo, kam bo želja odvedla enega in drugega. (Platon, 2009: 359b6-c3)

Če bi torej zdaj obstajala dva takšna prstana in bi si enega nataknil pravičen človek, drugega pa krivičen, se mi zdi, da nobeden ne bi bil tako adamanški, da bi vztrajal v pravičnosti /.../ (isto, 360b3-6)

Rezultat – bralec in drugi ljudje bi delali krivice, če bi bili nevidni – je opisan kot »dokazilo« ali »znak«:

Dejansko lahko rečemo, da je to pomemben dokaz (gr. *τεκμήριον*) za to, da ni nihče pravičen prostovoljno, ampak je takšen pod prisilo, saj (pravice) nima za dobro sam pri sebi – sleherni namreč tam, kjer vidi da, bi lahko ravnal krivično, tako tudi ravna. (Platon, 1999: 360c5-8)

Razmislite o naši dveh vprašanjih. Prvič, inferencialistični ugovor. Ali lahko pridemo do iste sodbe samo z opazovanjem dejanskega vedenja in izhajajoč iz njega? Morda, vendar ne zlahka, ali pa sploh ne. Zdi se, da zamišljen scenarij vpo-ključuje našo zmožnost vživljanja: simuliramo Gigesovo napredovanje in kvazi-doživljamo (vsaj nekatere) skušnjave, ki ga obhajajo: zlato (ne evri), lepa kraljica,

⁴ Rawls ni mislil, da potrebujemo bolj splošno in abstraktno zamenjavo za tančico; v knjigi *Political Liberalism* kritizira tančico, da je morda preveč abstraktna (2005: 1, § 4). O tem tu ne bo govora. Zahvaljujem se Lukasu Meyerju, ki me je opozoril na to, da je mogoče razpravo, ki jo poznam iz filozofije znanosti, uporabiti tudi v politični filozofiji.

možnost maščevanja osovraženemu kolegu itd. Tako pridemo do sodbe. Kakšno opazovanje in argument bi lahko privedla do istega »tekmèriona«?

Zdi se, da se vprašanje spremenljivk (ali podobnosti) razcepi. Na bolj mehanični stopnji nadzora je scenarij sprejemljiv s trohico domišljije: napravi le telo nevidno, ostalo pa pusti tako, kot je. Glede uporabnosti pa so bile reakcije na prstan kar takojšnje, torej že pri naslednjih generacijah. Irwin omenja, da so epikurejci trdili, da zahteva domišljjski scenarij preveč od človeka (Irwin, 1995: 186 isl.). Cicero ga je v knjigi *O dolžnostih* branil: iz navadnega človeka spravi resnico tako kot mučenje (!) (Cicero, 2011: III.). Carnead odgovarja, da če bi se znašli v situaciji, ki bi bila enaka Gigesovi, bi bilo noro odločiti se za pravico. Irwin se pridružuje strani, ki podpira ME: Glavkonova protidejstvena domneva ni pretirana, le »z abstrahiranjem in pretiravanjem preprosto pojasni pomislek, ki je očitno relevanten za našo odločitev v dejanskih okoliščinah« (Irwin, 1995: 187). Razprava še vedno poteka po več kot dva tisoč letih.

3. Država: opazovanje polisa v logosu

3.1 Sestavljanje mikro-ME

Omenili smo, da je Gigesov prstan nekoliko netipičen ME znotraj konteksta moralne in politične filozofije. Tipična vprašanja tu zadevajo normativne lastnosti dejanj ali ureditev, predlaganih v scenariju. V politični filozofiji je v ME naslovljeno predvsem vprašanje pravičnosti. In prav s tem se bodo ukvarjali sogovorniki v Državi.

Platonov znan predlog se prične z vizualno metaforo: lažje je prepoznati pravico v polisu kot v osebnosti posamezne osebe, saj je v polisu zapisana »z večjimi črkami«. Kje pa najdemo pravičen polis? Lahko ga, kot predlaga Sokrat, opazujemo v našem logosu:

Če bomo torej v misli [logosu, dodal N. M.] videli, kako nastaja polis,⁵ bomo vsekakor videli nastajati tudi njegovo pravičnost in krivičnost? /.../ In če se to zgodi, obstaja upanje, da bomo lažje uzrli, kar iščemo? (Platon, 2009: 369a5-10)

⁵ V grškem originalu: »ἀρ οὖν, εἰ γιγνομένην πόλιν θεασαίμεθα λόγῳ«.

»Logos«, kot je splošno znano, pomeni veliko stvari in v tem stavku je bil preveden kot »govor«, »diskurz«, celo »teorija«.⁶ Poleg tega Sokrat drugje pojasni to početje s sklicevanjem na pripoved: ἐν μύθῳ μυθολογοῦντες (376d9).⁷

Kar Sokrat potem dejansko stori, je, da prosi svoje sogovornike, naj si predstavljajo različne ureditve in jih presodijo glede na njihovo pravičnost.⁸ Na primer, Adejmant bo omenil temo družine. Sokrat opiše ureditev, ki vključuje »skupne« otroke, nato pa vpraša, ali je pravična. Adejmant pravilno razume Sokrata. Poskuša si zamisliti tako ureditev in je pri tem uspešen. Njegovo početje bi opisal kot poskusno izdelovanje oziroma modeliranje. Vključuje lahko njegove izkušnje z družinskim življenjem, tako lastnim kot življenjem znancev, razmišljanje o določenih okoliščinah, o katerih se je razpravljalo pri obravnavi predhodnih Sokratovih predlogov, in morda nekatere druge razmisleke.

Na koncu tega zamišljanja si tudi Adejmant zastavi vprašanje o pravičnosti: v celoti gledano, bi bila taka skupnostna ureditev pravična? Zdi se mu, da bi bila.⁹ Morda občuti nekaj čustvene topline ob tem in njegov takojšnji spontani odgovor »v misli« je pozitiven: »Ja, skupnost otrok je v tem primeru pravična.« (To je standardni primer »moralne intuicije«.) Zatorej pove, kar misli. Ureditev je bila potemtakem izdelana »v logosu«, tako v pogovoru kot v Adejmantovih mislih, ter je skrbno opazovana. Notranje opazovanje je spodbudilo intuicijo-sodbo, da je odgovor da – ta ureditev je pravična. Seveda bo filozofsko izurjen sogovornik ali bralec takoj poskusil spreminjati in posploševati poanto: ali jo je mogoče uporabiti v kateremkoli pravičnem polisu? In na koncu bi lahko končal/a s splošnim prepričanjem, da bi morali biti otroci v vsakem idealno pravičnem polisu skupni vsem državljanom.

Ker je to le eden od mnogih ME, ki tvorijo politično jedro *Države*, bi ga lahko opisali kot »mikro-ME«, za razliko od velikega sestavljenega zamišljanja in dokazovanja, ki ga knjiga predstavi svojim bralcem. Poglejmo si torej dve naši vprašanji. Najprej inferencialistični ugovor – bi lahko prišli do istega sklepa le z uporabo neproblematičnih premis, normativnih načel in empiričnega opazovanja ter

⁶ Naj omenim nekaj primerov iz angleških prevodov *Države*. Grube (1992) prevaja »v teoriji«, Alan Bloom (1968) »v govoru«, Sydenham in Taylor (1906) pa »v našem diskurzu«. Jowett (1991) zapiše: »In če si predstavljamo državo v nastajanju, bomo prav tako videli, kako nastaja pravica in krivica države.« Rouse (1999) izpusti besedo »logos« in pusti zgolj »videti mesto v nastajanju«. V starem prevod Daviesa (1866) je »samo predstavljajte si« in kasneje v besedilu (na str. 369) »v ideji«. Dodatek prevajalke: slovenski prevod Gorazda Kocjančiča – »v misli« – se dobro ujema z avtorjevo trditvijo, da gre tu za miselni eksperiment.

⁷ Predlagana povezava s pripovedjo je vredna nadaljnje raziskave, ki pa jo bomo morali pustiti za drugič.

⁸ Catherine Rowett (Osborne) lahko pove veliko zanimivega o vlogi posamečnega zastopanja v metodi *Države*; govori o ikonski metodi, s katero »motrimo jasno opisano posamezno stvar, za katero vemo, da je pravična, in poskusimo prepoznati, kaj je tisto, kar jo napravi pravično, in sicer s pomočjo ali brez primerjave z drugimi podobnimi primeri pravice« (Rowett 2012, 13).

⁹ Nekateri filozofi, npr. Sosa (2007: 60), označijo takšno dozdevanje kot intuicijo.

sklepanja? Utilitarka bi morda odgovorila pritrdilno; tukaj je njen, skrajno poenostavljen, odgovor. Upoštevati moramo količino uživanja (in morda trpljenja), ki ga ureditev prinaša vsem vpletenim (ali kakšno drugo ustrezno posledicovrednoto). Če je ta količina večja kot pri alternativnih ureditvah, potem je ureditev pravična.

Konsekvencialistka splošnih vrednot bi morala nekako oceniti vrednote, pomembne za situacijo, in bi se najbrž z veseljem sklicevala na ME in intuicije. Deontologinja bi bila prav tako zadržana glede inferencializma. Verjetno bi sprejela potrebo po takojšnji moralni sodbi in po neke vrste intuiciji, s čimer bi zavrnila inferencialistični ugovor. Torej je inferencialist daleč od tega, da bi bil zmagovalec.

Kaj pa nadzor spremenljivk? Je svet, v katerem so otroci skupni, preveč oddaljen od našega? Aristotel je, kot je splošno znano, odgovoril pritrdilno (Aristoteles, 2010: II.1). Ne moreš ljubiti tisoč otrok tako, kot ljubiš svoja dejanska otroka. Poleg tega ima večina normalnih staršev svoje otroke rada tako zelo, da ne bi nikoli pristali na to, da bi si jih delili, kot je bilo predlagano; še manj pa bi pristali na to, da bi jim prikrivali, kateri otroci so njihovi. Doda pa še veliko nadaljnjih pomislekov: tveganje incesta, razžalitve dejanskega starša v prepiru in tako naprej. Sodobni bralec se bo nedvomno strinjal z njim. Naj na kratko oblikujem te misli tako, kot je vseč avtorjem, ki govorijo o protidejstvenih situacijah. Če uporabimo znak »>« za protidejstveni če–potem, bi prvotna trditev v grobem imela naslednjo obliko:

otroke vzgaja celotna skupnost > otroci in njihovi starši so obravnavani pravično.

Nato si ogledamo najbližji svetu, v katerem je ta ureditev vpeljana. Tam so stvari večinoma take kot v domačem svetu, razen tega, da otroke vzgaja celotna skupnost. Toda zmožnost posamezne osebe za ljubezen ostaja konstantna. Kako potem oseba X uspe ljubiti na stotine otrok, kakor bi ljubila svoje? Tudi spolna privlačnost deluje kot običajno v tem bližnjem možnem svetu. Oseba Y se zaljubi v starejšo kolegico in spi z njo. Smola, da je ta mati osebe Y. Bi živeli v takem svetu? Bi bile te stvari pravične? Zagotovo ne, sugerira Aristotel.

Klasični platonski (in kasneje utopični) odgovor na tovrstne kritike je sklicevanje na vzgojo in prevzgojo. Taka ureditev vzgaja državljane, preoblikuje njihov značaj in jih pripravi do tega, da jo sprejmejo. Nevarnosti in obeti te vrste odgovora so dobro znane (pomislimo na Popperjevo kritiko Platona in platonske tradicije); tukaj sem le želel povezati razvpiti problem z kontekstom političnega miselnega eksperimentiranja.

Zdaj prehajamo k širši konstrukciji. Prava izgradnja Kalipolisa se začne z Glavkonovo zavrnitvijo minimalne ekonomije, »države svinj«, kot temu sam pravi, in z njegovim vztrajanjem, da dober polis potrebuje razkošje. Razkošje pomeni

uvoz, uvoz zahteva stvari za izvoz, to povečuje potrebo po dobrem kmetovanju, slednje pa po velikem območju, ki ga je nato treba zaščititi. Skratka, potrebujemo vojsko. Ta bo s seboj prinesla vrsto ureditev, nekaj res zelo znanih. Vsaka je preizkušena na že opisani način: zamislimo si jo, sledi vprašanje o tem, če je pravična, in pa odgovor. S tem je zaključen majhni ME, naš »mikro-ME«. Nato so sogovorniki in bralci povabljeni, da si zamislijo naslednjo ureditev. Bi bila ta pravična? Da. Nato ponovimo postopek. Vsak tak mikro-ME je korak znotraj makro-ME. Skratka, Sokrat vpraša: je ureditev a_i pravična? Če je odgovor za vsak a_i »da« in če so združljive, potem je rezultat popolnoma pravičen polis. *Država* je prva izmed takih makro-ME, medtem ko je Rawlsova *A Theory of Justice* (1971) najnovejša.

Seveda se bo bralcu zdelo zaporedje razprav o posameznih ureditvah (mikro-ME zaporedje) povsem naravno. Zato je koristno opozoriti, da ni vsak političen ME tako podroben. Rousseaujeva *Družbena pogodba* (2001) predlaga rešitev glavnega političnega problema v enem stavku, nato pa ne razpravlja o njem na način, kakršnega bi morda pričakovali, če vzamemo *Državo* ali Rawlsa za zgled, posledica tega pa je zmeda in resno nesoglasje glede celo najbolj osnovnih Rousseaujevih stališčih. Tudi dokaj pedantni Kant je manj zgovoren o svoji družbeni pogodbi, kot bi si bralci morda želeli. Posebna vrednost *Države* je njena dialoška zgradba, saj različni sogovorniki zahtevajo različne razlage o temah, ki so posebnega pomena za vsakega od njih. Glavkona zanima potrošništvo, Adejmanta podrobnosti vzgoje, medtem ko bi Polemarh rad čim več zvedel o družini in ženskah v Kalipolisu. Na ta način zvemo veliko podrobnosti. Po drugi strani pa je Platonu veliko do tega, da združi ureditve v večjo celoto. Na primer, razprava o vzgoji čuvarjev se prične z opombo, da se lahko pričakuje, da nam bo preučevanje te teme dalo uvid, ki nam bo prišel prav pri preučevanju, kako pravičnost in nepravičnost nastajata v državah (Platon, 2009: 376d). Prisotnost te motivacije je dober razlog, da obravnavamo celoten političen prikaz *Države* kot en velik, poenoten ME, na način, kot je to storjeno tukaj.

Naj omenim, da ima ME-*Država* še eno zanimivo lastnost. Izvaja se v dialogu aristokratskih intelektualcev, ki sami med sabo ugotavljajo, kaj je najbolje za vsak sloj prebivalcev: na primer, vsaka mlada in zdrava ženska v Kalipolisu mora iti skozi reproduktivni program, ki so ga zasnovali filozofi, ne sme pa ga kritizirati ali se odločiti, da v njem ne bo sodelovala. Če okleva, se ji pove plemenito laž in jo prisili, da ga sprejme. V tem smislu o usodi navadnih ljudi odločajo drugi, tj. filozofi, in zanje je ME-*Država* tretjeosebni postopek. Glavni sodobnejši tekmecev tovrstnemu poskusu je družbena pogodba, kjer se vsak udeleženec vpraša ključno vprašanje: Bi ti sam podpisal ta dokument? Proudhon izrazi to jasno in nazorno: »O družbeni pogodbi bi se moralo svobodno razpravljati in vsak udeleženec posebej bi jo moral sprejeti in lastnoročno podpisati« (Proudhon, 1969: 115). V tem smislu gre torej za prvoosebni politični ME; v ozadju pa je verjetno predpostavka osnovne enakosti in dostojanstva vsakega udeleženca.

3.2 Interludij: utopični vidik *Države*

Preden preidemo h kritikam in težavam, naj skrenem k bolj zgodovinski temi. Popolnoma pravičen polis, Kalipolis, je dober kraj za življenje, dejansko najboljši. Thomas More je pomagal pri imenu: rezultat uspešnega makro-ME (ali celo mega-ME) je utopija. To nas pripelje do zanimive karakterizacije *Države*. Dejansko bi lahko trdili, da se je iz nje rodil celoten žanr politično-filozofskih pripovednih izdelkov – utopij. Še več, Schofield je pred kratkim dokazoval, da je *Država* delo političnega utopizma: »*Država* /.../ je prvo veliko delo političnega utopizma vseh časov /.../« (2006: 205), pri čemer utopično razmišljanje razume kot

zamišljanje načrta za želen svet, ki pa je vseeno določen z današnjimi skrbmi, pri čemer so vprašanja o praktičnosti in legitimnosti (...) le sekundarnega pomena. Utopizem bo vedno z nami. /.../ Človeška misel se vedno ukvarja z možnim kot tudi z dejanskim..., zato pogosto nimamo druge izbire, kot da se lotimo raziskovanja alternativ. (prav tam)

Skratka, utopija je zamišljanje alternativ. Videli pa smo, da je zamišljanje alternativ, kot je izpeljano v *Državi*, politični ME. Ali med njima obstaja kaka razlika?¹⁰

Tu je moj predlog. Pri prikazu alternativne družbene ureditve je mogoče zasledovati različne cilje; za naš namen pa je glavna razlika med epistemskim in motivacijskim ciljem. V primeru filozofskega ME je izrecni cilj epistemski: ugotoviti, kaj je pravica. Postopek je bil pravkar opisan: poišči ali zgradi sistem ureditev, pri čemer naj se sogovorniku vsaka od njih zdi pravična, potem naj ta sam sklene, da je sistem pravičen, in ugotovi, kaj je značilnost sistema – in to je potem »tisto pravično«.

Enkrat ko obstaja načrt, pa bi nadaljnji cilj lahko bil motivacijski: vpeljati sistem, kolikor je le to mogoče. (Pojdi na Sicilijo in poskušaj prepričati lokalnega vladarja, da ga vpelje. Ali o sistemu pouči odpadniškega generala, Diona, in ko bo ta postal vodja, ga morda vpelje. Ali pa na ustrezen način uri elito svojega polisa; njen član bi lahko nekega dne postal vladar itd.) Toda dokler se ukvarjamo s filozofijo, je primarni cilj epistemski. Alternativno bi lahko oblikovali utopijo na enak način kot se oblikuje makro-ureditev za ME, vendar s primarno praktičnim ciljem. To je potem resnično domena utopične literature. Seveda je isto besedilo včasih tolmačeno bolj epistemsko včasih pa bolj motivacijsko.¹¹ Razliko lahko dodatno opišemo s pomočjo načina določitve; v teoretičnih ME gre ta od sveta k misli, v političnih ME pa v veliki meri od misli k svetu (tudi če je cilj epistemski: denimo, da bi našli objektivno Dobro ali Pravično itd.).

¹⁰ Hvala Tei Logar za to vprašanje.

¹¹ Npr. Lockov prikaz družbene pogodbe je v tradicionalnem branju epistemski, če pa sledimo Ashcraftu (1986) in beremo Drugo razpravo kot politični pamflet, potem se bo pokazal kot primarno motivacijski.

Na žalost imajo napake, zagrešene v teoretičnih ME, drugačne vrste posledic od napak v motivacijskih ME, in sicer prinaša vpeljava utopije s seboj veliko resnejše probleme.¹²

3.3 Skepticizem glede miselnega eksperimentiranja: Aristotel in rojstvo tradicije

Nazaj k *Državi*. Začnimo z našima dvema vprašanjema, saj nas bosta pripeljala do alternative za ME v politični filozofiji. V pričujočem kontekstu nima ravno smisla, da bi razpravljali o inferencialističnem ugovoru. Mega-ME je le prevelik, da bi lahko ponudili predlog za inferencialistično zamenjavo; če je ta mogoča, mora priti iz rekonstrukcije vsakega mikro-ME kot argumenta in sicer brez sklicevanja na domišljijo. Naša druga težava – nadzor spremenljivk – pokaže zobe že na samem začetku gradnje Kalipolisa.

Omenili smo že ta posebej dramatičen primer, in sicer prevzem potrebe po določenem razkošju iz realnega sveta, in pa potrebe po močni, stalni vojski čuvarjev (ki je poglavitna za celotno zgodbo). Slednjo potrebo bi morali postulirati (θετέον, 373a5) kot dano, da pridemo iz »države svinj« v bolj razkošno državo. Mnogi bralci bi se s tem morda strinjali, ampak vsekakor ne vsi. Zgodnji stoiki, denimo Zenon, ne bi bili navdušeni nad potrošništvom kot motivacijo za popoln polis. Zagovornica stališča lahko ponudi kar nekaj odgovorov. Kot prvo, lahko vztraja pri svojem stališču in trdi, da normalno dostojno življenje zahteva razširjeno proizvodnjo in da vsaka ureditev, ki ga omogoča, mora vzeti v obzir tudi varnost. Drugič, lahko bi rekla, da je potrošništvo le ena od motivacij za to, da imamo čuvarje: barbarstvo sosedov bi bila enako dobra motivacija za zaščito celo dokaj primitivne ureditve (predstavljajte si, da imate za sosedo stare Rimljane, ki prežijo na ženske vašega sabinskega polisa). Rousseau je preučil možne motive za novo družbeno ureditev,¹³ razširjeno porabo je zavrnil in predlagal potrebo, da bi zaščitili sebe in skupnost pred naravnimi nesrečami in grožnjami.

¹² Tukaj je zelo preprosta pogojna delitev političnih ME in utopij, upoštevajoč dejstvo, da so včasih le implicitni (npr. družbena pogodba, predstavljena kot zgodovinsko dejstvo) včasih pa eksplisitni. Stolpca ustrežata dvema zgoraj omenjenima ciljema – epistemskemu in motivacijskemu.

	EPISTEMSKI	MOTIVACIJSKI
EKSPLICITNI	Platon, Rawls	Platon in ????
IMPLICITNI	Hobbes, Locke	večina utopistov

¹³ Glej Fetcher (1975, 114).

Omenili smo tudi Aristotelovo kritiko skupnosti otrok. Dejansko je to prva in paradigmatična kritika domnevno nerealnih političnih ME. V *Politiki* II predlaga sledeče načelo kritike: »Nedvomno je prav, da privzamemo idealne pogoje (ὑποτίθεσθαι κατ' εὐχὴν), ni pa prav, če gremo čez vse meje možnega (ἀδύνατον)« (1265a17-18).

Kakšna je torej možnost, na katero se sklicujemo? Besedilo predlaga bližnje možnosti: komunizem glede premoženja ni metafizično nemogoč, tudi skupnost otrok ni. Vprašanje pa je, če bo posledična ureditev stabilna in če jo bo mogoče vpeljati brez uporabe nasilnih in nepravilnih sredstev.

Značaj Aristotelove kritike je predmet razprave, ki je lepo povzel Miller (2009). Po pregledu zgodovine razprave je opozoril na »vlogo idealov v Aristotelovem državništvu, ki se jih da pojasniti s pomočjo dveh načel. Prvo je načelo približevanja: 'Čeprav je očitno najbolje za vsako bitje, da doseže dejanski cilj, če pa to ni mogoče, potem bliže, ko je temu najboljšemu, boljše bo njegovo stanje' (*Cael.* II.12 292b17-19; gl. tudi *GC.* II.10 336b25-34). /.../ Drugo je načelo vzročne konvergence (gl. *Top.* III.5 119a17-18). Če V povzroča U, potem velja: bolj ko vzrok spominja na V, bolj je njegov učinek podoben U. Bolj ko je vroča voda, bližje je vrelišču. Ti dve načeli združeno pomenita, da lahko ideal služi kot standard, četudi ni dosegljiv« (Miller, 2009: 542). Prav tako ugotavlja, da nas oddaljene možnosti lahko naučijo, kaj je moralno zahtevano (kakor nas idealizacije lahko učijo o fizični realnosti). Na žalost se te zanimive teme tu ne moremo lotiti in prav tako ne vprašanja končne ocene o *Državi*.

Naj vsaj omenimo, da je Aristotelova kritika začela dolgo tradicijo skepticizma o političnih ME. Kasneje so platonisti, kot denimo Proklus, zagovarjali Platona in razprava se je nadaljevala. Slavni napad na idealizacijo je izrazil Machiavelli v *Vladarju*: kaj bi ljudje morali storiti, ne pa, kaj res počnejo (poglavje XV.). Sodobni čas nam je prinesel še eno pomembno vrsto političnih ME, in sicer hipotetično družbeno pogodbo (na način, ki ga Glavkon predvidi v *Državi*, I). In odziv kritikov: Locke in Rousseau so se postavili na eno stran, Hume pa na nasprotno; tudi Hegel kasneje nadaljuje s kritiško tradicijo. Možno je trditi, da je nasprotje med zagovorniki in kritiki političnih ME osrednje nasprotje v zgodovini metodologije politične filozofije od Platona in Aristotela do Rawlsa in Morgenthaua. In to nas pripelje prav do našega zaključka.

4. Zaključek

Platonova *Država* je gotovo temeljno (če ne najbolj temeljno) besedilo zahodne politične filozofije; zato je zelo zanimivo, da je organizirano okoli obsežnega in dobro izoblikovanega miselnega eksperimenta. Aristotelova *Politika* je še eno tako temeljno delo, zato je zanimivo, da se predstavlja kot alternativa Platonovi igri z možnostmi, ki ne morejo biti udejanjene.

Kar smo pravkar ugotovili, opozarja na dejstvo, da je veliko kasnejših del s tega področja mogoče razumeti kot – kot bi se izrazil Whitehead – niz opomb k temu prvemu in najbolj izvirnemu krogu razprave.

Da bi bolje razumeli temeljni status *Države*, smo omenili dve možni vlogi za politične ME: njihovo prvotno in intrinzično epistemsko vlogo, v kateri politični ME služi kot sredstvo za preverjanje moralnih vprašanj (kot pri Platonu, Locku in Rawlsu), in motivacijsko vlogo, v kateri postane utopija, ki ji je potrebno slediti, jo posnemati in morda s političnem delovanjem udejaniti. *Država* začneja obe tradiciji: bolj filozofsko epistemsko in bolj aktivistično ter literarno utopično.

Prav tako smo omenili staro vprašanje verodostojnosti, ki smo ga postavili v kontekst razprav o miselnih eksperimentih in povezali z dvema sodobnima različicama. Prva različica je izražena kot vprašanje o nadzoru spremenljivk, druga pa kot vprašanje o bližini protidejstvenega sveta (svetov), ki ga (jih) upodablja miselni eksperiment. Glede vprašanja nadomestljivosti smo bili bolj dogmatični: je politični ME mogoče nadomestiti z argumentom, ki vsebuje samo norme in empirične podatke? Zagovarjali smo negativen odgovor, ki spoštuje miselno eksperimentiranje: ME je izvirno in nenadomestljivo »sredstvo predstavljanja«, ki ga ni mogoče zreducirati na zgolj ilustracijo določenega abstraktnega znanja ali teorije. Kot vedno pri Platonu bi lahko povedali veliko več, vendar upamo, da tudi povedano lahko izzove koristno nadaljnjo razpravo.

Literatura

- Aristoteles (1932). *Politics* (prev. H. Rackham). Cambridge, Mass.: Harvard University Press.
- Ashcraft, R. (1986). *Revolutionary Politics and Locke's Two Treatises of Government*. Princeton: Princeton University Press.
- Baggini, J. (2007). »On Thought Experiments«. Dostopno na: <http://philosophybytes.com/2007/12/julian-baggini.html> [18. 8. 2012].

- Cicero, M. T. (2011). *O dolžnostih*. Ljubljana: Študentska založba.
- Cohen, M. (2005). *Wittgenstein's Beetle and other Classical Thought Experiments*. Oxford: Blackwell.
- Fetcher, I. (1975). *Rousseaus politische Philosophie*. Frankfurt am Main: Suhrkamp.
- Ierodiakonou, K. in Roux, S. (ur.) (2011). *Thought Experiments in Methodological and Historical Contexts*. Leiden: Brill.
- Irwin, T. (1995). *Plato's Ethics*. Oxford: Oxford University Press.
- Miller, F. D. (2009). »Aristotle on the Ideal Constitution«. V Anagnostopoulos, G. (ur.), *A Companion to Aristotle*, Oxford: Blackwell, str. 540–554.
- Miscevic, N. (v tisku). »Political Thought Experiments from Plato to Rawls«. V Frappier, M., Meynell, L. in Brown, J. R. (ur.), *Thought Experiments in Philosophy, Science, and the Arts*, London: Routledge.
- Norton, J. D. (2004). »Why Thought Experiments Do Not Transcend Empiricism«. V Hitchcock, Ch. (ur.), *Contemporary Debates in Philosophy of Science*, Oxford: Blackwell, 44–65.
- Platon (2009). *Država*. V Platon, *Zbrana dela IV*, Ljubljana: KUD Logos.
- Proudhon, P. J. (1969). *General Idea of the Revolution in the Nineteenth Century*. New York: Haskell House Publishers Ltd.
- Rawls, J. (1971). *A Theory of Justice*. Cambridge, Mass.: Harvard University Press.
- Rawls, J. (2005). *Political Liberalism* (razširjena izdaja). New York: Columbia University Press.
- Rousseau, J.-J. (2001). *Družbena pogodba*. Ljubljana: Založba Krtina.
- Rowett, C. (2012). »Do We Know What Justice Is? Method in Plato's *Republic*«. Dostopno na:
http://eastanglia.academia.edu/CatherineOsborne/Papers/1139260/Do_we_Know_what_justice_is_Method_in_Platos_Republic [18. 8. 2012].
- Santas, G. (2006). »Methods of Reasoning about Justice«. V Santas, G. (ur.), *Blackwell Guide to Republic*, Oxford: Blackwell.
- Schofield, M. (2006). *Plato: Political Philosophy*. Oxford: Oxford University Press.
- Smith, N. (1999). »Images, Education and Paradox in Plato's *Republic*«. *Apeiron*, 32, str. 125–142.

Smith, N. (2007). »Plato's Book of Images«. V Scott, G. A. (ur.), *Philosophy in Dialogue: Plato's Many Devices*, Chicago: Northwestern University Press, str. 3–14.

Sosa, E. (2007). *A Virtue Epistemology*. Oxford: Oxford University Press.

Williamson, T. (2007). *Philosophy of Philosophy*. Oxford: Blackwell.

Andraž Ivančič

Filozofska fakulteta Univerze v Ljubljani

Popperjev falzifikacionizem: kritični pregled

Zdi se, da v sodobnem svetu položaj na čelu naših prizadevanj po doseganju spoznanja zaseda znanost. Ta velja za zanesljivo in verodostojno, dvomi v njene temelje pa so (vsaj zunaj filozofije znanosti) redki. Med filozofskimi utemeljitvami zanesljivosti znanstvene prakse izpostavljam Popperjev falzifikacionizem, ki se je v 20. stoletju uveljavil med vodilnimi normativnimi teorijami na področju filozofije znanosti. Po kratkem orisu temeljev falzifikacionizma predstavim nekaj glavnih pomislekov, ki so jih Popperju namenjali že njegovi sodobniki – kritično gre oceniti predvsem njegovo pojmovanje zanesljivosti znanstvene prakse ter ustreznost razmejitve med znanostjo in neznanostjo. Takšno preučevanje falzifikacionizma postavi teorijo pod vprašaj, s tem pa odpira tudi možnost za nadaljnje razmisleke – o sodobni znanosti, o alternativnih spoznavnih metodah in o dosegljivosti resnice nasploh.

Ključne besede: filozofija znanosti, falzifikacija, Popper, razmejitve, znanost

Upravičevanje verodostojnosti sodobne znanosti je izziv, s katerim se soočajo tako znanstveniki kot filozofi. Vprašanje, kaj daje znanosti njeno verodostojnost, je v osrčju posebne filozofske veje, ki jo imenujemo filozofija znanosti, svoj razcvet pa je ta doživela skladno z bliskovitim razvojem znanosti v 20. stoletju. V tem času so svoj pečat v znanstveni teoriji in praksi pustili tudi nekateri filozofi, med katerimi gre izpostaviti Karla Popperja, ki je utemeljil načelo falzifikacije, brez katerega se zdi, da znanost le težko ponudi kakšno gotovo spoznanje. Popper je v *Logiki znanstvenega odkritja* (in drugih spisih¹) zavračal tedaj prevladujočo pozitivistično strujo, ki je zagovarjala načelo verifikacionizma – postavljanje hipotez, izvajanje eksperimentov in potrjevanje teorij po njegovem mnenju ne vodi do spoznanja. Po lastnem prepričanju je uspel rešiti pomanjkljivosti verifikacionizma z oblikovanjem načela falzifikacije, ki spričo asimetrije med možnostjo potrditve in možnostjo ovržbe teorije omogoča uporabo deduktivnega sklepanja v znanstveni praksi – na ta način doseže zahtevano zanesljivost, z enostavnim pravilom

¹ Predvsem Popper, 2002.

pa loči znanost od neznanosti oziroma psevdoznanosti. Popperjev model kot odličen normativni opis znanosti še danes pri mnogih znanstvenikih uživa velik ugled, seveda pa nanj vseskozi letijo tudi kritike.

Popper je bil s svojo domisljico nedvomno zadovoljen, o čemer kar nekoliko humorno priča tudi samozavestni zapis iz leta 1963: »Kritike svojih pogledov do zdaj še nisem srečal«² (Popper, 2002: 54). V verodostojnost takšne samohvale lahko upravičeno podvomimo. Popper je zamolčal skoraj tri desetletja rednih kritik na račun *Logike znanstvenega odkritja*, ki je izšla leta 1934, ter tudi sicer pestrega dogajanja na področju filozofije zgodovine v tistem času. Nenazadnje je leta 1962 izšla tudi knjiga *Struktura znanstvenih revolucij* Thomasa Kuhna, ki ne skriva svojega neodobravanja Popperjevih tez. Po drugi strani pa je samozavest v Popperjevem nastopu razumljiva – načelo falzifikacije nudi elegantno utemeljitev zanesljivosti znanstvenega raziskovanja in v očeh nekaterih še danes bolj ali manj uspešno konkurira alternativam na področju filozofije znanosti. A kritik vseeno ne gre prezreti.

Popper je predhodnikom očital dvoje: nezanesljivost induktivnega sklepanja in nezmožnost razmejitve med znanostjo in neznanostjo. Zanimivo je, da je kljub očitnim izboljšavam, ki jih uvede Popper z načelom falzifikacije, njegova teorija dovzetna za enake očitke. Tako veljavnost sklepov kot nedvoumnost razmejitve lahko postavimo pod vprašaj tudi v kontekstu falzifikacionizma.

Dedukcija kot zagotovilo gotovosti

Nadomestitev indukcije v temelju znanstvenega spoznanja z dedukcijo je vsekakor dobra poteza, če želimo zagotoviti zanesljivost znanstvenih sklepov. Morda se zdi tovrstna posodobitev celo samoumevna, posebej po branju Popperjeve kritike indukcije, ki jo lahko povzamemo s citatom: »Po mojem mnenju indukcija ne obstaja. Torej je sklepanje teorij iz posameznih stavkov, ki jih 'preverimo z izkus tvom' (karkoli že naj to pomeni), logično nedopustno. Teorije zato niso nikoli empirično preverljive« (Popper, 1998: 38). Sama ugotovitev, da indukcija znano sti ne nudi nič več kot vprašljivo oporo, seveda ni nova – kritike indukcije imajo dolgo brado, ki sega od antičnih pironikov preko mnogih uglednih filozofov (posebne omembe si zasluži David Hume) pa vse do Popperja. Popperjeva inovacija je bila implementacija *dedukcije* v znanstveno prakso, in sicer na mesto, ki ga je poprej zasedala indukcija. Načelo falzifikacije – to je bilo Kolumbovo jajce, s katerim se je proslavil avstrijski filozof. Zdi pa se, da se je nekoliko prena glil, ko je deduktivno sklepanje v kontekstu falzifikacionizma sprejel kot zanesljivo.

² Bolj posrečena formulacija v izvirniku: »I have yet to meet a criticism of my views« (Popper, 2002: 54).

Samemu deduktivnemu logičnemu sklepanju seveda ne moremo pripisati nezanesljivosti. Toda glavni problem je implementacija te rešitve, ki se ni izkazala za kaj več kot obliž. Zaverovan v učinkovitost svoje rešitve je Popper morda pozabil, da znanost ne sestoji iz niza enoznačnih premis in deduktivnih sklepov, temveč gre za bistveno bolj kompleksno mrežo, v kateri bi težko našli en sam primer izoliranega deduktivnega sklepa, ki ne bi bil odvisen ali ne bi vplival na druge, povezane informacije, vsak proces v tej mreži pa je nadalje odvisen od številnih drugih dejavnikov.

Sam proces falzifikacije je v bistvu zelo preprost:

- (1) Velja teorija T.
- (2) Rezultat R bi teorijo T ovrgel.
- (3) Ugotovimo, da R.

Sklep: $\neg T$

Zgornji poenostavljen primer pokaže, kako preprosta je falzifikacija, pa tudi, kako zanesljiva je v tem primeru dedukcija. Z logičnega vidika je ovržba teorije T na podlagi rezultata R nujnost, sklep pa je povsem gotov. Toda če želimo ta model prenesti v znanstveno prakso, se hitro pojavijo težave. Dedukcija resda zagotavlja resnične sklepe, a le, če so resnične tudi vse premise. Izkaže se, da preverjanje teh v znanosti ni tako preprosto.

Rezultati opazovanj in eksperimentov

Težave lahko povzroči že premisa, ki je videti najpreprostejša in najbolj nedvoumna. »Ugotovimo, da R« je na papirju morda videti zelo zanesljiva trditev, v znanstveni praksi pa rezultata opazovanja ali eksperimenta ne moremo vedno izolirati in natančno opredeliti. Okoliščin, ki bi lahko privedle do zmote pri pridobitvi rezultata opazovanja ali eksperimenta, je ogromno.

Recimo, da želimo odčitati temperaturo, pri kateri je zavrela voda – termometer pokaže 90° C. Opazimo vrenje vode ter indikacijo na termometru, ugotovitev pa se glasi, da je voda zavrela pri 90° C. Smo res ugotovili, da je voda zavrela pri tej temperaturi? Zdi se, da je tako, a ob natančnejšem premisleku odkrijemo vrsto neznank, ki zamajajo gotovost tega sklepa oziroma že zanesljivost pridobljenih podatkov.

Prva neznanka je zanesljivost termometra – morda je instrument pokvarjen, morda za pravilen prikaz temperature potrebuje več časa ali pa gre preprosto za termometer slabše kakovosti, ki pri merjenju ni zelo natančen. Z lahkoto bi navedli še vrsto drugih težav s termometrom, ki postavljajo meritev pod vprašaj.

Druga neznanka pa je zanesljivost opazovalca. Lahko, da smo se pri merjenju zmotili in odčitali napačno temperaturo, na primer zaradi nepozornosti. Morda smo odčitali pravo temperaturo, a se zmotili pri zapisovanju tega podatka. Ali pa smo temperaturo odčitali prezgodaj oziroma smo storili še kakšno drugo napako pri uporabi termometra. Možnosti je ogromno, ugotovitev pa še manj zanesljiva.

Neznank je seveda še veliko več, prav tako je delitev v kakršne koli skupine arbitrarna. Če bi na primer namesto vode nevede merili temperaturo kakšne druge tekočine, te napake ne bi mogli pripisati ne instrumentu ne opazovalcu, temveč bi bile za napačne rezultate meritve krive drugačne lastnosti same tekočine.

Jasno pa je, kaj razkriva zgornji primer. Še tako enoznačna ugotovitev zahteva vrsto predpostavk, da jo lahko sprejmemo kot gotovo. Kolikor se znanstveno raziskovanje nanaša na stvarnost, je dejavnikov, ki lahko prispevajo k zmoti, več: sam predmet raziskovanja, uporabljeni instrumenti in zmotljivost človeka, ki eksperiment oziroma opažanje izvaja. Zgornji primer opisuje zelo enostaven eksperiment, medtem ko se zavoljo kompleksnosti možnost zmote pri zahtevnejših eksperimentih še poveča. Seveda lahko možnost zmote pri pridobivanju rezultatov eksperimentov zmanjšamo ali skoraj povsem izničimo. A takšen približek za Popperja nikakor ne bi bil sprejemljiv, saj bi moral sicer priznati veljavnost tudi indukciji.

Morda ni odveč opomniti, da smo doslej govorili le o pridobivanju rezultatov iz eksperimenta oziroma opazovanja. Tudi s človeškim faktorjem v tem kontekstu merimo le na napake in zmote pri pridobivanju rezultatov, saj prav in samo o tem govori tretja premisa poenostavljene logične sheme falzifikacije, kateri smo najprej namenili pozornost.

Pokazali smo, da je že na videz enostavno opravilo v znanosti lahko pogubna past za vsakogar, ki želi svoje početje okronati z zagotovilom gotovosti. Preprosta ugotovitev, kateri bi v idealnem svetu izoliranih logičnih sklepov lahko pripisali le resničnostno vrednost, je v znanstveni praksi pod vprašajem zaradi neštetihih faktorjev, ki jih ni mogoče povsem eliminirati.

Interpretacije in povezave

Vidimo torej, da že površna kritika tretje premise v poenostavljeni logični shemi procesa falzifikacije zamaje zanesljivost celotnega sistema. A tudi če zgornjo kritiko povsem odmislimo, bo falzifikacionizem težko obstal. Če pozornost namreč preusmerimo na drugo premiso, postanejo pomanjkljivosti le še izrazitejše.

Druga premisa pravi, da moramo v primeru ugotovitve R teorijo T označiti za napačno. V znanstveni praksi pride v poštev sporočilo, ki je nekoliko širše od te preproste implikacije: pod kakšnimi pogoji R se teorija T izkaže za napačno, tj. katere ugotovitve bi T ovrgle? To vprašanje predstavlja srčiko Popperjeve teorije,

saj je ključni nastavek za pridobitev gotovega spoznanja, tj. končne ovržbe. Vsaka teorija, ki se hoče predstavljati kot znanstvena, mora podati tudi pogoje za lastno ovržbo. V praksi pa moramo vsako teorijo prav na podlagi teh pogojev izpostaviti neusmiljenemu zasliševanju. Jasno je torej, da bi bila lahko vsakršna ranljivost v tem elementu falzifikacije pogubna za zanesljivost celotne teorije. In prav te ranljivosti bomo sedaj skušali prikazati.

Pravkar razložena in razširjena premisa vsebuje dva člena: interpretacijo podatkov, ki naj bi teorijo ovrgli, in dejansko povezavo med temi podatki in teorijo. Vprašamo se torej dvoje:

1. Ali iz pridobljenih podatkov res sledi določena ugotovitev?
2. Ali ugotovitev zares pomeni ovržbo teorije?

Že iz formulacije vprašanj je morda razvidno, da pritrdilen odgovor ni (ali vsaj ne bi smel biti) samoumeven. Stvar pa postane še očitnejša, če jo ponazorimo s primerom.

Prej opisani eksperiment merjenja temperature ni bil zgolj kratkočasna dejavnost nadobudnega znanstvenika, temveč smo z njim skušali falzificirati teorijo, po kateri voda zavre pri 100° C. Sklep se glasi takole:

- (1) Velja teorija »Voda zavre pri 100° C«.
- (2) Teorijo bi ovrgli, če bi voda zavrela pri temperaturi, drugačni od 100° C.
- (3) Ob vrenju vode smo zabeležili 90° C.

Sklep: Teorijo »Voda zavre pri 100° C« ovržemo.

Predpostavimo, da smo pri meritvah temperature naš star, ceneni termometer nadomestili z nezmotljivim instrumentom, tako da so pridobljeni podatki popolnoma zanesljivi. Ko smo zaznali vrenje vode, smo odčitali temperaturo, ki je znašala 90° C.

Najprej se vprašamo, ali iz pridobljenih podatkov res sledi pričakovana ugotovitev. V tem enostavnem primeru bi temu težko oporekali, saj odčitek ne zahteva posebne interpretacije. Če bi bili strožji,³ pa bi se morda vendarle našel kak pomslek. Morda smo sočasnost brbotanja vode in odčitka temperature narobe razbrali in voda ob 90° C pravzaprav še ni zavrela ali pa je zavrela že prej. Ali pa morda

³ Če bi bili strožji, bi bile rigoroznejše tako premise kot pridobljeni podatki. Rezultat R bi razdelili na dve komponenti: pridobljene podatke in ugotovitev oziroma interpretacijo. V tretji premisi bi nastopali le pridobljeni podatki, v drugi pa le interpretacija, medtem ko bi morali seveda vključiti tudi implikacijo med obema komponentama. Pridobljeni podatki pa bi bili v »strogem jeziku« navedeni zgolj kot neposredno opažanje – »ob brbotanju segrevane vode je indikator na termometru dosegel oznako, ki določa vrednost 90° C.« Za predstavitev splošne kritike takšna rigoroznost ni potrebna.

ne vemo, kdaj točno nastopi vrenje. Na tak način lahko prikažemo zmotno interpretacijo, ki ohranja tretjo premiso resnično,⁴ pod vprašaj pa postavi upravičenost ovržbe teorije po drugi premisi.

Veliko bolj očitno bo odgovor na drugo vprašanje, ki smo si ga zastavili. Ali ugotovitev zares pomeni nujno ovržbo teorije? Deduktivna logika vsekakor zagotavlja gotovost tega sklepa, a znanstvena praksa ponovno postavi veljavnost premise pod vprašaj. Ponovno si lahko zamislimo scenarij, v katerem bi pokazali, da ovržba teorije ob še tako prepričljivi falzifikaciji ni samoumevna: Podatka, da voda zavre pri 100° C, ne moremo označiti za lažnega niti na podlagi eksperimenta, ki poda drugačno temperaturo vrelišča – te rezultate smo namreč pridobili v drugačnih pogojih, ki ne konkurirajo tistim iz izvirne teorije. Smo z meritvijo vrelišča vode pri 90° C dokazali, da voda ne zavre pri 100° C? Ne, dokazali smo le, da lahko voda zavre tudi pri drugih temperaturah. Dokončna ovržba bi bila v tem primeru prenapljena. Boljša rešitev, se zdi, bi bila nadgradnja obstoječe teorije z opisom pogojev, pri katerih teorija velja oziroma ne velja. Popper bi bil s takšnim zaključkom vse prej kot zadovoljen, saj je močno nasprotoval tako imenovanim *ad hoc* teorijam⁵, ki so po njegovem mnenju služile le ohranitvi propadle teorije pri življenju. A očitno je, da bi z ovržbo v tem primeru storili napako – dopolnitev teorije se zdi ne le smiselna, temveč tudi bolj zanesljiva in konstruktivna za naše spoznavne napore.

Tak zaključek pa seveda še dodatno zamaje že vprašljivo verodostojnost falzifikacionizma. Očitno brezkompromisna ovržba falzificiranih teorij v znanstveni praksi ni vedno najboljša rešitev – niti s pragmatičnega niti z epistemološkega vidika. Ko teorijo razširimo na kompleksno mrežo dejstev in povezav med njimi, načelo falzifikacije močno izgubi na svoji eleganci. Izkaže se, da ovržba teorij ni tako enoznačna, kot bi morda lahko sklepali iz te poenostavljene logične sheme. In ker filozofija znanosti vedno meri na dejansko znanstveno prakso, je ne moremo soditi zgolj na podlagi idealiziranega zapisa – upoštevati moramo *prakso*, ki dejansko izziva vsako filozofsko teorijo s tega področja.

⁴ Tretja premisa je namenoma zapisana kot kompromis med pretirano rigorozno in pretirano poenostavljeno različico. Ne pravi, da je voda zavrela pri 90° C, tako da interpretacijo prepušča drugi premisi, kar omogoča prikaz pomanjkljivosti Popperjeve teorije na tem mestu. V dejanski znanstveni praksi bi bila tretja premisa bolj rigorozna, kar bi dodalo pričujoči kritiki večjo vrednost, a na račun jasnosti in jednatosti.

⁵ Z *ad hoc* teorijami je možno pri življenju obdržati praktično vsako ovrženo teorijo – gre za ovinek, ki onemogoča falzifikacijo. Popper se je tega problema zavedal in ga je naslovil že v Popper (1998); v Bamford (1993) pa je problem razložen še z navezavo na konkretno znanstveno prakso.

Duhem-Quinova teza

Pomisleki, kakršni so zgornji, seveda niso ušli pozornosti filozofov, ki so se lotevali kritike falzifikacionizma v letih po formulaciji Popperjevih tez. Vsebina do sedaj zapisanih očitkov Popperjevi teoriji izhaja vsaj posredno iz dela dveh filozofov, in sicer Pierra Duhema in Willarda Van Ormana Quina, ki sta vsak v svojem času in vsak na svoj način podala kritiko, ki izpostavlja bistveno *poddoločenost* znanstvenih hipotez.⁶ Gre za to, da znanstvene hipoteze nikoli ne moremo povsem izolirati – vedno privzamemo še celo vrsto predpostavk, ki onemogočajo enoznačno falzifikacijo. Ker lahko operiramo le z bolj ali manj obširnimi zbirkami povezanih hipotez, se nam gotovost vedno izmika. V praksi je nemogoče oblikovati hipotezo tako, da bi bila sama zadostni pogoj za ovržbo teorije, saj ne moremo imeti izoliranega sistema, v katerem bi lahko imeli le eno teorijo in eno hipotezo. To posledično pomeni, da tudi eksperimenti ne morejo neposredno odločati o veljavnosti teorij, saj lahko relevantnost podatkov kvečjemu predpostavimo, nikakor pa ta ni samoumevna: »Za skrajno nespametno se namreč izkaže, da podatki brez vseh zadržkov odločajo o teorijah« (Mulej, 2008: 172).

Duhem-Quinova teza bi si nedvomno zaslužila podrobnejšo obravnavo, a naj v pričujoči kritiki služi le kot povzetek in utemeljitev dozdašnje rdeče niti.⁷ Dodati pa gre še to, da teza odpira širok spekter drugih pomislov, ki služijo kot strelivo za nasprotnike falzifikacionizma in kot zanimivost za tiste, ki se skušajo pri raziskovanju filozofije znanosti distancirati od vrednostnih sodb. Predvsem Quine se v svojem razmišljanju poda v epistemološke globine – poddoločenost hipotez v kontekstu znanosti namreč razširi na celotno védnost, ki je zdaj predstavljena kot kompleksna vseobsegajoča mreža medsebojno odvisnih (spremenljivih in zamenljivih) izjav (od »empiričnih« na periferiji do »logično-matematičnih« v središču).

Padec zanesljivosti

Do te točke je kritika merila zgolj na zanesljivost Popperjeve implementacije dedukcije v znanstveno prakso. Neoporečnost veljavnosti deduktivnega logičnega sklepanja lahko brezskrbno sprejmemo, ne moremo pa mimo težav, na katere naleti takšna logika v kontekstu dejanskega znanstvenega raziskovanja. Izkaže se, da zanesljivost same logične dedukcije še ni zagotovilo za zanesljivost falzifikacionizma, četudi ta temelji na deduktivnem načelu – pokaže se namreč več šibkih

⁶ Duhem in Quine sta podala precej različne pomisleke. Duhem izpostavlja zgolj poddoločenost hipotez pri fizikalnih eksperimentih (Duhem, 1954), medtem ko Quine odkriva poddoločenost v celoti našega vedenja in prepričanj (o čemer npr. piše v Quine, 2001).

⁷ Za nadrobnejšo predstavitev Duhem-Quinove teze glej zbornik Harding (1975), ki poleg že omenjenih temeljnih besedil Duhema in Quina združuje še razmišljanja drugih mislecev.

točk, ki prinašajo upravičene dvome v gotovost spoznanj, pridobljenih skozi takšen proces.

Prepričan sem, da zagovorniki falzifikacionizma ne bi zlahka priznali poraza, a upam, da pričujoča kritika vzbudi vsaj premislek in tehten odgovor. Poleg tega pa je bil namen premisleka izzvati zanesljivost sodobne znanosti, ki se pogosto predstavlja kot samoumevna. Na tem mestu bi veljalo navesti besede razvpitega Paula Feyerabenda, gorečega kritika Popperja in znanstvene metode nasploh: »Odlična kakovost znanstvenih rezultatov in metod se *predpostavlja*, ne pa *pokaže*«⁸ (Feyerabend, 2007: 93).

Vprašanje razmejitve

Ker smo, upam, pokazali dvomljivost zanesljivosti implementacije deduktivnega logičnega sklepanja v znanstveno prakso, lahko pozornost posvetimo še problemu razmejitve v kontekstu falzifikacionizma. Popper svoje mnenje o neustreznosti razmejitve pri verifikacionistih utemeljuje z nezanesljivostjo indukcije: če je znanost le tisto, kar je možno preveriti in potrditi, takšna razmejitev ni ustrezna, saj zaradi nezanesljivosti indukcije ostanemo le pri razmejitvi bolj ali manj (ne)zanesljivih praks brez nekega jasnega in trdnega kriterija. Popper svoje novo pravilo za razmejitev razvije iz načela falzifikacije: znanost se od psevdoznanosti razlikuje po tem, da podaja falzifikabilne (ovrgljive) napovedi, medtem ko se psevdoznanost možnosti falzifikacije izmika.

Popperjev kriterij razmejitve vzdrži kritiko, ki je pokopala razmejitev pri verifikacionistih, za kar poskrbi asimetrija med možnostjo potrditve in možnostjo ovržbe – pri (induktivni) potrditvi ne moremo nikoli doseči popolne gotovosti, pri (deduktivni) ovržbi pa že en primer zadošča za zanesljivo ovržbo. Tako na Popperja ne moremo nasloviti istih očitkov, ki so se izkazali kot pogubni za verifikacioniste, kar pa ne pomeni, da je problem razmejitve z njegovim kriterijem rešen.

Ena izmed prednosti Popperjeve razmejitve je enostavna formulacija edinega pravila: znanstvene teorije so falzifikabilne, psevdoznanstvene pa ne. Prav ta enostavnost pa odpira vrata kritikom, saj mora biti v svoji jedrnatosti in odločilnosti vsakomur že takoj sumljiva. Lahko tako elegantno pravilo postavi med znanost in psevdoznanost nesporen mejnik? Verjetno ne. Kaj kmalu namreč ugotovimo, da je kriterij razmejitve po eni strani prestrog in kot psevdoznanost izloči tudi tiste teorije, ki bi vsekakor morale veljati za znanstvene, po drugi strani pa je preveč ohlapen in dopušča, da se psevdoznanost v mnogih primerih izdaja za legitimno znanost.

⁸ Spet se mi zdi ustrežnejša angleška formulacija: »The excellence of science is *assumed*, it is not *argued for*« (Feyerabend, 1978: 73).

Prestroga razmejitev

Primer prestroge razmejitve nam ponudi kar Popper sam, ki v avtobiografiji ugotavlja, da »darwinizem ni preverljiva znanstvena teorija, temveč metafizični raziskovalni program – možen okvir za preverljive znanstvene teorije«⁹ (Popper, 2005: 195). Morda ni odveč izpostaviti, da Popper tu ne meri na celotno teorijo evolucije temveč na *darwinizem*, specifično na načelo naravne selekcije, ki ga lahko v neposrečenem slovenskem prevodu povzamemo kot *zakon močnejšega*.¹⁰ Tega zakona, pravi Popper, preprosto ne moremo falzificirati. Zakon se potrjuje sam – in to v vsakem primeru. V naravi bodo vedno preživela tista živa bitja, ki so bolj prilagojena na življenje v takšnem okolju (in so torej »močnejša«). Vsak primerek, ki preživi, je torej *že po definiciji* »močnejši«. Če bi preživel »šibkejši«, bi morali pač tega oklicati za »močnejšega« oziroma za bolj prilagojenega, kar onemogoča smiselno falzifikacijo. Lahko sicer postavimo kriterij, ki bi zakon močnejšega ovrgel: »preživel je primerek, ki je bil v primerjavi z drugimi slabše prilagojen za življenje v določenem okolju«, kar pa se ne bo nikoli zgodilo, saj se za vsak preživeli primerek mora izkazati, da je bolj prilagojen, sicer ne bi preživel.

Moramo darvinizem zaradi Popperjevih pomislov odpisati kot psevdoznanost? Vsekakor se zdi, da tudi zakon močnejšega sodi na področje znanosti. Tudi s to teorijo se namreč operira tako kot z drugimi znanstvenimi teorijami, poleg tega se zdi, da njena epistemološka vrednost ni nič manjša. Po drugi strani pa težko oporekamo Popperju zgolj na podlagi intuicije ali zdravega razuma, ne da bi postavili lasten kriterij demarkacije (ali kakšno *ad hoc* rešitev).

Zagovorniki Popperjevega kriterija razmejitve bi darvinizem verjetno zvesto zavrnili kot psevdoznanost, na enak način pa bi obravnavali tudi druge teorije, ki po Popperjevi razmejitvi z znanstvenega področja padejo na psevdoznanstveno.¹¹ In četudi zgornji primer ne ponuja nesporne ovržbe omenjenega kriterija razmejitve, služi vsaj temu, da pod vprašaj postavi njegovo navidezno enostavnost in enoznačnost. Sledeč kriteriju falzifikabilnosti moramo namreč mnoge mejne primere iztrgati s področja znanosti, kamor jih uvrščata zdravi razum in morda stoletna tradicija, ter jih postaviti na neslavni odpad, ki mu pravimo psevdoznanost.

⁹ Za zanimiv odgovor na Popperjevo pozicijo glej Halstead (1980).

¹⁰ V izvirniku »survival of the fittest«, kar seveda ne implicira razlike v moči, temveč v prilagojenosti okolju. Za komentar Popperjevih pogledov ta distinkcija niti ni ključnega pomena.

¹¹ Popper je sicer kasneje priznal, da se je s kritiko darvinizma prenačelil – na podlagi dodatnih informacij in natančnejšega razmisleka o teoriji in njenih temeljih je našel pogoje za falzifikacijo in ugotovil, da je bila teorija že izpostavljena ustreznim poskusom ovržbe. Svojo zmoto priznava v Popper (1978).

Preohlapna razmejitev

Zanimivo pa je, da lahko praznino, ki je nastala po pravkar omenjenem »čiščenju« znanosti, zapolnijo teorije, za katere se zdi, da na področju znanosti nimajo kaj početi. Razmejitev, ki temelji zgolj na falzifikabilnosti, podeljuje verodostojnost teorijam na podlagi njihove formulacije, ne pa vsebine. Psevdoznanost izkorišča prav to – v svojih napovedih posnema znanost, s tem pa želi pridobiti tudi njeno verodostojnost. Po Popperjevem sistemu jim jo morda celo moramo priznati.

Če za kriterij razmejitve sprejmemo falzifikabilnost, ne moremo enostavno zavrniti niti astrologije, ki velja za eno izmed bolj »očitnih« psevdoznanosti. Edina naloga znanosti po tem kriteriju je, da podaja falzifikabilne napovedi, kar astrologija z lahkoto doseže. Moramo vsako, še tako neverjetno hipotezo sprejeti kot znanstveno, dokler je ne ovržemo? Recimo, da astrolog v položaju planetov prepozna znamenje bližajoče se katastrofe in napove svetovni potop čez tri dni. Moramo to hipotezo obravnavati kot znanstveno? Popper bi verjetno odgovoril pritrdilno, saj gre za falzifikabilno napoved.

Podobno velja za homeopatijo, ki podaja številne drzne napovedi brez ustrezne znanstvene podlage. Hipoteza, da bo voda s svojim »spominom« po vsakodnevem uživanju v petih letih ozdravila raka, je na videz znanstvena, saj ustreza kriteriju falzifikabilnosti. Naj bo pozitiven izid še tako (ne)verjeten, to ne vpliva na »znanstvenost« teorije, saj moramo biti pozorni le na to, da ta podaja pogoje za lastno ovržbo.¹²

Za oceno epistemološke vrednosti praks, ki bi po ohlapnem kriteriju razmejitve prevzele naziv »znanost«, bi morali žrtvovati več časa in pozornosti, a na prvi pogled se zdi, da se v tem oziru nova »znanost« ne more kosati s tisto, ki razmejitvenemu kriteriju navkljub ostaja trdno na svojem mestu. To pomeni, da bi zgovornik Popperjeve razmejitve zaman vztrajal pri tem kriteriju – če zaradi falzifikabilnosti napovedi med znanost sprejmemo astrologijo, homeopatijo in podobno, nismo s tem dosegli ničesar več od trivialne razmejitve, ki je zdaj nesmiselna.

Zaključek

S kratko kritiko ugotavljamo, da se je Popper nemara prenagril v svojem zmago-slavju. Načelo falzifikacije, ki naj bi enkrat za vselej ustoličilo znanost na čelu prizadevanj po doseganju spoznanja, se očitkom zoperstavi le malo bolje kot

¹² Nadalje pa tudi dejanska ovržba določenih hipotez ne more ovreči celotne astrologije oziroma homeopatije – po Popperju lahko z ovržbo hipotez ovržemo le posamezne teorije, tako da bi na podlagi tega status psevdoznanosti težko pripisali celotni panogi.

predhodno načelo verifikacije. Pravzaprav pade prav v točkah, ki jih je Popper kritiziral pri svojih predhodnikih in domnevno izboljšal v svoji teoriji.

Uvedba deduktivnega logičnega sklepanja, ki je glavno orožje Popperjeve teorije, je na prvi pogled odlična domislica, ki pa ne zdrži kritike v kontekstu znanstvene prakse. Načelo falzifikacije se pri konkretnih znanstvenih teorijah ne izkaže za zanesljivo, saj ne zmore zagotoviti preciznosti, s katero bi lahko res falzificirali posamezne hipoteze oziroma teorije. Še tako neposredna ovržba se ne reši pomislov, ki sta jih izpostavila že Duhem in Quine. Popperjeva teorija bi vzdržala kvečjemu v nedosegljivih idealnih okoliščinah, ne pa v realnosti nezanesljivih merilnih naprav, zmotljivih znanstvenikov in poddoločenih znanstvenih teorij.

Podobno velja za kriterij razmejitve, ki prav tako zadovolji samo tiste, ki falzifikacionizem le ošvrknejo brez poskusa kritičnega premisleka. Preprosta formulacija razmejitvenega kriterija je vsekakor privlačna, a ne zagotavlja zanesljivega ločevanja med znanostjo in psevdoznanostjo. Falzifikabilnost teorij ne more biti zadostni pogoj za razmejitev, saj bi v tem primeru na meji med znanostjo in psevdoznanostjo zavladovala le zmeda, na eni in na drugi strani pa bi našli teorije, ki tam ne bi smele biti. Znanost ne bi bila več znanost, če bi se bila primorana ukvarjati z vsako načeloma ovrgljivo teorijo, medtem ko bi morali po Popperju kot psevdoznanstvene zavrniti nekatere teorije, ki morda skrivajo neprecenljiv prispevek k bogastvu našega spoznanja.

Resna kritika falzifikacionizma bi vsekakor morala vključevati neposrednejše so-
očenje s temeljnimi postavkami Popperjeve teorije, nemara pa tudi predlog vzpostavitve sistema, ki bi bil ustrezen nadomestek. Moj poskus pa naj služi za ponazoritev nekaterih ranljivosti falzifikacionizma, nenazadnje pa tudi kot opozorilo na samoumevnost zanesljivosti sodobne znanosti, kolikor se ta v nekaterih segmentih še vedno opira na temelje, ki jih je postavil Popper.

Popper's Falsificationism: A Critical Review

It seems that the pole position in the pursuit of knowledge in the modern world belongs to science. Modern science is considered to be reliable and credible while its foundations are rarely doubted (other than in the philosophy of science). One line of philosophical argumentation for certainty of scientific practice is provided by Popper's theory of falsification, which became one of the most prominent normative theories in the philosophy of science in the twentieth century. After a brief overview of the basic postulates of falsificationism, I present some of its shortcomings that were already expressed by Popper's contemporaries. I aim to examine the reliability of falsification-based scientific practice, and review the adequacy of Popper's criterion of demarcation between science and pseudoscience. Findings of such an examination question the theory of falsification and open the path for further considerations on modern science, alternative methods of acquiring knowledge, and the attainability of truth altogether.

Keywords: philosophy of science, falsification, Popper, demarcation, science

Literatura

- Bamford, G. (1993). »Popper's Explications of Ad Hocness: Circularity, Empirical Content, and Scientific Practice«. *The British Journal for the Philosophy of Science*, 44(2), str. 335–355.
- Duhem, P. (1954). *The Aim and Structure of Physical Theory*. Princeton: Princeton University Press.
- Feyerabend, P. (1978). *Science in a Free Society*. London: New Left.
- Feyerabend, P. (2007). *Spoznanje za svobodne ljudi*. Ljubljana: Krtina.
- Halstead, B. (1980). »Popper: good philosophy, bad science?«. *New Scientist*, 1210(87), str. 215–217.
- Harding, S. (ur.). (1975). *Can Theories be Refuted?: Essays on the Duhem-Quine Thesis*. Dordrecht: Springer.
- Kuhn, T. (1998). *Struktura znanstvenih revolucij*. Ljubljana: Krtina.
- Mulej, L. (2008). »Kontraindukcija kot vir novitet«. *Anthropos*, 40(1-2), str. 163–180.
- Popper, K. R. (1972). *Objective Knowledge: An Evolutionary Approach*. Oxford: Oxford University Press.
- Popper, K. R. (1978). »Natural Selection and the Emergence of Mind«. *Dialectica*, 32(3-4), 339–355.
- Popper, K. R. (1998). *Logika znanstvenega odkritja*. Ljubljana: Studia humanitatis.
- Popper, K. R. (2002). *Conjectures and Refutations: The Growth of Scientific Knowledge*. London: Taylor & Francis Ltd.
- Popper, K. R. (2005). *Unended Quest: An Intellectual Autobiography*. London: Taylor & Francis Ltd.
- Quine, W. (2001). »Dve dogmi empirizma«. *Analiza*, 5(4), str. 67–85.

Nika Škof

Filozofska fakulteta Univerze v Ljubljani

Tadej Pirc

Inštitut za filozofijo Univerze v Gradcu

Nedosegljivost ideala znanstvene objektivnosti

Človek kot bitje perspektive na svojo kontekstualno vpetost odgovarja z idejo objektivnosti, ki je odraz njegove želje, da bi se odvezal določenosti s perspektivo. V članku dokazujeva, zakaj je odprava perspektive le utopična želja in zakaj iz tega nujno sledi, da je objektivnost ideal, ki je nedosegljiv. Da bi nama to uspelo, članek pričneva s prikazom družbenozgodovinskega konteksta razvoja in ukoreninjenja znanosti kot izbrane tradicije zahodne kulture. Zatam obravnava problem, ki ga perspektiva predstavlja za ideal znanosti, ter vprašanje, kaj ostane od znanosti, če ji odvzamemo objektivnost in vrednotno nevtralnost. Premislek gradiva skozi prizmo udeležnosti, določenosti in vezanosti človeka na njegove lastne spoznavne in zaznavne sisteme. Nadaljujeva z opisom pojmovanja ideala objektivnosti, pojasniva, čemu je treba pojem objektivnosti nujno ohranjati v rigidnosti njegovega prvotnega pomena, in problematizirava objektivnost *zmerne objektivnosti*. V zaključku predstaviva prizadevanje po neperspektivnosti kot nekaj zaželenega, ob tem pa pokaževa, v čem se razlikuje od objektivnosti.

Ključne besede: znanstvena objektivnost, ideal, perspektivnost, zmerna objektivnost

Uvod

Temeljni predpostavki pričujoče razprave sta zelo preprosti. Prvič, človek je bitje perspektive, ki je zaznamovana z informacijami in oblikovana s preteklimi izkušnjami, delujoči subjekt pa je z njo neločljivo povezan. Kljub temu ne zagovarjama determinizma, temveč se v dilemi med združljivostjo¹ in nezdružljivostjo² determinizma s svobodno voljo nagibava k stališču kompatibilizma, in sicer v njegovi zmerni teoretski zasnovi, ki izbiro med potencialnimi alternativami dojema kot

¹ Klasik tega stališča je Dickinson Miller (1934).

² Za najvplivnejšo obrambo tega stališča glej van Inwagen (1983).

možno posledico svobodne volje, a v vnaprej določenih okvirih. Drugič, ideja objektivnosti je odgovor na bôl človekove perspektivnosti. Natančneje, je želja, da bi se človek odvezal določenosti s perspektivo in tako spoznal neperspektivno Resnico.

Objektivnost je skladno s tema izhodiščema nujno in izključno boleče nedosegljiv ideal, ki človeka s svojo nedosegljivostjo žene k naporom, potrebnim za asimptotično približevanje nikoli dosegljivi *neperspektivni perspektivi*. To danes precej razširjeno, čeravno v sodobnih znanostih pogosto potlačeno tezo je najverjetneje prvi zapisal Victor Hugo v svojem delu o Shakespearu iz leta 1864: »Znanost je asimptota resnice; neprestano se ji približuje in nikoli dotakne« (Hugo, 1899: 114). Kljub temu ji je pripisal izjemen pomen, saj »vendarle poseduje vso veličino. Ima voljo, natančnost, zanos, globoko zbranost, prodornost, pretkanost, moč, potrpežljivost do konca, stalen pregled nad pojavi, strast do napredka in celo trenutke hrabrosti« (prav tam).³

Nobena perspektiva asimptotične resnice torej nima nič opraviti z Resnico samo, saj ni zvedljiva na zgolj eno in edino možno ter nujno resnično perspektivo. Ta od človeka namišljena »Resnica«, ki je po njegovi definiciji ena, je prav zaradi človeku lastne perspektivne zamejenosti tako nemisljiva, da tudi njej, namreč Resnici, dozdevno približevanje ne predstavlja drugega kot v subjektivnost ujet pogled, ki se približuje neki ne le nejasni in nedosegljivi, temveč tudi od človeka neodvisno neobstoječi tvorbi. Naše dozdevno približevanje Resnici je človekov lov na svojo lastno izmišljijo, na meglico ali privid, ki se ob približevanju (temu niču) naposled razblini v to, kar je. Kljub vsemu svojemu prizadevanju namreč nismo nikoli zares objektivni. Asimptotično približevanje neki namišljeni točki Resničnosti ne more biti Resnica sama, saj kot ujetniki perspektive ne moremo ciljati v objektivno Resničnost kot točko, ki bi bila merilo našega približevanja. Približevanje je torej resnično le toliko, kolikor je Resnična točka sama – fiktivno, s krhkim referenčnim okvirjem in samo-ohranjujoče.

Članek pričinja s prikazom družbenozgodovinskega konteksta razvoja in ukočeninjenja znanosti kot izbrane tradicije zahodne kulture, ki je pridobivala na svoji avtoriteti od začetkov racionalizma, s pozitivizmom pa tekom 19. stoletja dosegla svoj vrhunec. Obravnavala bova perspektivnost in problem, ki ga ta predstavlja za omenjeni ideal znanosti, ter vprašanja, kaj ostane od znanosti, če ji odvzamemo objektivnost in vrednotno nevtralnost, zakaj izbiramo znanost kot najadekvatnejši način spoznavanja sveta oziroma bivanja, kaj nam ponuja in kaj vzame. Premislek bova gradila skozi prizmo udeleženosti opazovalca, raziskovalca, znan-

³ Ta ideja se pogosto pojavlja tako v filozofiji (znanosti) kot v razpravah znotraj znanosti. Med drugimi je kmalu po Hugoju *asimptotično stališče* leta 1879 povzel Léo Errera (1908) in ga apliciral na takratno znanost botanike, sredi dvajsetega stoletja ga je med filozofe razširil Thomas Kuhn (1962), ob koncu stoletja pa ga je v srenjo astronomov skušal vnesti Carl Sagan (1995), in sicer s tezo, da »imamo metodo, a nam ta metoda ne pomaga doseči absolutne resnice, temveč le asimptotična približevanja resnici«.

stvenika in filozofa, njegove določenosti in vezanosti na lastne spoznavne in zaznavne sisteme.

Nadaljevala bova z opisom pojmovanja ideala objektivnosti, pojasnila, čemu je treba pojem objektivnosti nujno ohranjati v rigidnosti njegovega prvotnega pomena, navedla parametre, ki jih prepoznavava kot njen nujni pogoj, hkrati pa pokazala, zakaj je objektivnost zgolj ideal. Zatem bova predstavila zamisli o zmerni objektivnosti, ki predstavljajo enega izmed možnih orodij za reševanje dosegljivosti in prisotnosti ideje objektivnosti v filozofiji znanosti, ter problematizirala *objektivnost* takšne »objektivnosti«.

Čeprav odločno nasprotujeva možnosti objektivnega spoznanja, bova naposled vendarle namenila pozornost slutnji, da je prizadevanje po neperspektivčnosti nekaj zaželenega. Pogledala bova, ali in kako se ta razlikuje od objektivnosti, ter izpostavila vprašanje, kako je možno ohranjati motivacijo za prizadevanje po neperspektivčnosti in iskanje nečesa nedosegljivega.

Družbenozgodovinska klima in ideal objektivnosti

Ocenjujeva, da v sodobni zahodni družbi prevladuje prepričanje, ki objektivnost pojmuje kot nekaj dosegljivega. To je moč sklepati na podlagi kontradiktorne in silovite vere v znanost, katere vodilo naj bi bila objektivnost. Vera v dosegljivost objektivnosti se kaže v avtoriteti, ki jo ima znanost na področju vednosti, natančneje v šoli, ki po mnenju večine uči Resnice ter dejstev. Četudi izobraževalni sistem postopoma uvaja kritični ali perspektivni pogled, je ta osredotočen zlasti na problematiziranje družbenih, kulturnih in umetniških stališč, medtem ko je znanost iz tega nabora izvzeta. V znanost se ne dvomi. Predstavlja se jo kot tisto, ki se je okovi epistemološke krajine družbenega ne tičejo. Kakor opaža tudi Feyerabend: »Odlična kakovost znanstvenih rezultatov in metod se *predpostavlja*, ne pa *pokaže*« (Feyerabend, 2007: 93).⁴ Širjenje in ohranjanje takšnega prepričanja je mogoče spremljati vse od vznika ideje pozitivistične empiristične znanosti, njegove nastavke pa lahko razbirmo skozi vso zgodovino človeštva. Zlasti ob paradoksalni predpostavki, da gre v znanosti primarno za normativno iskanje sorazmerno soglasnega odgovora na vprašanje, kako je *smiselno* biti.

Znanost je, tako trdiva, svojo moč pridobila, ker je izkoristila in zapolnila vrzel, ki se je pojavila v osmišljanju človekovega bivanja po razkroju primata religije.⁵

⁴ Ali v precej bolj povednem izvirniku: »The excellence of science is *assumed*, it is not *argued for*« (Feyerabend, v Kidd, 2010: 2).

⁵ Gre za klasično teorijo, ki so jo različni avtorji (od Marxa in Durkheima do Webra in Freuda) kovali na prelomu stoletij. A tudi ta teorija ni brez zanimivih protiuteži. Med zadnjimi je svoj pogled, s katerim skuša misliti zgodovino religije in znanosti hkrati (kot komplementarna koncepta), predstavil zgodovinar Peter Harrison (2015).

Sekularizacija je ljudi in celotne družbe pustila polne vprašanj in brez odgovorov nanje. Izguba garanta smisla in Resnice je povzročila iskanje nadomestnega ali novega odgovora. Takrat je pozitivistična znanost še močneje poprijela za vajeti kulturnega življenja in obljubila, da bo s svojo domnevno objektivno držo ponudila Resnico, ki je ne bo mogoče vreči na zemljo, ker je zemlja sama. Eksistencialni vakuum je zapolnila z gotovostjo, ki ji to pot (domnevno) ni bilo potrebno zgolj verjeti, temveč je bila dokazljiva, preverljiva, vrednostno nevtralna, neperspektivna. Skratka, bila je *objektivna*. Na to je opozarjal tudi Nietzsche (1988: §10), ki je jasno napovedal, da s tem, ko smo ubili boga, postiljamo znanosti kot novi veri.⁶

Kar sta pozitivistična znanost in racionalizem kot njen predhodnik ponudila kot možen odgovor, je močno resoniralo s potrebami družbe tistega časa. Vse večji dvom v možnost zanašanja na boga je človeka usmerjal v zanašanje nase. Rahljanje družbenih hierarhij se je povezovalo z idealoma preverljivosti in ponovljivosti ter občutkom, da ni potrebno verjeti na besedo, saj je mogoče vse preveriti. Hkrati se je uveljavila ideja človeka kot racionalnega bitja, ki je sposobno sodelovati in domneve eksperimentalno potrditi. Seveda je bila družbena realnost sprva drugačna, hierarhije niso razpadle čez noč, saj sta poslušnost in verjetje pomembni komponenti vsakršne tradicije, še zlasti, če sprejmemo predpostavko, da je ta izmišljena (Hobsbawm in Ranger, 2006). Dostopnost znanosti ni bila vsesplošna, toda njen ideal in ideja, ki ju je nosila s seboj – človek je tisti, ki zmore razumeti in nadzorovati svoje okolje⁷ –, sta bila osvobajajoča in potencialno prenosljiva na vsakega posameznega racionalnega pripadnika družbe.⁸

Z znanostjo je človek sam postal raziskovalec Resnice, ki je bila »tam zunaj«. Pričakovanje, da je skozi opazovanje in empirične podatke možno spoznati temeljne zakone ter usmerjevalna načela, ki se bodo z zadostno količino zbranih podatkov kar sami razprostrli pred znanstvenikom – ta je sicer človek, a vendar dovolj izjemen, da kaj takšnega zmore, oziroma dovolj učen, da mu lahko zaupamo –, je dajalo upanje in zaupanje v nov, svoboden način spoznavanja sveta. Na-

⁶ Prim. tudi Badiou (2003: 165 isl.) in Valadier (2010).

⁷ Glej Gadamer (2001: posebno 356 isl.) za zanimivo hermenevitično analizo znanstvene objektivnosti in njenega cilja raziskovati Resnico – nasebstvo –, njene pojmovne napake in popolnoma nasprotnega učinka od zelenega absolutnega spoznavanja. Slednje je posledica osnovnega ontološkega nerazumevanja absolutnega ali nasebnega znanstvene usmeritve, ki absolutno in nasebno zmotno ter pomanjkljivo enači s predmetnostjo in svojo domnevno objektivnost s tem popolnoma relativizira, saj načrtno izpušča dele realnosti. To se pokaže v nagnjenosti znanosti, da dognano uporabi za lasten smoter obvladovanja stvari, ki je v nasprotju z nasebstvom, saj to ne more biti predmet volje.

⁸ Tudi skupina racionalnih pripadnikov družbe oziroma tistih, ki so veljali za ljudi, se je s časom širila, vedno pa je zajemala trenutno najmočnejšo ali najrelevantnejšo družbeno skupino, ki je pomembno oblikovala družbeno realnost (od belih aristokratov do belih premožnih moških, do belih delavcev, do belih moških, do vključevanja drugih spolov in ras). Proces vključevanja oziroma upoštevanja pa je v teku še danes, ko si prizadevamo k družbeni ureditvi, ki bi bila neizključujoča do posameznih marginalnih in manjšinskih družbenih skupin.

čela francoske revolucije, ki so izjavljala hrepenenja množice – svoboda, enakost, bratstvo –, so svojo realizacijo prav tako uspešneje našla v znanstveni objektivnosti in racionalizmu kot v drugih tekmujočih tradicijah (Cranston, 1989). Človeku je objektivnost ponujala svobodo od družbenih in religioznih vezi, enakost na ravni pripadnosti razumni in racionalni vrsti, ki pred tem ni imela skupnega imenovalca, ter bratstvo v uporabi dognanj za nadzor okolja.

Znanstvena tradicija je po Feyerabendu (2007: 57) le ena izmed mnogih možnih tradicij, prepričanj, načinov življenja, ki sama po sebi ni ne dobra ne slaba, ampak preprosto je in zato ne more biti razsodnik med tradicijami, saj je le ena izmed njih. Takšno pojmovanje znanosti in zavzemanje stališča do načina bivanja zastopava skozi celotno besedilo, s čimer pritrjujeva še eni Feyerabendovi misli, da je izbira znanosti kot vredne in uporabne pravzaprav odgovor na eksistencialno vprašanje o tem, kakšno življenje naj živimo, kakšno življenje je vredno, zaželeno, dobro živeti (Feyerabend, v Kidd, 2010: 3). Znanost nam ponudi zadovoljiv odgovor, ki je v skladu s predhodno obstoječimi vrednotami in pojmovanji vrednega življenja, zato jo v nekem določenem družbenozgodovinskem kontekstu izberemo kot najboljši odgovor. Ne zaradi njene inherentne vrednosti, temveč zaradi skladnosti z našimi prepričanji.

Omenjena arbitrarnost tradicije in njena vrednotna nevtralnost, ki sta posledica konsenzualne in naključne gradnje družbenih norm ter vrednot, svojo nekoliko drugačno, zlasti pa sociološko utemeljitev dobi v delu *Družbena konstrukcija realnosti* (Berger in Luckmann, 1988), ki prikaže realnost kot skovano in arbitrarno določeno in ne kot Resnico ali njen odsev. Dogovorno in družbeno sooblikovanje našega sveta v njem določa smisle, pomene, vrednote in norme. Prav skupnost je tista, ki istočasno podeljuje pomene in avtoriteto. Omenjena razprava je tudi razprava o znanju, vednosti, ki izhaja iz fenomenološke pozicije posameznikovega zaznavanja in doživljanja. Teoretika zaključita, da je objektivna vednost tista, ki je institucionalizirana, ta pa se lahko s časom spreminja. Pri razpravi o objektivni vednosti torej ne gre za vprašanje Resnice ali Dobrega, saj »teoretske formulacije realnosti – filozofske, znanstvene ali mitološke – ne izčrpajo pomena tistega, kar je 'realno' za pripadnike neke družbe« (prav tam, 1988: 23).

Predstavljene relativistične in konstruktivistične pozicije služijo kot izhodišče premišljanja o znanosti, njenem primatu in objektivnosti. Iz tega razloga sva prikazala tudi družbenozgodovinski kontekst, v katerem je vzniknila in na moči pridobila znanost, saj lahko skozenj pojasnimo njeno časovno prostorsko določenost in specifične, na katere se je odzivala in prilagajala. V nadaljevanju bova govorila o po najinem mnenju bistveni komponenti in obljubi znanosti – objektivni Resnici, o njenih možnostih in iluzornosti.

Udeleženi opazovalec in perspektivnost njegovega pogleda

Idejno-historični pregled permutacij pristopov k znanosti v zadnjem stoletju pri-vede do nekaterih ključnih, čeravno splošnih sklepov. Relativizem in kritična teorija sta v intelektualne debate vnesli zavedanje o temeljni perspektivni določenosti človeka. Fenomenologija je postavila trdne teoretske temelje, ki jih ni možno ignorirati. Svet je s konstrukcionizmom in konstruktivizmom postajal vse bolj majav, kibernetika prvega in drugega reda strukturalistične teorije v šestdesetih pa je vse slutnje še dodatno potrdila. Človek je vpet v družbo, ki mu daje perspektivo, in v svoj subjekt, ki mu daje možnost kontakta z družbo ter sposobnost ustvarjanja lastne zgodovine. Teorije učenja, osebnosti, motivacije so vse močnejše kazale na družbene in odnosne faktorje, od katerih je odvisna oseba, njen razvoj in rast. Nič več solipsistični individuum, vedno bolj determinirani družbeni člen, ki se sam sebe zaveda – subjektivna pozicija – in ima možnost zmerne premika v levo ali desno, vseskozi pa ohranja referenčni okvir. Skozi te procese je človek postajal subjekt.

Razprava v tem članku temelji na pojmovanju človeka kot zaznavajočega, kot bitja, ki je v stiku s svetom preko svojih zaznav. Te so edine, ki mu omogočajo kontakt. Ne more se ločiti od svojega pogleda, svoje zgodovine, svoje umeščenosti. Vsako *objektivno* znanstveno teorijo sestavi človek, prebere človek, kritizira človek. Tako je vsako raziskovanje raziskovanje raziskujočega, dognanje dognanje dognojučega, mišljenje mišljenje mislečega. Akcija in delovanje imata zmeraj svojega nosilca, ki se od svoje pozicije ne more ločiti, saj ga brez nje ni. Ta nujna in morda celo nasilna vpetost v lastno okolje, feyerabendovsko rečeno – v tradicijo, s Husserlom (2005: §33, §34) rečeno pa – v življenjski-svet, onemogoča resnično objektivno spoznanje. »Zaničevanje, s katerim znanstvenik, ki sledi novoveškemu idealu objektivnosti, obravnava vse 'zgolj subjektivno-relativno', v ničemer ne spremeni njegovega posebnega načina biti, prav tako pa ne tudi tega, da mora to tudi samemu znanstveniku biti dovolj dobro, in sicer ne glede na to, kje se nanj sklicuje in neizbežno mora sklicevati« (Husserl, 2005: 156). Svoj svet zmeraj nosim s seboj, zato nanj ne morem gledati od nikoder. Vedno sem v njem in vedno sem nekje.

Tudi domnevno najbolj *objektivna* med znanostmi je vse bolj perspektivna. V kvantni fiziki na potek eksperimenta vpliva sama navzočnost raziskovalca. Kako bi bilo, če bi ne bil tam, kako bi bilo, če ne bi opazoval, ni jasno (Davies in Betts, 1994: 79). Pogled je nujen za to, da opazimo, vidim pa lahko le skozi lastne oči, ki so pred tem videle točno določeno zaporedje slik, ki je samo moje. Posebna in splošna relativnostna teorija sta znatno pripomogli k razgradnji modela brezperspektivnosti, saj so meritve in zakoni fizike v pomembni odvisnosti od perspek-

tive oziroma umeščenosti raziskovalca (Dolenc, 2000). Relativnostna teorija je dokončno razbila absolut in tako smo ostali sami s seboj ter lastno perspektivo.⁹

Na perspektivno zasidranost človeka, ne glede na to, v kateri vlogi se znajde, so med drugim opozarjale feministične teoretičarke, denimo Donna Haraway, Sandra Harding in Alison Wylie.¹⁰ Uveljavljen izraz za takšno pozicijo je *situated knowledge*, zagovarja pa jo tudi sama. Gre za to, da se v posameznikov kognitivni aparat »vpraskavajo nepozabne poteze individualnosti« (Sloterdijk, 1999: 14), določene z družbenimi vzorci, vednostjo in idejami. Človek je vedno človek, vrednotno in kulturno vpeti jaz, ki se ne more izgubiti do objektivnosti.

Vrednotno nevtralni pogled od nikoder

Sklepali bi torej lahko, da je prav objektivnost tisto, kar je znanosti zagotovilo avtoriteto na področju vednosti. Beremo lahko, denimo, da je znanstvena objektivnost značilnost znanstvenih trditev, metod in rezultatov, ki izraža prepričanje, da znanstveno delovanje ni in ne sme biti odvisno od perspektive, vrednot, družbenih pritiskov ter družbenega statusa raziskovalca, da je garant za zaupanje in razlog za spoštovanje (Reiss in Sprenger, 2014). Longino (1990: 62) navaja, da je objektivnost možno razumeti skladno z znanstvenim realizmom, ki objektivnost razume kot Resnico. Ta se referira na zunanje objekte, ki so tam tudi brez in neodvisno od subjekta. Skladno s tem stališčem so objektivni pogledi znanosti natančni opisi naravnih dejstev. Še določneje bi to pomenilo, da gre pri objektivni znanosti za pravilen in resničen pogled na predmete ter stvari v okolju in njihove medsebojne relacije, epistemološko pa odgovarja resničnosti in informira o dejstvih. Slednja sporoča na nearbitraren in nesubjektiven, torej neperspektivčen način. V preglednih besedilih lahko najdemo tudi enostavno in jasno zapisano, da je objektivnost povezana z Resnico, zanesljivostjo in realnostjo, zaradi česar je vredna zaupanja. Objektivna realnost je skladno s tem *zaupljivim* pogledom tisto, kar obstaja neodvisno in nespremenjeno tudi ob odsotnosti zavestnega opazovalca (Mulder, 2016).

Prav takšna pojmovanja objektivnosti in vera v možnost takšnega načina spoznavanja sveta sta znanost postavila na čelo vednosti. Ideal, ki ga je v svojih povojih predstavljala znanost, ko še ni naletela na mnoge nepremostljive prepreke in ne-načrtovane stranske in slepe ulice, je dajal upanje za možnost ločevanja od svoje lastne perspektive in prisvajanja božje pozicije, do katere je Zahod izgubil dostop z ukinitvijo boga samega (Nietzsche, 2005: §125). Objektivna pozicija je obljubila sigurnost in človek jo je pograbil, saj je ždenje v eksistencialni negotovosti, če

⁹ V Uršič (2002) beri, denimo, kako Einsteinova relativnostna teorija za zmeraj opravi z idejami absolutnega prostora in časa.

¹⁰ Prim. Reiss in Sprenger (2014).

je le bilo možno, bilo vredno premagati za vsako ceno. Vedeti, kaj je res, nadomestiti eno in edino Resnico krščanstva z eno in edino Resnico znanosti, je bilo manj zahtevno kot sprejeti omejenost in nedosegljivost slednje, saj bi to pomenilo, da se ji je potrebno odreči, namesto spremeniti jo. Hkrati bi to pomenilo priznanje omejenosti samega uma. Predvidevava, da je na tisti točki bilo *razumneje* oziroma manj nesmiselno in bolj povezano z dotedanjim razumevanjem sveta Resnico spremeniti, a jo kljub temu ohraniti, torej ohraniti pojmovanje obstoja od človeka neodvisnega in ločenega sveta – zdaj objektivnega, takrat božjega –, kot bi se ji bilo popolnoma odpovedati in s tem prevzeti odgovornost za svojo perspektivo, tako družbeno kot individualno, jo zrelativizirati in detronizirati.

Objektivnost je tako pomenila sigurnost, pogled od nikoder, torej brezperspektivnost, vrednotno nevtralnost, družbeno neodvisno Resnico, ki je v skladu z od človeka neodvisnim zunanjim svetom. Absolutna vednost in teorija vsega sta skrajna točka ideala objektivnosti. To je objektivnost, ki jo tudi sama pojmujeva kot objektivno. Objektivnost lahko nastopa izključno kot absolutna objektivnost, zato predlagava, do jo obravnavamo kot ideal, ki lahko služi za merilo oziroma kot neusahljiva želja in skrajni cilj. Ta je zaradi človekove perspektivnosti nedosegljiv.

V takšni konceptiji gre za strogo *teoretsko* pozicijo, ki vzdrži, dokler ni seznanjena z materialno razsežnostjo človekovega bivanja, realnostjo njegove nebožanskosti, njegove *praxis*. Gre za ideal, h kateremu si lahko prizadevamo, ne moremo pa ga doseči, zato ga iz njegove teoretske umeščenosti, skladno z najinim izvajanjem, neupravičeno skušamo premestiti v prakso. V praksi lahko poskušamo biti čim manj perspektivni, a nikoli nismo objektivni, saj ne bivamo v teoretskem vakuumu. Objektivnost je pravilčnost čistega uma, ki je v svoji realnosti vedno nujno praktičen.

Zmerna objektivnost

Takšna delitev na teoretski koncept objektivnosti in na prakso subjektivnosti oziroma perspektivnosti se v trenutni filozofiji znanosti ne uporablja. Prav obratno. Definicije in razumevanje objektivnosti se prilagajajo ter upravičujejo tako, da jih lahko pripišemo svojim vsakodnevnim praksam in delovanjem. Objektivnost tako slačimo iz njene idealno tipske podobe, prakso in teorijo pa združujemo kot še nikoli doslej, saj je glavno vprašanje možnost aplikacije. Čistemu umu smo odrekli idealizacijo in ga podvrgli praktičnemu. Sanjanje in spekulacija sta postala neproduktivna, zato je tudi teorija le še ena izmed praks. Objektivno tako ni več vprašanje ideala, ampak vprašanje izvedljivosti. Z relativizmom in konstrukcionizmom se je zavedanje o človekovi nujni perspektivnosti in vrednotni vpetosti v družbena razmerja razširilo toliko, da je vplivalo na postopno opuščanje ideala Resnice, kar se odraža tudi v nekaterih tokovih sodobne filozofije znanosti.

Kljub temu da je v medijih in pedagoških ustanovah znanost še vedno pojmovana kot objektivna, se na nekaterih točkah filozofije znanosti, sociologije in drugih družbenokritično naravnanih ved, pa tudi znotraj znanosti same, krepi skeptični duh perspektivizma. Dvojnost situacije je zanimiva in povedna: znanost v družbenih razmerjih ohranja svojo avtoriteto in moč, družbeno realnost skozi ideološke aparate reproducira, istočasno pa ne more biti popolnoma gluha za rastoče kritike in utemeljene ugovore. Zato za vrati, zaprtimi pred širšo javnostjo, dela na novi, sprejemljivejši podobi, ki bo pomagala pri nadaljnjem ohranjanju domene vednosti, ko in če nastopi trenutek širše družbeno zahtevane deobjektivacije in počlovečenja znanosti ter znanstvenika.

V filozofiji znanosti se zato govori o objektivnosti, ki je dosegljiva. Objektivnosti, za katero meniva, da nikakor ni v skladu s svojo izhodiščno idejo, ki je zagotovila primat znanosti. Objektivnost zgodnje znanosti je bila brezperspektivna, vrednotno nevtralna in brezinteresna. Objektivnost trenutnih spisov¹¹ stremi k temu, da bi bila čim manj perspektivna, vrednotno informirana in usmerjana na podlagi interesa (pa naj je to interes trga, družbe, epistemskih vrednot ali česarkoli drugega). Objektivnost je postala praksa, a je s tem izgubila vse svoje določujoče lastnosti. Vprašanje v tem primeru je, kaj ostane od objektivnosti. Zdi se, da gre le za to, da pozicija ni skrajno subjektivna.¹² Pa je to res zadosten približek objektivnosti?

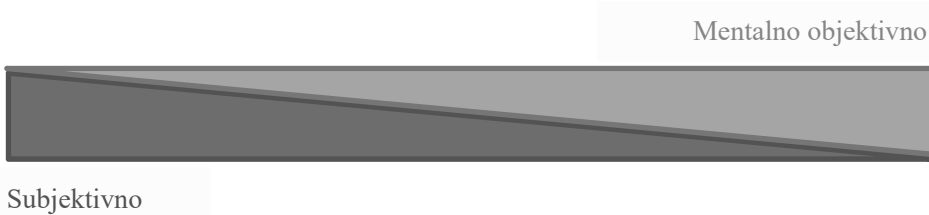
Heather Douglas (2007) je našela vsaj šest vrst *objektivnosti*, ki nam ostanejo tudi, ko se odpovemo njenemu pojmovanju skozi vrednotno nevtralnost. Douglas je ena izmed predstavnikov zmerno pozicije do znanosti in njene metodičnosti, ki zagovarjajo in priznavajo zmerno prisotnost perspektivnosti, kulturne in vrednotne določenosti znanosti. S tem pritrjujejo kritiki relativistov, a jo imajo za manj kritično, saj istočasno pritrjujejo scientistom glede upravičenosti primata znanosti na področju vednosti, pri čemer pa njihove argumente obravnavajo z enako mero zadržkov kot pozicije kritikov znanosti. Meniva, da če so scientisti skrajni teoretiki, ki se ne zavedajo omejitev, ki jim jih postavlja njihova žal gotova vpetost v prakso sveta, na drugi strani pa so relativisti teoretiki prakse, ki jih kompleksnost in številčnost perspektiv omejujeta pri postavljanju teorije ali pa jih ženeta v skrajno metafiziko, so zmerni teoretiki prej popolni praktiki kot nekaj vmesnega. Zmerni teoretiki so torej tisti, ki jim teorija predstavlja prakso.

V prispevkih Douglas (2007) in nekaterih drugih zmernih teoretikov (denimo Longino, 1990), pa tudi Thomasa Nagela (1986), lahko beremo, da objektivnost ni nujno brezperspektivna, kar pomeni, da je neka perspektiva objektivna, kadar ni *izključno* subjektivna. Meril za to, ali je neka perspektiva objektivna, je več.

¹¹ Denimo Longino (1990), Douglas (2007) in drugi *zmerni* teoretiki iz nadaljevanja.

¹² Glej razpravo o Nagelovem pojmovanju mentalne objektivnosti v nadaljevanju tega besedila ali primerjaj Nagel (1986: zlasti 3–19).

Nagel, denimo, razlikuje med mentalno in fizikalno objektivnostjo,¹³ prvo pa postavlja na kontinuum, ki bi, predlagava, izgledal nekako takole:



Subjektivno je za Nagela interno, notranje doživljanje posameznika, mentalno objektivno pa je pogled, ki ga posameznik zavzame, sicer nujno perspektivčen, a korak oddaljen od svojega notranjega doživljanja. Iz te točke mentalne objektivnosti lahko oseba ponovno naredi en korak vstran od svojega doživljanja te točke, kar ponovi znova in znova, tako da se po kontinuumu postopoma premika k točki vse bolj prevladujoče mentalne objektivnosti in vse manj subjektivnega doživljanja. Vendar tudi ko se znajde v zelo skrajnem delu mentalne objektivnosti, je ta še zmeraj vezana na njegovo perspektivo, na njegov korpus znanja, na dejstvo, da je človek, z lastno identiteto, vpet v polje družbenega in podobno. Mentalna objektivnost je stremljenje k čim bolj oddaljeni, od lastne subjektivnosti ločeni poziciji.

Kot primer in za lažjo dojemanje ideje Nagel povabi bralca, da si predstavlja lastno smrt. Če gleda s subjektivne, notranje, investirane pozicije, je to nekaj nezaželenega in najverjetneje strašnega. Če gleda z zunanje, ločene, mentalno objektivne pozicije, vidi, da je umrljivost del življenja in da je njegovo življenje le eno izmed mnogih, njegova lastna smrt pa ne nosi neke posebne in velike vrednosti, saj je v svetovnem merilu nekaj ne le vsakdanjega, temveč »vsakominutnega«, če ne »vsakosekundnega« (Nagel, 1986: 17–19).

Zmerne pozicije objektivnost sprejemajo kot praktično možno in zato prilagajajo njeno definicijo, da jo lahko umestijo v naš svet in nekatere pojave nazivajo z objektivnimi. Jezik in tudi teorija sta kakopak živa in arbitrarno določena, vendar domnevava, da gre v tem primeru za temeljno preoblikovanje, ki iz izhodiščnega termina vzame prav vsa sidrna določila in na njihovo mesto vsadi njim nasprotujoče pojme. Že res, da gre za prilagajanje praksi, a dvomiva, da ji delamo kakšno uslugo s tem, ko močno zaznamovan termin podvržemo plastični operaciji zgolj zato, da bi ga lahko ohranili, namesto da bi se mu odpovedali. S tem ko stare koncepcije nasilno prilagajamo trenutni praksi, namesto tega, da bi rezali v prakso samo, ponavljamo zgodovino.

¹³ Za razlago razlike glej Nagel, 1986: zlasti 13–19.

Zato iluzorično govorimo o objektivnosti, čeprav smo spoznali, da ne gre niti brez perspektive niti brez vrednot niti brez zgodovine. In ker beseda objektivnost nosi staro slavo, jo poskušamo ohranjati in jahati na njej tudi takrat, ko več ne stoji.

Plodnost prizadevanja po nepristranskosti

Iz doslej zapisanega sledi, da je pojmovanje človekove *prakse* kot kakorkoli objektivne nesprejemljivo. Kljub temu pa trdiva, da je prizadevanje po nepristranskosti pomembno in konstitutivno za družbo. Prav razlika med objektivnostjo in prizadevanjem po nepristranskosti se zdi pomembna točka, ki jo je potrebno izpostaviti.

Objektivnost razumeva kot dejavnost, ki išče in zagovarja eno Resnico, o kateri se lahko, če se *civilizirano* in *razumno* pogovorimo, strinjamo in jo dosežemo. Prav tako zapostavlja vpliv družbeno kulturnih dejavnikov oziroma jih želi eliminirati in ne pripoznati, niti ne nameniti pomembne vloge. Objektivnost je tako nek ostank pozitivism, kjer se ni možno strinjati, da se ne strinjamo, ker obstaja samo *da* ali *ne*, ki se glede na logiko zahodne znanosti nujno medsebojno izključujeta.

Na drugi strani je prizadevanje po nepristranskosti relativistično usmerjeno prepoznavanje in poznavanje lastne tradicije. Je trud, da meril svojega ozadja ne vsiljujemo drugim in je pripravljenost sobivanja in upoštevanja tudi v primerih nestrinjanja. To so točke, kjer se vprašanje *svobode za* in *svobode od* (Berlin, 1969) v družbenih razmerjih in ureditvah kar samo zastavlja. Nepristranskost zahteva dvom v lastno pozicijo in v potrebo po imeti prav ter vedeti. Je trening postavljanja meja in razumevanja nerazumevanja ter skrajno izpostavljanje eksistencialni negotovosti in stalno iskanje moči za pritrjevanje življenju v dvomu.¹⁴

Zaključek

Pomemben poudarek, ki sva ga želela izpostaviti, je razločitev prizadevanja po nepristranskosti od zahteve po objektivnosti ali brezperspektivnosti. Hkrati sva želela vzbuditi vprašanje o produktivnosti popolne pomenske alteracije pojma za voljo ohranjanja nekega starega občutka zaupanja in kredibilnosti, ki mu ga pripisujemo. S prikazom družbenozgodovinske vpetosti in sentimentov, na katere je zelo dobro odgovorila znanost v času svojega vzpona, pa sva skušala ponovno nakazati na nenormativnost življenjske pozicije neke kulture, s čimer sva pričela pričujočo relativistično razpravo.

¹⁴ Podobno Feyerabend (2007) in tudi Nietzsche s konceptoma *amor fati* (2005: §276) in *ja sagen* (1989: §10).

Tudi najina perspektiva je le perspektiva, ena izmed vseh možnih, in skladno s to se slutnja perspektivizma in relativizma vse jasneje kaže in seli tudi v pore vsakdanjega življenja. *Anything goes* je vse bolj prisotna forma bivanja, nek Nietzscheju povsem tuj svetobolni nihilizem tudi. Zdi se, da se slutnja po izgubljanju ene in edine Resnice plazi zaenkrat le po kletah zahoda. Kljub temu znanost ohranja vajeti v svojih rokah in si lasti avtoriteto na področju vednosti, medtem ko se poslužuje *ad hoc* reševanja svojih zdaj šibkih točk, ki so prav te iste, ki so ji omogočile vzpon. V ozadju se morda postopoma plete popolna redefinicija objektivnosti, Resnice, nepristranskosti, nevtralnosti, a se za zdaj lahko sprijaznimo le s sklepom, da »čeprav je resnica o resnici v tem, da ni resnice o resnici, človek mimo resnice ne more« (Hribar, 1981: 163).

Morda smo, kuhnovsko rečeno, na prelomu paradigem, morda smo v obdobju porajanja *nenormalne* znanosti. Takšno opažanje naju usmerja k iskanju alternative sodobni znanosti. Bo nastopil čas za fenomenološko perspektivo resničnosti? Kakšno formo bi ta prevzela? Kakšen vpliv na družbeno ureditev, razmerja, mišljenje ljudi bi imela intersubjektivno usmerjena avtoriteta? Je subjekt 21. stoletja pripravljen sprejeti eksistencialno negotovost, ne izbrati Resnice in sprejeti svojo relativno resnico? Ali nas čaka nova, radikalnejša oblika Resnice, temelječa na trdni veri v nekaj ali nekoga, ki nudi varnost in se zdi v zadnjih desetletjih nekoliko podhranjena?

Unattainability of the Ideal of Scientific Objectivity

A human as a being of perspective responds to his/her contextual integration with the idea of objectivity, which is the expression of his/her desire to be absolved from determination by a perspective. In the article, we demonstrate why the elimination of perspective is merely a utopian wish, and why this necessarily results in objectivity as an unattainable ideal. To achieve this, we start by demonstrating the social-historical context of the development and establishment of science as the chosen tradition of Western culture. Then we discuss the issue which the perspective presents for the ideal of science and the question of what is left of science if objectivity and value neutrality are taken from it. We build our position through the prism of participation, determination and dependence of a human being on one's own epistemic and perceptual systems. We continue by describing the conception of the ideal of objectivity, explain why it is necessary to maintain the notion of objectivity rigidly in its original meaning, and challenge the objectivity of *moderate objectivity*. In conclusion, we present the aspiration for non-perspectiveness as something desirable, and also demonstrate how it differs from objectivity.

Keywords: scientific objectivity, ideal, perspectiveness, moderate objectivity

Literatura

- Badiou, A. (2003). *Infinite Thought: Truth and the Return to Philosophy*. New York: Continuum.
- Berger, P. in Luckmann, T. (1988). *Družbena konstrukcija realnosti*. Ljubljana: Cankarjeva založba.
- Berlin, I. (1969). »Two Concepts of Liberty«. V Berlin, I., *Four Essays on Liberty*, London: Oxford University Press.
- Cranston, M. (1989). »The French Revolution: Ideas and Ideologies«. *History Today*, 39(5).
- Davies, P. in Betts, D. (1994). *Quantum Mechanics*. Boca Raton: CRC Press.
- Dolenc, S. (2000). »Prostor-čas in gravitacija: Uvod v splošno teorijo relativnosti«. Kvarkadabra. Dostopno na: <http://www.kvarkadabra.net/2000/12/relativnost-splosna/> [28. 10. 2016].
- Douglas, H. (2007). »Rejecting the Ideal of Value-Free Science«. V Kincaid, H., Dupre, J. in Wylie, A. (ur.), *Value-Free Science?: Ideals and Illusions*, Oxford: Oxford University Press, str. 120–139.
- Errera, L. (1908). »Botanique Générale«. V Errera, L., *Recueil d'oeuvres de Léo Errera*, Bruselj: H. Lamertin.
- Feyerabend, P. (2007). *Spoznanje za svobodne ljudi*. Ljubljana: Založba Krtina.
- Gadamer, H.-G. (2001). *Resnica in metoda*. Ljubljana: Literarno-umetniško društvo Literatura.
- Harrison, P. (2015). *The Territories of Science and Religion*. Chicago: University of Chicago Press.
- Hobsbawm, E. in Ranger, T. (2006). *The invention of tradition*. Cambridge: Cambridge University Press.
- Hribar, T. (1981). *Resnica o resnici*. Maribor: Založba Obzorja.
- Hugo, V. (1899). *William Shakespeare*. Chicago: A. C. McClurg & Co.
- Husserl, E. (2005). *Kriza evropskih znanosti in transcendentalna fenomenologija*. Ljubljana: Slovenska matica.
- Kidd, I. (2010). »The True, the Good, and the Value of Science«. *Philosophical Writings: Proceedings of the Thirteen Durham Bergen Conference*, Durham: Durham University Press, str. 1–9.
- Kuhn, T. (1962). *The Structure of Scientific Revolutions*. Chicago: University of Chicago Press.

- Longino, H. (1990). »Values and Objectivity«. V Longino, H., *Science and Social Knowledge: Values and Objectivity in Scientific Inquiry*, Princeton: Princeton University Press, str. 62–82.
- Miller, D. (1934). »Free Will As Involving Determination and Inconceivable Without It«. *Mind*, 43 (169), str. 1–27.
- Mulder, D. (2016). »Objectivity«. V Fieser, J. in Dowden, B. (ur.), *Internet Encyclopedia of Philosophy*. Dostopno na: <http://www.iep.utm.edu/objectiv/> [29. 10. 2016].
- Nagel, T. (1986). *The View from Nowhere*. New York, Oxford: Oxford University Press.
- Nietzsche, F. (1988). *Onstran dobrega in zlega*. Ljubljana: Slovenska matica.
- Nietzsche, F. (1989). *Ecce homo*. Ljubljana: Slovenska matica.
- Nietzsche, F. (2005). *Vesela znanost*. Ljubljana: Slovenska matica.
- Reiss, J. in Sprenger, J. (2014). »Scientific Objectivity«. V Zalta, E. N. (ur.), *Stanford Encyclopedia of Philosophy*. Dostopno na: <http://plato.stanford.edu/entries/scientific-objectivity/> [29. 10. 2016].
- Sagan, C. (1995). »Wonder and Skepticism«. *Skeptical Inquirer*, 19(1).
- Sloterdijk, P. (1999). *Prihajati k svetu – prihajati k jeziku*. Ljubljana: Društvo Apokalipsa.
- Uršič, M. (2002). *Štirje časi: filozofski pogovori in samogovori*. Pomlad: prvi čas. Ljubljana: Cankarjeva založba.
- Valadier, P. (2010). »Science as New Religion«. V Babich, B. (ur.), *Nietzsche, Epistemology, and Philosophy of Science*, Dordrecht: Kluwer, str. 241–252.
- van Inwagen, P. (1983). *An Essay on Free Will*. Oxford: Oxford University Press.

Peter Lukan

Filozofska fakulteta Univerze v Ljubljani

Temeljni pojmi Boškovičevega atomizma v perspektivi vznikanja moderne znanosti

Namen članka je podati oris atomske teorije Ruderja Boškovića, nakazati vpliv in tehtnost nekaterih njegovih najbolj prodornih teoretičnih znanstvenih dosežkov ter jih umestiti v odnosu do razvoja moderne znanstvene misli in metode. V uvodu na kratko predstavim Boškovićevo življenje in delo, nato orišem njegovo pojmovanje atomov, ki ga vzporejam z manj dodelanimi pojmovanji Newtona in Leibniza, po čigar teorijah se je deloma zgledoval. V nadaljevanju predstavim njegov zakon univerzalne sile, po čemer je najboljše poznan, ter njegov odnos do naravnih zakonov nasploh. Posebno pozornost namenjam Boškovićevo izvirni predelavi aristotelskih pojmov potencialnega in aktualnega, ki ju aplicira na pojma zveznega in diskretnega v povezavi s pojmovanjem prostora in časa ter sta implicitni temeljni sestavini njegovega atomizma. Na koncu omenjam še Boškovičev prispevek k znanstveni metodi in navajam nekatere dolgoročne vplive njegovega dela.

Ključne besede: Bošković, atomizem, naravni zakoni, prostor, potencialnost, znanstvena metoda

Uvod

Ruder Josip Bošković¹ (1711–1787), dubrovniški fizik, astronom, matematik, geodet, filozof, inženir, hidrograf, diplomat, pesnik in jezuit je razmeroma dobro poznana zgodovinska figura na območju Balkana – po njem se na primer imenuje tudi osrednji hrvaški znanstveni inštitut v Zagrebu. Kljub temu je v svetovnem merilu relativno slabo poznana figura glede na njegov prispevek k znanstveni misli in metodi, čeprav za časa njegovega življenja ni bilo tako. Po mladosti, ki jo

¹ Njegovo italijansko poimenovanje se glasi Ruggero Giuseppe Boscovich, angleško pa Roger Joseph Boscovich. Ker je živel v času pred vznikom nacionalnih držav, se ni posebej izrekal za državljanstvo v današnjem smislu, čeprav je sebe vsekakor štel za Dubrovničana, saj je Dubrovnik bil samostojna republika. Danes si ga lasti več držav: Hrvaška, Srbija in Italija.

je preživel v Dubrovniku, se je 14-leten odpravil v Rim na študij pri jezuitih. Večino svojega življenja je deloval na območju Italije in Francije, veliko je tudi potoval po Evropi z namenom raziskovanja. Med svojim daljšim obiskom v Londonu je postal tudi član angleške Kraljeve družbe.

V slovenskem prostoru se je o Boškoviću pisalo relativno malo, izjemi sta Stanislav Južnič in Sandi Sitar,² ki sta o njem napisala več člankov, tako da je Boškovićevo življenje in delo vsaj v osnovi že bilo predstavljeno. Južnič se v svojih strokovnih člankih osredotoča predvsem na njegove dosežke na področjih astronomije, kemije in tehnike ter na njegov vpliv na ožjem prostoru Ljubljane ter širše v slovenskem prostoru, tudi z njegovo fizikalno teorijo. Kot jezuit je Bošković namreč večkrat tudi predaval v Ljubljani ter s tem deloma pustil svoj pečat tudi pri nas, ki pa je majhen v primerjavi z njegovim mednarodnim vplivom, o čemer bom pisal kasneje. O Boškoviću je na balkanskem prostoru najbolj poglobljeno pisal že pokojni znani hrvaški jedrski fiziki in humanist Ivan Supek, na čigar monografijo o Boškoviću se tudi opiram (Supek, 1989). Sam se osredotočam na Boškovićevo teorijo iz problemsko-konceptualnega vidika, ki je pomemben za sočasni razvoj znanosti.

Boškovićeve atomi

Bošković je dandanes med zgodovinarji znanosti najbolj poznan po svojem pojmovanju atomov in z njim povezano t.i. krivuljo sil, o kateri več nekoliko kasneje. Tako kot velja za druge pojme pri njem, se tudi glede razumevanja samih atomov zgleduje tako po Leibnizu kot po Newtonu, ki sta sicer bila zagrizena nasprotnika na več področjih, zato bom najprej na kratko predstavil ti dve pojmovanji.

Newtonovo razumevanje atomov, ki ga je podal na precej skop način, je mogoče povzeti predvsem iz nekaterih odlomkov iz njegove *Optike* (1952). Tam med drugim zapiše, da je verjetno Bog ustvaril

/.../ snov kot trdne, masivne, trde, nepredirne, gibljive delce takšnih velikosti in oblik in s takšnimi drugimi lastnostmi ter prostorskimi razmerji, da so najprimernejši za namen, zaradi katerega jih je ustvaril; in da so ti prvotni delci, ki so trdni, neprimerno trši od kateregakoli poroznega telesa, ki je iz njih sestavljeno, celo tako trdi, da se nikoli ne obrabijo in razlomijo na koščke. (Newton, 1952: 375-376)

Johnson in Wilsonova na podlagi tega odlomka iz tretje knjige *Optike* trdita, da Newton na tem mestu pravzaprav parafrazira Lukrecija in njegovo delo *O naravi sveta*, ki je imelo vpliv že na Pierra Gassendija, enega prvih evropskih mislecev

² Pri Sitarju gre za dva krajša članka bolj opisne narave, ki se ne ukvarjata z Boškovićevo naravoslovno teorijo iz problemsko-konceptualnega vidika.

atomizma, ter da njegov atomizem lahko v grobem razumemo kot lukrecijevski tip atomizma, čeprav mu je dodal tudi teološke elemente kreacionizma in voluntarizma (Johnson in Wilson, 2007: 141–142). Če nekoliko povzamem, Newtonovi delci imajo razsežnost, maso, inercijo, ki je pasivna lastnost, sila med njimi pa je neživa (t.i. *vis mortua*), kar se je tedaj razlikovalo od tako imenovane žive sile (*vis viva*), in po Newtonu ne deluje na razdaljo.

Leibniz po drugi strani pristaja na *monade*, ki zanj pomenijo neminljive, nespremenljive in nedeljive enostavne umske substance (Leibniz, 2004a: 133). Monade so »resnični Atomi Nature in, z eno besedo, Elementi stvari« (ibid.). Njihova pomembna značilnost je dušni princip, ki ga povzemata lastnosti percepcije in apetitije. Percepcija ni nič drugega kot »menljivo stanje, ki zaobjema in predstavlja neko množstvo v celoti ali v enostavni substanci« (Leibniz, 2004a: 135), torej stanje, ki odraža mnogoteri vpliv zunanosti v monadi sami. To moramo razlikovati od zavesti oziroma apersepcije.³ Apetecija je »dejavnost notranjega principa, po kateri se dogaja spreminjanje ali prehajanje od percepcije k percepciji« (ibid.), pri čemer apetit teži k percepciji in je ne more povsem doseči. Z nekoliko ustvarjalnega razumevanja bi v sodobnem fizikalnem smislu apetecijo morda lahko vzporejali z variacijskim načelom minimalne akcije, ki opisuje prehod v ravnovesno stanje, percepcijo pa kot to ravnovesno oziroma stacionarno stanje. Pri vsem omenjenem ni odveč poudariti, da Leibniz ločuje med dušo in monado:

Če rečemo Duša vsemu, kar ima *percepcije* in *apetecije* /.../, potem bi lahko vse enostavne substance ali ustvarjene Monade imenovali Duše; toda ker je čutenje nekaj več kot enostavna percepcija, sem sam za to, da naj zadostuje obče ime Monada ali Entelehija za enostavne substance, ki imajo samo le-to, medtem ko imenujemo *Duše* samo tiste, katerih percipiranje je bolj določno in ga spremlja pomnljivost. (Leibniz, 2004a: 135)

Bolj korektno je torej reči, da pojem monade vključuje princip duše in da gre za umsko substancu v tem smislu besede.

Monade so lahko aktivne ali pasivne, mera njihove aktivnosti pa je odvisna od mere (ne)jasnosti njihove percepcije: »Tako pripisujemo Monadi Aktivnost, kolikor so njene percepcije razločne, trpnost pa, kolikor so zmedene« (Leibniz, 2004a: 141). V tem smislu inercija monade ni njena pasivna lastnost, njena pritliklina, temveč je rezultat dogajanja v smislu zmedenih predstav. Monade tako tudi nimajo mase, masa je šele rezultat enostavnih substanc. Monade nimajo razsežnosti, med sabo ne interagirajo na razdaljo, sila med njimi pa je živa (*vis viva*). Morda ni odveč omeniti, da se je Leibniz glede atomizma zelo verjetno navdahnul tudi pri piscu arabske srednjeveške šolske filozofije Kalāma Maimonidesu (Jammer,

³ Leibniz je npr. mnenja, da so kartezijanci zgrešili prav v tej točki, »ko so za nič šteli percepcije, katerih se ne zavemo« (Leibniz, 2004a: 133).

1954: 64).⁴ Do svojega pojmovanja atomizma je torej prišel po nekoliko drugačni poti kot Newton.

Bošković svoje razumevanje atomov poda predvsem v *Theoria philosophiae naturalis* (1758), ki predstavlja njegov *opus magnum*. Supek (1989) in Child (1922) povzemata nekatere osrednje lastnosti njegovih atomov. Njegove »snovne točke«, kot jih tudi imenuje, so med seboj enake, nedeljive, nepredirljive, telesa, ki so iz njih sestavljena, pa nič od tega. Bošković tako torej ne priznava neskončne deljivosti, temveč samo neskončno sestavljenost. Atomi niso niti togi niti elastični. Sami po sebi nimajo mase, čeprav je ta sorazmerna številu »snovnih točk«, torej množini snovi. Zaradi takšnega pogleda je Bošković bil nekoliko v konfliktu s tedaj prevladujočim materializmom, ki se je najbolj učvrstil prav s pojmom mase. Kljub temu so »snovne točke« inertne, kar jih dela sorodne Newtonovim atomom, pri katerih pa je inercija neposredno povezana z maso. V vseh teh lastnostih so Boškovičevi atomi podobni Leibnizevim monadam (Child, 1922: XII). Bistvena razlika med obema je, da so Boškovičevi atomi neživi, ne poznajo apeticije in percepcije, njihova inercija pa je – za razliko od Leibnizevih monad – pasivna lastnost. Kot omenja Supek, je še ena od pomembnih lastnosti, ki jih loči od monad, ta, da so atomi zaznavni samo v medsebojni interakciji, medtem ko za monade to ne velja, saj je percepcija vsebovana v njih samih (Supek, 1989: 110). Tudi Child poudarja, da »posamična točka ne bi dala čutnih podatkov [*sense datum*], če je ločena od druge točke; zato je lahko posamezne točke razumemo kot same na sebi nezaznavne, vendar kot zaznavne skozi odnose do ostalih snovnih točk« (Child, 1922: XIV).

Sam Bošković je trdil, da njegova *Teorija* stoji med Newtonovo in Leibnizevo. Supek piše, da Bošković

od Newtona odstopa, ker zavrača trdne atome, vendar ohranja dominantnost sile, Leibnizu se približuje s sliko breztežnih atomov, vendar zavrača njegovo predestinirano harmonijo, ki upravlja vsako monado posebej. Tako Leibniz kot Bošković se držita načela kontinuitete, vendar to Dubrovničan najprej aplicira na gibanje, medtem ko materija sestoji iz diskretnih točk. (Supek, 1989: 110)

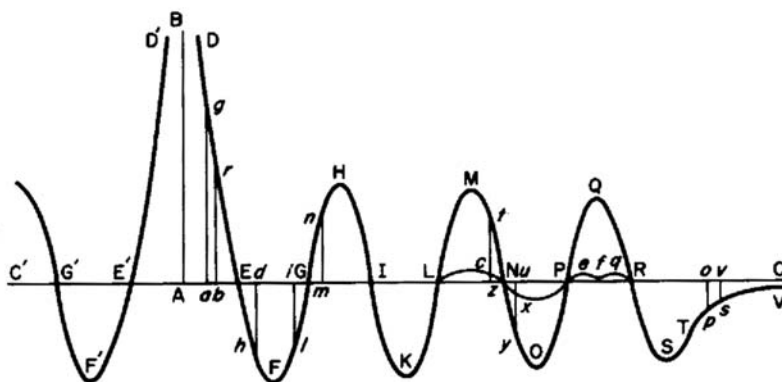
Boškovičeva novost v odnosu do klasičnega atomizma je prav to njegovo priseganje na princip kontinuitete, kljub temu da se to sliši protislovno. Antični atomizem Demokrita in Lukrecija je namreč poleg obstoja atomov postuliral tudi obstoj praznine, to je njegov bistveni sestavni del, medtem ko Bošković zavrača atomistično predpostavko praznega prostora (ibid.: 85). V tem smislu je zanj glavna on-

⁴ Ena od zanimivih posebnosti te šole v odnosu do atomizma je, da je gibanje atomov razumela kot *diskretno*, kar je zanimiva podobnost z nekaterimi interpretacijami kvantne mehanike. Kljub temu da se je ta arabska srednjeveška »sholastika« izdatno zgledovala po Aristotelu, *al failosoufu*, je vseeno prisegala na atomizem, ki ga je on sam zavračal.

tološka danost »sila«, materiji odvzame razsežnost, kar je bila bistvena značilnost materialne substance po Descartesu (ibid.: 80). V naslednjem razdelku zato obravnavam njegovo krivuljo sil.

Boškovičeva »krivulja sil« in razumevanje naravnih zakonov

Na koncu prejšnjega razdelka je bilo nakazano, da Boškovićevo pojmovanje atomov ni zgolj nekakšna kombinacija Newtonovega in Leibnizevega pojmovanja, temveč prinaša tudi pomembne izvirne elemente. Te najbolj povzema njegova »krivulja sil« (slika spodaj), ki predstavlja prvo matematično artikulacijo teorije atomov, ki so vse do tedaj bili razumljeni na nematematični način. S tem se teorija atomov prvič resno sreča z razvijajočo se moderno znanostjo.



Boškovičeva krivulja sil, kot jo najdemo v njegovi *Theoria philosophiae naturalis* (vir: internet). Snovni delec se nahaja v izhodišču v točki A, krivulja sil pa nakazuje, kakšna sila deluje na drug delec, ki se nahaja na določeni oddaljenosti od točke A.

»Krivulja sil« prikazuje, kako se sila na neko snovno točko spreminja z oddaljenostjo v smeri vzdolž premice od snovne točke, ki je postavljena v izhodišče koordinatnega sistema (točka A). V veliki oddaljenosti (od točke S dalje) sila dokončno preide v privlačno gravitacijsko silo, ki se zmanjšuje s kvadratom razdalje in gre proti 0 z večanjem razdalje med snovnima točkama. Iz grafa je razvidno, da je sila na nekaterih odsekih pozitivna, na drugih pa negativna, kar z drugimi besedami pomeni, da atomi na neki razdalji čutijo odbojno, na drugi privlačno silo, na nekaterih območjih pa so prehodi med obema režimoma. To ima za posledico, da obstajajo določene medsebojne ravnovesne lege, v katerih je sila med atomoma enaka nič in v katerih atomi obmirujejo (Bošković, 1996: 167). Gre torej za medsebojna ravnovesna stanja atomov, s kakršnimi se danes opisuje povezavo dveh atomov v molekuli.

Kar se pri tem pojmovanju zdi pomembno, je to, da »krivulja sil« predstavlja neko zvezno entiteto, ki zapolnjuje prostor. Ravno v tem pogledu je Boškovičeva krivulja sil prekurzor kasnejšega pojma električnega polja, ki ga je uvedel Michael Faraday, in kasneje tudi vseh drugih variant fizikalnega pojma polja. Takšno pojmovanje je bilo novost v razmišljanju o atomih in silah nasploh. Sicer je tudi Newton zagovarjal stališče, da sila ne deluje preko praznega prostora, da pri gravitaciji ne gre za neko »strašljivo silo na daljavo« (*spooky action at a distance*). Golo delovanje na daljavo bi zanj pomenilo absurd, kot je zapisal v pismu Bentleyju, čeprav se ne izjasni, ali gre za snovni ali nesnovni medij. Kljub temu Newton ni določene oblikoval tega medsebojnega delovanja delcev, nikdar ni omenjal možnosti sile, ki bi prehajala iz privlačne v odbojno. Bošković sicer citira mesto iz njegove *Optike*, kjer piše o odbojnih silah za razliko od gravitacije, ki je privlačna sila, na drugem mestu pa omenja silo kohezije, tj. takšno, ki drži telesa skupaj, tako da mu takšni možnosti delovanja sil nista tuji. Bošković je takšne različne variante sil zvedel na eno univerzalno silo, ki jo je razumel kot rezultanto vseh sil.

Za bolj celovito razumevanje Boškovičevega atomizma se zdi pomembno natančneje secirati tudi njegov pojem sile. Kot poudarjajo nekateri proučevalci, Boškovičeva »krivulja sil« pravzaprav ne opisuje sil v Newtonovem smislu. Sicer drži, da je svoj zakon poimenoval *legis virium*, vendar izraz *vis* v njegovem času ni imel enoličnega pomena sile. Kot piše J. M. Child v svojem predgovoru k *Teoriji*, je Boškovičeva teorija *kinematična*, kar pomeni, da uporablja zgolj intuitivne in empirične pojme, kot so pot, hitrost, pospešek, čas. Za razliko od tega je Newtonova teorija *dinamična*, saj uporablja tudi pojem sile in njej pridružen pojem mase, ki sta aksiomatizirana v 2. Newtonovem zakonu. »Zgleda, da je Boškovičeva krivulja graf pospeška-intervala [*acceleration-interval graph*]; napačno je nanašati se na njegov kozmični sistem kot na sistem 'centrov sil'. /.../ Kljub temu je besednjak 'privlačnosti' in 'odbojnosti' toliko bolj priročen kot ta o 'pospešku hitrosti približevanja' in 'pospešku hitrosti oddaljevanja', da bo uporabljen v nadaljevanju – prav tako, kot je bil uporabljen tudi vseskozi prevod razprave« (Child, 1922: XIV).

Max Jammer v svoji monografiji *Concepts of Force – a Study in the Foundations of Dynamics* (1957) prav tako potrjuje takšno razumevanje pojma sile pri Boškoviću:

Prav gotovo ne more iti za Newtonovo definicijo [sile], saj Boškovičeve materialne točke, čeprav so navdane z inercijo, nimajo mase v newtonovskem pomenu besede. /.../ Torej je za Boškovića »sila« določenost, težnja približevanja in oddaljevanja in se meri s proizvedenim pospeškom. Nikjer v svojih spisih Bošković ne poskuša razložiti vzroka za te določenosti; zgolj postulira njihov obstoj. Torej, strogo rečeno, ordinate /.../ na krivulji sil predstavljajo zgolj pospeške. (Jammer, 1957: 177)

Takšno razumevanje pojma sile lahko nadalje podpremo tudi z Boškovičevo izrecno in matematično natančno trditvijo, da »ploščina krivulje [tj. pod krivuljo],

katere abscise predstavljajo razdalje in ordinate sile, predstavlja povečanje ali zmanjšanje kvadrata hitrosti»⁵ (Bošković, 1922: 105). Pojmi, ki jih Bošković uporablja, torej označujejo osnovne opazljive količine, ki opisujejo gibanje. Nikjer izrecno ne omenja mase, kot je že bilo rečeno. Takšen pojem sile torej ni ekvivalenten Newtonovemu pojmu sile, temveč mu je zgolj premo sorazmeren v matematičnem smislu, tako kot pospešek. Child zato argumentira, da to, kar resnično obstaja v Boškovičevi teoriji, ni sila, temveč gibanje, ki ga lahko kvantificiramo kot produkt pospeška in mase v smislu množine snovnih točk. Privlačnost teles, kot tudi Bošković izrecno zapiše, je neka določenost oziroma težnja: »To določenost imenujem 'sila' <vim> /.../; ta izraz ne označuje načina delovanja, temveč izražanje določenosti same [*ipsam determinationem exprimens*], karkoli že je njen izvor, katere magnituda se spreminja s spreminjajočimi razdaljami« (ibid.: 39). Takšno razumevanje pojma sile nakazuje, da njegov pojem univerzalnega zakona implicira nekaj takšnega kot univerzalno težnjo in se ne pretvarja, da lahko pojasni vse z uporabo pojma sile.

Bošković sam je na svoj zakon gledal kot na razširitev Newtonovega univerzalnega zakona gravitacije tudi na mikroskopsko domeno. Podobno kot Newton je tudi sam za cilj imel postavitev naravnega zakona, ki pa mora vključevati tudi dogajanja na atomskem nivoju. V času Boškovičevega pisanja *Teorije*, okrog 1750, je postalo znano Franklinovo delo o elektriki, ki je uvajalo električno silo. Bošković, ki se je zavzemal za univerzalno silo, je zato hotel električno in gravitacijsko silo zvesti na eno skupno silo, ki deluje na osnovne gradnike. Svoje delo je v tem oziru razumel kot »mogočen napredek v filozofiji narave«, saj je dogajanje razlagalo na podlagi maloštevilnih osnovnih principov. V predgovoru k svoji knjigi *Teorija naravne filozofije* piše:

V njej je vsebovana nova zvrst univerzalne filozofije narave, takšna, ki se močno razlikuje od katerekoli, ki je splošno sprejeta ali prakticirana v sedanjem času; kljub temu pa so ključne točke vseh najbolj zaslužnih teorij današnjega časa, prepletene in utrjene na resnično čudovit način, kombinirane v njej; ravno tako to velja za enostavne nerazsežne elemente pristašev Leibniza kot tudi za newtonovske sile, ki proizvajajo medsebojno približevanje na nekaterih razdaljah in medsebojno oddaljevanje na drugih in ki ju običajno imenujemo privlačnost in odbojnost. (Bošković, 1922: 9)

Bošković se torej javlja kot ta, ki spravlja dve različni teoriji o naravi, Newtonovo in Leibnizevo, ki sicer ni bila matematično formulirana, v okviru nove teorije, ki opisuje interakcije med telesi na poljubnih medsebojnih razdaljah. V tem smislu jo lahko imamo za zgodovinski primer teorije poenotenja, ki povezuje več narav-

⁵ Ploščina pod krivuljo (in nad absciso) v grafu namreč ustreza integralu količine, ki je predstavljena na ordinati. Če z integracijo po koordinati x na abscisi dobimo količino »kvadrat hitrosti«, v^2 , potem mora integrirana količina biti pospešek, a , saj velja: $\int a \, dx = \int \frac{dv}{dt} \, dx = \int dv \, \frac{dx}{dt} = \int v \, dv = v^2 + \text{const.}$

nih zakonov. Tovrstna težnja je sorodna današnjim težnjam po poenotenju teorije relativnosti in kvantne teorije. Pri govoru o tej sintezi pa je le potrebno biti nekoliko bolj natančen. Boškovičeva sinteza sestoji iz združitve Newtonovega gravitacijskega zakona in njegovega novega zakona o medatomskih silah. Pri tem je zelo pomembno poudariti, da tukaj ne gre za enakovredno sintezo, saj je v tem času gravitacijski zakon že bil empirično verificiran, medtem ko je bil zakon o medatomskih silah povsem spekulativne narave in eksperimentalno nepreverjen. V današnjem smislu bi lahko govorili prej o matematični spekulativni razširitvi zakona izven eksperimentalno dosegljivega območja.⁶

Poleg omenjenega je potrebno dodati, da Boškovičeva univerzalna sila ni bila matematično enolično artikulirana. To je gotovo tudi posledica njegove želje, da bi se približal bralcu: »Odločil sem se, da je zame najpomembnejše, da čim bolj poskrbim, kolikor je to bilo mogoče, da naj bo vse jasno razloženo in da tukaj zato ni potrebe po višji geometriji ali računanju« (ibid.: 13). Že v uvodu omenja, da v *Dodatku III*, naslovljenem *Analitična rešitev problema določitve narave zakona sil* (ang. *Analytical Solution of the Problem to determine the nature of the Law of Forces*), ponuja tudi analitično formulo za svojo krivuljo (ibid.: 427); gre za polinom n -te stopnje, ki ima n (realnih) ničel. Iz tega dodatka je razvidno, da njegova analitična formulacija zakona sil ni matematično enolična, saj tukaj razpravlja bolj ali manj zgolj o nujnosti, da ima krivulja več izmenjujočih se lokalnih minimumov in maksimumov, vendar ne ve, koliko jih je v resnici. Njegov pristop ostaja v splošnem geometrijski, tako kot Newtonov slabo stoletje pred tem (Supek, 1989: 106). Kljub temu sam poudarja edinstvenost svojega naravnega zakona:

Čeprav, kot sem že poudaril, je mogoče ta zakon sil miselno razdeliti v več zakonov, te pa bi bilo mogoče predstaviti z več ustreznimi krivuljami, je mogoče ta zakon, ki je unikatni, sestaviti iz vseh teh s pomočjo unikatne [*unique*], zvezne [*continuous*] in enolične [*one-fold*] krivulje, ki jo podajam. (ibid.: 23)

Pri Boškoviću lahko torej govorimo kvečjemu o kvalitativni, predvsem pa o konceptualni sintezi naravnih zakonov in ne o znanstveni sintezi teorij v modernem smislu. Tudi Supek piše, da v *Teoriji* Bošković ne podaja nobenega temeljnega fizikalnega zakona ali eksperimenta, temveč poskuša konceptualno spremeniti pogled na naravo v tem smislu, da pojmovanje materije, prostora in časa spremeni s pomočjo univerzalne sile.

Na podlagi Boškovičevega razumevanja vloge matematike pri opisu naravnih zakonov in njegovega razumevanja pojma sile lahko zaključimo, da je naravne zakone razumel kot univerzalne, vendar ne nujno kot univerzalno formulirane. Uni-

⁶ Čeprav je ta oznaka morda slišati slabšalno, je treba dodati, da tudi v moderni znanosti to ni redka praksa – kot takšne lahko označimo različne teorije strun, teorijo kvantne gravitacije, tudi razne modele črnih lukenj, čeprav slednji vseeno temeljijo na empirično trdnih temeljih splošne relativnosti.

verzalnost naravnih zakonov ni pogojena z njihovo matematično enoličnostjo. Njegova formulacija zakona opisuje gibanje snovnih točk, ki bi načeloma lahko bilo matematično izraženo na več načinov. V modernem smislu je njegov pogled na naravne zakone torej neredukcionističen, zakonov ne reducira na to ali ono opredelitev.

Načelo zveznosti in dihotomija potencialni-aktualni prostor

Prehajam k še eni od posebnosti Boškovičeve teorije atomizma in sicer k njegovemu pojmovanju prostora in časa. To je tesno povezano z vlogo zveznosti v njegovi teoriji, ki ima status načela oziroma zakona, o njej je napisal tudi posebno delo, *De continuitatis lege* (1754). Zakon neprekinjenosti je pravzaprav *apriorni* temelj celotne Boškovičeve teorije, kar sam jasno poudari (Bošković, 1996: 159). Prevzel ga je od Leibniza in je posebni pečat njegovega atomizma, kot sem že omenil. Med misleci, ki dopuščajo neprekinjenost v naravi, Bošković med drugim navaja Newtona in škotskega matematika Maclaurina. Ta dva jo dopuščata predvsem v tem smislu, da gre pri trdih telesih za kontinuum, manj pa v smislu razmerja med atomi in prostorom, kar je ravno Boškovičev poudarek. Drugi za vsa telesa trdijo, da so elastična – to v prvi vrsti velja za vse pristaše Leibniza.

Bošković načelo neprekinjenosti v svoji *Teoriji* dokazuje na dva načina, induktivno in metafizično (Bošković, 1922: 55–70). Njegov metafizični argument je ta, da je v prostoru, če ga razdelimo na neskončno majhne intervale, lahko neskončno mej, te pa niso deljive in tudi po Aristotelu povezujejo to, kar delijo – za Aristotela je prostor namreč »prva negibna meja obdajajočega« (Aristotel, 2004: 201). V tem smislu lahko govorimo o neprekinjenosti mej. Še drugače rečeno, če ves prostor razparceliramo na neskončno majhne dele, od njih ostanejo le še meje, te pa dele prostora povezujejo in ne delijo. S tem miselnim limitnim procesom dobimo (potencialno) zvezen prostor, ki je poln povezujočih mej. Aristotel sam naj sicer ne bi imel povsem jasne predstave neprekinjenosti, pripominja Supek, čeprav loči med zaporednim, stičnim in neprekinjenim. Po njegovem mnenju Aristotel zavrača atomistiko zato, ker jo razume v fizikalnem in ne matematičnem smislu, ravno to slednje pa je značilna poteza Boškovičevega atomizma. Lahko bi rekli, da je Boškovičev atomizem matematični atomizem, saj je atom zanj nerazsežna točka, torej matematični objekt, ki povzema lastnost nedeljivosti.

Omenil sem že, da je ena od Boškovičevih novosti znotraj atomizma bila ta, da prostora ni dojemal kot zgolj praznega (Supek, 1989: 85), temveč kot nepraznega in zveznega. Za razliko od Leibniza Bošković ni zavračal diskretnosti, temveč je poiskal drugačen način, na katerega lahko ta sobiva s pojmom zveznosti. Da bi to bolje razumeli, je treba poudariti, da je ta njegov pogled strogo vezan zgolj na gibanje: »V nadaljevanju pokažem, da je zveznost sama v resnici lastnost zgolj gibanj in je v ostalih stvareh bolj ali manj zgrešen privzetek« (Bošković, 1922: 25). To pomeni,

da so zanj vse bivajoče stvari diskretne, le gibanje je zvezno. Supek v svoji prodorni analizi argumentira, da Bošković uspe najti pomiritev med pojmom kontinuuma in diskretne točke v Aristotelovi dihotomiji potencialno-aktualno (Supek, 1989: 85). Ta dihotomija leži pravzaprav v samem srcu njegove atomske teorije. Zanj predstavlja zadostni temelj, da lahko uvede atomizem kljub temu, da je ta tradicionalno nasprotoval pojmu zvezne snovi, saj je predpostavljal prazen prostor.

Ta aristotelska dihotomija je tuja Newtonu, ki razume čas in prostor kot absolutna, ter ju označi za božja atributa – prostor je »božji senzorij«. Pri njem »[a]bsolutni, resnični ali matematični čas teče enakomerno, sam po sebi in iz svoje lastne narave, brez odnosa do česar koli izven sebe« (Newton, 1995: 6). Dogajanje, ki je vezano na praktično mero časa, pa povzema pojem relativni čas: »/.../ relativni, navidezni ali navadni čas [pa] je neko čutno in zunanje (bodisi natančno bodisi nenatančno) merilo trajanja z gibanjem, ki ga ponavadi uporabljamo namesto resničnega časa: na primer dan, ura, mesec, leto« (Newton, 1995: 6). Ta se relativno dobro ujema z Aristotelovo definicijo časa iz *Fizike* – ta je »število gibanja glede na preje in kasneje« (Aristotel, 2004: 229) in je neločljivo povezan z gibanjem oziroma spreminjanjem. Nekaj podobnega kot za čas velja za Newtonovo razumevanje prostora: »Absolutni prostor ostaja vselej isti in negiben po svoji lastni naravi, brez odnosa do česar koli izven sebe. Relativni prostor [pa] je neka gibljiva razsežnost ali merilo absolutnih prostorov, ki ga naši čuti določijo s svojim položajem do teles in katerega navadno smatramo za negibnega« (ibid.). Newton torej pozna ontološko temeljni absolutni prostor ter čutni relativni prostor, ki ima ontološko podrejeni status.

Prav tako je omenjena dihotomija tuja Leibnizu, ki je zavračal realnost prostora in časa – njegova znana trditev iz pisemske polemike z Newtonovim zagovornikom Clarkom se glasi: »Kar se mene tiče, sem nič kolikokrat dal vedeti, da mi velja prostor za nekaj povsem relativnega, prav tako kot čas; da mi pomeni neki red koeksistenc, tako kot je čas neki red zaporedij« (Leibniz, 2004b: 165). Relacijska narava prostora in časa pri njem tema dvema pojmom odvzema temeljni ontološki status, kakršnega imata absolutni prostor in čas pri Newtonu.

Bošković razume prostor in čas v okviru aristotelske dihotomije potencialno-aktualno. V dodatku k *Philosophiae Recentioris* (1760) klerika Benedikta Staya piše o dveh vrstah prostora. Prvi je imaginarni zvezni matematični prostor, ki je neskončen in večni, v njem se izvajajo temeljne matematične operacije. Drugi je realni in povezan s časom, o njem lahko razmišljamo kot o množici vseh snovnih točk oziroma atomov, ki v nekem trenutku v času pridejo iz virtualnega prostora, in je v tem smislu diskreten. Gibanje v tem kontekstu pomeni stalno manifestacijo realnega prostora in časa: »Pri Boškoviću je nujno razumeti, da se virtualna prostor in čas neprestano transformirata v realne položaje in čase njegovih točk in to je tisto, kar se kaže kot gibanje« (Supek, 1989: 87).

Kar se pojma prostora tiče, ima Boškovičev potencialni prostor drugačen status od Newtonovega absolutnega in Leibnizevega relativnega, čeprav so vsi trije matematizirani. Boškovičev pojem potencialnega prostora lahko sicer v sebi vključuje pojem privilegiranega univerzalnega koordinatnega izhodišča, kot velja za Newtonov absolutni prostor (ne pa za Leibnizev relativni), vendar ima drugačen ontološki status, kar nakazujeta tudi pridevka »absolutni« in »potencialni«. Boškovičev aktualni prostor je bistveno relacionalen, tako da je blizu temu, kar Newton označuje kot relativni prostor. Boškovičeva dihotomija pojmovanja prostora prinaša v primerjavi z Newtonovo drugačno ontološko postavitev, zanj je novoveški dualizem nekoliko pretesen,⁷ medtem ko za Newtonovo pojmovanje tega ne bi mogli reči. V tem sam vidim največjo Boškovičevo osvežitev v tedanjem mišljenju pojma prostora.

Potencialni čas je ravno tako matematiziran in s tem soroden Newtonovemu absolutnemu času. Aktualni čas je po drugi strani ta, v katerem se dogaja spreminjanje in ki ga lahko merimo oziroma je mera spreminjanja, kar se ujema z Aristotelovo definicijo časa oziroma z Newtonovim relativnim časom. Boškovičeva aristotel-ska dihotomija mimogrede velja tudi za sile – tudi te so bodisi potencialne bodisi aktualne: prve ne povzročajo pospeškov, medtem ko jih druge povzročajo.

Pomembno je poudariti, da je Bošković aristotel-ska pojma potencialnosti in aktualnosti transformiral. Aktualno, za katerega Aristotel uporablja izraza *entelechéia* oziroma *énérgeia*, je tisto stanje, ki pomeni dopolnitev, doseganje *télosa*, in ne tisto, ki je prezentno, kot se pojem aktualnost pogosto tolmači danes. V *Metafiziki* izrecno povezuje aktualnost oziroma dejanskost in dovršenost: »delo (*êrgon*) je namreč smoter (*télos*), dejanskost (*énérgeia*) pa je delovanje, zaradi tega pa se tudi ime dejanskost izreka po delu, in se zaradi tega razteza proti dovršenosti (*èntelécheia*)« (Aristotel, 1999: 1050 a 20-23). V tem smislu je npr. otrok, ki ta hip stoji pred nami, zanj zgolj potencialno človek, aktualno človek pa postane, ko razvije dušo in s tem doseže svoj smoter. Snov, ki je čista potencialnost (*dýnamis*), se razvije v dovršene forme s posredovanjem finalnega vzroka (lat. *causa finalis*).

Bošković je eksplicitno zavračal uporabo finalnih vzrokov pri tolmačenju narave (Supek, 1989: 114), kot je to značilno za vso moderno znanost, in ravno v tej točki je njegov odmik od Aristotela. Poleg tega slednji daje prednost aktualnosti pred potencialnostjo:

Ker pa je bilo že zgoraj opredeljeno, na koliko načinov se izreka »prej«, je očitvidno, da je dejanskost pred možnostjo. Menim pa, da je prej ne zgolj pred že opredeljeno zmožnostjo, ki se izreka kot spremljevalno počelo v

⁷ Na to kaže tudi Boškovičevo mišljenje povezave med dušo in telesom, ki ga predstavlja v enem od dodatkov na koncu *Teorije*. V ta namen uporablja dvojico aktualno-potencialno. Na nekem drugem mestu tudi izrecno zapiše, da njegov sistem ne razlikuje med telesnim in duhovnim, torej ni dualističen (Bošković, 1996: 173).

drugi stvari ali v stvari sami, kolikor je druga, marveč nasploh pred vsakim gibalnim in ustalitvenim počelom. (Aristotel, 1999: 1049 b 4-9)

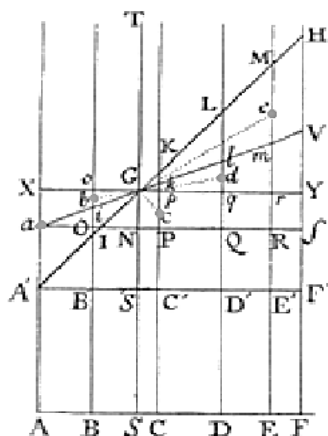
Bošković po drugi strani ne daje prednost nobenemu od obeh vidikov. Da je torej lahko uporabil to dihotomijo, jo je moral transformirati tako, da jo je ločil od pojma finalnega vzroka. Kar se vzročnosti tiče, Bošković namreč prisega na moderno zakonsko (deterministično) kavzalnost, kot jo je kasneje poudarjeno zagovarjal Laplace. Poleg omenjenega je bistven Boškovičev odmik od Aristotelovega pojma potencialnosti v tem, da je svoj pojem potencialnosti bistveno vezal na prostor in čas, medtem ko je pri Aristotelu vezana na snov (*hýle*). Na nek način je s tem aristotelsko potencialno snov »križal« s pojmom prostora, ki je za Aristotela objektivni in ima svojo kozmološko strukturo, vendar je neločljiv od predmetov, saj je njihova nepremična zunanja meja, ki telo vsebuje.

Napetost med diskretnim in zveznim ponovno izrazito oživi v moderni znanosti z vznikom kvantne mehanike, v kateri lahko vidimo moderno znanstveno področje, za katerega se zdi Boškovičeva zastavitev pojmov filozofsko zelo relevantna (Lukan, 2015: 73–74). Njegova »krivulja sil« je tako kot atomski potencial v kvantni fiziki struktura, ki kaže na potencialne interakcije med vsaj dvema delcema, je nekakšna struktura potencialne aktualizacije, bi lahko rekli. V pojmu potenciala oziroma pojmu polja se poudarek z interakcije dveh delcev prenese na samo en delec (za razliko od pojma sile, kjer je ta vedno definirana kot sila med dvema telesoma), polje povzema strukturo potencialnosti interagiranja enega delca z drugim. Sam sem mnenja, da je Boškovičeva transformirana pojmovna dvojica lahko zelo ploden koncept tudi za interpretacijo pojma verjetnosti, kot ga najdemo v moderni fiziki (tj. tako imenovana mersko-teoretska verjetnost).

Prispevek k znanstveni metodologiji

Če govorimo o Boškovičevem prispevku k razvoju moderne znanosti, se zdi pomembno opozoriti še na en vidik. Ta je za razliko od njegovega atomizma, ki je teoretičnega pomena, bolj metodološkega pomena za znanost. Boškoviću pripisujejo prvo rabo matematične metode za enolično določitev najbolj optimalne krivulje prilagajanja množici merskih točk, čemur se dandanes v žargonu reče »fitanje« (ang. *best fit to data*) (Eisenhart, 1961: 200-212). Gre za to, da množici meritev poiščemo optimalno krivuljo, za katero verjamemo, da najbolje povzema meritve. Metodo je do popolnosti razvil francoski matematik Legendre in je znana pod imenom metoda najmanjših kvadratov (ang. *least square method*) ter se danes uporablja v vseh obdelavah podatkov, v katerih nastopa prilagajanje krivulje. Ta pristop omogoča tudi oceno merske napake in torej bolj rigorozno operiranje z merskimi natančnostmi.

Bošković je ta pristop prvič uporabil na problemu določitve eliptičnosti Zemlje. Na podlagi petih tedaj znanih meritev z različnih koncev Zemlje je prvi izračunal ta parameter in s tem empirično potrdil hipotezo, da Zemlja ni povsem okrogla. Njegov pristop je bil geometrijski, tako kot Newtonov v *Principih*, za razliko od nekoliko kasnejšega Legendrovega, ki je bil analitičen.



Boškovičeva aplikacija metode najmanjših kvadratov na geometrijski način pri problemu določitve eliptičnosti Zemlje na podlagi petih meritev (vir: Stigler, 1986: 48). Pri tej metodi se izračuna »težišče« podatkov – matematika je povsem enaka kot pri računanju fizičnega težišča.

Metoda najmanjših kvadratov je igrala pomembno vlogo v zgodovini razvoja teorije meritev in razvoja statistike. Skupaj z združitvijo uporabe normalne krivulje (oziroma krivulje napak, kot so jo tedaj imenovali) v tako imenovani Gauss-Laplaceovi sintezi je tvorila teoretično hrbtnico znanstvene metode 19. stoletja, ki je v nadaljnjem razvoju privedla tudi do razvoja statistike kot samostojne discipline (Stigler, 1986: 139–158). V okviru t.i. galilejske paradigme, ki naravoslovno znanost razume kot kombinacijo uporabe matematike v teorijah in izvajanja eksperimentov za njihovo podkrepitev, je teorija meritev pravzaprav tisto presečno področje, v katerem sta oba elementa te paradigme najbolj trdno povezana. Za metodo najmanjših kvadratov lahko zagotovo rečemo, da je kot del Gauss-Laplaceove sinteze teorije meritev pomagala postaviti trden temeljni steber galilejske paradigme.

Bošković sam je s svojim pristopom uvajanja rigoroznih matematičnih principov v znanstveno metodologijo implicitno izkazal privrženost galilejski paradigmi, še več, s svojim dosežkom je celo pomembno prispeval k njeni utrditvi. S tem je prispeval tudi k razvoju teorije merskih metod, ki so več kot stoletje kasneje pomagale utrditi moderni atomizem na podlagi več dobro zasnovanih poskusov. Bošković sam za takšne poskuse ni imel nikakršne možnosti, saj pravzaprav vsi

temeljijo na uporabi elektronske merilne tehnike, ki v njegovem času še ni obstajala. Njegov atomizem je zato ostal pri statusu matematične spekulacije. V tem smislu lahko vidimo nekaj ironije v tem, da je njegovi atomistični teoriji – poleg konkretnih poskusov – umanjkal splošni metodološko-eksperimentalni temelj, ki ga je pravzaprav on sam šele pomagal zgraditi ter je kasneje bistveno prispeval h končni uveljavitvi atomizma v fiziki.

Vpliv in pozaba

Boškovićevo delo se je po evropskih univerzah, predvsem znotraj Avstro-Ogrske, predavalo v drugi polovici 18. in prvi polovici 19. stoletja, vendar se kasneje ni obdržalo. Postopna pozaba Boškovičevega doprinosa se je odvijala tudi kasneje v 20. stoletju in je – kot poroča Supek – na primer razvidna tudi skozi zaporedne izdaje *Encyclopedie Britannice*, v katerih se je geslo pod njegovim imenom čedalje bolj krajšalo, dokler ni v izdaji iz leta 1977 povsem izpadlo.

Glavni razlog za to, da se Boškovičev atomistični nauk ni ohranil, Supek vidi v tem, da to ni bilo matematizirano znanje. Na tem mestu je potrebno dodatno pojasnilo, saj sem zgoraj zapisal, da gre pri Boškoviću za prvo matematično artikulacijo atomistične teorije. Kot že omenjeno v razdelku o krivulji sil, pri Boškoviću ne gre za matematično enolično artikulacijo zakona o medatomskih silah, ta bi lahko bil tudi drugače oblikovan. Njegova krivulja ni podana kot rešitev matematičnega problema, tj. diferencialne enačbe, temveč kot rešitev konceptualnega problema. V modernem fizikalnem smislu so naravni zakoni običajno pojmovani kot diferencialne enačbe (pogojno morda kot krivulje, ki predstavljajo njihove rešitve), ne pa kot poljubne matematične krivulje. Ustrezni matematični problem, tj. ustrezne in precej zapletene diferencialne enačbe, je rešil šele Friedrich Bessel v prvi polovici 19. stoletja.

Boškovićevo delo je ponovno oživil ameriški eksperimentalni fizik Michael Faraday, ki je znan predvsem po svojih poskusih z električnimi pojavi in po uvedbi pojma električnega polja. Faraday je trdil, da mu je pri uvedbi pojma polja pomagala Boškovičeva teorija atomov, ki razume sile kot prvotno realnost, čeprav, kot sem pokazal zgoraj, ni šlo za newtonovsko pojmovanje sil. Zanj je ena glavnih potez Boškovičeve misli v tem, da je dala univerzalno silo, na katero se dajo zvesti vse ostale znane sile. Faraday se od Boškoviča razlikuje samo po tem, da slednji še vedno ohranja aristotelska pojma aktualnega in potencialnega (Supek, 1989: 172). Pri obeh polje – čeprav Bošković ne uporablja tega izraza – ustvarja maso atoma, sile so bolj temeljne od atoma.

Pojem polja je v moderni fiziki ključnega pomena, saj je sestavni del obeh osrednjih fizikalnih teorij, splošne relativnosti in kvantne teorije polja. Lahko bi rekli, da predstavlja tisto entiteto, ki najbolje povzema predstavo moderne znanosti o odnosu med prostorom in snovjo: polje lahko razumemo tudi kot ime za to, da

prostor nikoli ni povsem prazen, ker je v njem vedno polje, čeprav v njem ni obstojnih delcev, hkrati pa polje interagira z delci. Gledano skozi prizmo antične dileme med atomizmom in t.i. plenizmom, h kateremu štejemo tudi Aristotela, polje funkcionira kot nekakšen mediator med pojmom snovi in prostora, s čimer presega atomistično dihotomijo, po drugi strani pa gre za brezmasno entiteto in torej ne za snov v modernem fizikalnem smislu. Boškovičev pojem univerzalne sile je zgodovinski predhodnik pojma polja tudi v tem, da kot pojem – za razliko od pojma sile – ne potrebuje dveh entitet, med katerima deluje. S tem Boškovičev atomizem že prinaša konceptualna vprašanja, ki so relevantna tudi za razumevanje statusa modernega pojma polja.

Tudi James Clerk Maxwell, ena od največjih fizikalnih eminenc vseh časov in oče elektromagnetizma, ki je sprva zagovarjal atomizmu konkurenčno teorijo etra in vrtnicev, je naposled leta 1877 v reviji *Nature* zapisal, da je najbolje, kar se lahko naredi, to, da se namesto togega jedra uvede Boškovičeve atome (ibid.: 176). Te je priznaval tudi Lord Kelvin, še ena fizikalna eminenca, po kateri se med drugim imenuje enota za absolutno temperaturo. Tako Niels Bohr kot Werner Heisenberg, dva izmed očetov kvantne mehanike, sta Boškoviču priznavala pomembno mesto v razvoju znanosti in atomistike še posebej. John Henry Poynting, angleški fizik, znan po svojem doprinosu k elektromagnetizmu, je zapisal, da je Boškovičevo premoščanje meje med silo in maso pomenilo trajen doprinos k filozofskim idejam (ibid.: 178). Po Supku naj bi Boškovič naredil podoben konceptualni korak kot kasneje Einstein s teorijo relativnosti, v kateri je uvedel svojo znano enačbo za mirovno energijo $E = mc^2$. S tem naj bi se Boškovičev program začel širiti na atom kot masno točko – začela se je pretvorba pojma mase v energijo, medtem ko je Boškovičev pojem mase »pretopil« v pojmu sil oziroma polja (ibid.: 179).

Basic Concepts of Boscovich's Atomism in Light of the Emerging Modern Science

The aim of this article is to give an account of Roger Joseph Boscovich's atomic theory, to show the influence and relevance of some of his most penetrating scientific achievements, and to put them into the context of the development of modern scientific thought and method. In the introduction, I briefly present Boscovich's life and work, then I provide a sketch of his conception of atomism, which I compare with the less elaborated conceptions of Newton and Leibniz whose theories he took as exemplifying achievements. Further, I present Boscovich's law of universal force, for which he is best known, and his attitude towards the laws of nature in general. I pay special attention to his original transformation of the Aristotelian concepts of the actual and the potential, which he applies to the concepts of the continuous and the discrete with regard to his conception of space and time, and which are the implicit cornerstones of his atomism. In the last section, I also discuss Boscovich's contribution to the scientific method, and mention certain long-lasting influences of his work.

Keywords: Roger Boscovich, atomism, laws of nature, space, potentiality, scientific method

Literatura

- Aristotel (2004). *Fizika: knjige 1, 2, 3, 4*. Ljubljana: Slovenska matica.
- Aristotel (1999). *Metafizika*. Ljubljana: Založba ZRC, ZRC SAZU.
- Bošković, R. J. (1996). *O zakonu neprekinutosti*. Zagreb: Školska knjiga. Hrv. prev.
- Talanga, J. (Izvirnik: Boscovich, R. J. (1754), *De continuitatis lege et ejus consecretariis pertinentibus ad prima materiae elementa eorumque vires*. Rim: Facolta di Roma.)
- Bošković, R. J. (1922). *A Theory of Natural Philosophy*. Chicago, London: Open Court Publishing Company. Ang. prev. Child, J. M. (Izvirnik: Boscovich, R. J. (1763), *Theoria philosophiae naturalis redacta ad unicam legem virium in natura existentium*. Venezia.)
- Child, J. M. (1922). »Introduction«. V Bošković, R. J., *A Theory of Natural Philosophy*, Chicago in London: Open Court Publishing Company, str. XI–XVIII.
- Eisenhart, C. (1961). »Boscovich and the Combination of Observations«. V Whyte, L. L. (ur.), *Roger Joseph Boscovich*, London: Allen & Unwin, str. 200–212.
- Leibniz, G. W. (2004a). *Principi filozofije ali monadologija*. V Leibniz, G. W., *Izbrani filozofski spisi*, Ljubljana: Slovenska matica, str. 131–150.
- Leibniz, G. W. (2004b). *Kontroverza med Leibnizom in Clarkom*. V Leibniz, G. W., *Izbrani filozofski spisi*, Ljubljana: Slovenska matica, str. 151–230.
- Jammer, M. (1954). *Concepts of space: the history of theories of space in physics*. Cambridge: Harvard University Press.
- Jammer, M. (1957). *Concepts of force: a study in the foundations of dynamics*. Cambridge: Harvard University Press.
- Johnson, M. in Wilson, C. (2007). »Lucretius and the History of Science«. V Gillespie S. in Hardie P. (ur.), *The Cambridge Companion to Lucretius*, Cambridge: Cambridge University Press, str. 131–148.
- Lukan, P. (2015). »Roger Boscovich and the quantum mechanical combination of dynamic and statistical laws«. *Almagest*, vol. 6(1), str. 65–78.
- Newton, I. (1952). *Opticks: or a treatise of the reflections, refractions, inflections & colours of light*. New York : Dover Publications.
- Newton, I. (1995). *Mathematical Principles of Natural Philosophy and System of the World* (ang. prev. Andrew Motte). Amherst, N.Y.: Prometheus Books.
- Stigler, J. (1986). *History of Statistics: The History of Uncertainty Before 1900*. Cambridge [etc.]: The Belknap Press of Harvard University Press.
- Supek, I. (1989). *Ruđer Bošković, vizionar u prijelomima filozofije, znanosti i društva*. Zagreb: Školska knjiga.

Nika Češnovar

Filozofska fakulteta Univerze v Ljubljani

Etični vidiki prekinitve aktivnega zdravljenja

Članek obravnava temo evtanazije ter prekinitve (opustitve/odtegnitve) aktivnega zdravljenja, predvsem s filozofsko-etičnega, deloma tudi pravnega vidika. Začnem s kratkim pregledom zgodovinskega razvoja pojmovanja evtanazije. Osrednji del zajema razlago nekaterih pomembnih pojmov: podrobnejšo analizo samega pojma ter vrst evtanazije. Nadaljujem z razlago prekinitve aktivnega zdravljenja (PAZ), kamor spada opustitev (najpogostejša oblika opustitve je izraz volje o ne-reanimaciji) ter odtegnitev aktivnega zdravljenja (najpogostejše gre za vprašanje glede odtegnitve hidracije in hranjenja). Pomemben vidik oskrbe in nege ob koncu življenja predstavlja tudi paliativna oskrba. Predstavim še nekatere pravne argumente glede omenjene tematike – predvsem, kako je vprašanje evtanazije in prekinitve aktivnega zdravljenja urejeno v Sloveniji –, sklenem pa z analizo etičnih priporočil za odločanje ob koncu življenja.

Ključne besede: bioetika, prekinitve aktivnega zdravljenja, evtanazija, oskrba ob koncu življenja, zakonodaja s področja zdravljenja

1. Uvod

V sodobnem svetu, svetu, vpetem v vsakodnevni razvoj in napredek na tehnoloških področjih, so čedalje bolj prisotna vprašanja in dileme glede etične sprejemljivosti dejanj, postopkov in politik na teh področjih. S prvimi velikimi misleci ter zdravilci pride tudi do etičnih vprašanj, povezanih z oskrbo bolnika, ter vprašanj, kaj je za bolnika dobro in kaj ni. In to temeljno vprašanje zdravilcev, vprašanje glede oskrbe bolnika, se skozi različna zgodovinska obdobja razvija, o tem se teoretizira ter ustvarja nova in nova stališča. Razvoj tehnologij je prav tako botroval razvoju novih, izpopolnjenih in v vseh pogledih boljših tehnik in načinov zdravljenja bolnikov, pa vendar tudi danes ni mogoče pomagati vsem. Bolniki, katerih smrt je neizogiben »končni udarec« njihove bolezni, nemalokrat trpijo bolečine, tako fizične kot psihične – sploh v zadnjih stadijih, v zadnjih dneh svojega življenja. Kako jim pomagati? Kako jim olajšati bolečine? Kako jim omogočiti čim mirnejšo smrt? Vse to so vprašanja, s katerimi se zdravstveni delavci srečujejo

vsak dan. Na eni strani imajo obvezo do bolnika – omogočiti mu čim več lagodja, čim manj boleče in čim mirnejše zadnje dni življenja, na drugi strani pa morajo slediti zakonom in pravilom, ki jih nemalokdaj omejujejo, predvsem zaradi svoje togosti in neprilagodljivosti.

Osnovne filozofske razmisleke o evtanaziji lahko umestimo že v antiko – takrat je pomenila *dobro smrt*. Bodisi dobro smrt v smislu hitre, častne smrti (npr. želja vsakega vojaka), ali pa dobra smrt v smislu pravičnosti (takšen pogled na dobro smrt so imeli npr. modreci). Nekoliko drugačno stališče in razmišljanje je bilo razširjeno med zdravniki, filozofi in ostalimi misleci; že takrat so se namreč porajala vprašanja in dileme glede filozofskih ter etičnih vidikov (prenehanja) aktivnega zdravljenja pri osebah, ki so bile neozdravljivo bolne. Tako je npr. Hipokrat svetoval: »ne zdraviti bolnikov, ki jih je bolezen premagala, zavedajoč se nemoči medicine v takih primerih« (Grosek et al., 2013: 590). Ta razmislek se nato nadgrajuje in dopolnjuje, pri čemer se je pojmovanje in razumevanje pojma evtanazije in z njo povezanih tematik spreminjalo. V dvajsetem stoletju pride do hitrega napredka na področju medicinskih tehnologij, med drugim izuma naprav, ki omogočajo mehansko predihavanje in podporo srca, defibrilacijo srca in številna zdravila, ki lajšajo pacientove bolečine. Z vsem tem pa hkrati pride do še več etičnih dilem ter moralnih vprašanj glede oskrbe bolnikov. Prav tako je to obdobje, ko se, skladno s tehnološkim razvojem, začnejo pojavljati globlje dileme glede smiselnosti/nemiselnosti nadaljnjega zdravljenja. Pojavijo se vprašanja glede prekinitve aktivnega zdravljenja (v nadaljevanju PAZ); bodisi dileme glede dodajanja novih terapij, ali zgolj odvzema kakšne že uvedene terapije. Razpravo okoli teh vprašanj so oblikovali tudi posamezni primeri, ki so postali središče bioetičnih in pravnih razprav.

Primer Karen Ann Quinlan

Je eden ključnih za razpravo o prekinitvi zdravljenja in vprašanj glede tega, kaj razumemo kot običajna in kaj neobičajna sredstva zdravljenja. Karen Ann je bila v trajnem vegetativnem stanju po zaužitju prevelikega odmerka drog. Njeno stanje je bilo označeno kot »trajna koma«, pri življenju pa jo je ohranjal respirator, ter umetno hranjenje in hidracija. Ta primer je botroval temu, da se je v okviru PAZ vzpostavilo razliko med »navadnimi« in »nenavadnimi« sredstvi za vzdrževanje življenja, »s tem [pa] je bila tudi postavljena jasna ločnica med namerno usmrtnitvijo bolnika na eni strani ter smrtjo po PAZ na drugi strani« (Grosek et al., 2013: 591).

Pomemben vidik omenjene problematike je predstavljen tudi v smernicah bioetičnega raziskovalnega inštituta, Centra Hastings. Glede odločitev o oskrbi, terapijah in negi ob koncu življenja zavzamejo stališče, da v principu nobena terapija ali oblika nege na smrt bolnega sama po sebi ni »navadna« ali »nenavadna«. Šele

bolnikov¹ pogled na določeno terapijo pravzaprav določi stališče do »navadnih« in »nenavadnih« oblik terapij, ne glede na kompleksnost in tehnično sofisticiranost določenega posega/terapije (od najpreprostejših in rutinskih posegov do najkompleksnejših).

Sam pomen in ločevanje med »navadnimi« in »nenavadnimi« oblikami pomoči na smrt bolnim pa se je, ne glede na vse, skozi leta prakse spreminjal. Tako je lahko razumljen kot razlika med moralno priporočljivimi ter moralno opcijskimi posegi, razlika med standardno in nestandardno nego, ali celo kot razlika med preprosto in kompleksno terapijo. Zaradi zgoraj naštetih različnih pojmovanj, katere terapije spadajo pod »navadne« ali »nenavadne«, je velikokrat prišlo do nerazumevanja ter zmede pri odločanju o le-teh. Danes se zato raje uporabljata izraza »prednost«, ki se nanaša na terapije, ki imajo korist za bolnika, in »breme«, ki se nanaša na terapije, ki bi bolniku potencialno lahko povzročile večje breme kot korist. V vsakem primeru pa ima pacient, oz. njegov zakoniti zastopnik, pravico odločanja o tem, katere terapije dovoljuje in katerih ne, ne glede na zahtevnost posameznih zdravljenj (Berlinger et al., 2013: 4).

Primer Nancy Cruzan

Tudi naslednji primer je močno vplival na razvoj smernic glede PAZ, posebej z vidika razumevanja etičnega statusa bolnikov v vegetativnem stanju v navezavi s prekinitvijo hranjenja in hidracije. Nancyjini starši so po osmih letih nespremenjenega stanja hčerke v ireverzibilni komi želeli odstraniti prehranjevalno sondo, a se zdravniško osebje bolnišnice s tem ni strinjalo. Tudi ta primer je zato prišel na sodišče, kjer pa je bila prošnja zakoncev Cruzan zavrnjena s pojasnilom, »da se sama Nancy Cruzan ne more odločiti in zavrniti postopka za ohranjanje življenja, in ker je ohranjanje življenja državni interes, bi sodišče lahko izdalo dovoljenje za prekinitve le, če bi obstajal jasen in prepričljiv dokaz, da bi to želela Nancy« (Singer, 2004: 79). Tako kot Karenini starši iz zgornjega primera, so se tudi Nancyjini starši odločili obrniti na vrhovno sodišče z argumentom, da ima njuna hči »ustavno pravico do smrti«. V razsodbi je bilo ugotovljeno, da ustava ZDA sicer dopušča tovrstno dejanje, vendar pa se lahko vsaka država ZDA odloči po svoje in zavrne takšno prošnjo, če ni nobenega dokaza o tem, da si je/bi si pacient resnično želel tako skrajne odtegnitve aktivnega zdravljenja. Torej, ustava ZDA »dopušča prekinitve hranjenja in hidracije (in ju [je] s tem izenačilo z ostalimi načini zdravljenja, ki jih je mogoče v določenih primerih prekiniti). Pogoji je, da to zahteva oseba z ohraneno zmožnostjo odločanja oz. da obstajajo priče, ki to lahko nedvoumno potrdijo« (Grosek et al., 2013: 591).

¹ V primeru, da bolnik zaradi kakršnegakoli razloga ni sposoben sam odločati o sebi in svojih željah, to vlogo prevzamejo njegovi bližnji oz., izjemoma, tudi bolnikov zdravnik.

Po še nekaj podobnih primerih v zadnjih dvajsetih letih, ki so se znašli na tem istem vrhovnem sodišču, je to odločilo še naslednje: »bolnik [ki] zavrne zdravljenje, ki ga vzdržuje pri življenju, umre od osnovne bolezni; ko pa prejme smrtni odmerek zdravila od zdravnika, pa je usmrčen od tega zdravnika« (Grosek et al., 2013: 591). Iz zgornjega zapisa je tako razvidno, da obstaja jasna razlika med »evtanazijo«, torej aktivnim končanjem bolnikovega življenja, ter med odtegnitvijo aktivnega zdravljenja, ki ima lahko za nenačrtovano posledico tudi smrt; v tem primeru »bolniku pustimo umreti«.

2. Evtanazija

Pri evtanaziji »gre za olajšanje trpljenja ali smrtnega boja umirajočega ali neozdravljivo bolnega na tak način, da umirajoči oziroma neozdravljivo bolni umre« (Bošnjak, 1996: 7). Torej gre za »umetno izzvano smrt pri neozdravljivo bolnih, namerno skrajšanje življenja, da bi bolnika razbremenili bolečin in trpljenja« (Dolenc, Medicinska etika in deontologija: 419). Urbančič o evtanaziji zapiše: »evtanazija pomeni zavestno povzročeno smrt po drugi osebi na izrecno želje osebe, ki umre. Ravno ta želja oz. soglasje loči evtanazijo od umora« (Urbančič, 2006: 206).

Ugotavljam, da neka univerzalna, vsesplošna definicija evtanazije ne obstaja, niti ni mogoča, zato se tudi pri razlikovanjih, morda celo načinih evtanazije, avtorji ne morejo poenotiti. V deklaraciji o evtanaziji iz leta 1987 je zapisano: »dejanje, s katerim se namenoma konča bolnikovo življenje, ne glede na bolnikovo lastno prošnjo ali na prošnjo bližnjih sorodnikov, je neetično« (Dolenc, 1993: 110). Lucas zagovarja stališče, da »evtanazija ni nikoli dopustna, tudi takrat ne, ko jo naredimo iz usmiljenja in na željo bolnika. Gre za uboj človeka, za kršitev načela varovanja življenja. /.../ Gre za kršitev človekovega dostojanstva, zločin proti življenju, za napad na človeka« (Lucas, 2005: 159–160).

Pri aktivni evtanaziji zdravnik z nekim izrecnim dejanjem, torej aktivno ter s tem seveda namerno, usmrti bolnika. Ponavadi gre za smrtonosno injekcijo, npr. injekcijo kalijevega klorida (Rachels, 1987: 116), ali dajanje prevelikega odmerka analgetikov in sedativov v obliki tablet. Bolnik po uvedbi tovrstne terapije umre takoj oz. zelo kmalu po prejemu le-te. Pri tej obliki evtanazije tako zdravnik igra glavno vlogo, saj je pravzaprav on tisti, ki pacienta »odreši muk« – v tem primeru zdravnik izvrši uboj pacienta. Torej je aktivna evtanazija »tista, pri kateri se povzroči smrt s storitvijo nekega dejanja, ki meri na povzročitev smrti na željo tistega, ki želi umreti« (Urbančič, 2006: 206). Aktivna evtanazija je, globalno gledano, z zakoni (za enkrat še) prepovedana, zdravnik, ki jo vseeno izvede, velja za

morilca. So pa nekatere države že spremenile svoje zakone ter tako v določenih pogojih dovoljujejo izvedbo tudi aktivne evtanazije.²

Pasivna evtanazija je drugačna od aktivne. Tu pravzaprav ne gre za uboj bolnika, ampak gre za prekinitev aktivnega zdravljenja. Bodisi za opustitev bodisi za odtegnitev aktivnega zdravljenja ali pa za obe dejanji hkrati. »Pomeni, da se preprosto odrečemo vsaki dejavnosti, ki bi bolnika obdržala pri življenju. /.../ zadržimo zdravljenje ali druge življenje ohranjujoče terapije oz. zavrnemo operacijo in tako naprej ter pustimo bolnika umreti 'naravno', za posledicami bolezni, ki jih ima« (Rachels, 1987: 116). Torej je »pasivna evtanazija tista, pri kateri pospešimo smrt z opustitvijo določenih dejanj, npr. izključitvijo/spremembo režima delovanja kakšnega aparata ali zdravila, ki vzdržuje življenjske funkcije pri bolniku, in ga prepustimo naravnemu toku stvari. Primeri za to so izključitev respiratorja, prenehanje dajanja vitalno pomembnih *medikamentov*, ipd.« (Urbančič, 2006: 207). Ta vrsta evtanazije je med zdravniki bolj sprejemljiva. Prav tako je tudi javnost bolj naklonjena tovrstnim ukrepom, torej opustitvi in odtegnitvi zdravljenja, kot pa »direktnemu uboju« bolnika. Še več, »opustitev ali prekinitev aktivnega zdravljenja je v mnogih okoliščinah splošno sprejeta kot moralno pravična« (Hope, 2014: 23). In če povzamemo osnovno pojmovanje aktivne – pasivne evtanazije, ugotovimo, da gre na teoretični ravni predvsem za razlikovanje med »storiti« in »dopustiti« (*doing in allowing*) ter med storitvijo in opustitvijo (*act of commission in act of omission*).

Poglejmo še pojme voluntarne, nevoluntarne ter prisilne evtanazije. Voluntarna oz. prostovoljna evtanazija predstavlja bolnikovo željo po smrti. Bolnik sam je pobudnik in izraža željo po prekinitvi lastnega življenja, pri čemer je popolnoma priseben in razsoden. Prostovoljna evtanazija je tako pomoč pri umiranju osebe, ki trpi zaradi zadnjih stadijev terminalnih bolezenskih stanj. Preden zdravnik izvede evtanazijo – zadnje dejanje je namreč tu v rokah zdravnika, pacient pa zgolj poda soglasje za izvedbo postopka evtanazije, mora biti zadoščeno naslednjim petim pogojem (Young, 2015):

- bolnik trpi za posledicami neozdravljive, terminalne bolezni;
- v času, ki je bolniku še ostal, so možnosti za odkritje zdravila za izboljšanje njegovega zdravstvenega stanja zanemarljivo majhne;
- bolnik trpi neznosne bolečine zaradi bolezni oz. je popolnoma odvisen od pomoči drugih ali od tehnične pomoči za ohranjanje življenja;
- bolnik je zmožen pri polni zavesti izraziti željo po smrti (oz. je to željo jasno izrazil že v preteklosti za primer, če bi prišlo do določenih »pogojev«);
- bolnik je nezmožen končati svoje lastno življenje brez pomoči.

² Na primer, Albanija (1999), Belgija in Nizozemska (2002), Luksemburg (2009); v Švici je dovoljena pomoč pri samomoru.

Nevoluntarno oz. neprostovoljno evtanazijo je moč še naprej razčleniti na dve obliki evtanazije in sicer na neprostovoljno (*nonvoluntary*) in na prisilno (*involuntary*) evtanazijo.

Neprostovoljna evtanazija predstavlja uboj človeka, ki ni sposoben odločati sam zase, tu gre predvsem za bolnike v trajnem (ireverzibilnem) vegetativnem stanju; pri prisilni evtanaziji pa gre za uboj osebe, ki je popolnoma prisebna, se pravi, da lahko aktualno odloča o sebi – torej pomeni usmrtitev (bolnika), dejanje proti njegovi volji.

3. Prekinitev aktivnega zdravljenja (PAZ)

»Pri odločanju o PAZ gre za tehten premislek med pričakovano učinkovitostjo in koristni zdravljenja ter bremenom, ki ga le-to predstavlja brezupno bolnemu« (Oražem in Kanič, 2012: 13). Pri neozdravljivo bolnih osebah nemalokdaj pride do trenutka, ko bi nadaljevanje uvedenega zdravljenja ali celo dodajanje novih terapij pravzaprav predstavljajo večje breme za bolnika kot pa bi lahko npr. nadaljevanje s terapijami slednjemu koristilo. Gledano z etičnega vidika je zato popolnoma dopustno, celo priporočljivo razmisliti o PAZ. Vendar pa ima odločitve o prenehanju aktivnega zdravljenja »pogosto za neposredno posledico smrt brezupno bolnega« (Grosek et al., 2013: 590). Tako ugotavljam, da je PAZ nekako etično sprejemljiv, vendar zgolj v primerih, ko je zdravljenje brezupno. Tega pa največkrat opredelimo s pomočjo naslednjih štirih točk (Grosek et al., 2013: 592):

- nesmiselno zdravljenje (*pointless/not resanoble*);
- cilji zdravljenja niso dosegljivi (*no goal of care attainable*);
- breme zdravljenja je večje od koristi, ki ga prinaša;
- brezupnost zdravljenja bi se morala definirati celostno, za vsakega bolnika posebej.

Tudi smernice Centra Hastings glede odločitev o oskrbi, terapijah in negi ob koncu življenja podpirajo dejanja PAZ. Hkrati pa opozarjajo, da tovrstne odločitve morajo stati na etično trdnih tleh, da postopek odločanja zajema sodelovanje in soodločanje tako na smrt bolnega, njegovih svojcev, kot tudi izvajalcev zdravstvene terapije in nege tega bolnika. Pri vsem tem pa je potrebno upoštevati več vidikov, recimo religijo, etičnost dejanja, smiselnost/nesmiselnost nadaljevanja ali prenehanja določene terapije glede na pretekle terapije in potek zdravljenja (Berlinger et al., 2013: 1).

Opustitev zdravljenja je dejanje, ko se pri bolniku ne uvaja dodatnega ali ne intenzivira obstoječega zdravljenja, prav tako se ne začneja s posegi za ohranitev življenja (kot so npr. diagnostični postopki, hemofiltracija/hemodializa, dajanje antibiotikov, vazoaktivnih zdravil, krvnih nadomestkov, umetno predihavanje, ki-

rusško zdravljenje itd.). Iz dejanja opustitve so izvzeti analgetiki, ki služijo kontroliranju in znosnosti bolečin(e).

DNR oz. naročilo »ne oživljaj« velja za eno izmed najpogostejših in hkrati tudi najbolj sprejetih oblik opustitve (aktivnega) zdravljenja. Nekateri teoretiki sicer DNR uvrščajo k odtegnitvi aktivnega zdravljenja, vendar jih večina DNR še vedno uvršča med opustitvena dejanja zdravljenja. Že z razvojem metod oživljanja in električnega defibriliranja srca so se začela porajati vprašanja o tem, ali je takšno oživljanje res vedno primerno. Seveda je število preživelih, npr. pri srčnih zastojih, z vpeljavo defibrilatorja v klinično prakso močno naraslo, vendar se samo vprašanje o primernosti uporabe tovrstne metode ni nanašalo na omenjene bolnike. Vprašanja in dileme so se porajale predvsem pri terminalno bolnih ljudeh v zadnjih stadijih, ki bi jim, denimo, oživljanje zgolj podaljševalo predsmrtno agonijo in ki prav gotovo ne bi več imeli koristi od takšne skrajne terapije. Njihovo življenje bi se podaljšalo morda za minute, ure, dneve, kakovost teh zadnjih dni pa zaradi oživljanja zagotovo ne bi bila nič boljša. Prav nasprotno, kakovost življenja bi se lahko še občutno poslabšala. Danes se tem vprašanjem posvečajo bolnišnične etične komisije, seveda pa ima tudi bolnik možnost soodločanja o tem, kdaj, torej v katerih primerih ter do katere mere želi nadaljevati z vsemi možnimi sredstvi za ohranjanje življenja ter kdaj ne želi več prejemati določenih vrst (aktivne) terapije. Temu danes pravimo vnaprejšnja volja bolnika, saj je s tem posamezniku omogočeno, da se sam odloči za potek zdravljenja, četudi mu je med napredovanjem bolezni odvzeta zmožnost samoodločanja. Tako je v zakonu o pacientovih pravicah (ZPacP) zapisano, da ima posameznik pravico do samostojnega odločanja o zdravljenju ter pravico do upoštevanja njegove vnaprej izražene volje. Tu velja morda samo še omeniti dejstvo, da je pri nas »v praksi uporaba instituta vnaprejšnje volje zelo redka in zakonsko tudi precej omejena, zato je pomanjkljiva tudi praksa na tem področju« (Grosek et al., 2013: 594).

Pri odtegnitvi zdravljenja, za razliko od opustitve, se sprejme odločitev, da se ukinejo zdravljenja, ki jih bolnik že prejema, torej gre za aktivno prenehanje terapij (npr. prenehanje hemofiltracije/hemodialize, umetnega predihavanja, dajanja antibiotikov, vazoaktivnih zdravil, krvnih nadomestkov itd.). Tu gre v večini primerov za odklop umetnega predihavanja bolnikov v nereverzibilni komi, prekinitve mehanskega prečiščevanja krvi itd., v nekaterih primerih pa tudi za prenehanje z umetnim hranjenjem in hidracijo, vendar pa je slednje precej pereča tema tako z zdravstvenega kot tudi z etičnega vidika.

Umetno hranjenje in hidracija neozdravljivo bolnega sta precejšnji etični in moralni vprašanja v navezavi z oskrbo bolnika ob koncu življenja. Nekateri so tako mnenja, da ob morebitni prekinitvi hranjenja in hidracije kršimo osnovne človekove pravice, saj bi s tem povzročili smrt zaradi stradanja in dehidracije, ki nikakor ne more biti »lepa in mirna«. Spet drugi pa se sklicujejo na opravljene študije, ki se nagibajo k priznavanju teh dveh skrajnih odtegnitev, saj naj bi človeku v

terminalni fazi bolezni ravno odtegnitev hidracije in hranjenja omogočilo »lažjo in mirnejšo« smrt (Oražem in Kanič, 2012: 16; Grosek et al., 2013: 595).

Vprašanj odtegnitve umetnega hranjenja in hidracije se dotaknejo tudi avtorji smernic Sveta Evrope, in sicer zavzamejo stališče, podobno zgornjemu. V nekaterih državah sta umetno hranjenje in hidracija obliki aktivnega zdravljenja, zato se ju, v izjemnih okoliščinah, lahko na željo pacienta ukine, če je ta priseben in sposoben samoodločanja (drugače lahko to vlogo praviloma prevzamejo bolnikovi bližnji svojci). Lahko se za dejanje opustitve umetnega hranjenja in hidracije odloči tudi skupina zdravstvenih delavcev vpletenih v dotični primer, pri čemer pa mora biti odločitev sprejeta v okviru kolektivnega postopka. Na drugi strani pa imamo države, ki verjamejo, da umetno hranjenje in hidracija ne predstavljata (aktivnega) zdravljenja, ki ga je moč omejiti ali celo popolnoma odtujiti, saj sta hranjenje in hidracija osnovni človekovi potrebi. Vendar tudi države s takšnim nazorom dovoljujejo bolnikovo samoodločanje, kar pomeni, da je v končni fazi bolezni posameznika lahko ugodeno tudi takšni skrajni želji (Council of Europe, 2014: 12–13).

Kakorkoli že, slovenska strokovna javnost ima o tej problematiki dve popolnoma ločeni mnenji, vendar pa Komisija Republike Slovenije za medicinsko etiko (KME RS) za enkrat še priznava etično sprejemljivost ukinitve hidriranja in umetnega hranjenja (predvsem pri bolnikih v trajnem, ireverzibilnem vegetativnem stanju).

3.1 Oskrba ob koncu življenja

Med zdravniki in etiki, ki se ukvarjajo s problematiko evtanazije na sploh, velja prepričanje, da ne obstaja nikakršna etična razlika med opustitvijo ter odtegnitvijo aktivnega zdravljenja, da sta obe vrsti omenjene oskrbe na smrt bolnega etično enako sprejemljivi. Vsak pacient ima tako pravico zavrniti kakršnokoli nezaželeno obliko terapije, bodisi ukiniti že uvedeno terapijo ali zavrniti novo. Se pa kaže, verjetno predvsem psihološka, razlika med odtegnitvijo in opustitvijo zdravljenja, zdravstveni delavci se namreč večkrat in lažje odločajo za slednjo (Berlinger et al., 2013: 4). Potrebno se je zavedati, da ob morebitni prekinitvi aktivnega zdravljenja nikakor ni, in hkrati ne sme biti, prekinjena tudi skrb za bolnika in njegova zdravstvena nega. Pri vsakem bolniku, za katerega se odločijo, da je določena vrsta prekinitve zdravljenja zanj najbolj primerna, je potrebno najprej narediti celosten načrt oskrbe bolnika ob koncu življenja in mu s tem zagotoviti kakovostno paliativno oskrbo.

3.2 Paliativna oskrba

»Paliativna (blažilna) oskrba preprečuje, odpravlja in blaži tegobe neozdravljive bolezni. Zagotavlja nepretrgano, aktivno oskrbo do smrti« (Onkološki vikend 18, 2005: 9). Paliativna oskrba je pravzaprav neke vrste multidimenzionalen, celosten

pristop k negi neozdravljivo bolnega. Vloga paliativne oskrbe ni zgolj lajšanje fizičnih simptomov in neugodja, ki so posledica bolezni, temveč bolniku nuditi tudi psihično oporo, da se lažje spopada s položajem, v katerem se je znašel zaradi napredujoče bolezni. Glavni cilj tovrstne oskrbe pa je spoštovanje človekove avtonomije ter njegovih pravic. Kot že samo ime pove, je paliativna oz. blažilna oskrba v prvi vrsti namenjena lajšanju zadnjih dni življenja, predvsem gre ponavadi za lajšanje neznosnih bolečin, saj prav »nelajšana« bolečina lahko privede do povečanega neugodja bolnika, ker se počuti zapuščenega. »Kontroliranje in zmanjševanje bolečine je pacientova pravica in ne privilegij« (Markwell, 2005: 1132).

3.3 Terapija z dvojnim učinkom

Načelo dvojnega učinka uvrščamo med enega izmed temeljnih načel bioetike, ki pravi, da ima namerno dejanje »lahko poleg neposredno želenega učinka tudi posreden, stranski učinek, ki je lahko povsem nehotena (ne kot cilj, ne kot sredstvo) posledica, vendar pa jo dopuščamo, ker je neizogibno vezana na to, kar želimo doseči« (Lucas, 2005: 27). Avtor v nadaljevanju poda preprost primer za lažje razumevanje samega koncepta tega načela. Neka oseba se zdravi zaradi levkemije, terapija, katere neposredni učinek je zdravljenje bolezenskega stanja, pa hkrati povzroča tudi izgubo las – plešavost je v tem primeru posledica (posredni učinek) zdravljenja (Lucas, 2005: 27).

Zaradi obsežnosti teorije načela dvojnega učinka se bom v tem članku dotaknila zgolj načela dvojnega učinka z medicinske plati, natančneje, terapije z dvojnim učinkom, s katero se praviloma srečujejo na smrt bolni v zadnjih dneh življenja. Terapija oziroma zdravljenje z dvojnim učinkom je pravzaprav neke vrste paliativna oskrba bolnika. Gre za dejanje, ki bo imelo tako dobre, te so nameravane, kot tudi slabe, nenameravane posledice. Takšno dvostransko dejanje je, kot pravi Markwell (2005: 1134), moralno dovoljeno zgolj v primerih, ko:

- »dejanje samo po sebi ni amoralno;
- je dejanje izvršeno samo zaradi želje po doseganju pozitivnih plati dejanja in nikakor ne zaradi morebitnih slabih učinkov, čeprav vemo, da so le-ti možni;
- so posledice dejanja lahko vsaj toliko pozitivne kot negativne;
- je dobro (kot posledica koristnega učinka) res dovolj dobro, da odtehta slabe (nenamerne) učinke.«

V primeru paliativne oskrbe se načelo dvojnega učinka pojavlja predvsem v primerih paliativne sedacije³ ter povečevanja odmerka protibolečinskih zdravil, »pri katerih je osnovni namen odpravljanje bolečine, medtem ko bi se morebitno skrajšanje življenja smatralo kot slaba in nenamerna posledica« (Grosek et al., 2013: 595). V smernicah Centra Hastings o oskrbi, terapijah in negi ob koncu živ-

³ »Paliativna sedacija je premišljeno povzročeno in ohranjeno stanje globokega spanja /.../ V tem stanju oseba ne trpi zaradi težav, ki jih ne moremo več obvladovati« (Onkološki inštitut Ljubljana, 2008: 40).

ljenja je glede terapije z dvojnimi učinkom zapisano še, da je pravzaprav smrt, povzročena zaradi dodajanja protibolečinskih zdravil v praksi precej redka. Ob pravilnem dovajanju opiatov, ki služijo za lajšanje dispneje ali bolečine, ki jo povzroča rak, je sicer možen pojav respiratorne motnje, vendar te, le izjemoma, dejansko vodijo tudi v popolno respiratorno odpoved in posledično smrt (Berlinger et al., 2013: 5).

4. Etična priporočila za odločanje o zdravljenju in paliativni oskrbi bolnika ob koncu življenja v intenzivni medicini

Ta priporočila predstavljajo nekakšen zbir tako etičnih kot tudi pravnih dolžnosti vseh zdravstvenih delavcev povezanih z vidikom odločanja. Dokument temelji na štirih temeljnih načelih.

Načelo spoštovanja avtonomije

To načelo temelji na ustavi RS (34. člen, 35. člen in 51. člen), na ZpacP (17. člen, 32. člen, 33. člen). Svobodo pri odločanju o zdravljenju (posledično tudi o prekinitvi le-tega) zagovarja tudi Evropska konvencija o človekovih pravicah in Oviedska konvencija (Grosek et al. (ur.), 2015: 23–24).⁴ Načelo spoštovanja avtonomije ne smemo razumeti kot absolutno avtonomijo pacienta, gre zgolj za pravico bolnika, da privoli v ali zavrne predlagan ukrep ali predlagano medicinsko oskrbo. Hkrati je pomembno, da zdravnik bolnika pred takšnimi odločitvami seznani z vsemi informacijami, le-te pa morajo biti podane tako, da so bolniku popolnoma in v celoti razumljive. Pojasnila morajo tako vsebovati vse ključne informacije; predvsem pa o bolezni, ki jo bolnik ima, o možnih razpletih in razvoju bolezni, o poteku in možnih zapletih oz. koristih določenega predlaganega posega ali terapije. Tu so pomembne tudi informacije o morebitnih posledicah zavrnitve določenega zdravljenja, posega ali terapije. Bolnik ima pravico biti informiran tudi o drugih možnostih zdravljenja; seveda pa mora pacient prejeti informacije tudi o tem, kateri ukrepi, ki bi si jih morda želel, v Sloveniji niso na voljo (prav tam: 24–25). Ne glede na vse, pa je zdravnik vseeno dolžan vsaj poskusiti prepričati pacienta, da dovoli posege, ki bi bili zanj najboljši, mu stvari in situacijo ponovno, morda na drugačen način spet predstaviti, svetovati, ali mu vsaj ponuditi možnost drugega mnenja – skratka, zdravnik mora storiti vse, kar je v njegovi moči, da pacientu predstavi vse vidike bolezni, terapij in možnosti, saj se le na tak način lahko pacient resnično odloči, kaj si dejansko želi, s katerimi posegi soglaša in česa si v nobenem primeru ne želi in ne dovoli. Stvar je malo drugačna pri bolnikih, ki so zaradi kakršnegakoli razloga izgubili zmožnost samoodločanja. Tu

⁴ Tu povzemam skupno izjavo Slovenskega združenja za intenzivno medicino in KME RS z naslovom *Etična priporočila za odločanje o zdravljenju in paliativni oskrbi bolnika ob koncu življenja v intenzivni medicini*.

nastopi pomembnost pacientove vnaprejšnje privolitve, če pa privolitev ni znana, namesto njega lahko odloča najožji družinski član, ki ima pravico »govoriti« v imenu bolnika, skupaj z bolnikovim zdravnikom in zdravniškim timom. V tem primeru pa mora biti bolnikova korist vedno na prvem mestu (prav tam: 25–27).

Načelo neškodljivosti in dobrodelnosti oziroma koristnosti.

Zdravnik bolniku v nobenem primeru ne sme namerno povzročati škode. Zato je pri vsakem zdravljenju potrebno primerjati razmerje med koristmi posamezne terapije in škodo, ki jo ta terapija povzroča. Zaradi tega je v določenih primerih moralno in etično dopustno prekiniti aktivno zdravljenje, bodisi gre za dejanje odtegnitve bodisi dejanje opustitve, če je to v bolnikovem najboljšem interesu. Nikakor pa ni dopustno prekiniti z oskrbo ob koncu življenja, torej s paliativno oskrbo. »Pravica do paliativne oskrbe je absolutna; ni je dopustno niti odreči niti kakor koli omejevati, in kot taka sodi v notranje jedro ustavne pravice do zdravstvenega varstva« (Grosek et al. (ur.), 2015: 28). Tudi ta pravica je namreč »zavarovana« v ustavi RS (s členi 17., 18., 34., 35. in 51.).

Načelo pravičnosti

To načelo temelji na enakih pravicah vseh, v našem primeru bolnikov. »Formalno bistvo pravičnosti terja, da so primerljivi bolniki v primerljivih okoliščinah obravnavani na enak ali čim bolj podoben način« (Grosek et al. (ur.), 2015: 29). Jasno je namreč, da se z razvojem tehnologije razvijajo tudi vedno novi in novi, dražji načini zdravljenja bolnih, te metode pa niso vedno in pri vsakem bolniku učinkovitejše od preprostejših, bolj uveljavljenih (cenejših) oblik terapije. Torej, po načelu pravičnosti, tudi dobrodelnosti, se je smiselno vprašati, ali so te oblike (dražjega) zdravljenja resnično najprimernejše za dotičnega bolnika v primerih, ko le-temu prinašajo zgolj »marginalno« korist. Država je namreč dolžna svojim državljanom zagotoviti kvalitetno, za vse enako, pravično obravnavanje ter zdravljenje bolezni (50. in 51. člen Ustave RS). Javna sredstva za zdravstvene storitve pa so seveda omejena. Za upravičenost nudenja zdravstvene oskrbe iz teh sredstev morajo biti istočasno izpolnjeni naslednji trije pogoji (Grosek et al. (ur.), 2015: 29–30):

- »Zdravstvena storitev je potrebna po pravilih medicinske stroke
- Glede na sodobno medicinsko doktrino se upravičeno pričakuje, da bo bolniku koristila, in
- Pričakovane koristi posega so za bolnika večje od tveganj ter obremenitev (načelo sorazmernosti).«

Pri odločanju o ukrepih in terapijah ob koncu življenja je potrebno upoštevati še naslednje vidike, naslednja načela (prav tam: 30–31):

- Bolnik mora biti vedno v središču pozornosti, njegove pravice do primerne in kakovostne oskrbe se v nobenem primeru, v nobeni fazi bolezni ne sme zanemariti. Prav tako je za kvalitetno zdravljenje potreben celosten pristop do bolnega, to pa zajema tudi upoštevanje različnih družbenih (kulturnih, etničnih, družinskih) situacij in okoliščin;
- Bolniku mora biti omogočena tudi verska oskrba, v kolikor izrazi željo po tem;
- Odločitve za ukrepe o prenehanju (aktivnega) zdravljenja morajo biti natančno dokumentirane;
- Nujno je zagotoviti možnost skupnega odločanja (zdravniški tim in osebe oz. družinski član(i), ki imajo pravico odločati namesto bolnega), kadar se izkaže, da bolnik ni več sposoben jasno odločati o sebi in o terapijah ter načinih zdravljenja;
- V primeru, ko bolnik lahko sam zase odloča glede postopkov ob koncu življenja, mu mora biti zagotovljena odprta komunikacija z vsemi vpletenimi v omenjene postopke;

In – nenazadnje, a zato nič manj pomembno, prej ravno obratno – vsi zdravstveni delavci, ki so tako ali drugače vpleteni v odločanje in procese prenehanja aktivnega zdravljenja, se morajo (oz. bi se morali) stalno izobraževati in strokovno izpopolnjevati o tej problematiki, saj le tako lahko dejansko nudijo najboljšo terapijo/ najboljšo pomoč in podporo umirajočemu.

5. Zaključek

Mislim, da beseda evtanazija v današnjem času povzroča veliko več zgražanja, zavračanja in neodobravanja kot si ga morda v resnici »zasluži«. Glavni razlog za to, po mojem mnenju predstavlja napačno pojmovanje ter razumevanje tako obširnega pojma. Sam pojem evtanazije (omejila se bom zgolj na Slovenijo) ni dobro zasnovan, saj je njegov pomen preširok, etično in moralno sporno je pravzaprav samo dejanje aktivne evtanazije. Pasivna evtanazija je, tako rečeno, popolnoma sprejemljiva, še več – tudi v Sloveniji se tovrstna oblika evtanazije izvaja dnevno (enote intenzivne medicine, domovi za ostarele, hospici, neonatalni oddelki...). Ta vrsta evtanazije ni nezakonita, niti ni moralno ali etično sporna, saj predstavlja prekinitev aktivnega zdravljenja – PAZ. Pomeni, da se ne dodaja terapij, ki bi bile pacientu v večjo škodo kot korist in da se odtegne terapije, ki jih pacient morda že prejema, pa mu niso v korist oz. pacient ne kaže znakov izboljšanja, ali pa so možnosti za izboljšanje nične.

Obstaja torej znatna razlika med dejanjem aktivne evtanazije, ki je pri večini označena za neetično dejanje in se smatra kot uboj bolnika, ter med dejanjem pa-

sivne evtanazije, kamor spada PAZ. Dejanja slednjega pa so, kot že rečeno, tudi v Sloveniji etično povsem sprejemljiva in dopustna, celo zapovedana.

Delitev in razčlenitve tako obširnega pojma, kot evtanazija nedvomno je, je torej preskopa. Zato bi bilo potrebno popolnoma ločiti ter jasno razlikovati med dejanjem (aktivne) evtanazije ter med PAZ (torej opustitvijo in odtegnitvijo zdravljenja na smrt bolnega). Prva, aktivna evtanazija, bi bila tako lahko zakonsko še vedno prepovedana, vsaj dokler se v zdravstvu in pravu ne najde dovolj dobrih (skupnih) rešitev in zaščite pred zlorabo tovrstnega instituta. PAZ z vsemi svojimi »navadnimi« in »nenavadnimi« terapijami pa ne bi več predstavljal neke tanke meje med etično in moralno dopustnostjo in nedopustnostjo. Nenazadnje se zdravstveni delavci že sedaj dnevno srečujejo z dejanji pasivne evtanazije, pomislimo zgolj na domove za ostarele, intenzivne enote itd.

Tako pridem do zaključka, da je situacija v Sloveniji pravzaprav precej protislovna. Evtanazija je namreč z zakoni prepovedana, obenem pa je dejanje PAZ popolnoma etično in moralno dopustno. Kako je torej mogoče, da evtanazijo absolutno zavračamo, čeprav v isti sapi dodamo, da so metode PAZ poponoma dopustne, tako z etičnega kot tudi moralnega vidika?

V »Kodeksu medicinske deontologije Slovenije« je tako zapisano: »Zdravnik odklanja in obsoja evtanazijo in jo ima za lažni humanizem« (43. člen). Tudi stališče KME RS glede evtanazije je izrazito negativno, o čemer priča naslednji citat: »aktivni ukrepi da se bolnikovo življenje konča (evtanazija), pa so prepovedani po zakonu in po načelih medicinske deontologije tudi takrat, ko je bolnik tik pred smrtjo /.../ Zdravnikova pomoč pri samomoru je etično nesprejemljiva in po zakonu prepovedana« (Mlinar, 2005: 297). Prav tako isto stališče zavzema tudi Kazenski zakonik RS, v svojem 120. členu namreč predvideva zaporno kazen od šest mesecev do petih let, za napeljevanje ter pomoč pri samomoru (tudi v primerih, ko bi šlo za dejanje iz usmiljenja).

Na drugi strani pa imamo člene zakonov, za katere lahko trdimo, da zagovarjajo PAZ.

Na primer Ustava RS, ki v svojem 51. členu zagovarja pravico do zdravstvenega varstva, vendar pa je hkrati zapisano, da se nikogar ne sme prisiliti k zdravljenju, razen v primerih, določenih z zakonom. Tudi v zakonu o pacientovih pravicah (ZPacP) je zapisanih nekaj, za pričujočo tematiko pomembnih členov; npr. pacientova pravica do samostojnega odločanja o poteku zdravljenja in o terapijah, ki jih želi prejemati ter tistih, ki jih ne želi; pravica do vnaprej izražene volje, kar pomeni, da se pacient lahko že pred napredovanjem bolezni odloči, katere terapije in ukrepe želi, ter katerih ne. Vsak bolnik ima tudi pravico do preprečevanja in lajšanja trpljenja. O oskrbi in negi na smrt bolnega so smernice zapisane tudi v Kodeksu medicinske etike in deontologije in sicer, da je potrebno spoštovati pacientove potrebe in želje; omenjeno je tudi, da mora zdravnik odkloniti vsakršen poseg, ki bi utegnil biti neetičen, torej takšen, ki bi bil bolniku v večje breme kot

korist. Absolutno pa mora biti pacientu nudena primerna oskrba, ter nega, tudi v zadnjih dneh, urah in minutah njegovega življenja.

Menim, da so to dovolj jasna dejstva, ki kličejo po ponovnem razmisleku in o potrebi po novih, natančnejših in bolj dodelanih zakonih. Zakonih, v katerih bi bilo specifično prepovedano in hkrati družbeno moralno nedopustno zgolj dejanje aktivne evtanazije, PAZ pa bi bil določen s svojimi zakoni, v nikakršni navezavi s prepovedano (aktivno) evtanazijo.

Ethical aspects about active treatment discontinuation

This work deals with euthanasia and with active treatment discontinuation, mainly from philosophical-ethical, but partly also from juridical perspective. The first part presents historical development in understanding of euthanasia. The central part explains some main concepts: detailed analysis of euthanasia concept, including its branching. Then I explain the active treatment discontinuation, which includes letting die (with desire not to reanimate as its main form) and active healing refusal (the most common issue is the discontinuation of hydration and of nourishment). Another important aspect of the end of life care is palliative care, which I discuss next. This is followed by presentation of some juridical arguments concerning this topic; in particular, how the issues of euthanasia and active treatment discontinuation are regulated in Slovenia. I conclude with an analysis of the ethical recommendations for decision-making in end of life care.

Keywords: bioethics, active treatment discontinuation, euthanasia, end-of-life-care, laws regarding medical treatment

Literatura

Berlinger, N., Jennings, B. in Wolf, S. M. (2013). *The Hastings Center Guidelines for Decisions on Life-Sustaining Treatment and Care Near the End of Life* (revised and expanded second edition). Oxford: Oxford University Press.

Bošnjak, M. (1996). »Nekaj misli o problemih evtanazije«. *Obzornik zdravstvene nege*, 30(1/2), str. 7–11. Dostopno na: <http://www.dlib.si/?URN=URN:NBN:SI:doc-EK958H3V> [10. 4. 2015].

Council of Europe (2014). *Guide on the decision-making process regarding medical treatment in end-of-life situations*. Council of Europe.

Dolenc, A. (1993). *Medicinska etika in deontologija: dokumenti s komentarjem*. Ljubljana: Založba Tangram.

Grosek, Š., Grošel, U., Oražem, M. in Trontelj, J. (2013). »Etične dileme ob koncu življenja v intenzivni medicini«. *Zdravniški vestnik*, 82(9), str. 589–601.

Grosek, Š., Grošelj, U. in Oražem, M. (ur.) (2015). *Etična priporočila za odločanje o zdravljenju in paliativni oskrbi bolnika ob koncu življenja v intenzivni medicini: skupna izjava Slovenskega združenja za intenzivno medicino in Komisije RS za medicinsko etiko*. Ljubljana: Univerzitetni klinični center.

Hope, T. (2014). *Medicinska etika: zelo kratek uvod*. Ljubljana: Založba Krtina.

»Kazenski zakonik Republike Slovenije« (KZ-1) (2008/2009). *Uradni list RS*, št. 55/08 in 39/09.

Zdravniška zbornica Slovenije (1992). »Kodeks medicinske deontologije Slovenije«. (Spremembe kodeksa sprejete na 27. redni seji skupščine Zdravniške zbornice Slovenije 1997.) Dostopno na: <http://www.zdravniskazbornica.si/zzs.asp?FolderId=386> [13. 5. 2015].

Lucas, R. (2005). *Bioetika za vse*. Ljubljana: Družina.

Markwell, H. (2005). »End-of-life: a Catholic view«. *The Lancet*, 366(9491), str. 1132–1135.

Mlinar, A. (2005). *Evtanazija: Zgodovinski pregled, današnji položaj in etična refleksija*. Ljubljana: Študentska založba.

Onkološki inštitut Ljubljana (2008). *Pogosta vprašanja v paliativni oskrbi*. Ljubljana: 2008.

Oražem M. in Kanič M. (2012). *Etična načela pri odločanju o prenehanju aktivnega zdravljenja pri brezupno bolnih v intenzivnih enotah*. Ljubljana: Medicinska fakulteta Univerze v Ljubljani.

Onkološki vikend 18 (2005). *Paliativna oskrba bolnikov z rakom: zbornik / 18. onkološki vikend, Laško, 10. in 11. junij 2005*. Ljubljana: Kancerološko združenje Slovenskega zdravniškega društva, Onkološki inštitut Ljubljana in Zveza slovenskih društev za boj proti raku.

Rachels, J. (1987). *Pravica do smrti: evtanazija in morala*. Ljubljana: Cankarjeva založba.

Singer, P. (2004). *Razmislimo znova o življenju in smrti: sesutje naše tradicionalne etike*. Ljubljana: Studia Humanitatis.

Urbančič, V. (2006). »Pogled pravnika na evtanazijo«. *Medicinski mesečnik*, 2(6-7), str. 206–213. Dostopno na: http://www.medicinski-mesecnik.com/MM_06_06-07/MM_06_06-07_pogled-pravnika-evtanazija.pdf [11. 9. 2015].

»Ustava Republike Slovenije« (1991). *Uradni list RS*, št. 331/1991-I, Spremembe: *Uradni list RS*, št. 42/1997, 66/2000, 24/2003, 69/2004, 69/2004, 69/2004, 68/2006. Dostopno na: http://www.uradni-list.si/_pdf/1991/Ur/u1991033.pdf#!u1991033-pdf [12.5.2015].

Young, R. (2015). »Voluntary Euthanasia«. V Zalta, E. N. (ur.), *The Stanford Encyclopedia of Philosophy* (Pomlad 2015). Dostopno na: <http://plato.stanford.edu/archives/spr2015/entries/euthanasia-voluntary/> [9. 5. 2015].

»Zakon o pacientovih pravicah« (ZPacP) (2008). *Uradni list RS*, št. 15/2008. Dostopno na: http://www.uradni-list.si/_pdf/2008/Ur/u2008015.pdf#!/u2008015-pdf [10.5.2015].

Sandra Vajda

Filozofska fakulteta Univerze v Ljubljani

Abolicionizem in moralna dopustnost uporabe živali v raziskovalne namene

Članek govori o etičnih razsežnostih poskusov na živalih, predvsem z abolicionističnega vidika. Najprej se posvetim kratki zgodovini moralnega odnosa do nečloveških živali in razvoju le-tega. Dotaknem se nekaj najpomembnejših zgodovinskih mislecev in njihovega pogleda na živalsko etiko. Nadalje se dotaknem etičnih razsežnosti poskusov na živalih, kjer vključim nekaj statistike in stališča nekaterih pomembnih sodobnih bioetikov in filozofov o živalski etiki. Vprašam se, ali so živali res nujno potrebne v raziskovalne namene. Nato opišem, kaj abolicionizem sploh je, in pa stališče, ki ga zagovarja v odnosu do živali. Abolicionizem se je prvotno zavzemal za odpravo suženjstva in za odpravo smrtne kazni. Skozi zgodovino pa se iz zagovarjanja pravic za vse ljudi razširi na živalsko etiko. Abolicionistični pristop se zavzema za popolno odpravo uporabe živali, tako v prehrani kakor tudi pri poskusih v medicini ter drugje. Na koncu še na kratko opišem slovenske in evropske zakonske direktive.

Ključne besede: bioetika, poskusi na živalih, abolicionizem, zakonodaja s področja poskusov na živalih

Izbrani vidiki predzgodovine abolicionizma

Pričnimo zgolj z nekaterimi koraki iz zgodovine našega ravnanja z nečloveškimi živalmi, kakor jo vidi abolicionizem. V zgodovini je bila razširjena misel, da je neenakost »naravna« in potrebna za vzpostavitev družbenega reda. Aristotel je menil, da so sužnji orodje lastnikov, ter je suženjstvo upravičeval s sklicevanjem na psihološke oziroma biološke lastnosti. Kljub temu da nečloveške živali uporabljamo za doseganje naših ciljev, se je etične premisleke o uporabi živalih, njihovem statusu ali obravnavi izključevalo iz etičnih razprav (Cavalieri, 2006: 13).

Aristotel je razumel človeka kot racionalno žival in ga je zato tudi postavil na vrh hierarhične lestvice, na katero je glede na racionalnost razporedil vsa živa bitja. Čeprav živali lahko občutijo bolečino in žalost ter ju celo sporočajo, je menil, da ima samo človek občutek za dobro in zlo oziroma pravičnost in nepravičnost.

Descartes je, nasprotno od Aristotela, menil, da nečloveške živali niso sposobne čutiti ničesar, vključno z občutkom bolečine. Nečloveška bitja je uvrstil k goli materiji, zaradi česar nečloveške živali nimajo nobene subjektivne vrednosti. Verjel je, da so živali popolni avtomati, ki jih je lahko ustvaril le Bog, in so »narejeni neprimerljivo bolje od kateregakoli stroja, ki bi ga lahko ustvarili ljudje« (Cavalieri, 2006: 59). Pri Kantu srečamo podobno stališče do nečloveških živali, kot ga je zagovarjal Aristotel. Živali niso nič drugega kot sredstvo za doseganje potreb človeka in so zaradi slabega moralnega statusa izključene iz moralnih razprav. Moralnim akterjem pa kljub temu stališču ni dovoljeno vsakršno ravnanje z živalmi. Odsev našega ravnanja z živalmi se odraža v našem ravnanju z edinimi moralnimi »pacienti«, torej z drugimi človeškimi bitji (Cavalieri, 2006: 65). Kant je, nasprotno od Descartesa, menil, da nečloveške živali imajo zavest, vendar tudi on zavrača neposredne dolžnosti do nečloveških živali. Poziva zgolj, da naj s »stvarmi«, v katere je uvrstil živali, ne ravnamo kruto, ker bi to lahko posledično pripeljalo do krutega ravnanja drug z drugim.

Za spremembo v odnosu do živali sta v veliki meri zaslužna Bentham in Mill, ki sta pripomogla, da se v 19. stoletju nečloveškim živalim začne pripisovati večjo moralno vrednost. Bentham preusmeri etično pozornost na vsa bitja. V središču etike se znajdejo moralni »pacienti«, to so vsa bitja, ki imajo interes občutiti srečo in se ogniti bolečini. Kot utemeljitelj modernega utilitarizma, ki se drži formule »največje sreče za največjo količino bitij«, je bil zagovornik najvišjega moralnega načela, tj. načela koristi. Torej, ko izbiramo med različnimi dejanji, se odločimo za tistega, ki omogoči največ dobrega za največje število udeležencev (Cavalieri, 2006: 81). Poglejmo primer, kjer imam na eni strani pravico in interes do življenja, na drugi strani pa bi me lahko uporabili v bolečih biomedicinskih raziskavah, katerih posledica bi bila moja smrt, hkrati pa tudi zdravilo za raka. Večina utilitaristov ne bi imela težav s sodelovanjem ljudi v biomedicinskih raziskavah, v kolikor bi bile posledice ugodne za večino (Francione, 2010a: 20, 21).

Bentham je bil mnenja, da je zmožnost čutiti trpljenje, pglavitna lastnost, ki je hkrati dovolj, da daje bitju pravico do enakega upoštevanja glede na ozir trpljenja. Znan je po misli, da pravo vprašanje ni, ali lahko nečloveške živali mislijo, niti ne to, ali lahko govorijo, temveč to, ali lahko trpijo. Bentham in Mill sta oba menila, da tudi če živali ali druga možna bitja niso ali ne bi bila racionalna ali kako drugače podobna ljudem, te razlike še vedno ne bi zadoščale za to, da ne bi upoštevali njihovega trpljenja ter si ga prizadevali zmanjšati (Francione, 2010b: 26). Kljub vsemu je Bentham dopuščal uporabo živali, v kolikor jim ne zadajamo nepotrebnih bolečin (Francione, 2010a: 9).

Etične razsežnosti poskusov na živalih

V večini držav zakonodaja določa, da je potrebno zdravilo podvreči številnim testom, preden le-to pride na trg. Testiranje zdravil na živalih za človeka ni najbolj uporabno niti dovolj varno, kajti nobena živalska vrsta ne more biti dovolj dober model za neko drugo živalsko vrsto. Število živali, ki jih na letni ravni uporabijo v laboratorijih, je ocenjeno na okoli 500 milijonov. Največkrat, v 70% se poskusi opravljajo brez anestezije, v ostalih 30% pa le z delno anestezijo (Stop Vivisection!, 2016). V Sloveniji je bilo leta 2013 v poskusih uporabljenih 11.446 živali, leta 2012 – 10.035 in leta 2011 – 11.874. »Določene vrste živali – kunci, miši, podgane – morajo te neboleče teste normalno preživeti, v nasprotnem primeru bi bilo z zdravilom nekaj hudo narobe. Nekateri testi pa so zahtevnejši in takrat so živali v globoki narkozi. Čeprav smo v zadnjih letih marsikateri test ukinili ali ga nadomestili z alternativnimi metodami in vitro, na žalost za nekatere teste takšne metode še ne obstajajo. V takih primerih moramo za analizo relacije zdravilo–človek še vedno uporabiti laboratorijske živali,« pravijo v Leku. V Krki želijo sicer poskuse na živalih čim prej tudi povsem opustiti (Videmšek, 2014).

Albert Schweitzer je menil, da tako kot obsega moja volja do življenja hrepenenje po nadaljnjem življenju in bojazen pred uničenjem, zajema to hrepenenje tudi volja do življenja okoli mene, naj bo ta nema ali pa se ima moč izraziti. Etika obstaja v tem, da nujno vsej volji do življenja izrekam enako spoštovanje do življenja kot svoji. »Dobro je življenje ohranjati in ga spodbujati, zlo je življenje uničevati in življenje ovirati« (Schweitzer, 1991: 1203).

Človek je etičen samo, če se zavzema za pomoč vsemu življenju in se upira temu, da bi škodoval čemurkoli živemu. Če pa škodujemo kateremukoli bitju, nam mora biti jasno, ali je to neizogibno, in ne smemo iti dlje, kot je to nujno potrebno. Tisti, ki na živalih delajo operacije, oziroma jih okužijo z raznimi boleznimi, se nikoli ne smejo sprijazniti z argumentom, da njihovo dejanje opravičuje dragocen smoter. Če je bolečino moč preprečiti, je ne bi nihče smel dovoliti (prav tam: 1203).

Singer kot temeljno moralno načelo postavlja načelo enakega upoštevanja interesov za vse ljudi, ne glede na razlike, kot so spol, rasa ipd., hkrati pa »nas sprejetje načela enakosti kot trdne moralne osnove za odnose s pripadniki naše vrste zavezuje k sprejetju tega načela, kot trdne moralne osnove za odnose z osebkami zunaj naše vrste - z nečloveškimi živalmi« (Singer, 2008: 63).

Načelo enakosti terja, da se ne glede na naravo trpečega bitja njegovo trpljenje upošteva v enaki meri kakor podobno trpljenje vsakega drugega čutečega bitja. »To načelo implicira, da naša skrb za druge ne sme biti odvisna od tega, kakšni so in kakšne so njihove sposobnosti (čeprav je odgovor, na vprašanje, kaj točno ta skrb od nas zahteva, odvisen od lastnosti tistih, na katere z našim delovanjem vplivamo)« (Singer, 2008: 64).

Specizem pomeni diskriminacijo na podlagi vrste in zagovarja stališče, da imajo ljudje pravico do privilegiranega moralnega statusa zgolj zato, ker so ljudje. Specisti, ki bi jih lahko enačili z rasisti oziroma seksisti, menijo, da bolečina miši ali prašiča ni enaka bolečini človeka. Lahko bi sicer trdili, da miš v primerjavi s človekom ne more čutiti enako globoke in močne bolečine. Ljudje imajo drugačno zavest ter predvsem jasnejše samozavedanje, kar lahko njihovo trpljenje še potencira. Čeprav človek, ki umira za rakom, čuti večje trpljenje kot miš z enako boleznijo, to še vedno ne ovrže načela enakega upoštevanja interesov. Človek lahko trpi bolj zaradi sposobnosti predvidevanja, boljšega spomina in, ne nazadnje, boljšega dojemanja o tem, kaj se dogaja okoli njega. »V nekaterih situacijah bo pripadnik neke vrste bolj trpel kot pripadnik druge vrste. Tudi v tem primeru moramo uveljaviti načelo enakega upoštevanja interesov, toda uveljavljanje tega načela daje seveda prednost lažšanju večjega trpljenja« (Singer, 2008: 66).

Kognitivne razlike med ljudmi in nečloveškimi živalmi odpirajo nadaljnje vrste težav. Če izvedemo izredno boleče znanstvene raziskave na normalnih odraslih ljudeh, ki smo jih naključno ugrabili v parku, to povzroči, da bi bili med sprehajanjem v nenehnem strahu, kar pa bi povzročilo še dodatno trpljenje. Če izvedemo enak poskus na živalih, te občutijo manjše trpljenje, kajti ne občutijo strahu pred samo ugrabitvijo. Toda, ali bi bilo moralno sprejemljivejše, če bi namesto živali uporabljali bodisi novorojenčke ali umsko zaostale ljudi? Ti se prav tako ne bi zavedali, kaj se je zgodilo, in bi tako kot živali čutili »le« trpljenje, ki ga povzroča poskus. Slednji argument umesti umsko prizadete ljudi, novorojenčke in živali v isto kategorijo in, če dopuščamo uporabo živali v raziskovalne namene, bi potemtakem morali sprejeti tudi raziskovanje na ljudeh. Če primerjamo nekatere nečloveške živali in nekatere kategorije umsko prizadetih ljudi, so mnogi izmed slednjih manj upravičeni do tega, da jim priznamo samozavedanje oziroma avtonomijo.

Prav tako bi lahko ugovarjali, da primerjave trpljenja med različnimi vrstami ne morejo biti natančne. To seveda drži, vendar pa tudi primerjava zgolj med ljudmi ni mogoča. Vsak posameznik se lahko na isto količino trpljenja odziva drugače. Povzročanje bolečine je v vsakem primeru slabo, pa naj gre za ljudi ali živali. Kako slaba je bolečina, je odvisno od njene količine in trajanja (Singer, 2008: 67, 68).

Žival, ki trpi bolečine, se obnaša zelo podobno kot človek, kar je dovolj dober razlog da verjamemo, da živali čutijo bolečino. Dojenčki oziroma mlajši otroci svoje bolečine ne morejo izraziti z besedami tako, kot je ne morejo živali, pa navkljub temu verjamemo bolečini, ki jo ob poškodbi čutijo. Z znanstvenega vidika vemo, da je živčni sistem vretenčarjev, tako sesalcev kot ptic, več ali manj podoben, kar je v dovoljšnji meri sprejemljiv argument v prid zgornjim trditvam. Nasprotniki sposobnosti zavedanja živali so podali protiargumente, ki se sklicujejo na

to, da živali ne morejo misliti, zaradi česar niso sposobne sklepati, to pa pomeni, da nimajo predstave o samih sebi.

Misel, ki smo jo opazili tako pri Kantu kot pri Aristotelu, da si samozavedajoča bitja zaslužijo prednostno obravnavo, je združljiva z načelom enakega upoštevanja interesov. To bi držalo v primeru, da je z njo »mišljeno samo to, da lahko nekaj, kar se pripeti samozavedajočim se bitjem, nasprotuje njihovim interesom, medtem ko podoben pripetljaj ne bi nasprotoval interesom bitij, ki nimajo samozavedanja« (Singer, 2008: 79). Samozavedajoče bitje se veliko bolj zaveda dogajanja, ki ga je sposobno umestiti v neko daljšo časovno obdobje, ima drugačne želje in podobno. Če stališče pripeljemo do absurdnih skrajnosti, sprejmemo tezo, da medtem ko ima človek samozavedanje, ga svinja nima, iz česar sledi, da bi prikrajšanost za svinjetino pripeljala do več trpljenja, kot pa bi ga svinji povzročila nezmožnost prostega gibanja. Nekaj povsem drugega bi bilo misliti, da bi bilo trpljenje samozavednega bitja, četudi ne bi nič bolj trpelo od zgolj čutečega bitja, pomembnejše, ker naj bi bila vrednost določenih vrst bitij večja od vrednosti drugih. Na mestu bi si bilo postaviti vprašanje, zakaj so samozavedajoča bitja bolj pomembna in imajo večjo vrednost od drugih? Ali, če gremo še v večjo skrajnost, zakaj bi trdili, da imajo manj pomembni interesi samozavedajočega bitja prednost pred bolj pomembnimi interesi »samo« čutečega bitja, če samozavedanje prvega bitja ni ogroženo? Kajti v tem primeru življenje samozavedajočih bitij ni ogroženo in se njihovo samozavedanje ne bo zmanjšalo, ne glede na to, kako se odločimo. Interesi tako človeka kot živali so še vedno interesi in si zatorej zaslužijo enako obravnavo.

Če bi sprejeli zgornjo trditev, da imajo samo samozavedajoča bitja pravico do popolnega moralnega statusa, bi umsko ne dovolj zreli ljudje in novorojenčki potemtakem spadali v isto kategorijo kot nečloveške živali. Seveda se že sama misel, da bi opravičevali poskuse na ljudeh, zdi absurdna. Kljub temu pa so nekateri filozofi podali argumente, ki zatrjujejo, da uporaba lastnosti samozavedanja in avtonomije za razločevanje ljudi od nečloveških živali ne bi imela posledic.

Singer obravnava tri takšne premisleke. V prvem predlogu opiše, da so ljudje, četudi so ti umsko zaostali, še vedno del naše vrste in bi z njimi bilo potrebno ravnati enako. Čeprav sami teh lastnosti oziroma sposobnosti, ki jih ima normalen umsko dorasel človek, nimajo, jih večina članov njihove skupine ima, in je to dovolj dober razlog za enako obravnavo. Ta argument bi bili pripravljeni zagovarjati samo v razpravi ravnanja z drugimi vrstami, medtem ko se večina tega ne bi posluževala pri zagotavljanju pravic žensk oziroma pri rasni diskriminaciji. To bi pripeljalo do sklepa, da lahko v izbranih primerih bitja obravnavamo kot posameznike, spet v drugih pa kot skupine. Sprejetje takšnega sklepa pa se zdi nerazumno. Drugič, čeprav nekateri umsko zaostali ljudje velikokrat nimajo višjih, če sploh, sposobnosti od nekaterih živali, so še vedno ljudje in imamo z njimi drugačen odnos, kot s pripadniki nečloveških vrst. »Pristranskost do naše lastne vrste in

znotraj nje do znatno manjših skupin je, kakor vesolje, nekaj, kar moramo pač sprejeti... Poskus izničenja pristranskih čustvenih vezi je nevaren, ker lahko privede do izničenja izvora vseh čustvenih vezi« (recenzent *Animal Liberation*, v Singer, 2008: 81). Ta argument poudarja moč čustvenih vezi, ki jih čutimo do nekaterih drugih ljudi. Singer meni, da je argument precej slab, saj imajo nekateri ljudje rajši oziroma so jim čustveno bližje nečloveške živali, kot pa na primer njihovi sosede, pa to še vedno ni dovolj dober razlog, da bi v odnosu do teh ljudi ravnali manj moralno. Tretji argument lahko povežemo z zmoto spolzke strmine. Če sprejmemo, da umsko bolj zaostali ljudje zadostujejo moralnim pravicam živali, bi nam na drsalnici hitro lahko spodrsnilo. To bi privedlo do zmanjšanja ali pa celo ukinitve pravic ljudem, ki so v družbi posebni, na koncu pa bi prišli celo do totalitarne vlade, za katero je značilno netoleriranje drugačnih. Singerjev namen pa nikakor ni bil zniževati moralni status živali ali celo zagovarjati poskuse na ljudeh (Singer, 2008 : 8).

Abolicionizem kot pristop k živalski etiki

Abolicionizem se je prvotno zavzemal za odpravo suženjstva v ZDA, danes pa je močno prisoten tudi pri zagovornikih odprave smrtni kazni ter pri zagovarjanju reforme kazensko-pravnega sistema v smeri nadomestitve zaporne kazni s civilno pravnimi alternativami (Levičar, 2012: 17).

Skozi zgodovino se je abolicionizem razširil iz zagovarjanja pravic za vse ljudi na živalsko etiko, kjer se razširi misel, da je uporaba živali v katerikoli namen nedopustna.

Abolicionistični pristop v grobem pomeni zagovor popolne odprave uporabe živali, tako v smislu uporabe v prehrani kakor tudi za namene poskusov v medicini ter drugje. Cilj abolicionizma je dosega popolnega moralnega statusa živali, za katerega je pogoj sposobnost čutenja. Abolicionizem zavrne tako imenovano »humano obravnavo« živali, ki ostaja zgolj v iluziji in pravzaprav služi ekonomskim interesom prehranske industrije. Vsako čuteče bitje ima svoje interese in posledično svojo subjektivno zavest, ter ima pravico ne biti uporabljeno kot sredstvo za doseg človekovega cilja (Strahovnik, 2013: 6).

Abolicionistični pristop v odnosu do živali je v največji meri utemeljil in razširil Gary L. Francione. Zavzema se, da živali ne bi več obravnavali kot našo lastnino. Testi in vivisekcija na živalih se še vedno izvajajo za dobiček podjetij in ne več toliko zaradi izboljšanja zdravstvene oskrbe, kot to mogoče še vedno želijo prikazati podjetja. Treba pa je opozoriti, kako pomembno je mišljenje v sodobni družbi, da z uporabo živali, namenjenih za testiranje in razvoj zdravil, pridobimo veliko koristi in da ima živalska uporaba vsaj navidezni namen, ki služi višjemu cilju in ne samo zabavi in užitku. Tisti, ki zagovarjajo uporabo nečloveških živali v ta namen, trdijo, da jih uporabljajo samo, če alternativ ni na voljo, uporabijo pa

najmanjše možno število živali ter te živali izpostavijo najmanjši možni bolečini in trpljenju. Pravijo, da je uporaba živali v te namene – ter bolečina in trpljenje, ki smo ju s tem povzročili, – nujno potrebna, kar pa ne moremo trditi za druge vrste uporabe živali. Mogoče je zanimivo omeniti, da je bil glavni poudarek britanskega gibanja za zaščito živali v 19. stoletju in ameriškega gibanja za živali v 20. stoletju namen oziroma nujnost uporabe živali v raziskovalne namene. To ne pomeni, da zagovorniki živali niso promovirali vegetarijanstva ali nasprotovali lovu. To je samo ugotovitev, da je bila glavna skrb zagovornikov živali vivisekcija, ki vključuje manj živali od prehranske industrije, zaradi česar pa se zdi to vprašanje bolj udobno. Rezultat tega je več zakonov, ki naj bi urejali vivisekcijo tako v ZDA kot v Veliki Britaniji, vendar pa je učinkovitost le-teh vprašljiva. Pri vivisekciji se v primerjavi z uporabo živali v prehranske namene uporablja razmeroma majhno število živali, zaradi česar je kritika vivisekcije manj pereča tema za širšo javnost kot na primer zagovorništvo vegetarijanstva.

Raziskave, izvedene na živalih, so drage, in, kot poudarja Francione, bi lahko prišli do boljših rezultatov, če bi ta denar porabili na drugačen način. Na primer, pri raziskavah AIDSa poskusi na živalih niso bistveno pripomogli k izboljšanju življenja obolelim za AIDSOM (do leta 1984 je bilo več kot 150 šimpanzov okuženih s HIV, ki so kazali zgolj AIDS-u podobne simptome, nobena izmed opic pa ni razvila bolezni). V tem primeru so boljše rezultate prinesla poskusna zdravljenja, izvedena na ljudeh. Prav tako bi z denarjem, ki je bil namenjen za raziskave na živalih, lahko dosegli večjo učinkovitost z boljšo preventivo, na primer s promoviranjem pomembnosti zaščite pri spolnih odnosih, uporabe čistih, neuporabljenih igel in podobno. Živali se skoraj vedno uporablja pri razvoju medicine, zato je težko podati natančne ocene o uporabnosti, ki so jih imele živali v medicinskih raziskavah. Ker so živali vedno uporabljene za modele za raziskovanje bolezni oziroma zdravil, ne moremo z gotovostjo vedeti, ali je bila njihova vloga ključna pri odkritju ali bi do odkritja prišlo tudi brez raziskovanja na živalih (Francione, 2007: 243). Podatki, ki jih pridobimo pri poskusih na živalih, so pogosto nezanesljivi. Tako se, na primer, testi toksičnosti, ki so bili izvedeni na živalih, med seboj močno razlikujejo glede na vrsto uporabljene metode raziskovanja. Tudi če sprejmemo, da so nekatere raziskave na živalih potrebne, je velik del poskusov prej ovira kot pomoč pri odkrivanju bolezni oziroma zdravil.

Odnos do živali se je skozi zgodovino le malo spremenil. Etiologija priča o zaznavnih, senzornih lastnostih in lastnostih, ki so skupne ljudem in vsaj nekaterim živalim. Prav tako je dokazano, da živali enako kot ljudje čutijo bolečino, strah, lakoto, neugodje, osamljenost, so sposobne občutiti sočutje in se žrtvovati za druge, zameriti storjene krivice ter druga podobna občutja, katera radi pripisujemo samo ljudem (Klampfer, 2010: 233). Kognitivni etologi so tako potrdili, da imajo živali, vključno s sesalci, ptiči in ribami, vsaj enakovredne kognitivne lastnosti, ki so jih včasih pripisovali samo ljudem.

Nečloveške živali so prav tako sposobne predelati informacije na zapleten način. Poskusi, da bi na podlagi pomanjkanja kognitivnih lastnosti opravičevali uporabo nečloveških živali kot sredstev za doseg naših ciljev, narekujejo, da so nekatere človeške lastnosti boljše in zaslužijo posebno mesto. Tudi če lahko določene stvari naredi samo človek (čeprav te določene stvari ne zna napraviti vsak človek), so na drugi strani stvari, ki so jih zmožne narediti zgolj nečloveške živali. Na primer, samo človek lahko piše simfonije, računa, vozi avto in podobno, ampak nasprotno lahko samo žival leti ali plava pod vodo brez pomoči. Torej so »posebne« lastnosti samo tiste, za katere mi menimo, da so. Te razlike pa niso dovolj, da bi si lastili nečloveške živali in jih uporabili v raziskavah. Če omenimo neko lastnost, ki jo označimo za edinstveno človeško in se pri določenih ljudeh sploh ne pojavi, ne bo moralni status človeka zato nič manj vprašljiv. Mnogi trdijo, da ker živali nimajo občutka za pravičnost in ker se lahko samo človek odzove na moralne dolžnosti, nimamo nobenih moralnih dolžnosti do nečloveških živali ter jih lahko izločimo iz moralnih razprav. To pomeni, da ker se nečloveške živali ne odzivajo na naše moralne dolžnosti, tudi same niso upravičene do le-teh. Peter Carruthers denimo trdi, da živali ne morejo sodelovati pri nastajanju družbene pogodbe in zato nimajo moralnega statusa. Dotakne pa se tudi ljudi, ki prav tako ne zadovoljujejo potrebe družbene pogodbe. Kot razlog, zaradi katerega lahko tem ljudem priznamo moralni status, ne da bi ga razširili tudi na živali, poda, že prej omenjeni, argument spolzke strmine. Če zanikamo moralni status nekaterih ljudi na podlagi njihove neracionalnosti, smo stopili na spolzko strmino, ki se lahko razširi na manjše otroke in barbare. Carruthers sicer ne trdi, da vsi ljudje zaslužijo moralne pravice, temveč, da zaradi varnosti raje ne stopajmo po tej poti. Pravi, da je njegov pristop odvisen od dejstva, da ni ostre meje med manj inteligentnimi posamezniki in nekom, ki je mentalno bolan (Francione 2007: 246).

Priznanje moralnega statusa vključuje več vidikov. Bitju pripadajo določene moralne pravice, zaradi česar imamo do njega moralne dolžnosti. Naše ravnanje z bitjem z moralnim statusom mora biti skladno s sistemom moralnih pravil. V moralni presoji in odločanju moramo upoštevati morebitne negativne posledice, ki bi jih to bitje utegnilo občutiti (Klampfer, 2010: 234).

Regan meni, da zgolj s prizadevanjem za dobrobit živali, ne pa tudi za uveljavitev njihovih pravic, ne bomo dosegli trajnostnih sprememb pri izboljšanju njihovega moralnega statusa. Posveti se pojmu krutosti in ugotavlja, da se le-ta kaže na različne načine. Ljudi je mogoče upravičeno soditi bodisi zaradi njihovih dejanj in početij ali pa zaradi stvari, ki jih ne počno. Ni nujno, da so kruti ljudje tisti, ki povzročajo bolečino, kajti tudi zobozdravnik jo lahko povzroči, kar pa ga še ne naredi krutega, temveč so kruti tisti, ki v bolečini uživajo. Nekateri kruti ljudje ob povzročanju bolečine ne čutijo zadovoljstva, vendar pa se zdi, da pravzaprav ne čutijo ničesar. Njihova krutost se kaže v pomanjkanju občutka, kaj je primerno čustvo za dejanje, ki ga povzročajo, oziroma v pomanjkanju sočutja do trpečega. Enako kot, recimo, lev nima nobenega sočutja in občutka za bolečino, ki jo izkuša

njegov plen. Človek se loči od živali prav zaradi zmožnosti sočutja. Prav zaradi tega pa so ljudje, ki povzročajo bolečine, pogosto označeni za »žival« (Regan, 1980: 534).

Regan loči med aktivnim in pasivnim delovanjem. Če vzamemo za primer nekoga, ki brez razloga pretepa psa do nezvesti, je ta aktivno krut, medtem pa bi pasivni delovalec na primer ne hranil psa do točke, ko je ogroženo njegovo zdravje. Slednji ni krut zaradi dejanja, ki ga je storil, temveč zaradi opustitve le-tega (Regan, 1980: 535).

Živali imajo moralno vrednost in njihove pravice bi morali uravnati do te mere, da bi popolnoma odpravili izkoriščanje nečloveških živali in jih ne bi skušali zgolj regulirati. Zavzema se torej, da se ne bi zgolj osredotočali na samo ravnanje do živali, temveč tudi na vprašanje, zakaj jih sploh uporabljamo. Zavrne tako utilitarizem kot teorijo živalske dobrobiti. Meni, da nimamo moralne utemeljitve za ravnanje z živalmi kot s sredstvi za doseg naših ciljev. Če bi se sicer znašli v situaciji, ko smo na rešilnemu čolnu in moramo izbirati med življenjem psa ali človeka, Regan pravi, da vseeno rešimo človeško življenje. Kajti smrt človeka bi vseeno povzročila več »škode« kot smrt psa (Francione, 2010b: 30). Da je živalska smrt manjše zlo v primerjavi s človeško, utemeljuje s trditvijo, da imajo živali manj priložnosti za zadovoljstvo. Tu se pojavi vprašanje, kako lahko z gotovostjo trdimo, da imamo več možnosti za zadovoljstvo, kot pa jo ima rešen pes, ki si deli dom z nami? Ne moremo, prav tako kot ne moremo trditi, da imamo več priložnosti za zadovoljstvo kot neko drugo človeško bitje (Francione, 2010b: 33).

Abolicionistično stališče trdi, da je kakršnakoli uporaba živali nedopustna, ter celo zagovarja odpravo živali v smislu hišnih ljubljencev. Zagovarja pristop k moralnemu statusu živali glede na sposobnost občutenja, kajti vsako čuteče bitje ima svoje interese, želje, potrebe. Tudi tako imenovana »humana obravnava« se zdi iluzorična, kajti mejo povzročanja nepotrebnega trpljenja, na katerega se sklicuje humana obravnava, je težko jasno zarisati. Prav tako pa se zdi, da takšna obravnava zgolj sledi prehranski ekonomski politiki, kajti zmanjša se strošek za zdravljenje in kupci so večinoma pripravljeni plačati več (Strahovnik, 2013: 6).

Zagovorniki dobrobiti živali se opirajo na trditev, da obstaja razlika med razvojem mišljenja pri ljudeh in razvojem mišljenja pri vsaj večini nečloveških živali. Te razlike med človeškim in nečloveškim živalmi pa niso dovolj dober razlog, ki bi upravičeval, da imajo živali manjšo željo po življenju. Čutečnost kot taka ni lastnost, ki se je razvila, da bi služila kot cilj, temveč omogoča bitjem prepoznati nevarnost, ki bi jim lahko ogrozila njihovo preživetje. Enako kot vsa bitja, ki vidijo, nimajo interesa postati slepa oziroma imajo interes še naprej ohranjati svoj vid, tako imajo čuteča bitja željo po izogibu smrti. Griffin zanika samozavedanje nečloveških živali, kajti ne morejo misliti takšnih misli, kot je na primer »jaz plezam na drevo oziroma jaz tečem«. Ko žival zavestno zazna delovanje (plezanje, lovljenje), se mora tudi zavedati, kdo počne te stvari. Toda če vzamemo ljudi, ki

imajo globalno amnezijo, torej poškodbe možganov, ti pogosto izgubijo spomin in se niso zmožni projicirati v prihodnost. Njihov »občutek jaza« je na podobni ravni kot pri zgoraj omenjenih nečloveških živalih. Prav tako občutek za samoza-vedanje ni na isti ravni kot pri zdravem odraslem posamezniku, vendar pa jih še vedno ne bi uvrstili na nivo nezavedajočih bitij. Mogoče jih ne bi postavili na mesto kirurga, vendar jih tudi ne bi uporabili kot sredstvo za doseg naših ciljev (Francione, 2010b: 32, 33). Zagovorniki živalske dobrobiti menijo, da imamo ljudje boljše mentalne sposobnosti, s katerimi lahko ocenimo živalsko bolečino oziroma ocenimo, ali človekovo zadovoljstvo opravičuje bolečino in trpljenje nečloveških živali (Prav tam: 33).

Pristop, ki ga zavzamejo predstavniki dobrobiti živali, zagovarja spoštovanje življenja nečloveških živali, ki imajo manjšo moralno vrednost kot ljudje. Francione zavrača stališče, da so življenja nečloveških živali vredna manj od življenja ljudi. Tako kot tudi duševno zaostali ljudje niso pod prisilo za darovanje organov in se jih ne sili v boleča znanstvena raziskovanja. Zagovorniki živalske dobrobiti menijo, da je sprejemljiva živalska uporaba v namene, ki bi se nam sicer v primeru uporabe ljudi zdeli nesprejemljiva, dokler je uporaba »humana« in ne povzroča nepotrebnega trpljenja. V tej poziciji se lahko razbere, da je živalsko življenje vredno manj od človeškega (Francione, 2010a: 5).

Francione zavrne tudi stališče, ki nekaterim nečloveškim živalim zaradi njihove večje podobnosti človeku pripisuje večje moralne pravice kot drugim. Vsaka žival je človeku podobna v tem, da ne želi trpeti. Dokler si lastimo pravico do živali, ne morejo biti nikoli del moralne skupnosti. Interesi nečloveških živali vedno štejejo manj kot človeški interesi (Francione, 2010b: 35).

Iz povedanega lahko razberemo, da je abolicionizem stališče, ki predstavlja skrajno obliko zagovarjanja zaščite živali. Abolicionisti namreč zagovarjajo popolno odpravo uporabe živali, tudi za hišne ljubljence. Čeprav se je dandanes skrb za domače živali močno povečala in se zdi, da je nemalo nečloveškim živalim veliko bolje kot v naravnem okolju. Abolicionistični protiargument temu je, da s povzročanjem odvisnosti živali od ljudi povzročamo več škode kot koristi. Le-te so namreč povsem odvisne od nas, zaradi česar pa je njihovo preživetje v naravi nemogoče. Vidimo torej, da se abolicionizem močno zavzema za pravice živali, vendar pa bi bilo potrebno prilagoditi stališče vsaj glede hišnih ljubljencev, ki v večini primerov živijo bolj lagodno življenje od tistega v naravi.

Zakonodaja

Na zakonskem področju lahko sicer opazimo, da se dogajajo pomembni premiki, ki prepovedujejo testiranje na živalih. Leta 2013 je Evropska unija sprejela predpis, ki prepoveduje uvoz in prodajo kozmetičnih izdelkov, ki so testirane na živalih. Na policah sicer še vedno vidimo izdelke, testirane na živalih, to pa zaradi

dovoljenja za prodajo, katerega so podjetja pridobila v preteklosti. Od marca 2013 naprej velja, da niti izdelki niti njihove sestavine ne smejo biti testirane na živalih. Rezultati testiranj ne zastarajo, zato ponovna testiranja za že preizkušene izdelke niso potrebna. Prepoved tako velja le za nove teste oziroma za testiranje novih izdelkov oziroma novih sestavin (Miklavc, 2013). V tem okviru sicer lahko vidimo zametke, ki jih zagovarja abolicionistično stališče, vendar pa se predvsem farmacevtske družbe še vedno povečini poslužujejo raziskav izvedenih na živalih.

Slovenska zakonodaja je nekoliko ohlapnejša. 4. člen pravilnika o pogojih za izvajanje poskusov na živalih prepoveduje testiranje na zapuščenih oziroma prosto živečih živalih, vendar takoj v nadaljevanju dopušča izjeme, če so znanstveno utemeljene. Raziskovalci so dolžni izvajati poskuse na živalih pod anestezijo, bodisi lokalno ali splošno, razen če bi le-ta ovirala namen raziskovanja. Torej ta zakon dopušča izvajanje poskusov na živih živalih in dovoljuje njihovo mučenje, če je to v »dobrobit« raziskave. Vsak laboratorij je sicer dolžan imeti Komisijo za dobrobit živali, vendar pa se zdi, da ta žal ne igra pomembne vloge, če sploh kakšno. Etična komisija za dobrobit živali mora sicer določiti, ali je raziskava utemeljena z znanstvenega vidika in ali jo je mogoče izvesti na human način. Ob tem se poraja vprašanje, kaj sploh je izvajanje poskusa na human način oziroma kaj določa human način. Izvajalec projekta je sicer dolžan prekiniti nepotrebno bolečino ali trpljenje živali in ji bolečine po koncu testiranj kar se da lajšati.

V 4. členu Pravilnika o pogojih za izvajanje poskusov na živalih je določeno, katere vrste živali je dovoljeno uporabiti v znanstvene in raziskovalne namene. Lahko se uporabi živali, ki so bile vzrejene ali gojene oziroma skotene in rejene kot laboratorijske živali. Z zakonom je prepovedana uporaba prosto živečih živali, razen v primeru, ko namena poskusa ni mogoče doseči z uporabo laboratorijskih živali, ki so z zakonom predpisane in gojene v te namene. Zakon prepoveduje uporabo zapuščenih hišnih in rejnih živali. »Ne glede na prejšnji odstavek se lahko za izvajanje poskusov na rejnih živalih ali drugače vzrejenih živalih izda dovoljenje za poskus na podlagi znanstvene utemeljitve, da namena poskusa ni mogoče doseči z uporabo laboratorijske živali, in ocene etične komisije za poskuse na živalih (v nadaljnjem besedilu: etična komisija) iz 22. člena tega pravilnika« (Pravilnik o pogojih za izvajanje poskusov na živalih, URS št. 31/2013 živali v poskusih, 4. člen). Prav tako je v slovenski zakonodaji prepovedano izvajanje poskusov brez anestezije, razen kadar bi le-ta ovirala znanstveno raziskavo: »Če je to neprimerno glede na namen poskusa in se uporablja analgezija ali druga ustrezna metoda za zagotovitev, da sta trpljenje in stiska poskusnih živali čim manjši« (20. člen). O primernosti uporabe anestezije odloča etična komisija, ki mora upoštevati, ali je anestezija sama za žival bolj travmatična od poskusa ter skladnost anestezije z namenom poskusa.

Zakoni bi tako potrebovali dodatne omejitve in potreben bi bil ponoven premislek o tem, v kakšni moralni skupnosti želimo živeti. Takšni, ki bo dopuščala boleče biomedicinske raziskave, ali pa smo morda dovolj zreli, da popolnoma prekinemo z obstoječo prakso. Zaradi sodobne tehnologije bi bilo prav mogoče potrebne raziskave opraviti kako drugače in ne na živalih. Britanski znanstveniki namreč trdijo, da bodo nove računalniške tehnike v modeliranju in bioinženiringu kmalu toliko napredovale, da bi lahko nadomestile testiranje zdravil in drugih sestavin na živalih (L. P., 2014). Torej, bi z zakonom lahko prepovedali uporabo živali v raziskovalne namene na katerikoli način, denar, ki sedaj izteka v raziskovalne laboratorije, pa namenili za nakup tehnologije in drugih tehnik učenja, kjer ni potrebna uporaba živali. Nemara bi bile tehnike tudi bolj zanesljive, kajti, kot rečeno, nečloveške živali zaradi drugačne biološke strukture velikokrat ne ustrezajo kot modeli oziroma so rezultati popolnoma neuporabni za ljudi.

Zaključek

Odnos do nečloveških živali se je skozi zgodovino spreminjal in četudi so nekateri misleci verjeli, da si živali sploh ne zaslužijo mesta v moralni obravnavi, pa je bilo to vprašanje vseskozi prisotno. Predvsem Bentham in Mill gre zasluga za to, da so nečloveške živali začele pridobivati status moralnega pacienta.

Nečloveške živali so sicer v preteklosti, v primerjavi s sodobnim časom, imele le malo moralne vrednosti, so pa številke ubitih živali danes nepredstavljivo večje kot nekoč. Vprašanje, ali je lahko nečloveška žival enako moralno obravnavana kakor človek, tako ostaja. Nečloveške živali morda res nimajo tako kompleksne zavesti, kot na primer človek, pa vendar nikoli ne vidimo, da bi neka žival mučila drugo žival samo iz užitka, če odmislimo naravni gon po preživetju. Kaj je torej tisto, kar dela človeka, kot najbolj zavedajoče bitje, bolj plemenitega in moralnega od živali?

Da bi človek spoznal, da njegovo življenje ni vrednejše od življenje drugih bitij, bi bilo potrebno sprejeti Singerjevo moralno načelo enakega upoštevanja interesov. Vsako čuteče bitje ima interes, da ne bi trpelo, kar je dovolj dober razlog, da se to načelo razširi iz odnosov med ljudmi na odnose do živali. To, da nekdo ni pripadnik naše vrste, ni dovolj, da bi dovoljevali izkoriščanje živali za dosego naših ciljev. Regan se zavzema za uveljavitev pravic nečloveških živali, kajti zgolj gibanja za dobrobit živali ne bodo prinesla trajnostnih sprememb pri izboljšanju njihovega moralnega statusa. Razlogi, zakaj se poskusi še vedno izvajajo, kljub vsemu povedanemu in kljub dokazom, da žival prav tako čuti kot človek, so različni. Takšne poskuse je razmeroma enostavno objaviti, kar pa ne drži za raziskave na ljudeh, ki so veliko bolj zamudne. V primeru neželenih učinkov jemanja zdravila ali pa celo smrti, se farmacevtska združenja sklicujejo na to, da so izvedle zakonsko določene »varne« teste na živalih. Abolicionizem, kot skrajna oblika

zagovora odprave uporabe živali, se močno zavzema za pravice živali in ima tudi najbolj konsistentna in izdelana stališča glede uporabe nečloveških živali v poskusih.

Abolitionism and the Moral Admissibility of Using Animals for Research Purposes

The article deals with the ethical dimensions of animal experimentation, particularly from the abolitionist perspective. First, I focus on the history of moral attitudes towards non-human animals and the development thereof. Some of the most important historical thinkers and their views on animal ethics are reviewed. The next section discusses the ethics of animal experiments, and includes some statistics and relevant views of animal ethics by certain important contemporary bioethics thinkers and philosophers. I touch upon the issue of whether animals are genuinely necessary for research purposes. Then I characterize abolitionism and describe its position in relation to animals. Abolitionism originally advocated the abolition of slavery and that of the death penalty. Over the course of history it extended to animal ethics. It argues for the complete abolition of the use of animals for food as well as in medical experiments. I conclude with a short description of the Slovenian legislation and European directives in this regard.

Keywords: bioethics, experiments on animals, abolitionism, legislation on animal experiments

Literatura

- Cavaleri, P. (2006). *Živalsko vprašanje - za razširjeno teorijo človekovih pravic*. Ljubljana: Založba Krtina.
- Francione, G. L. (2007). »The Use of Nonhuman Animals in Biomedical Research: Necessity and Justification«. *The Journal of Law, Medicine & Ethics*, 35(2), str. 241–248.
- Francione, G. L. (2010a). *The Animal Rights Debate: Abolition or Regulation?* New York: Columbia University.
- Francione, G. L. (2010b) »Animal Welfare and the Moral Value of Nonhuman Animals«. *Law, Culture and the Humanities*, 6(1), str. 24–36.
- Stop Vivisection! (2016). »Kaj so poskusi na živalih: vprašanja in odgovori«. Dostopno na: <http://www.stopvivisection.eu/sl/content/kaj-so-poskusi-na-%C5%BEivalih-vpra%C5%A1anja-odgovori> [15. 6. 2016].
- Klampfer, F. (2010). *Cena življenja: razprave iz bioetike*. Ljubljana: Založba Krtina.

- L.P. (2014). »Testiranje zdravil na živalih se bliža koncu«. Dostopno na: <http://www.delo.si/znanje/znanost/testiranje-zdravil-na-zivalih-se-bliza-koncu.html> [pridobljeno 22.5.2016].
- Levičar, M. (2012). *Nasilje med zaporniki v času prestajanja zaporne kazni*. Diplomsko delo. Maribor: Univerza v Mariboru, Fakulteta za varnostne vede.
- Miklavc, N. (2013). »Naprej k mednarodni prepovedi testiranja na živalih«. Dostopno na: <http://www.dzzz.si/sl/naprej-k-mednarodni-prepovedi-testiranja-na-zivalih.html> [26. 7. 2016].
- Prijatelj Videmšek, M. (2014). »Poskusi na živalih so slaba znanost«. Dostopno na: <http://www.delo.si/novice/okolje/poskusi-na-zivalih-so-slaba-znanost.html> [26. 7. 2016].
- Regan, T. (1980). »Cruelty, Kindness and Unnecessary Suffering«. *Philosophy*, 55(214), str. 532–541.
- Schweitzer, A. (1991). »Etika spoštovanja do življenja«. *Nova revija*, 10(113-114), str. 1202–1205.
- Singer, P. (2008). *Praktična etika*. Ljubljana: Založba Krtina.
- Strahovnik, V. (2013). »Živalsko vprašanje: med trpljenjem in statusom nečloveških živali«. *Dignitas*, št. 57-58, str. 206–223
- »Pravilnik o pogojih za izvajanje poskusov na živalih« (2013). *Uradni list RS*, št. 37/2013. Dostopno na: <http://www.uradni-list.si/1/objava.jsp?urlid=201337&stevilka=1429> [12. 5. 2016].

Šemsudin Džopa

Filozofska fakulteta Univerze v Ljubljani

Naturalistični vidiki svobodne volje in moralne odgovornosti

Svobodno voljo in moralno odgovornost lahko obravnavamo na dveh nivojih. Na absolutnem nivoju sta težko združljiva s sodobno znanostjo in mogoče celo nekonsistentna sama na sebi. Na kompatibilističnem nivoju pa ju lahko opredelimo kot evolvirani zmožnosti človeka kot biološkega bitja. Determinizem osnovnega nivoja ne samo, da ne nasprotuje tako pojmovani svobodni volji in moralni odgovornosti, ampak celo pomaga pri evoluciji vedno boljših izogibalcev na višjem nivoju. Z aplikacijo zapornikove dileme na evolucijo je možno utemeljiti moralno obnašanje kot evolucijsko stabilno strategijo, tudi ob odsotnosti močnih pojmov svobodne volje in moralne odgovornosti, ki ju kompatibilizem ni zmožen zagotoviti. Za kompatibilistično svobodo je pomembno, da je človek kot racionalni akter zmožen razumeti in reflektivno ovrednotiti razloge, po katerih želi ravnati, ter zmožen nadzirati svoje delovanje v skladu s tako reflektiranimi razlogi. Sodobne psihološke in nevroznanostvene raziskave kažejo na probleme pri obeh zgornjih zmožnostih ter na človekovo veliko dovzetnost za različne kognitivne disonance, samoprevare in podobno.

Ključne besede: svobodna volja, moralna odgovornost, kompatibilizem, evolucija, racionalni akter

1. Uvod

Pojma svobodne volje in moralne odgovornosti lahko obravnavamo na dveh različnih nivojih, absolutnem in kompatibilističnem (Smilansky, 2000). Absolutna oziroma močna pojava sta težko združljiva z determinizmom in sodobno znanostjo, mogoče pa celo nekonsistentna sama na sebi. Kompatibilistični nivo obravnave pa utemeljuje njuno šibko različico, ki se bistveno razlikuje od njenega močnega, metafizičnega pojmovanja. Ravno iz nerazlikovanja in nepriznavanja teh nivojev, ki bistveno vplivata na samo definicijo in vsebino pojmov, izhajajo pogosti nesporazumi med kompatibilisti in inkompatibilisti.¹

¹ Iz predpostavke dveh različnih nivojev argumentacije sledi, da je inkompatibilistični spor med libertarci in trdimi deterministi utemeljen, spor med inkompatibilisti in kompatibilisti pa ne.

V pričujočem članku bomo absolutni, (meta)fizični nivo pustili ob strani in se posvetili kompatibilističnemu pojmovanju tako, da bomo svobodno voljo in moralno odgovornost utemeljevali kot evolvirani zmožnosti človeka kot biološkega bitja. Pomagali si bomo predvsem z evolucijsko teorijo, s katero bomo najprej poskusili dokazati konsistentnost pojma izogibljivosti v determinističnem vesolju. Determiniranost in izogibljivost bomo obravnavali kot lastnosti dveh različnih nivojev, za razliko od tradicionalnega, inkompatibilističnega enačenja determiniranosti in neizogibnosti. Determinizem osnovnega nivoja se bo izkazal celo kot pozitiven za evolucijo vedno boljših izogibalcev (ang. *avoiders*) na višjem nivoju.

V drugem delu bomo s pomočjo zapornikove dileme iskali pogoje, v katerih sodelovanje, moralno obnašanje in celo altruizem postanejo evolucijsko stabilne strategije, kar nasprotuje razširjeni predpostavki, da darvinizem implicira skrajnje sebično, sovražno in anti-socialno obnašanje.

V tretjem, zadnjem delu članka pa bomo raziskali razum kot človekovo glavno biološko *differentio specifico* in njegov odnos do zmožnosti svobodne volje in moralne odgovornosti. Razum omogoči »ujetje« razlogov delovanja, ki so v celotnem dotedanjem poteku evolucije bili »prosto lebdeči«. Ravno zmožnost racionalne refleksije in zmožnost samonadzora predstavljata najpomembnejši vidik kompatibilistične svobode. Ta vidik je praktično celo veliko bolj pomemben kot metafizična razprava na absolutnem nivoju. Zanimalo nas bo namreč, v kakšni meri so človeku dostopni resnični razlogi njegovega delovanja in v kakšni meri je zmožen reflektiranega samonadzora ter koliko je po drugi strani dovzeten za različne kognitivne disonance, samoprevare, *ad hoc* prilagajanje svojega vrednostnega sistema in racionaliziranje kot opravičevanje svojega početja. Velika dovzetnost za naštetu bi namreč predstavljala pomembno oviro človekovi svobodi in ideji racionalnega akterja.

2. Evolucija znotraj determinističnega vesolja

2.1 Evolucija in igra *Življenje* Johna Conwaya

Za prikaz tega, kako lahko v strogo determiniranem svetu pridemo do pojava »izogibnosti« ali »izogibljivosti« (ang. *avoidance*), uporabi Daniel Dennett primer iz igre *Življenje* (*Life*), ki si jo je zamislil britanski matematik John Conway. Kot že samo ime igre pove, je igra simulacija pojava življenja. Po enostavnih, determinističnih pravilih igre, ki so prispodoba naravnih zakonov, se v dvodimenzionalnem prostoru, glede na začetno postavitev, oblike spreminjajo, izginjajo, rastejo, jejo, razmnožujejo itd. Nekatere oblike, na primer *glider* (jadralec), se izkažejo za zelo obstojne, nekatere druge pa zelo hitro izginejo. Vse pa je seveda zelo odvisno od začetne postavitve, naprej se vse dogaja samo od sebe, torej determinirano s preteklim oziroma začetnim stanjem in tremi enostavnimi »fizikalnimi« zakoni.

Na *osnovnem* nivoju, recimo mu fizikalni, je vse, kar se dogaja, zgolj prižiganje in ugašanje »osnovnih delcev« oziroma kvadratov, ni nobenega gibanja. Na *višjem* nivoju, ko spremljamo, kaj se dogaja z oblikami sestavljenimi iz več kvadratov, pa vidimo, da se nekatere oblike obdržijo oziroma »preživijo« dlje kot druge. Nekatere oblike so bolj trajne, odporne in se lažje »izognejo uničenju« oziroma ugašanju kot druge (Dennett, 2003: 40). Na višjem nivoju spremljamo premikanje nekaterih oblik, kar sicer pomeni, da jih na osnovnem nivoju vsakič predstavljajo drugi kvadrati. Nekatere oblike imajo zmožnosti, ki jih druge nimajo, zahvaljujoč zgolj svoji *obliki* (prav tam: 42).

Čeprav torej na osnovnem nivoju ne moremo govoriti o nobenih zmožnostih, se nam na višjem zdi zelo smiselno primerjati različne oblike glede na njihove »zmožnosti preživetja« oziroma vztrajanja v času. Nekatere bolj kompleksne strukture so se tako zmožne bolj kot druge *izogniti* uničenju, tudi tako, da dobijo informacije iz okolja, »zaznajo in predvidijo« prihajajoče oblike. Temu pojavu Dennett pravi *rojstvo izogibljivosti* (Dennett, 2003: 43).

Kot smo omenili, samo ime igre nakazuje na avtorjevo tezo o nastanku in razvoju življenja.² Zato se prispodoba z evolucija že sama vsiljuje. V igri sicer igralec oziroma ustvarjalec vsakič nastavi začetno postavitev in tako pride do bolj zanimivih oblik in dogajanja. Če pa poskusimo odmisлити ustvarjalca in si zamislimo igro s skoraj neskončnim številom kvadratov, ki se *naključno* znajdejo v začetni postavitvi, brez nekega ustvarjalca (boga), se bodo zanimive oblike, zmožne učenja, predvidevanja, izogibanja in reprodukcije, razvile same od sebe. Vse to dobimo v dvodimenzionalnem prostoru s tremi čisto determinističnimi pravili in zgolj dvema vrednostma vsakega kvadrata. Lahko bi rekli, da je analogija z našim vesoljem celo preveč enostavna.

Če si predstavljamo skoraj neskončno število atomov oziroma osnovnih delcev v vesolju, ki se po strogo determinističnih zakonih povezujejo v molekule in tako naprej, ali pa razpadejo nazaj v osnovne delce, nam ni potrebno predpostaviti nobenega »inteligentnega načrta« za tako neverjetne stvaritve, kot jih najdemo v naravi. To je bil Darwinov nenavadni obrat v razmišljanju. Po besedah njegovega kritika: *absolutno modrost* je zamenjala *absolutna nevednost* (MacKenzie, v Dennett, 2003: 47). Pred pojavom teorije evolucije je teleološki dokaz za obstoj boga kot stvarnika veljal za enega najmočnejših dokazov za obstoj boga. Ljudje so si težko predstavljali, da bi tako popolna in okolju prilagojena bitja živalskega in človeškega sveta z vsemi neverjetnimi lastnostmi lahko nastala po naključju. Ko obrnemo način razmišljanja, pravi Dennett, bomo videli, kako se inteligenca raz-

² Igra je bila ustvarjena z namenom testiranja miselnega eksperimenta o samoreproducirajočih se avtomatih Johna von Neumanna. In jim je tudi uspelo »ustvariti« takšne avtomate pri številu kvadratov 10^{13} , kar pomeni, da če bi vsak kvadrat bil velikosti enega milimetra, bi celoten »svet« meril približno 3 km. Dennett komentira, da je najbrž (ni še dokazano) potrebna takšna razlika v redu velikosti od osnovnih delcev do strukture, ki se je zmožna reproducirati, kot je od atomov do organizma (Dennett, 2003, str. 49.)

vije iz »intelligence«, kako »slepi urar« ustvari čut vida, kako se izbira razvije iz »izbire« in kako je lahko celota bolj *svobodna* kot njeni deli (Dennett, 2003: 48).

2.2 Evolucija izogibalcev, od generacij do mikrosekund

Česar še nikomur ni uspelo simulirati v Conwayevi igri *Življenje*, je *mutacija*, kar je najbrž posledica tega, da »svetovi« niso dovolj veliki oziroma so premalo »bogati«. Dennettu pa je najbolj bistveno, da je v tem miniaturnem determinističnem svetu dokazal obstoj vseh nujnih gradnikov za evolucijo *izogibalcev* (Dennett, 2003: 51), kar po njegovem zadostuje za zavrnitev kognitivne iluzije, ki determinizem povezuje z neizogibnostjo.

Če se zdaj vrnemo k iskanju začetka izogibalcev v našem svetu, vidimo, da je na začetku evolucije maksimalna hitrost izogibanja določena s hitrostjo menjave generacij. Torej prva možna reakcija na problem oziroma nevarnost, ki se pojavi v času ene generacije, se zgodi v času naslednje generacije. Šele naslednje generacije lahko razvijejo nek odgovor na problem. Tista serija, vrsta oziroma rod potomcev, ki se bolj uspešno ubrani nevarnosti, preživi, ostali rodovi izumrejo. Seveda bi bilo to narobe razumeti v smislu, da so te osnovne življenjske oblike karoli *počele*. Tiste *srečne*, ki so preživele, so preživele zaradi uporabnih mutacij:

Te srečne življenjske oblike so potem imele potomce, katerih srečni potomci so spet imeli potomce in tako naprej, dokler ne pridemo do nas. Mi pa smo tisti srečneži, sestavljeni iz tako uporabnih delov, posebej narejenih zato, da se dobro obnesejo pri izogibanju, ampak tokrat v veliko krajšem času. (Dennett, 2003: 52)

To je bistvo evolucije oziroma naravne selekcije, tukaj ni potrebna nikakršna daljnovidnost niti predvidevanje. Naravna selekcija deluje s preizkusi skozi različne značilnosti mutiranih vrst in rodov, od katerih jih veliko ne preživi. Lahko bi rekli, da so ti preizkusi precej »potratni«, oziroma dolgotrajni.

Vendar pa ni potreben noben čudež, da bi kratkoviden proces ustvaril daljnovidnega. Slepi ali vsaj kratkovidni proces naravne selekcije si, kot pravi Dennett, pomaga z »žerjavi« in ne z »nebesnimi kljukami«, kjer »raziskovanje v času ene generacije (generacije) lahko vpliva na osnovni proces genske evolucije /.../« (Dennett, 2003: 53). Danes so človeška bitja zmožna zaznati stvari pri ogromni razdalji, prirojenim organom pomagajo protetične naprave, kar omogoča zaznavo in rešitev problemov s hitrostjo, ki se bliža maksimalni možni hitrosti v vesolju, svetlobni hitrosti (prav tam: 53). Zamislimo si možnost, da se bo človeštvu nekoč uspelo »izogniti se« nekemu uničevalnemu meteorju, podobnemu tistemu, ki naj bi bil odgovoren za izumrtje dinosavrov. Seveda bi tudi to bilo *popolnoma determinirano* dejanje na fizikalnem nivoju, ampak ali bi to res pomenilo, da pojma »izogniti se« niti v tem primeru ne bi smeli uporabiti izven navednic? Ker človeštvo v tistem trenutku ne bo imelo *resnične* izbire? Vsekakor pa bo imelo vsaj več *navidezne* izbire kot dinosavri ali človeštvo *sámo* v preteklosti.

2.3 Razlika med determiniranostjo in neizogibnostjo

Tezo o tem, da determinizem ne implicira neizogibnosti, postavi Dennett v obliki argumenta (2003:56):

- 1) V nekaterih determinističnih svetovih obstajajo izogibalci, ki se izogibajo škodi.
- 2) Torej se v nekaterih determinističnih svetovih izogibalci nekaterim stvarjem izognejo.
- 3) Torej v nekaterih determinističnih svetovih ni vse neizogibno.
- 4) Torej determinizem ne implicira neizogibnosti.

Ta argument nam utegne delovati sumljiv, za kar so po Dennettu najbolj odgovorne skrite predpostavke, ki jih ponavadi ne opazimo. Te predpostavke lahko osvetlimo, če primerjamo stavka (prav tam: 57):

- a) Če je determinizem resničen, potem je vse, kar se zgodi, *neizogiben* izid celotnega seznama vzrokov.
- b) Če je determinizem resničen, potem je vse, kar se zgodi, *determiniran* izid celotnega seznama vzrokov.

Če je neizogibnost samo sinonim za determiniranost, potem prvi stavek ne dodaja nobene vsebine v primerjavi z drugim. Če pa neizogibnosti ne razumemo zgolj kot sinonima za determinizem, potem mora vsebovati nek dodaten pomen. Mogoče se dodatek skriva v neizogibnosti izida. Ampak o neizogibnosti ali izogibnosti nekega izida lahko govorimo samo v primeru akterjev, ki imajo interes izogniti se nečemu. Vendar je sposobnost živih bitij oziroma določenih vrst, da se izognejo nekemu izidu, odvisna od okoliščin, ne od determinizma (prav tam: 57). V nadaljevanju bomo predstavili možne ugovore, za katere Dennett predvideva, da bodo doleteli njegovo teorijo izogibnosti v determinističnem svetu.

Prvi ugovor proti možnosti izogibalcev v determinističnem svetu, ki ga Dennett poskuša zavrnil, je, da izogibalca čaka točno določena usoda, ki mu je bila *od vedno* usojena. Izogibalci tako nimajo »boljše možnosti« za preživetje, možnost nikoli ne »vstopi v igro«: tisti, ki preživijo, preživijo in obratno; vse to je determinirano že od začetka. Na to Dennett odgovarja, da je možnost popolnoma kompatibilna z determinizmom. Med drugim z njo razlagamo evolucijo, ki je tudi kompatibilna z determinizmom. Zakaj vztrajati, da determinirano izogibanje ni resnično izogibanje? Dejstvo, da je celoten *proces* determiniran, nikakor ne zmanjšuje dejstva, da *s časom* ta proces ustvarja vedno več nečesa, kar se vsem zdi kot izogibanje (Dennett, 2003: 58).

Drugi ugovor se glasi, da resnično izogibanje pomeni izogniti se nečemu, kar se *bo zgodilo*. To je po Dennettu odvisno od tega, kaj pomeni »se bo zgodilo.« Nekaj se resnično *ne bo zgodilo* natančno zato, ker se bo sprožil sistem za izognitev. Mi bi pa še dodali, da kritik tega sistema izognitev seveda ne bo upošteval kot lastnosti izogibalca kot akterja, ki ga vsaj v primeru človeka najverjetneje obrav-

nava kot tisto neskončno majhno točko v možganih. Dennett nenehno poudarja, da lahko pozunanjimo vse, če si sebe predstavljamo kot neko matematično točko v možganih.

Tretji ugovor pravi, da izognitev pomeni *spreminjanje izida*. Iz katerega izida v kateri, se sprašuje Dennett. Sama ideja spreminjanja izida je nekonsistentna – razen če to pomeni *predvidevanega* izida, kar se ravno dogaja pri determiniranem izogibanju. Resničnega, dejanskega izida ni možno spreminjati ne v determinističnem ne v indeterminističnem svetu (Dennett, 2003: 59).

Četrti ugovor proti izogibljivosti v determiniranem vesolju trdi, da t.i. izogibalci *neizogibno* imajo samo tiste sposobnosti, ki jih imajo. Dennett odgovarja, da je to ravno definicija neizogibnosti, ki enači neizogibnost in determinizem, tako da ničesar ne dodaja. Če pa hočemo razložiti svet biologije, potrebujemo takšen pojem izogibanja, ki ga lahko svobodno nanašamo na dogodke v zgodovini življenja na Zemlji, ne glede na determiniranost te zgodovine. In ravno v luči zgodovine življenja lahko rečemo, da je determinizem zaveznik in ne sovražnik tistih, ki ne marajo neizogibnosti (Dennett, 2003: 60).

Na tej točki smo seveda še vedno daleč od nečesa, kar bi bilo podobno svobodni volji. Za Dennetta je bilo bistveno dokazati, da determinizem ne implicira neizogibnosti. Kot vidimo, lahko evolucijo razložimo kot popolnoma *determiniran* proces in istočasno ravno kot evolucijo *izogibnosti*. Evolucija je evolucija izogibnosti. Pojem neizogibnosti (ali izogibnosti) ne pripada fizičnemu, osnovnemu nivoju gradnikov, ampak nivoju dizajna,³ na katerem lahko primerjamo neke oblike, vzorce ali stvaritve ravno glede na njihovo zmožnost izogibanja (Dennett, 2003: 62). Prav tako o pojmu avtonomije lahko govorimo samo na nivoju dizajna.

3. Vznik morale

3.1 Začetki življenja

Da bi razumeli pojav človeške svobode, bomo v nadaljevanju z Dennettom sledili Darwinovemu »čudnemu obratu razlaganja«. Na začetku življenja na Zemlji ni bilo prave svobode, razuma, izbire, ampak samo njihove zgodnje oblike: proto-svoboda, proto-razum, proto-izbira (Dennett, 2003: 143).

³ Beseda dizajn (*design*) v besedni zvezi nivo dizajna (*design level*) bi se lahko nanašala na obliko, vzorec, pa tudi na namero in načrt, kot v besedni zvezi *Intelligent design*.

Na začetku življenja je samo ena izbira omogočala preživetje, torej »naredi A ali umri«, danes pa imamo možnosti »naredi A ali B ali C ali, D ali... umri«. Da se je to zgodilo, so morale oblike življenja postati bistveno bolj kompleksne, kot so bile na začetku. Prvi bistven napredek se je zgodil, ko so prve oblike življenja »ugotovile«, da lahko energijo potrebno za življenje, ki so jo do takrat dobile večinoma od sončne svetlobe, dobijo tako, da razgradijo svoje sosesde, in dobijo že »sestavljene rezervne dele«. Tako so prvi plenilci in prvi uplenjenci začeli »oboroževalno tekmo«, ki je v približno milijardi let pripeljala do zelo razvitih oblik enih in drugih, do velikega števila načinov preživetja. Tudi večcelični organizmi so se razvili kot posledica te milijardo let trajajoče »državljske vojne« (Dennett, 2003: 144).

3.2 Teorija iger in evolucija

V nadaljevanju si bomo pogledali, kako poteka evolucijska oboroževalna tekma, kaj sproži in kaj zadrži napad različnih strani v tej tekmi. Napad je seveda evolucijsko veliko bolj razumljiva taktika kot sodelovanje, ki ga je nujno potrebno utemeljiti.⁴ Na tem mestu si bomo pomagali z zelo uporabnim konceptom iz teorije iger. Zapornikova dilema je najbrž najbolj znana igra, ki ima velik razlagalni potencial na področju biologije, psihologije, ekonomije, etike, politologije in ostalih družboslovnih ved. Osnovna ideja je zelo enostavna. Dve osebi sta pripti, ker sta obtoženi nekega prekrška ali zločina. Tožilec ločeno obema obtožencema ponudi isto izbiro. Če noben ne prizna očitane dejanja, vsak dobi kratko kazen. Če en prizna in drugi ne, potem je prvi svoboden in drugi dobi dosmrtno kazen. Če oba priznata, dobita oba srednje dolgi kazni. Vprašanje je torej, kaj se najbolj splača narediti. Če oba ne priznata, bi to seveda bil najboljši scenarij za oba, ker bi v tem primeru dobila najmanjše skupno število let kazni. Če oba priznata, pa je to najslabši izid za oba, ker je skupno število let kazni v tem primeru največje. Seveda bi za vsakega posameznika bil najboljši izid, če prizna, medtem ko drugi ne prizna, in, obratno, najslabši izid, če ne prizna, medtem ko drugi prizna. Mogoče bi bilo najbolje obljubiti si *sodelovanje*, torej da noben ne bo priznal. Kljub obljubi pa bi bila seveda oba soočena z veliko skušnjavo, da priznata in se tako popolnoma izogneta kazni, seveda v primeru, da bo drugi »naivno« držal obljubo. Zlasti v primeru, da se drugi ne bo držal zaobljube, je seveda boljše, da tudi prvi prizna, da sam ne izpade naivnež in si tako »prisluži« dosmrtno kazen. Če oba priznata, dobita torej srednje dolgo kazen. To pa je seveda slabše, kot če bi sodelovala in noben ne bi priznal. Torej bi za oba bilo najboljšo, če bi jima le nekako uspelo preseči svoje sebično razmišljanje in strah pred »izdajo« oziroma sebičnim razmišljanjem drugega igralca.

⁴ To je najpomembnejši razlog (in napaka, kot bomo videli!) za idejo, da bi darvinizem, apliciran na družbo (socialni darvinizem), impliciral skrajne sebično, sovražno, militantno, anti-socialno obnašanje.

Za teorijo je pomembna logična oblika igre, ne njena vsebina. Zato je tako zelo uporabna na različnih znanstvenih področjih, kot smo omenili. Na primer zapornikove dileme naletimo kjerkoli, kjer srečamo takšno strukturo – enostransko odpadništvo/zavrnitev sodelovanja (*defection*) je bolj koristno kot obojestransko sodelovanje, ki pa je bolj koristno kot obojestranska zavrnitev sodelovanja, ki je še vedno bolj koristna od enostranskega sodelovanja. V evolucijski teoriji iger se koristi štejejo v potomcih. Modele iger se prireja na takšen način, da bi prišli do pogojev, v katerih se bolj splača sodelovati. Odpadništvo je izhodiščna zmagovalna teorija, ker se bolj splača od sodelovanja, ne glede na to, kaj drugi igralec naredi. Če torej dva igralca primerjamo po številu potomstva, vidimo, da je evolucijsko veliko bolj utemeljeno odpadništvo. Simulacije igre, kjer je recimo izhodišče takšno, da imamo konstantne odpadnike (*defectors*) in konstantne sodelavce (*cooperators*), pokaže, da odpadniki zelo hitro prevladajo v določeni populaciji, in sicer v času nekaj generacij. (Dennett, 2003: 149) Brez *dodatnih* okoliščin se žal sodelovanje izkaže za zelo slabo strategijo. Najbolj vplivno aplikacijo teorije iger na evolucijsko teorijo je razvil John Maynard Smith s konceptom *evolucijsko stabilne strategije* (ESS). Svet, v katerem vsi zavrnejo sodelovanje, je tako primer ESS, ker alternativne strategije nimajo možnosti, da bi uspešno spremenile trend.⁵ Če bi torej v takšen model postavili nekaj sodelavcev, bi jih takšen svet po zelo hitrem postopku iztrebil. Obstajajo pa posebne okoliščine, ki omogočajo drugačen izid. Ravno takšne okoliščine so bile potrebne za nastanek današnje družbe, kot bomo videli v nadaljevanju.

3.3 Možnosti evolucijske razlage pojava morale

Sledimo torej evoluciji življenja, ki poteka v popolnoma determinističnem okolju. Zgodovina razvoja življenja je zgodovina razvoja vedno bolj sposobnih in vztrajnih izogibalcev, ki jim deterministični fizični zakoni ravno pomagajo pri preživetju. Ko smo na takšen svet naravne selekcije nanegli strukturo zapornikove dileme, smo videli, da simulacija ne podpira sodelovanja kot stabilne strategije, tj. takšne, ki bi se lahko obdržala. V nadaljevanju bomo preverili, če določene specifikacije okoliščin lahko prevesijo tehtnico na stran sodelovanja, če lahko odkrijemo *rojstvo morale iz duha evolucije*.

Kaj bi iz evolucijske perspektive lahko rekli o altruizmu? Altruizem lahko opredelimo kot postavljanje dobrobiti drugih ljudi kot smotra na sebi. (Sober in Wilson, 1998, v Dennett, 2003: 194–195) Tukaj seveda *drugi ljudje* ne pomeni lastni otroci ali družina, za takšen »altruizem« ponuja evolucija zelo prepričljivo razlago.⁶ Če pa je dobrobit drugih *smoter na sebi*, ali potem ne govorimo o pravem al-

⁵ Še en primer ESS je naravna selekcija. Ko se naravna selekcija enkrat utrdi znotraj neke populacije, potem ni možno, da bi jo neka alternativna strategija uspešno premagala.

⁶ David G. Myers tako komentira, da se Carnegiejeva medalja za pogum ne dodeli tistim, ki rešijo sorodnike (Myers, 2008).

truizmu, če se zaradi altruističnega dejanja nekdo počuti dobro sam pri sebi, če mu to dejanje lajša vest? Predpostavimo, da je pojem altruizma vsaj malo bolj konsistenten od tradicionalnega močnega pojma svobodne volje. Ali so altruisti mogoče podobni mulam, ki so možen pojav, čeprav takšen, ki se ni zmožen obnavljati, reproducirati? (Dennett, 2003: 196). Če altruizem primerjamo s sodelovanjem, bi mu, glede na izsledke zgoraj omenjene simulacije, bil usojen propad, če se pojavi v odpadniškem okolju. Ampak tako kot mule, kljub temu da se ne morejo razmnoževati, lahko doživijo razcvet v *določenih* okoliščinah, na primer s pomočjo človeške vrste, morda obstajajo tudi okoliščine, v katerih altruizem ni slepa ulica, strel v nogo, temveč omogoča razvoj posameznika in družbe, v kateri se pojavi.

Najosnovnejši evlucijski model družbe, kot smo videli zgoraj, nam dokazuje, da konstantni sodelavci, recimo jim kar altruisti, neizogibno propadejo. Vendar pa, kolikor več realnih okoliščin dodamo v model, toliko bolj začenjamo prepoznavati obrise nečesa podobnega altruizmu ali vsaj kvazi-altruizmu. Razviti modeli prikazujejo »pogoje, pod katerimi /.../ lahko evolucija določi organizme za sodelovanje ali, bolj natančno, za takšno obnašanje, da so bolj naklonjeni dolgoročni dobrobiti skupnosti kot svoji trenutni dobrobiti« (Dennett, 2003: 197).

Na osnovnem modelu hitro vidimo, kako občutljiv je pojav sodelovanja. Če se znotraj celotne populacije sodelavcev, altruistov, pojavi odpadnik ali zastonjkar, bo imel vsaj enako ali večje število potomcev. Zastonjkar bo namreč deležen istih koristi kot vsi, ne da bi ob tem imel stroške, ki jih imajo sodelavci. Tako bo zelo kmalu znotraj nekoč sodelovalne populacije prevladalo zastonjkarstvo.⁷ Razen če seveda ne intervenira nekaj, kar bi preprečilo razmah zastonjkarjev.

Odmislimo trenutno neke nadnaravne korektive, kot je na primer božja kazen, in poskusimo poiskati razlago v naravi. Ali lahko evolucija proizvede neko sistematično zaščito pred invazijo odpadnikov? Modeli evlucijske zapornikove dileme *druge generacije* vsebujejo eno »konkretizacijo« več glede na svoje bolj enostavne predhodnike. Dodali so namreč dejavnik »gostote« tako, da se verjetnost stika z nekom, ki je bližje, povečuje z gostoto sveta. Tako se je izkazalo, da *soseščina* igra veliko vlogo v razvoju sodelovanja. Soseščina naredi bolj verjetno dejstvo, da bo nekdo stopil v stik s sebi podobnimi. Sodelovanje tako v povprečju prinaša večje koristi glede na to, da bo obnašanje na drugi strani verjetno recipročno, v primeru recipročnosti pa je seveda sodelovanje boljša strategija. Dodatni kriterij, ki je sledil, je bila *možnost izbire*, tako da so igralci lahko pod določenimi pogoji zavrnilo igro. To je povzročilo oblikovanje rojev podobnih igralcev. Sodelavci so

⁷ Takšen pristop je zlahka mogoče uporabiti pri analizi zgodovinskih okoliščin propada socialističnih sistemov, ki pa očitno lahko ob spremenjenih okoliščinah vzniknejo in se ohranijo tudi »od spodaj navzgor« (*bottom-up*). Takšen primer so med drugim skandinavske države, ki seveda »nominalno« ne štejejo kot socialistične, ampak za njih bi lahko rekli, da je odpadništvo manj pogosto. Naša »evlucijska« teza bi bila, da so bolj ostri pogoji preživetja ustvarili bolj sodelovalno skupnost in posameznike.

rajši poiskali druge sodelavce, tako da so tisti, ki so odklonili sodelovanje, obtičali z drugimi, sebi enakimi (Dennett, 2003: 200). Znotraj roja sodelavcev je torej sodelovanje postalo stabilna strategija. Ampak takšno obnašanje sodelavcev lahko opredelimo za kvazi-altruistično, ker bi od pravega altruističnega sodelavca pričakovali, da bo na svojo lastno škodo vztrajal pri odpadnikih. Seveda pa so tudi znotraj tega modela igralci zelo preprosti, z zelo malo »izbire«. Taktika, ki jo uporabljajo, je zelo podobna taktiki, ki jo najdemo na makromolekularni ravni intragenetskega spopada v prokariontski eri (prav tam). Za simulacijo makromolekule nam torej takšen model zadostuje, za simulacijo človeške družbe pa ga bo seveda potrebno izpopolniti.

3.4 Pomen kazni pri evoluciji sodelovanja

Če naše igralce opremimo s še nekaj sposobnostmi, predvsem možnostjo učenja iz izkušenj, postane model bolj zanimiv. Zastonjkarji so bili zmožni preplaviti neko skupino, ker jih ni nihče zaznal. To lahko popravimo tako, da dodamo možnost, da skupina pridobi informacije o njih in se odzove. Skupina lahko tako uvede določene sankcije proti zastonjkarjem *in* naivnežem, recimo jim resnični altruisti, ki se puščajo izkoriščati zastonjkarjem, saj s tem omogočajo njihov razcvet. Seveda to še ne pomeni utrditve sodelovanja kot ESS, v resnici se oboroževalna tekma s tem šele začne – ali se lahko mutirani zastonjkarji in naivneži dovolj dobro zakrinkajo, da jih nihče ne loči od sodelavcev, in ali lahko skupina razvije nove metode za odkrivanje mutantov. Za evolucijo sodelovanja oziroma družbe je torej bistven pojav kazni, v kakršnikoli obliki že. Rob Boyd in Peter Richardson sta v svojem članku o kulturni evoluciji dokazala, da če so stroški kaznovanja relativno nizki, če lahko tudi tisti, ki sami ne kaznujejo, zanesljivo pričakujejo kaznen, to lahko pripelje do neomejenega konformizma znotraj skupine. Naslov njunega članka je zelo zgovoren: »V dovolj velikih skupinah kaznovanje omogoča evolucijo sodelovanja (ali česar koli drugega)« (Dennett, 2003: 200).

Uspelo nam je torej utemeljiti pojav sodelovanja na skrajnje darvinističnih temeljih, na podlagi skrbi za lastno dobrobit in dobrobit svojega potomstva. Sebičnigen se torej lahko izkaže za ne tako skrajnje sebičnega oziroma ne *samo* sebičnega. V besedah Allana Gibbarda: »Oblike sodelovanja, ki so našim prednikom pomagale prenesti njihove gene, da bi ustvarili nas, je vredno hoteti – zaradi boljših razlogov. Darvinistične sile so oblikovale skrbi in čustva, ki jih poznamo, od katerih so nekatera v splošnem moralna« (Gibbard, v Dennett, 2003: 201). Ampak zgoraj opisani svet, kljub nekaterim »moralnim« potezam, deluje zelo hladno in robotsko, kvečjemu spominja na pojave iz živalskega sveta, kjer so nekatere vrste bolj »človeške«, bolj nagnjene k sodelovanju. Za pojav resničnega altruizma bomo našemu modelu seveda morali dodati kakšno značilno človeško potezo. Vsekakor pa smo videli, kako lahko v enostavnem modelu pridemo do družbenega sistema, ki se lahko sam ohranja in ga lahko vidimo kot nujen predpogoj za dolgoročni razcvet resničnega človeškega altruizma.

»Pošast odpadništva visi nad nami vsemi, evolucijski izvirni greh s svojim, večno mamečim, preudarkom: kako je lahko odpadništvo v tem primeru *neracionalno*?« pravi Dennett (Dennett, 2003: 202). Odmislimo primere, ko se ne splača sodelovati zaradi strahu pred kaznijo⁸ ali izobčitvijo. Na sled svobodni volji in moralni odgovornosti nas pripeljejo ravno težki primeri, ko nam notranji glas govori, da »saj nihče ne bo zvedel, pomisli kaj vse lahko s tem pridobiš« (prav tam: 203). Takšni najtežji primeri skušnjave nam lahko razložijo pojav tradicionalnega koncepta duše kot nesmrtnih in duhovnih entitet. Poskusili bomo slediti njeni evoluciji iz predhodnih oblik nadzornih sistemov, da bi razumeli, zakaj deluje tako kot deluje.

3.5 Pojav racionalnih akterjev

Ekonomist Robert Frank je postavil tezo, da se evolucijski napredek človeka in družbe zgodi z obratom: »biti dober zato, da bi *dajal videz dobrega*« zamenja »biti dober je bolj pomembno kot dajati videz dobrega«. S tem je človeštvo po Franku doseglo novo raven *evolucije svobode*, tako da so se ljudje naučili soočiti in rešiti *probleme zaveze*, ki »nastanejo, ko oseba iz lastnega interesa sklene obvezujočo zavezo do nekega obnašanja, ki bo v prihodnosti dajalo vtis, da nasprotuje lastnemu interesu« (Frank, 1988, v Dennett, 2003: 204). Vidimo, da ima problem zaveze isto strukturo kot zapornikova dilema – problem je seveda v tem, kako vedeti, kdo bo spoštoval obveznosti, ki izhajajo iz zaveze, in kdo ne. Prvoosebni problem, ki se pojavi z razvojem racionalnih akterjev, ki so zmožni reflektirati razloge za delovanje, pa je dvojne narave: kako naj *sam* spoštujem zavezo in kako naj prepričam *druge*, da jo nameravam spoštovati. Ker so drugi tudi racionalni akterji, ne morem pričakovati, da mi bodo verjeli na besedo. Po Franku pa ravno na tem mestu nastopijo *čustva*, in sicer v dvojni funkciji: preprečujejo nam, da bi bili *preveč* racionalni, in poskrbijo za naš sloves kot *ne preveč* racionalnih. Čustva preprečujejo neželjen izbruh kratkovidne oziroma lokalne racionalnosti, kar je zelo pomembno za razvoj družbeno odgovornega akterja, ki bo zaradi njih težje podlegel skušnjavam, ki bi jih sicer zgolj racionalno lahko upravičil (Frank, 1988, v Dennett, 2003: 204).

Sloves sodelavca kot »ne preveč pametnega« je pomemben predvsem iz vidika sprejema v skupino sodelavcev, ki je materialno veliko bolj preskrbljena.⁹ Takšna skupina bo hitro postala plen odpadnikov, če ne bo iznašla učinkovitega načina ločevanja med sodelavci in odpadniki. Če to prej na molekularnem nivoju ni bil problem, pa vsekakor postane s pojavom racionalnih akterjev. Verodostojen znak

⁸ Tukaj upoštevamo samo *tuzemsko* kazen. *Onstranska* kazen bi že bila dobra rešitev oziroma stabilna strategija, če bi le vsi verjeli vanjo. Sicer bi pa tej rešitvi lahko očitali isto, kar se očita evolucijskim modelom moralnosti – nemoralen, utilitarističen temelj.

⁹ Spomnimo se razlike med izidi zapornikove dileme v primeru sodelovanja obeh ali v primeru odpadništva obeh.

zanesljivosti nekega akterja ne sme biti takšen, da ga bo »racionalen akter zmožen enostavno ponarediti«¹⁰ (Dennett, 2003: 205). Torej poleg osebnega problema, kako postati zaupanja vreden akter, imamo še problem, kako *dokazati*, da smo vredni zaupanja.

3.6 Samonadzor–vest–altruizem

Po Franku je evolucijska lepota ravno v tem, da čustvom dodeli dvojno nalogo: pomagajo nam namreč pri samonadzoru in hkrati drugim priskrbijo verodostojen znak tega uspeha. Drugi se tako prepričajo, da smo čustvene osebe, ki *strastno* skrbimo za spoštovanje zavez, da smo za ohranitev svoje integritete, poštenosti, pripravljeni plačati iracionalno visoko ceno, seveda z vidika kratkovidnega kritika. Trik za pridobitev ugleda poštenjaka je, da *smo* resnično poštene (Dennett, 2003: 213).

Zakaj je to najbolj učinkovita rešitev problema, nam bo jasno, ko spoznamo, da je to cena, ki jo plačamo za samonadzor. Nad samonadzorom »nimamo nadzora« – ne poteka selektivno niti natančno. Ne moremo namreč v realnem času nadzirati vsakega ločenega dela svojega odločanja,¹¹ zato se zanašamo na malo bolj grob, nerafiniran pristop, ko se »opremimo z močno emocionalno naravo, ki se razlije čez svoj cilj, in zaradi katere se tresemo od jeze takrat, ko je jeza primerna, ali smo nezmožni brzdati svoje veselje, ko je veselje na mestu...« (Dennett, 2003: 214). Prav *ista* čustva, ki nam pomagajo pri samonadzoru, ko se odpovemo neki svoji kratkoročni koristi zaradi neke svoje dolgoročne koristi, so tista, zaradi katerih se odpovemo svojim kratkoročnim interesom v korist drugih ljudi (prav tam). Takšna nujna altruistična posledica pridobitve samonadzora je v resnici zelo koristna. Brez čustvenih izrazov bi nas obtožili pretvarjanja, pravi Dennett.

O odnosu samonadzora in vesti pravi Frank naslednje: »Lahko bi si predstavljali populacijo, kjer so ljudje z vestjo bolj uspešni kot tisti, ki so brez vesti. Ljudje brez vesti bi sicer varali manj, če bi lahko, ampak enostavno imajo večje težave pri reševanju problema samonadzora. Nasprotno pa so ljudje z vestjo zmožni pridobiti velik ugled in uspešno sodelovati s sebi podobnimi« (Frank, 1988, v Dennett, 2003: 215). Ali lahko tako pridemo do *pristnega altruizma*?

Ljudje s pristnimi moralnimi čustvi so bolj sposobni ravnati v svojem lastnem interesu /.../ Ljudje z visokim ugledom lahko tako rešijo celo nepovnljene zapornikove dileme. Na primer: lahko uspešno medsebojno sodelujejo v poslovnih podvigih, kjer je nemogoče odkriti prevaro. Z drugimi besedami – pristni altruizem lahko vznikne že na podlagi vzpostavitve ugleda zaradi preudarnega ravnanja. (prav tam)

¹⁰ Teza Amotza Zahavija (v Dennett, 2003: 205).

¹¹ Za več o tem glej Dennett (2003: 221) oziroma nevrološke raziskave o zavestnem odločanju.

Frankova teza je torej, da altruizem še zdaleč ni obsojen na propad. Tezo je testiral tudi s poskusom, kjer so študenti igrali večkratno zapornikovo dilemo, v kateri so lahko zaslužili določeno nezanemarljivo vsoto denarja.¹² Študenti so pred začetkom igre dobili možnost, da se na kratko spoznajo. Frank je ugotovil, da so na podlagi tega bili med 60 in 75 odstotkov sposobni napovedati, kdo bo »sodelavec, kdo pa odpadnik« (Frank, 1988, v Dennett, 2003: 216). To po njegovem potrjuje naši intuiciji, da lahko spoznamo neoportunistične ljudi, kar je glavna premisa, na kateri je utemeljen model zaveze:

Iz te premise logično sledi, da se bo neoportunistično obnašanje pojavilo in preživelo celo v neusmiljeno tekmovalnem materialističnem¹³ svetu. Lahko se torej strinjamo, da materialistične sile absolutno upravljajo naše obnašanje, vendar istočasno zavrnemo idejo, da so ljudje vedno in povsod motivirani z lastnim materialnim interesom.« (Frank, 1988, v Dennett, 2003: 216)

V material(ističnem) svetu seveda dominira tisto obnašanje, ki ima za posledico največji materialni uspeh. Ampak »vedno znova ugotavljamo, da najbolj prilagojeno obnašanje ne vznikne iz prizadevanja po materialni koristi. Zaradi pomembnih problemov z zavezo in izvedbo se bo to prizadevanje izkazalo za 'samoporazno'. Zato, da bi nam (materialno) šlo dobro, moramo včasih nehati skrbeti za maksimizacijo materialne koristi.« (Frank, 1988, v Dennett, 2003: 217)

Torej nam Frank ponuja razlago, kako altruizem oziroma moralo spraviti z evolucijo. Altruizem se dolgoročno materialno splača oziroma vsaj ne škodi:

Videti je, da gre altruistom tudi ekonomsko bolje: eksperimentalne študije vztrajno odkrivajo povezavo med altruističnim obnašanjem in socioekonomskim statusom. To seveda ne pomeni, da altruistično obnašanje *povzroča* ekonomski uspeh. Nakazuje pa, da altruistična drža ne more biti preveč resno breme v materialnem smislu. (Frank, 1988, v Dennett, 2003: 217)

¹² Zanimivo dejstvo je, da so bili študenti prvega letnika ekonomije bolj nagnjeni k sodelovanju kot študenti četrtega letnika ekonomije.

¹³ Materialno in materialistično je v tem primeru mišljeno v smislu gmotne koristi in ne filozofskega nauka.

3.7 Vznik ključnih značilnosti morale iz evolucijske oboroževalne tekme

Poglejmo torej, kam nas je pripeljal poskus naturalistične razlage morale:

Zapletenost družbenega življenja *homo sapiensa* kot bitja, ki ga bistveno določata jezik in kultura, ustvarja serijo evolucijskih oboroževalnih tekem, iz katere vzniknejo ključne značilnosti človeške morale: interes za iskanje pogojev, v katerih se lahko krepi sodelovanje; občutljivost do kaznovanja in groženj; skrb za ugled; nagnjenja na višji ravni k pozitivni samomanipulaciji, ki je ustvarjena zato, da okrepi samonadzor vpricho skušnjave; in sposobnost ustvariti zavezo, ki jo bodo drugi cenili in upoštevali. Takšne inovacije lahko uspevajo skupaj s posebnimi pogoji, s katerimi so se pojavile, ter lahko tako izpodrinejo kratkovidnost oziroma sebičnost bolj enostavnih organizmov. (Dennett, 2003: 218)

Skupaj z Dennettom in Frankom smo torej skicirali pot od začetkov življenja do pojava oseb kot akterjev, ki jim »svoboda predstavlja največjo moč in hkrati največji problem« (Dennett, 2003: 218). Dennett torej na naturalističnih temeljih ponudi precej prepričljivo razlago nastanka morale. Malo manj pa je prepričljiva njegova razlaga svobode. Tako kot se je njegove knjige »*Pojasnjena zaveza*« prijel naslov »*Odstranjena zaveza*«, so nekateri kritiki tudi knjigo »*Vznik svobode*« preimenovali v »*Vznik morale brez svobode*« (Mozes, 2010). Ne gre samo za negativno nastrojene kritike, ki ga vidijo kot dogmatičnega determinista. Številni avtorjev, ki sicer simpatizirajo z njegovim projektom, so mnenja, da Dennett svobodno voljo utemeljuje predvsem zaradi strahu pred negativnimi družbenimi posledicami, ki bi se po njegovem pojavile, če bi ljudje mislili, da nimajo svobodne volje. Dennett sam seveda ponuja ogromno priložnosti za takšno branje.¹⁴

Tudi Dennettova teza, da si ljudje želimo biti obravnavani kot moralno odgovorni, ker si želimo biti polnopravni člani družbe, kaže na pojmovanje svobodne volje in moralne odgovornosti kot neke družbeno dodeljene »zmožnosti«. ¹⁵ In ravno za takšen »družbeno realen« pojav gre – svobodna volja je po Dennettovih besedah pojav, ki se vzpostavi v družbi, tako kot denar ali glasba. Izven družbe je torej po Dennettu nesmiselno govoriti o svobodni volji in moralni odgovornosti.

V nadaljevanju si bomo pogledali, katere pogoje mora izpolniti posameznik, da bi mu družba lahko upravičeno pripisala svobodno voljo in moralno odgovornost, in

¹⁴ Tako na primer v svojem govoru ob sprejemu nagrade fundacije Erasmus Dennett pravi: »Fundacija Erasmus in jaz deliva mnenje, da bi nazor o svobodni volji kot iluziji lahko imel usodne posledice za družbo v primeru, da se ga argumentirano ne ovrže.«

¹⁵ Zdi se, da je to trditev težko spraviti z zgornjo trditvijo o usodnih družbenih posledicah nazora o iluzornosti svobodne volje. Ljudje, ki si *želijo* biti moralno odgovorni, bi torej, soočeni z iluzornostjo moralne odgovornosti in svobodne volje, v takšni meri spremenili svoje obnašanje, da bi to lahko imelo usodne posledice?

ali jih je sploh zmožen izpolniti. Glavna pogoja za takšno pojmovanje svobodne volje in moralne odgovornosti bosta racionalnost in samonadzor. Kriterijev seveda ne bodo zmožni izpolniti vsi posamezniki, vendar ne po lastni krivdi. Veliko vlogo igra namreč tudi sreča, namreč s kakšnimi predispozicijami se oseba rodi in kakšne vzgoje je deležna. Preiskali bomo torej človekovo glavno *differentio specifico* glede na njegovo biološko opredelitev: mišljenje oziroma razum (*sapientia*¹⁶). Iskali bomo možnost človekove avtonomije, ki jo lahko izpeljemo iz naturalistično razumljenega pojma razuma, iz tega, da človek pridobi oziroma ozavešti razloge, ki so pred njegovim nastankom bili zgolj »prosto lebdeči«, kar človeka naredi za *racionalno* žival.

4. Vznik razuma

4.1 Prosto lebdeči in ujeti razlogi

Do zdaj smo evolucijski proces in tako tudi razloge za sodelovanje ali odpadništvo utemeljevali brez kakršnega koli sklicevanja na razumevanje tega procesa in na razloge s strani akterjev, ki jih ta proces oblikuje. Razlogi lahko ostanejo popolnoma skriti svojim živalskim izvrševalcem in še vedno popolnoma učinkoviti v evolucijskem smislu. Lahko so prosto lebdeči oziroma *free floating rationale*, kot se pogosto izrazi Dennett. Čebeli, ki žrtvuje svoje življenje, ni treba vedeti, da je to naredila zato, da bi rešila preostalo populacijo čebelnjaka. Evolucija oziroma narava funkcionira na *need to know* osnovi (Dennett, 1984: 24). Do pojava *homo sapiensa* bi lahko rekli, da noben posameznik v živalskem svetu ni imel potrebe ali koristi od reflektiranja svojih razlogov ali interesov. Evolucija je nemoteno potekala naprej. Avtomatizirano, »programirano« obnašanje in odzivi na okolje so zelo učinkoviti, hkrati pa ne zahtevajo veliko energije, kot jo na primer zahteva pisanje tega stavka. Primeri, ko se nam kakšno živalsko obnašanje zdi nerazumno ali celo smešno,¹⁷ so popolnoma zasenčeni ob navdušenju nad razumnostjo celote stvarstva te »Absolutne nevednosti« (Dennett, 1984: 11). Vendar pa, kot smo videli zgoraj v primeru izogibalcev, predstavlja možnost za inovacije, ki se dogajajo *horizontalno* in s pomočjo »žerjavov«, za razliko od *vertikalnih*, ki potrebujejo več generacij, velik evolucijski napredek. Videti je, da se je investicija v razvoj energetske zelo zahtevnih možganov evolucijsko izredno izplačala.

Medtem ko smo zgoraj z lahkoto izpeljali skoraj celotno evolucijo brez sklicevanja na kakršnokoli vednost, pa vidimo, da že mladiči vrste *homo sapiens*, seveda ob posredovanju kulture in jezika, kažejo zelo veliko vnemo zlasti pri enem opra-

¹⁶ Čeprav se pojem *sapientia* nahaja v človekovi biološki opredelitvi, ga nikakor ne nameravamo samoumevno sprejeti kot dejstvo, ki bi utemeljilo možnost svobodne volje in moralne odgovornosti.

¹⁷ Ko osa iz rodu *Sphex* pred brlog prinese paraliziranega črčika, ga pusti zunaj in opravi pregled brloga. Znanstveniki so izvedli eksperiment tako, da so med njenim pregledom brloga črčika prestavili za nekaj centimetrov. Ko je osa prišla ven, je črčika zvelkla nazaj pred brlog in znova opravila svoj pregled brloga. Črčika so štiridesetkrat prestavili in »uboga« osa je štiridesetkrat, vedno znova, preverila brlog.

vilu: spraševanju »zakaj«, iskanju razlogov oziroma vzrokov. Otroke »že od majhnega spodbujamo, da se nam pridružijo v iskanju in podajanju razlogov ter utemeljevanju prihodnjega početja z razlogi« (Dennett, 2003: 273). Kultura, skupnost in zlasti jezik so najbolj botrovali evoluciji človekove zmožnosti *prepoznati, reflektirati* razloge svojega delovanja.

Jezik je izjemna evolucijska pridobitev, ki je omogočila izjemen razvoj številnih, če ne prav vseh, človekovih sposobnosti. In od nekdaj je jezik služil za komunikacijo razlogov. Če parafraziramo znani biblijski citat, lahko rečemo, da je na *začetku bil razlog*. Poglejmo si namreč, kaj vse so stari Grki razumeli pod besedo *logos*: beseda, govor; pa tudi razlaga, razlog; in prav tako razum. Latinski *ratio* pa obdrži pomena razum *in* razlog, prav tako iz njega nastali francoski *raison* in iz tega nastali angleški *reason*. Od latinščine dalje se izgubi povezava z jezikom, se pa obdrži povezava razlog – razum. Te povezave v slovenščini nimamo, je pa razum bistveno povezan z iskanjem oziroma podajanjem razlogov, z razlaganjem. Slovar slovenskega knjižnega jezika poda naslednjo razlago besede razum: sposobnost vključevati, sprejemati v zavest in ugotavljati vzročne, logične povezave.

4.2 Od razloga do razuma (»From reason to Reason«)

Če torej narava funkcionira na »*need to know*« osnovi, pa je človekov moto ravno obraten: pridobiti čim več relevantnih informacij glede okolja *in* glede sebe in drugih ljudi. Zlasti zaželeno so seveda informacije, med katerimi je možno vzpostaviti vzročno zvezo, torej poiskati vzroke oziroma razloge (»Zakaj?«). Pojem vzročnosti se nam kaže kot konstitutivna komponenta razuma, že če se ozremo na zgoraj navedeno slovarsko geslo, brez filozofskih teorij in kategorij.

Človeku je zlasti pomembna refleksija njegovih in tujih razlogov za neko delovanje. Zato je tudi bil skozi zgodovino svojega razvoja, kot smo zgoraj omenili, nagnjen pripisovati razloge in interese tudi naravnim silam in neživim stvarim.¹⁸ Tudi ko se je tega sčasoma odvadil, pa je bil še vedno prepričan, da pozna razloge svojega delovanja. Zato je bil najbrž najtežji od treh udarcev¹⁹ za njegov ego prav Freudov: večina razlogov, interesov, motivov, miselnih procesov, spominov ni dostopna introspekciji. Tisto, kar je dostopno, predstavlja samo vrh ledene gore.²⁰ Ne potrebujemo torej nobenih zlobnih nevroznanstvenikov iz filozofskih miselnih eksperimentov na temo svobodne volje in moralne odgovornosti. Tukaj imamo precej benevolentnega nevroznanstvenika, ki nam pravi, da ne poznamo večine vzrokov za svoja dejanja.²¹ Ampak to je samo pol problema, s katerim nas sooči

¹⁸ Zanimiva je komična »filozofska smer« rezistencializem, katere glavni slogan je »Stvari so proti nam«.

¹⁹ Prvi je Kopernikov udarec, da Zemlja ni središče vesolja, drugi pa Darwinov, da imajo človek, opica in vse ostalo življenje na Zemlji skupnega prednika, da človek ni narejen po podobi boga.

²⁰ John Bargh trdi, da se je Freud zmotil za vsaj decimalno, gre bolj za snežinko na vrhu ledene gore.

²¹ Še dva »benevolentna« nevroznanstvenika povezana s problemom svobodne in zavestne volje sta Benjamin Libet in Daniel Wegner.

Freud. Če je naše nezavedno dostopno psihoanalitikom, bi lahko torej z njihovo pomočjo dobili bolj pravilno in celovito sliko svojih lastnih vzgibov. Za možnost racionalnega delovanja in načrtovanja zahtevamo čimbolj natančno poznavanje vseh svojih razlogov, motivov in interesov, tako da lahko razumsko izbiramo med bolj in manj upravičenimi.

Druga polovica problema je, da je bil Freud prepričan, da tudi če smo seznanjeni z nezavednimi nagoni, ki vodijo naše ravnanje, nimamo možnosti bistveno spremeniti vzorcev svojega obnašanja. Vzorci, ki jim potem sledimo celo življenje, so po Freudu namreč bili fiksirani v zgodnjem otroštvu. Človek torej nima pomembne možnosti spremeniti navade, ki jih oceni za negativne, tudi če se jih zaveda. Navedimo primer, ki naj bi dokazoval, da so ozaveščeni razlogi kljub Freudovega mnenju lahko »praktični«. Ženska, ki se zaveda, da jo privlačijo nasilneži, se lahko zavestno začne izogibati krajem, kjer je večja verjetnost, da spozna nasilne moške. Freud bi najbrž odgovoril, da je to možno, ampak da bo najverjetneje kljub svojim frankfurtovskim željam drugega reda sledila svoji želji prvega reda (Frankfurt, 2007). Verjetnost, da spremenimo vzorce svojega obnašanja zaradi novih informacij, je po našem vsekakor večja, kot je bila pred pridobitvijo teh informacij. Vprašanje pa je seveda, ali je takšna verjetnost bistveno večja, zlasti v primeru najbolj trdovratnih vzorcev.

4.3 Svobodna volja in racionalnost

Ob zgoraj obravnavanih težavah se nam možnost navezati svobodno voljo na princip *popolne racionalnosti* zdi zelo problematična. In vendar je ideja podreditve volje nareku razuma zelo pogost pojav v literaturi o svobodni volji. Kant je mogoče res utemeljil svojo celotno moralno filozofijo na tej ideji, ampak nikakor ni bil osamljen niti ne prvi, ki je svobodo vezal na razum, med drugimi so podobno trdili Spinoza, Descartes in celo Locke. Povezavo med racionalnostjo in svobodno voljo pa na podoben način ugotavlja tudi John Searle. Po njegovem teorije racionalnosti skoraj neizogibno predpostavljajo neko stopnjo svobodne volje. Kakšen smisel bi namreč imela zmožnost racionalno narediti najboljši načrt za delovanje, če ne bi bili zmožni spremeniti svojega obnašanja tako, da sledimo temu najboljšemu načrtu. Brez tega bi, po njegovem, zavestni um (*conscious mind*) bil nemočen, depresiven opazovalec lastnih iracionalnih dejanj, medtem ko bi ugotavljal, da bi neko drugo dejanje bilo boljše (Searle, v Baumeister, 2008: 71). Vendar, kot smo videli pri Freudu, tudi takšna pesimistična razlaga ni izključena. Verjetno se dogaja, da se »opazujemo« pri ravnanju, za katerega smo ocenili, da ni najbolj racionalno, smotno. Ampak bi se tudi v tem primeru strinjali s Searlom, da bi *možnost* racionalnega ravnanja implicirala možnost svobode. Racionalno dejanje bi bilo bolj svobodno od iracionalnega, če naj svoboda karkoli pomeni, pravi Locke:

Če je resnična svoboda v osvoboditvi od vodstva razuma in v želji po omejitvi razmišljanja in presojanja, ki nam preprečujeta, da bi izbrali ali naredili slabše dejanje, potem so norci in neumneži edini svobodni ljudje. Menim, da se nihče ne bi odločil biti nor zavoljo takšne svobode, razen tistega, ki je že nor. (Locke, 1690, v Dennett, 1984: 51)

Po Freudu vsi kažemo nekatere psihopatološke znake v vsakdanjem življenju – vsi smo nekoliko nori, če se izrazimo bolj neposredno. Ampak ali ne nastopa ravno to kot omejitev naše svobode? Svobodo ravno vidimo v možnosti, tudi če popolnoma nerealni, ubežati ravnanju, za katerega ocenjujemo, da je manj racionalno in manj pod našim zavestnim nadzorom. Ali mogoče tudi ne?

Ideja svobode kot popolne racionalnosti je izražena v filozofski literaturi tako pogosto, pravi Dennett, da mora biti na njej nekaj resničnega. Ampak po drugi strani je izražena tako vehementno, da se v njej mora kot protiutež skrivati neka napaka, nekaj zaskrbljujočega in nasprotujočega naši intuiciji. Intuicija ali nekaj drugega očitno ustvarjata takšen odpor do te ideje, da se tem avtorjem zdi tako pomembno razglasati svoj vpogled²² (Dennett, 1984). Vzrok je najbrž v strahu pred avtomatizmom, izgubo spontanosti, individualnosti, ustvarjalnosti. Ampak ta strah predpostavlja nekatere zelo neverjetne lastnosti (človeške) racionalnosti. Ena takšna lastnost je, da bi v določeni situaciji bilo najbolj racionalno izbrati točno eno in isto dejanje, ne glede na ponovitve te situacije. Vprašanje je torej parafrazirano vprašanje svobodne volje: če popolna racionalnost obstaja, *ali bi lahko ravnal drugače?* Ali ne bi bilo najbolj racionalno *vsakič* ravnati enako, če predpostavimo identične okoliščine oziroma nekaj takega kot identično preteklost v primeru svobodne volje? Temu lahko ugovarjamo, da bi bilo nemogoče predvideti vse posledice vsakega od možnih dejanj, tudi če bi lahko poznali »vse relevantne podatke« (*all! things considered*). Odločitev se na koncu vedno zgodi tako, da premišljevanje oziroma racionaliziranje na neki točki prekinemo in se odločimo, ne pa da ga pripeljemo do absolutnega konca, ki bi zahteval neskončno časa. V takšni bitki s časom, in ob nepopolnih podatkih, bi se najbrž tudi najbolj popoln razum ob ponovljeni izbiri različno odločal. Takšna »spontanost« ima veliko prednosti. Največja je v tem, da na ta način lahko razvijamo vedno nove in bolj učinkovite načine odzivanja. Ljudje smo sicer popperjanska bitja, zmožna virtualno testirati različne scenarije in se tako izogniti posledicam slabih strategij, ki bi v velikih primerih bile skrajne pogubne. Ampak nekaterih »simulacij« enostavno ne moremo »zagnati« in izpeljati samo v možganih, oziroma vsaj celote njihovih posledic ne moremo predvideti samo z možgansko simulacijo.²³ Različnost odzivov na enake

²² Je v temelju te intuicije ravno strah, da nismo kompetentni racionalni akterji? Najbrž ne, spomnimo se Descartesove šaljive opazke, kako je očitno razum najbolj pravično razdeljena dobrina. Intuicija se vsekakor ne napaja iz Freuda.

²³ Podobno kot računalniške simulacije imamo namreč »težavo«, ker je vsakič potrebno okoliščine idealizirati, jih poenostaviti.

situacije je racionalna tudi zaradi večje nepredvidljivosti, kar je bistvenega pomena, glede na evolucijsko oboroževalno tekmo (Dennett, 1984: 71). Racionalnost torej nikakor ne nasprotuje spontanosti, individualnosti in ustvarjalnosti.

Torej skrb, da bi kot popolni racionalni akterji izgubili pristne človeške značilnosti in zdrsili na nižjo razvojno ali, človeku prav tako pomembno, moralno stopnjo,²⁴ ni utemeljena. Vseeno pa ima filozofska tradicija od Sokrata dalje težave, kako uskladiti človeka in razum. Tisti Searlov racionalni opazovalec lastnega iracionalnega početja ima po našem dve možni »racionalni« razlagi te diskrepance: *šibkost volje* ali *samoprevara*. Šibkost volje je torej prepoznanje diskrepance, »nemočno« opazovanje tega, kako naše želje prvega reda premagujejo želje drugega reda. Samoprevara pa je najbrž še večji udarec predpostavki racionalnega akterja. Beseda racionalizacija je že dolgo nazaj postala sopomenka upravičevanju tudi kot opravičevanju svojega početja drugim in samemu sebi. Mogoče pa samoprevara ne bi smeli tako kritično obravnavati. Možno bi bilo namreč tudi zagovarjati njene pozitivne evolucijske in psihološke vplive. Tudi če samo za trenutni namen predpostavimo, da je na primer občutek svobodne volje samoprevara, lahko trdimo, da gre za pozitivno samoprevaro, za katero smo že omenili, da ima veliko pozitivnih učinkov na človekovo življenje.

4.4 Neskončna lahkost avtomatiziranosti

Po zgoraj omenjenih iracionalnih »pošastih iz Ida«, kot Kihlstrom imenuje Freudove nezavedne motive naših dejanj, si zdaj pogledajmo novo verzijo *adaptivnega nezavednega* (*adaptive unconscious*) kognitivne psihologije (Kihlstrom, 2008: 165). John Bargh v svojem zagovoru *avtomatiziranosti* človekovega delovanja trdi, da se je Freud utegnil zmotiti za celo decimalko, da je zavestno delovanje, bolj kot vrh ledene gore, zgolj snežinka na vrhu le-te (Kihlstrom, 2008: 162) ali, kot sam nekoliko v šali pravi, avtomatično delovanje predstavlja 99,44% vsega človekovega delovanja (Bargh, 1997, v Kihlstrom, 2008: 162). Avtomatiziranost ne implicira nujno iracionalnosti delovanja, razen če seveda za racionalnost ne zahtevamo ozaveščenost, temveč delovanje na *avtopilotu*, brez vplivov zavestnega preudarjanja. Po Barghu namreč nismo ne tako zavestni ne tako svobodni, kakor smo si do zdaj mislili (Bargh, 1997, v Kihlstrom, 2008: 162), zato se moramo sprijazniti z *neskončno lahkostjo avtomatiziranosti*. Adaptivno nezavedno namreč po Barghu odlično opravlja različna pomembna dela, kot so zaznavanje okolja, opozarjanje pred nevarnostjo, zastavljanje ciljev in sprožanje dejanj na sofisticiran in učinkovit način, in predstavlja nujen in obsežen del visoko učinkovitega psihičnega aparata (Bargh, 1997, v Kihlstrom, 2008: 162). Bargh trdi, da je težko ubežati napovedi, da bo z napredkom znanosti o psiholoških pojavih vedno manj prostora za svobodno voljo in celo zavestno odločanje samo (Bargh, 1997, v Kihlstrom, 2008: 1).

²⁴ Glej Frankova izvajanja o racionalnosti popolne sodelovalne (moralne) države v zgornjem razdelku.

Ob takšnih raziskovalnih ugotovitvah, kot so zlasti Barghove in Wegnerjeve, in podobnih nadaljnjih napovedih, se zdi problem svobodne volje bistveno manj pomemben kot problem zavestnega odločanja. Če je namreč sama *zavestna volja* iluzija, kot se glasi naslov Wegnerjevega temeljnega dela (*The Illusion of Conscious Will*, 2002), ali ni s tem odpravljena tudi sama svobodna volja, ali pa bi lahko govorili o nezavedni svobodni volji? V primeru, da se Barghova teza o 99,44% avtomatiziranosti vsega človekovega delovanja in Wegnerjeva teza o iluzornosti zavestne volje izkažeta za resnični, s tem nikakor ne bo ogrožena oziroma odpravljena samo možnost svobodne volje. Obe tezi se bistveno dotikata tudi vprašanja mentalne vzročnosti. Zlasti Wegnerjeva teza o iluzornosti zavestne volje, ki implicira epifenomenalizem psihičnih stanj, ogroža možnost vzročnosti zavestne volje. Isti determinizem, ki ogroža možnost svobodne volje, predstavlja tudi problem za mentalno vzročnost zaradi predpostavke sklenjene vzročne verige vseh fizičnih pojavov. Vendar sta determinizem in vzročna sklenjenost fizičnega sveta skoraj neizogibni znanstveni predpostavki. Zdi se, da lahko tudi psihologija pridobi resen znanstveni status s pomočjo avtomatiziranosti in nezavednosti delovanja in mišljenja ter epifenomenalnosti psihičnih pojavov²⁵ (Kihlstrom, 2008: 172).

Poleg takšnih precej črnogledih, bolj teoretičnih vidikov so mogoče praktični vidiki tega, kar bi lahko imenovali kompatibilistična svoboda, ravno tako precej ogroženi s številnimi dognanji socialne psihologije, ki jih po Kihlstromu lahko razumemo kot trditev »Ljudje so neumni« (Kihlstrom, 2008). Če odmislimo tezo, po kateri bi nas popolnoma racionalno delovanje naredilo nesvobodne in nas ravno iracionalnost tako naredi bolj svobodne, predstavljajo ugotovitve o bistveno »zmanjšani« racionalnosti zelo resno grožnjo predpostavljeni svobodi racionalnega akterja. Poskusi kažejo, da se ljudje v vsakdanjem življenju močno zanašamo na predsodke, hevrstike in ostale podobne procese, ki pripeljejo do velikega števila napak v sodbah.²⁶ Nekatere primere bi sicer lahko razložili na podlagi omejene racionalnosti (*bounded rationality*), ki se zaradi kompleksnosti problema, časovne omejitve in omejenih informacij zateče k uporabi določene hevrstike, ven-

²⁵ Kihlstrom ugotavlja, da socialna psihologija kaže bistvene podobnosti behaviorizmu. Po Barghu je socialni psihologiji uspelo ravno tam, kjer behaviorizmu zaradi zanikanja nujnih znotrajosebnih psiholoških razlagalnih mehanizmov ni uspelo.

²⁶ Logične napake žal nikakor niso nujno moralno nevtralne, ampak so lahko zelo povezane s predsodki. V znanem poskusu Kahnemana in Tverskyja je 85% udeležencev zagrešilo »napako konjunkcije«. Udeleženci so dobili naslednje besedilo: »Linda je stara 31 let, je samska, odprta in zelo bistra. Diplomirala je iz filozofije. Kot študentko jo je zelo skrbela diskriminacija in družbena pravičnost. Prav tako je sodelovala v protijedrskih demonstracijah. Potem so morali odgovoriti, katera od spodnjih trditev je bolj verjetna:
Linda je blagajnik.
Linda je blagajnik in je aktivna v feminističnem gibanju.
V nekaterih izvedbah poskusa se je celo 85% udeležencev odločilo za drugo možnost. Stephen J. Gould je bil očaran ravno nad tem primerom. Zanimivo mu je bilo, kot pravi, ravno to, da kljub temu, da ve, da je konjunkcija dveh trditev manj verjetna kot vsaka posamična, nek človeček v njegovi glavi še vedno skače gor in dol in se dere nanj: »Saj ne more biti samo blagajnik – preberi si opis!«

dar obstajajo tudi dokazi, da se to dogaja zaradi avtomatiziranosti in nezavednosti teh procesov. »Ne gre samo za to, da ne razmišljamo dovolj intenzivno, ampak tudi za dejstvo, da nismo dovolj pozorni na dogajanje okrog nas in na naše lastno početje« (Kihlstrom, 2008: 173). Ob takšnih ugotovitvah in podobnih napovedih behaviorizma s kognitivnim obrazom, kot Kihlstrom imenuje socialno psihologijo, se nam iracionalnost zdi veliko večja ovira naši svobodi kot determinizem fizičnega nivoja.

5. Zaključek

V pričujočem članku smo na naturalističnih temeljih poskusili razviti kompatibilistično različico svobodne volje in moralne odgovornosti. V prvem delu smo raziskali evolucijo izogibalcev znotraj determinističnega sveta in dokazovali, da je smiselno govoriti o *determiniranem izogibanju*, da (ne)izogibnost ni lastnost osnovnega, fizičnega nivoja, ampak pojem znotraj višjega nivoja, s pomočjo katerega primerjamo zmožnosti preživetja različnih bioloških vrst. Takšna rešitev seveda ne bo prepričala tistih, ki determiniranost in neizogibnost obravnavajo kot sinonima in potemtakem tudi kot pojava z istega nivoja. Vendar v tem primeru ne smejo zahtevati, da neizogibnost pomeni karkoli več kot determiniranost.

V drugem delu smo s pomočjo aplikacije zapornikove dileme na evolucijo utemeljili možnost za naturalistično razlago morale, kar je prav tako dolgo časa veljalo za težko izvedljivo. Moralno obnašanje se je izkazalo za evolucijsko stabilno strategijo, ki omogoča večji in hitrejši napredek določene družbe. Za kompatibilistično razpravo o svobodni volji in moralni odgovornosti je to pomembno, ker dokazuje, da je možno utemeljiti moralo tudi brez močnega pojma svobodne volje in moralne odgovornosti.

Zadnji del članka je posvečen razumu, ki je najverjetneje najpomembnejši kompatibilistični vidik svobodne volje. Na razum je bistveno navezana *svoboda reflektivnega samonadzora*, kot Robert Kane (2005) imenuje zmožnost razumeti in reflektivno *ovrednotiti* razloge in motive, po katerih želimo ravnati ali naj bi ravnali, in zmožnost *nadzorovanja* našega vedenja v skladu s tako reflektiranimi razlogi. Kot smo pokazali, ima človek probleme pri obeh zahtevah te svobode.

Prvi problem se nanaša na zmožnost samoovrednotenja, zmožnost spoznanja razlogov in motivov, po katerih ravnamo ali naj bi ravnali. Prvi pomemben udarec predpostavki racionalnega akterja je zadal Freud s teorijo nezavednega. Socialna in kognitivna psihologija ter nevroznanost pa so, po nekaterih precej prepričljivih

interpretacijah obsežnega raziskovalnega materiala, pokazale na veliko stopnjo avtomatiziranosti in celo iracionalnosti delovanja psihičnega aparata.²⁷

V prvem delu obravnave smo videli, da se je zmožnost samoovrednotenja izkazala za problematično, kar lahko samo po sebi predstavlja problem za možnost reflektiranega samonadzora. Vendar bi tudi ob predpostavki popolne zmožnosti samoovrednotenja srečali probleme pri samem samonadzoru, kot je recimo primer odvisnika, ki ni zmožen delovati v skladu s sicer pravilnim samoovrednotenjem. V vsakdanjem življenju pa poznamo takšne primere pod imenom šibka volja, pa tudi primere, kjer zaradi neke nezavedne želje prilagodimo naš vrednostni sistem oziroma kognitivno disonanco blažimo s pomočjo samoprevare.

Vsekakor sta ta dva vidika zelo pomembna, in to ne samo za kompatibilistično svobodno voljo. Ne glede na obstoj neke višje, metafizične vrste svobodne volje, nam je namreč pomembno, da so naša spoznanja samih sebe in sveta pravilna in da je naše ravnanje skladno s temi spoznanji.

Naturalistic Aspects of Free Will and Moral Responsibility

Free will and moral responsibility can be assessed on two levels. On the ultimate level of evaluation they are difficult to reconcile with contemporary science and might even be incoherent. On the compatibilist level they can be defined as evolved capacities of humans as biological beings. The determinism of the basic level does not contradict such concepts of free will and moral responsibility. It even enhances the evolution of avoiders on the design level. Moral behavior can be established as the evolutionary stable strategy with the application of the prisoner's dilemma to the evolutionary theory, even without the strong metaphysical conceptions of free will and ultimate moral responsibility which compatibilism cannot provide. Compatibilist freedom requires the ability to understand and reflectively evaluate the reasons one wants to act upon, and to control one's behavior in accordance with such reflectively considered reasons. Contemporary psychological and neuroscientific research indicates that rational agents face problems with both capacities. It also shows that humans are highly susceptible to cognitive dissonance, self delusion and the like.

Keywords: free will, moral responsibility, compatibilism, evolution, rational agent

²⁷ Filozofija in naravoslovne znanosti že dolgo obravnavajo introspekcijo kot zelo sumljiv vir informacij o psihičnih procesih. Kljub umiku behavioristične psihologije pa je videti, da zgoraj navedene znanosti poskušajo popolnoma zaobiti zavest kot predpostavko, ki je ne potrebujejo, če parafraziramo slavni Laplacov komentar. Zavest tako ne samo da ni gospodar v hiši, pač pa je – če do konca potegnemo sklep iz visoke stopnje avtomatiziranosti in iracionalnosti – človek celo avtomat, pa še to bolj neumen.

Literatura

- Baumeister, R. F. (2008). »Free Will, Consciousness, and Cultural Animals«. V Baer, J. (ur.), *Are We Free? Psychology and Free Will*, Oxford: Oxford University Press, str. 65–85.
- Dennett, D. C. (1984). *Elbow Room: The Varieties of Free Will Worth Wanting*. Cambridge, Mass.: MIT Press.
- Dennett, D. C. (2003). *Freedom Evolves*. New York: Viking/Penguin.
- Frankfurt, H. G. (2007). »Svoboda volje in pojem osebe«. V Šuster, D. (ur.), *O svobodni volji: Od Leibniza do Libeta*, Maribor: Založba Aristej, str. 117–131.
- Kane, R. (2005). *A Contemporary Introduction to Free Will*. Oxford: Oxford University Press.
- Kihlstrom, J. F. (2008). »The Automaticity Juggernaut – or, Are We Automats After All«. V Baer, J. (ur.), *Are We Free? Psychology and Free Will*, Oxford: Oxford University Press, str. 155–180.
- Locke, J. (1690/1959). *An Essay Concerning Human Understanding* (A.C. Fraser Edition). New York: Dover.
- Mozes, E. (2010). »The Dogmatic Determinism of Daniel Dennett«. Dostopno na: <http://atlassociety.org/objectivism/atlas-university/deeper-dive-blog/3712-the-dogmatic-determinism-of-daniel-dennett>.
- Myers, D. G. (2008). »Determined and Free.« V Baer, J. (ur.), *Are We Free? Psychology and Free Will*, Oxford: Oxford University Press, str. 32–43.
- Smilansky, S. (2000). *Free Will and Illusion*. Oxford: Clarendon Press.
- Wegner, D. M. (2002). *The Illusion of Conscious Will*. Cambridge, Mass. in London: MIT Press.

Tine Kolenik

Pedagoška fakulteta Univerze v Ljubljani

Kaj imata skupnega Heidegger in Roomba: utelešena umetna inteligenca kot peskovnik kontinentalne filozofije

Robotski sesalec Roomba ne bi bil mogoč brez bogate zgodovine filozofskega izročila, ki predstavlja temelj klasične umetne inteligence. Načelo slednje je računska metafora za um, kjer je kognicija le manipuliranje arbitrarnih simbolov. A takšna paradigma ima številne težave, kot sta problem okvirja in ozemljitve simbolov. To so razkrile ideje Heideggerja, Merleau-Pontyja idr. kontinentalnih filozofov, ki so skozi robotika Brooksa, izumitelja Roombe, vodile nastanek utelešene umetne inteligence (UUI). UUI temelji na paradigmi nujnosti spoja agenta in okolja ter vznika kognicije iz senzomotorične zanke. A raziskovalci na tem področju ponujajo zelo različne interpretacije teh predpostavk. To ponovno prinaša težave, predvsem v obliki vprašanja avtonomije in nastanka pomena. Namen članka je predstaviti vzajemen odnos filozofija-UUI. Filozofijo vzpostavi kot ključno za razvoj (U)UI ter ponuja pregled trenutnih trendov in praktičnih primerov. Članek na koncu prikaže problematičnost sodobnih smernic ter ponudi morebitne rešitve, ki med drugim izvirajo prav iz filozofije.

Ključne besede: utelešena umetna inteligenca, utelešena kognicija, fenomenologija, klasična umetna inteligenca, enaktivizem

1. Uvod

Roomba, mali robotski sesalnik podjetja iRobot, je skupek nekaj senzorjev, s katerimi zaznava okolico (npr. odbijači, merilniki razdalj), in aktuatorjev, s katerimi deluje v okolju (npr. motorski pogon, sesalni mehanizem). Je tudi posledica velikega dela večstoletnega filozofskega izročila – najsibo to doktrina realizma (Ockham, Galileo, Hobbes, Locke, Hume, Moore, Russell, Kopernik, Descartes, Kepler...), idealizma (Berkeley, Kant, Schopenhauer, Nietzsche, Hegel...) ali

vmesne poti med njima (Heidegger, Merleau-Ponty, Wittgenstein...). Roombo je izumil avstralski robotik Rodney Brooks, oče utelešene umetne inteligence (UUI), ki je na prehodu iz 80. v 90. leta prejšnjega stoletja sprožil revolucijo na področju umetne inteligence (UI). Slogan brooksijanske revolucije je bil preprost: »najboljši model sveta je svet sam« (Brooks, 1997)¹. Brooks je na podlagi te ideje utemeljil svoj pristop k modeliranju inteligentnih robotov, kjer reprezentacije robota niso izhajale od znotraj, temveč so bile svet sam preko neprestanega zaznavanja. To se razlikuje od klasične UI, ki temelji na računski metafori uma kot računalnika in duševnih procesov kot manipuliranja arbitrarnih simbolov, kjer so bile reprezentacije in vse znanje vnaprej podani, programirani v robotih. Eden izmed glavnih vplivov na to revolucijo (in Roombo samo) je bilo sporočilo Heideggerja, da »pomenljivi predmeti, /.../ med katerimi živimo, niso model sveta, ki je spravljen v naših možganih ali umu, temveč svet sam« (Dreyfus, 1972: 265–266).

Heidegger, eden najpomembnejših sodobnih nemških filozofov, je predstavljal, podobno kot Brooks, reakcijo na takratni pogled na svet. Nasprotoval je temu, da je svet mogoče spoznati preko kvantifikacije vsakdanjega izkustva, zbiranja empiričnih podatkov. Trdil je, da tovrstna metodologija ne prinaša spoznavanja obstoječega sveta. Pomembno je znanje *kako* (npr. védenje, kako voziti kolo), ne znanje *da* (npr. védenje o fizičnih silah med vožnjo kolesa). Pomembno je naše delovanje v zvezi s kolesom. Heideggerjeva fenomenološka zapuščina je rodila nove pristope k raziskovanju prvoosebnega izkustva (Stolorow, 2006) in gibanja, ki so prinesla paradigmo utelešene kognicije (UK; Froese, 2007). Nemški fenomenolog je preko Dreyfusovih interpretacij na prehodu iz 80. v 90. leta prejšnjega stoletja močno vplival na Rodneyja Brooksa, velikega kritika dotedanje t.i. stare dobre UI (SDUI; ang. *Good Old-Fashioned Artificial Intelligence* – GOFAI; Dreyfus, 2007). Brooks je po več robotih, ki so utelešali novo, brooksijansko paradigmo, izumil Roombo, robotski sesalnik, ki se z neprestanim zaznavanjem sveta odziva z izredno učinkovitim sesanjem prahu. Ti roboti naj bi Heideggerjeve misli utelešali v tem, da spoznavajo svet z delovanjem, da pridejo preko tega delovanja do znanja *kako* in ne z znanjem *da*, kjer bi imeli že vse vnaprej programirano v sebi.

Za izumitev Roombe je bila torej potrebna dolga in naporna pot številnih mislecev, filozofskih pristopov, empiričnih dognanj in postavljenih paradigem iz najrazličnejših zgodovinskih obdobj. Povezava med filozofijo in UI je še bolj eksplicitna pri zagovornikih stališča, da mora robotika opravljati vlogo orodja za testiranje filozofskih stališč (Harvey, 2000; Froese in Ziemke, 2009). Članek poskuša, predvsem z vidika UI, na kratko opisati prav omenjeno pot, posebna pozornost pa je posvečena zanki med filozofijo in UUI, praktičnim primerom, težavam UUI (vprašanji avtonomije in pomena) in trendom, ki le-te poskušajo rešiti.

¹ Vse citate iz tuje literature je prevedel avtor.

2. Filozofija in stara dobra umetna inteligenca

Rojstvo UI (te, ki jo zdaj nekoliko pokroviteljsko kličemo SDUI, sicer pa klasična oz. simbolna UI) sega v leto 1956, ko se je odvila t. i. dartmouthska konferenca. Osrednji misleci le-te so bili McCarthy, Minsky, Newell, Samuel in Simon. Zaradi dosežkov njihovega dela, ki je navdušilo z navidez zelo človeškimi sposobnostmi (npr. igranje šaha in govorjenje angleščine), so si dartmouthovci zastavili ambiciozen cilj: odkriti algoritem, enačbo, s katero bi opisali človeško oz. splošno inteligenco, ki bi bila zmožna prepričljivo reševati med seboj zelo različne naloge (Crevier, 1993). Ta cilj je odseval močno predpostavko – da naša duševnost resnično operira na podlagi takšnega algoritma. Podobno je razmišljal tudi Turing, eden izmed najpomembnejših znanstvenikov preteklega stoletja. Računska metafora, um kot računalnik, je najmočnejšega in konkretnega utemeljitelja dobila prav v njem, ko je leta 1936 »definiral računanje kot formalno manipulacijo z neinterpretiranimi simboli, ki se izvaja s pomočjo formalnih pravil.« (Markič, 2010: 29). Prav na tem temelji Turingov stroj, ki lahko potencialno izvede prav vsako nalogo, če mu le lahko določimo neko končno število korakov, ki pripeljejo do rešitve. Turing je na podlagi tega umislil stroj, ki bi s takšno formalno strukturo posnemal mišljenje.

Na tem temelji metoda »zgoraj-navzdol« (»*top-down*«) raziskovalcev SDUI. Ti svojim programom podajo vse znanje, ki ga bodo potrebovali za reševanje danih nalog. Ko ti programi rešujejo naloge, se zanašajo na vnaprej vprogramiran repertoar znanja, na knjižnico, katere pot je od zgoraj, iz programa, navzdol, na nalogo iz okolja. Model sveta imajo že v sebi, rešitve za morebitne težave prav tako. Odgovore na vprašanja imajo že vnaprej definirane.

Dreyfus (2007) se spominja, kako so se študenti Newella in Simona v 1960-ih zgrinjali nanj z besedami, da filozofi v 2000 letih razmišljanja v naslanjačih še vedno ne razumejo inteligence, oni pa so v kratkem času v laboratorijih UI ravno v tem uspeli (dokazali naj bi, da inteligenca deluje na trditvah simbolne paradigme). A rojstvo UI je omogočilo prav bogato filozofsko izročilo. Pri tem izstopata dve doktrini: realizem in kartezijska filozofija, ki ju, čeprav se zdita na prvi pogled v nasprotju, SDUI obe predpostavlja. Realizem je doktrina, ki zagovarja stališče, da skozi čutila zaznavamo resnični svet, kakršen je izven nas; ta določa, kako iz njega izvirajoče predmete zaznavamo. Dualizem pa zagovarja, da sta duševnost in telo dve fundamentalno različni snovi, ki ju ne gre reducirati na eno ali drugo. A ravno to počne SDUI – če je misel le manipulacija simbolov, to pomeni, da je ločena od telesa in da le-to ne vpliva na kognicijo, inteligenco (Zebrowski, 2010). Sicer je o mišljenju kot računanju že v 17. stoletju pisal Hobbes, Descartes pa je v istem obdobju razmišljal o reprezentacijah. In to je le začetek – vplivi na UI so prisotni tudi v Leibnizevih univerzalnih karakteristikah, Kantovi univerzalnosti konceptov (koncepti kot pravila), formalni logiki Fregeja in Russella, logičnem atomizmu Wittgensteina itd. SDUI se torej zdi kot nekakšna aplikacija teh

idej (Froese, 2007). Ta aplikacija je še posebej uspešna v različnih ekspertnih sistemih (tj. sistemih, ki posnemajo človeško odločanje, npr. zdravstvena diagnostika; Russell in Norvig, 2003), programih za igranje strateških iger (npr. program za igranje šaha Deep Blue; prav tam) ipd. SDUI je sicer utemeljil tudi klasično informacijsko-procesno kognitivno znanost.

Zdi se torej, da SDUI ne bi bila mogoča brez filozofije. A prav filozofija je bila tista, ki je SDUI zadala tako močne kritike, da je bila paradigmatška sprememba nujna.

3. Filozofija in utelešena umetna inteligenca

Simon in Newell sta v 1950-ih izjavila, da bo računalnik v desetih letih svetovni šahovski prvak in da bo odkril ter dokazal nov matematični izrek (Russell in Norvig, 2003). V 1960-ih je Simon, potem ko računalnik teh dveh ciljev ni dosegel, nekoliko pompozno popravil izjavo in zatrdil, da bo računalnik v dvajsetih letih nadomestil človeka v opravljanju katerekoli naloge (Crevier, 1993). Isto desetletje je Minsky napovedal, da bo UI svoj prvotni cilj dosegla v eni generaciji, v 1970-ih pa, da bodo v nekaj letih zgradili stroj s človeško inteligenco (prav tam). Po takšnem vzorcu se je datum cilja, tj. obstoja stroja s človeško inteligenco, vedno znova prestavljal. A raziskovalci SDUI so še vedno verjeli, da je predvsem računska zmogljivost tista, ki jih loči od splošne UI.

Težav, ki so botrovale prestavljanju datuma splošne inteligence računalnikov, je bilo veliko več, kot so si jih sami raziskovalci upali priznati. SDUI je obtičala na mestu – in to relativno brez napredka ob siceršnjem tehnološkem razvoju. Programi, ki so temeljili na goli računski moči, so imeli velike težave z delovanjem, ki ga ljudje nenehno jemljemo za samoumevnega, predvsem pa ga ne enačimo z inteligenco. To je vključevalo predvsem vidno in slušno zaznavanje ter gibanje. Program, ki je premagal samega Kasparova, ni bil denimo sposoben nečesa, nam tako preprostega, kot je hoja. A to je bila le prva vrsta težav. Druga vrsta je bila bržkone bolj povezana s predpostavkami paradigme v zvezi s človeško inteligenco in kognicijo kot tako. Programi so imeli velike težave z učenjem novih reprezentacij, kategorizacijo, reševanjem nalog, za katere ni eksplicitnih pravil (torej delovanje s pomočjo ozadnjega znanja (*background knowledge*)) ipd. Prvi pogubni udarec je SDUI zadal problem okvirja. Programi so imeli težave z ugotavljanjem, katere informacije so relevantne za njihovo delovanje, saj programer enostavno ni vedel, v kakšni situaciji se bo program znašel. Drugi pogubni udarec je SDUI zadal filozof Searle z miselnim eksperimentom kitajska soba (Searle, 1980), katerega bistvo je, da stroj, ki operira z arbitrarnimi simboli, pač nima pra-

vega razumevanja.² Kitajska soba je vodila v prelomni Harnadov članek »The Symbol Grounding Problem« (1990). Ta prevprašuje ravno predpostavke SDUI. Iz kje vznikne pomen skozi simbole, ki so manipulirani le na podlagi njihove sintaktične, formalne oblike? Kako je lahko njihov pomen ozemljen kjerkoli drugje kot le v nadaljnjih brezpomenskih simbolih? Harnad ponudi potencialne rešitve problema, predvsem pa opozori na to, da se mora problem reševati od »spodaj-navzgor« (»bottom-up«) – simboli morajo biti ozemljeni v senzomotorični interakciji. A vse to je le odmev prvotne Dreyfusove kritike SDUI iz začetka 1970-ih, ki temelji na idejah Heideggerja.

Filozofija je torej, tako se zdi, igrala glavno vlogo pri pokopu SDUI. A takšno stanje v krogih SDUI je ustvarilo možnost za novo skupino mislecev in za novo paradigmo. Ta se je usidrala v kontinentalni filozofiji, predvsem v fenomenološkem izročilu Heideggerja, Husserla, Merleau-Pontyja, Sartra, kasneje pa tudi v idejah različnih sodobnikov paradigme UK, kot so Flores, Lakoff, Varela, Thompson, Noë, Gallagher, Clark, Haugeland, Wheeler idr.

Brooksijanska revolucija je neposredna posledica kritik SDUI, ki so izšle iz filozofije. Zdi se, da vse glavne Brooksove ideje nosijo misli Heideggerja in njegovih sodobnikov (Dreyfus, 2007). Heidegger je dokončno zavrnil dihotomijo med svetom in nami samimi, s tem pa postavil naš obstoj ali *Dasein* (tu-bit, biti-v-svetu) kot dejavnost znotraj te domene. To pomeni, da svet ne obstaja zunaj nas. Mi smo v njem in ne ločeni od njega. Iz tega izhajajo pomembne implikacije: kar zaznavamo, je odvisno od nas samih oz. tega, kar smo. Igramo netrivialno vlogo v oblikovanju resničnosti, predvsem dela, ki vpliva na nas same, na našo izkušnjo resničnosti. Svet seveda obstaja kot tak, a naše zaznavanje in konstrukcijo le-tega kujemo (tudi) mi. Posebno vlogo ima tu naše aktivno delovanje v svetu, preko katerega spoznavamo svet, pridobimo znanje *kako*. Primer tega je naš odnos do orodij. Heidegger (1953) pravi, da naša dejanja z različnimi orodji spremenijo zaznavno naravo samega orodja in našega odnosa z njim. Ko pišemo, smo tako zavzeti v tem dejanju pisanja, da preidemo v doživljajsko stanje nekakšnega zlitja s pisalom samim; pisalo postane utelešeno, ločnica med spoznavajočim subjektom in spoznavanim objektom se zabriše. Posledica tega je nekakšno izginotje materialnosti. Vse to vključuje preprosti slogan brooksijanske revolucije: »najboljši model sveta je svet sam«, ki razblinja dihotomijo med svetom, reprezentiranim v umu in svetom izven uma. A Brooksove ideje nosijo širšo zapuščino. Omeniti gre še Merleau-Pontyja, ki nadaljuje Heideggerjeve ideje. Vplivi so vidni v Brooksovi ideji inteligence brez reprezentacij in v njegovem grajenju robotov. Brooks nikoli ne gradi duha znotraj stroja, temveč utelešena bitja, ki delujejo v dialogu s svetom in nimajo potrebe po nekem ločenem umu za inteligenco, saj ta ni potreben za to, da utelešeno bitje izvaja kompleksne naloge v okolju.

² Tega sicer ne gre zamešati s še eno kritiko SDUI, ki jo izraža Blockov miselni eksperiment kitajski možgani (Bringsjord in Arkoudas, 2014).

Seveda brooksijanska revolucija ni potekala le znotraj področja UI, ampak tudi znotraj kognitivne znanosti, ki se je prav tako oddaljevala od simbolne paradigme in začela raziskovati predpostavke UK. Pomembnih mejnikov je znotraj kognitivne znanosti več, zato je težko govoriti o le enem poglavitnem akterju, kakršen je bil Brooks za UUI. Tako Lakoff (1987) pokaže, da jezik na vseh nivojih poteka skozi metafore, ki so ubesedene z dejanji (npr. prepir kot vojna, kjer braniš svoje mnenje, napadaš nasprotnikove argumente itd.). Maturana in Varela (1984) vpe-ljeta koncept strukturnega spoja (nujno medsebojno vplivanje dveh sistemov, npr. agent-okolje ali agent-agent, za katera je bistvena zgodovina interakcij). Tega kasneje Varela, Thompson in Rosch (1991) razširijo v močnejšo tezo utelešenosti in enaktivizma. Noč (2005) trdi, da zaznavanje ni nekaj, kar se nam zgodi, temveč je nekaj, kar mi počnemo. Razmah takšnih razmišljanj in raziskav, ki naj bi potrjevale teze UK,³ je precejšnji. Trenutni prevladujoči trend med raziskovalci UK (še posebej v nevroznanosti in UI) je, da »mišljenje ni manipuliranje simbolov, temveč notranje simuliranje s pomočjo senzomotorične kognicije«⁴ (Farkaš et al., 2012). Kognicijo obravnavajo kot namenjeno delovanju, ne računanju, ozemljeno v senzomotorični zanki in nerazdružljivo od telesa. Takšna stališča podpira kopica empiričnih dokazov. Nevrokomputacijski modeli kažejo, da se komunikacijske vrline razvijejo le, ko senzomotorično delovanje in razumevanje dejanj drugih delujeta skladno (Rizzolatti in Arbib, 1998); rezultati s področja nevroslikanja prikazujejo, da se ob misli na brco zasveti homunkularni del motoričnega korteksa za nogo (Pulvermüller, 2005); kognitivna neurofarmakologija pa izvaja izredno zanimive eksperimente, kjer omejitve telesa na nek način vpliva na naše čustveni odziv (na slednje je vplivalo vbrizganje botoksa v ustnice), presojanje in razumevanje (Havas et al., 2010).⁵

Omeniti velja, da je tovrstno razumevanje UK le najmočnejši trenutni trend, ki se kaže v kognitivni znanosti. UK kot raziskovalni krog nikakor ni enoten v predpostavkah in trditvah. Zdi se, da so se začetki, na katere je tako zelo vplivala fenomenologija in so nosili idejo močnejše utelešenosti, nekoliko omilili, če ne celo strivializirali (Froese, 2007). Kljub temu da vplivov fenomenologije ne gre zanikati, se zdi, da je vodilni trendi ne upoštevajo (Dreyfus, 2007). To je jasno tako v UK kot v UUI, kjer so vsi raziskovalci združeni le v imenu in v njihovi drži proti simbolni paradigmi. Gre torej bolj za to, proti čemu UK nastopa, in ne to, kaj UK

³ Za namen tega članka približno enačim utelešeno kognicijo (*embodied cognition*), ozemljeno kognicijo (*grounded cognition*), umeščeno kognicijo (*situated cognition*) in ostale vzporedne »paradigme«, ki med seboj niso strogo ločene. Razlike seveda obstajajo, a niso bistvene za ta članek.

⁴ Mišljenje torej poteka na podlagi našega zaznavanja in delovanja.

⁵ Pomembno je dodati, da tovrstne raziskave pritrjujejo le prevladujočemu trendu znotraj UK. Polje le-te ni enotno, kar pomeni, da takšne raziskave ne podpirajo vseh tez (npr. Varelove teze, da telo ne le vpliva na kognicijo, temveč jo konstituira). Polje predvsem družiti kritika simbolne paradigme.

zagovarja (Ziemke, 2004). Razlike se kažejo na več nivojih. Prvi primer bi bil odnos do reprezentacij – od tega, da jih imamo in da so ozemljene v našem senzomotoričnem sistemu, do radikalnejših smeri, ki trdijo, da reprezentacij nimamo. Drugi primer bi bil odnos do tega, za kaj sploh je kognicija – od informacijsko-procesne paradigme, ki trdi, da smo pasivni agenti, ki sprejemamo informacije iz okolja za reševanje nalog, do enaktivizma Varele, Thompsona in Rosch (1991). Slednji trdijo, da celotno bistvo organizma (njegov svet, kognicija, adaptacija, če jih lahko ločimo) vznikne v interakcijskem trikotniku uma, telesa in okolja. Različne vidike, ki se pojavljajo na področju UK, je združila Margaret Wilson: »1) kognicija je umeščena; 2) kognicija deluje pod časovnim pritiskom; 3) kognitivno breme prelagamo v okolje; 4) okolje je del kognitivnega sistema; 5) kognicija je za akcijo; 6) telo je osnova odklopljene kognicije« (Wilson, 2002: 625). Temu primerno je tudi raznoliko, razpršeno in veliko manj poenoteno polje UUI.

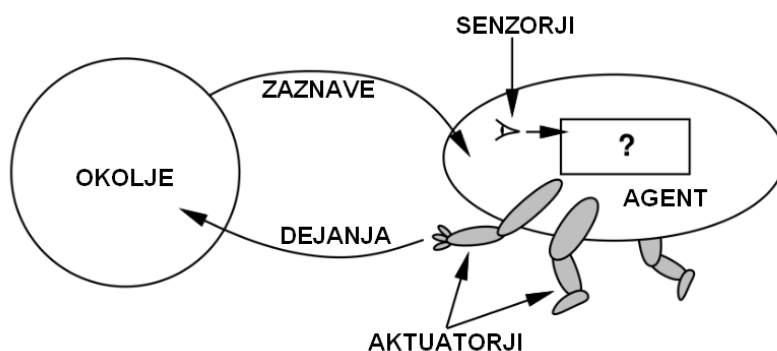
4. Utelesena umetna inteligenca v akciji

Znotraj UUI velja še močnejša težnja po postavljanju enotnih temeljev kot v kognitivni znanosti. Poskus slednjega predstavlja že Brooks v enem od dveh svojih prelomnih člankov, ki sta sprožila revolucijo. V »Intelligence without reason« Brooks (1995) našteje ključne značilnosti utelešenega sistema (str. 27):

1. [Umeščenost] Roboti so umeščeni v svet – ne ukvarjajo se z abstraktnimi opisi, ampak s *tukaj in zdaj* pričujočega sveta, kar neposredno vpliva na njihovo vedenje.
2. [Utelesenost] Roboti imajo telesa in dostopajo do sveta neposredno – njihova dejanja so del dinamike s svetom in imajo takojšnji zankasti učinek na njihovo zaznavanje.
3. [Inteligenca] Roboti izgledajo inteligentno – a izvor te inteligence ni le njihova koda. Inteligenca izhaja tudi iz njihove umeščenosti v svet, iz njihove pretvorbe dražljajev znotraj senzorjev in iz njihovega fizičnega spoja s svetom.
4. [Emergenca] Inteligenca sistema vznikne iz sistemovih interakcij s svetom in iz posrednih interakcij med njegovimi komponentami – včasih je težko pokazati na točen dogodek ali mesto znotraj sistema in zatrditi, da je to vzrok za neko dejanje.

Ko pride do modeliranja višjih kognitivnih procesov, se našteje značilnosti poskušajo realizirati z rabo umetnih nevronske mreže (UNM). UNM so se v disciplini ustalile zato, ker so osnovane na nevronske mreže človeških možganov, ker se lahko učijo, ker delujejo od »spodaj-navzgor«, ker so robustne, ker znanje ni diskretno shranjeno itd. UNM, dane v robotsko telo, tako za raziskovalce predstavljajo rešitev težav in potešitev kritik SDUI (Pfeifer et al., 2007). A večkrat kot ne je uporabljen hibridni pristop (s simbolnimi metodami) – »spodaj-navzgor« in

»zgoraj-navzdol«, nekaj podanega eksplicitnega znanja (sploh v primerih, ko gre za modeliranje zelo specifičnih kognitivnih procesov), diskretne vrednosti nevronov⁶ ipd. Raziskovalci skozi modeliranje nižjih kognitivnih procesov iščejo predvsem vznik obnašanja skozi spoj preprostih robotov in okolja. To poglavje bo zato razdeljeno v dva dela: prvega bodo predstavljali primeri modelov nižjih kognitivnih procesov (npr. senzomotorika), drugega pa primeri modelov višjih kognitivnih procesov (npr. abstrakcija). Tovrstno razdeljevanje izvira iz tega, da prevladujoči trendi na področju kognicijo obravnavajo kot kontinuum med nižjimi in višjimi kognitivnimi procesi; raziskovanje (in uspeh) z modeliranjem prvih je veliko pogostejše (Mirolli in Parisi, 2011). Ne glede na predmet modeliranja so utelešeni sistemi (oz. agenti), abstraktno gledano, funkcija iz zgodovine zaznav v dejanja ($f: P^* \rightarrow A$), izgledajo pa takole:



Slika 1: Prikaz senzomotorične povratne zanke med agentom in okoljem (Russell in Norvig, 2003: 32)

4.1 Primeri modeliranja nižjih kognitivnih procesov

Prvi sklop UII v praksi predstavljajo primeri modelov nižjih kognitivnih procesov. Pred UII se je na tovrstne procese, predvsem na motoriko in gibanje, gledalo kot na problem nadzora – telo je bilo le orodje, ki mora od »zgoraj-navzdol« prejeti pravilne ukaze. UII prikaže, da nižja kognicija (npr. gibanje) vznikne skozi dinamičen odnos interakcije telesa in okolja, ki zahteva zelo malo, če sploh kaj, ukazov od »zgoraj«.

Tega se je seveda zavedal že Brooks. Zaradi neuspeha SDUI pri izdelavi sistemov, ki bi bili sposobni opravljati preproste naloge v okolju, je že v zgodnjih 1980-ih v svojem laboratoriju na MIT začel z delom na robotu, ki bi se bil sposoben prilagajati in spreminjati z okoljem. Nastal je Herbert. Herbertova naloga je

⁶ To pomeni, da en nevron predstavlja točno eno stvar (npr. modro barvo).

bila navidez preprosta: tavati je moral po laboratoriju in metati prazne pločevinke pijač, na katere je naletel, v koš za smeti. Slednje mu je omogočila Brooksova posebna subsumpcijska arhitektura (*subsumption architecture*), ki so jo sestavljale med seboj povezane mreže, vsaka odgovorna za svojo nalogo (razlikovanje med polnimi in praznimi pločevinkami, izogibanje mizam itd.). Herbertova dejanja so vzniknila iz različnih aktivacij in inhibicij – tako je recimo miza aktivirala mrežo za izogibanje oviram, ta pa je posledično inhibirala mrežo za premikanje naprej. To naj bi bila fizična realizacija Brooksove ideje, da je svet sam najboljša reprezentacija. Herbert ni posedoval vnaprej sprogramiranega notranjega zemljevida prostora, ni bil opremljen z diskretnim znanjem o vseh možnih pločevinkah, ki jih laboranti ne bodo vrgli v smeti, njegovo obnašanje je vzniknilo od »spodaj-navzgor«, ne od »zgoraj-navzdol«. S tem je bil rešen problem okvirja. Vseskozi se je neposredno odzival na okolje, tukaj in zdaj (Clark, 1997).

Podobni poti so sledili nadaljnji modeli nižjih kognitivnih procesov. Nekateri so bili še bolj preprosti od Herberta – ne le, da niso imeli vprogramiranega kontrolnega centra »tam zgoraj«, od koder bi prejeli ukaze, imeli niso niti aktuatorjev. S tem so raziskovalci želeli še radikalneje prikazati, kako je lahko svet sam najboljša reprezentacija. Tovrstni pristop temelji na različnih eksperimentih, ki raziskujejo dinamiko med telesom in okoljem, iz katere vznikne določeno (vsaj navidezno) inteligentno obnašanje. Primer tega je raziskava, kjer so mrtvo ribo položili v umetno generirane vodne tokove, telo ribe pa se je nemudoma prilagodilo okolju ter se obnašalo in obvladovalo tokove (telo se je samodejno premikalo proti toku), kot da bi bila riba živa (Liao et al., 2003). Tovrstne raziskave so že v začetku 1990-ih med drugim navdihnile t.i. pasivne dinamične sprehajalce (*passive dynamic walkers*). Slednji so sposobni hoje po različnih vzdolžnih klančinah brez kakršnihkoli aktuatorjev in vgrajenega nadzora. Nimajo ne motorjev ne senzorjev ne mikroprocesorjev. Uspešna hoja navzdol, tudi ob spreminjajočem se, nepoznanem in nejasnem terenu, vznikne preprosto iz dinamike med klancem in telesom robota (McGeer, 1990).

Emergentno vedenje, ki ni bilo predvideno in vprogramirano v robota samega, ki ni eksplicitno podano v sistemu in ne poteka od »zgoraj-navzdol«, je bilo eno glavnih vodil raziskovalcev UUI. Preprost primer takšne emergence iz 1990-ih je kvadrasti robot, ki je imel na vsaki strani po en infrardeči senzor (Wilson in Golonka, 2013). Preko njega je robot ob zaznavi predmeta spremenil smer – robot se je torej izogibal oviram. Dan je bil v okolje s kockami, lažjimi od sebe, razmetanimi po prostoru. Ko so mu raziskovalci odklopili sprednji senzor, se je začel zaletavati v kocke in jih potiskati naprej. Ko so med potiskanjem neprednji senzori zaznali oviro (zid ali kocko), je robot nehal potiskati kocko pred seboj in se obrnil. To je povzročilo, da je robot praktično pospravljaval kocke, ki so končale zložene v kupih. Pospravljanje je torej vzniknilo skozi zelo preproste elemente prostora in samega robota.

Raziskovalci UUI se pri modeliranju večkrat še neposredneje zatečejo v naravo. Primer tega je robotski črček (Webb, 1995). Črčki se reproducirajo tako, da samice iščejo samce, ki se najglasneje oglašajo. Med seboj povezane bobenčke imajo na sprednjih nožicah, ki so povezane z vmesnimi nevroni za obračanje. Celoten sistem črčka se odziva le na specifične frekvence. Samice se torej vedno obrnejo proti samcu, ki je najglasnejši. Ravno to so raziskovalci reproducirali v robotskem črčku z majhnim številom zelo osnovnih mehanizmov.

Kompleksnost modelov s takšnimi pristopi seveda variira. Tako poskušajo raziskovalci odkriti različne rezultate emergence s kompleksnim oblikovanjem človeških mišic (emergenca sistema tetiv, ki uravnavajo kopico dejanj, za katera se ne potrebuje možganov *per se*), oči (emergenca paralakse gibanja), koordinacijskih vzorcev (emergenca preživetja v Sahari) itn. (Pfeifer et al., 2007).

Modeliranje nižje kognicije v UUI je prek koncepta emergence pokazalo, da lahko kompleksno obnašanje vznikne iz le nekaj preprostih komponent robota, ki je postavljen v pravo okolje (prav tam). To je doseženo brez eksplicitnih reprezentacij ali neposrednega programiranja takšnega vedenja. Modeli torej kažejo na to, da lahko koncepta utelešenosti in umeščenosti nadomestita kompleksne notranje algoritme in vnaprej podano znanje.

4.2 Primeri modeliranja višjih kognitivnih procesov

Področje modeliranja višjih kognitivnih procesov je precej drugačno kot prejšnje. Raziskovanja je veliko manj, čemur botrujejo tudi težave pri modeliranju takšnih procesov (na kar opozarjajo tudi kritiki, večinoma pristaši SDUI) – enega izmed razlogov za to navaja Clark (1997). Takšni procesi naj bi zahtevali več reprezentacij (*representation-hungry problems*). Emergenca ni tako očitno v ospredju, velikokrat je prisotna kompleksna kognitivna arhitektura (iz UNM), vedenje modelov je manj očitno, manj vidno. Zaradi specifičnosti vseh najrazličnejših procesov je modeliranje izredno osredotočeno, razpršeno in raznoliko.

Med modeli višjih kognitivnih procesov se najde nekaj takšnih, ki poskušajo svoj cilj doseči s pristopi modelov nižjih kognitivnih procesov (emergenca skozi zelo osnovno mehansko delovanje robota v povezavi z okoljem). Primer tega je bil ambiciozen projekt Brooksa, ki je skupaj s filozofom Dennettom ustvaril robota Coga (Dennett, 1997). Njegov cilj je bil nivo človeške inteligence (predvsem govor, manipulacija predmetov, raziskovanje, učenje itd.), do katere bi prišlo popolnoma od »spodaj-navzgor«. Na odpadu je končal leta 2003, brez pretiranega pompa okoli neuspeha. Vznikanje kompleksnega obnašanja, ki bi se približalo višjim kognitivnim procesom, najdemo sicer že v 1980-ih, in sicer v poznanih Braitenbergovih vozilih (Braitenberg, 1984). Braitenberg je bil psiholog, ki je trdil, da lahko z roboti simuliramo miselni eksperiment, kjer čustveno (psihološko) obnašanje vznikne skozi interakcijo med zelo preprostimi telesi več agentov (ali med agentom in okoljem) z le enim senzorjem in enim aktuatorjem. Njegovo na-

čelo je bilo, da je sinteza lažja kot analiza – ustvariti nekaj, kar izgleda kompleksno, je lažje, kot analiza nečesa, kar izgleda kompleksno. Tako je robota s senzorjem, ki ob zaznavi svetlobe sorazmerno požene kolesa, dokler svetlobe ne zaznava več, strah svetlobe. V njej je viden, kar ogroža njegov obstoj. Zelo preprost sistem simulira torej strah. Če bi senzor in motor robota povezali tako, da bi se robot premikal proti svetlobi in v njegovo okolje postavili še enega robota, ki je enako zgrajen, bi dobili ljubezen. Obema bi dodali le še en aktuator – lučko. S tem bi se privlačila. Tako nekako je razmišljal Braitenberg.

A takšni modeli delujejo naivno in ne oponašajo procesov, ki se dogajajo v ljudeh. Večina raziskovalcev vseeno sledi temu, da višji kognitivni procesi izvirajo iz nižjih. Velikokrat je le vprašanje, kaj je tisto, kar premosti ta most med obema vrstama procesov. Kaže se močan trend, ki trdi, da je jezik tisti, ki nižjo kognicijo potisne do višje (Mirolli in Parisi, 2011). Jezik naj bi nižjo kognicijo vodil do učenja, kategorizacije, abstrakcije, spomina, samokontrole in bogatega mentalnega življenja. Veliko raziskovalcev višjih kognitivnih procesov v UUI prav tako trdi, da so rešili problem ozemljitve simbolov (*symbol grounding problem*; Steels, 2008) – nekateri pri tem povzdigujejo prav vlogo jezika kot kognitivnega orodja, ki pripelje do vznika pomena s senzomotoričnem delovanjem v okolju (Farkaš et al., 2012). Gre za modele, ki v senzomotorični interakciji z okoljem, predmeti v njem in jezikovnimi signali, uspejo povezati jezikovni signal z nekim dejanjem. Do tega pride z učenjem (konkretno gre za spreminjanje uteži (umetnih sinaps) v UNM) od »spodaj-navzgor«.

Druge raziskovalne skupine jeziku ne pripisujejo takšne pomembnosti in modelirajo agente, ki naj bi konstruirali pomen brez jezika. Pomembno jim je, da agent skonstruirajo pomene, ki so pomembni zanj – do tega pride s pomočjo kognitivnih procesov, preteklih izkušenj in njihovih interakcij z okoljem. S tem naj bi prišlo do utelešenega znanja iz zaznavanja in delovanja v okolju. Primer takšnega raziskovanja je model, ki je postavljen v okolje z igračami, sadjem in pohištvo (Takáč, 2008). Agent lahko med raziskovanjem v vsakem koraku naključno izbere predmet in ga ali dvigne ali pa položi, to pa povzroči spremembe različnih parametrov (sila, stopinje premikanja agentovih aktuatorjev). V primeru dvigovanja se predmetu poveča vertikalna komponenta, če je sila agenta večja od teže predmeta, sicer dejanje nima efekta. V primeru spuščanja se predmetu zniža vertikalna komponenta glede na spremembo stopinje premikanja, a le, če predmet ni na tleh. Po enem dejanju agent zapiše povzročene spremembe v obliki kavzalnih razmerij med dejanji, predmeti in dinamiko med obema. Tako ustvari različne kategorije, ki so med sabo različno povezane, pomen pa tvori predmet, povezano dejanje, kavzalno razmerje med njima itd. Gre torej za neke vrste »danosti« (*affordances*) delovanja v prostoru, zgodovine tega delovanja, odnos med agentom in predmeti ... torej širokim diapazonom odnosov, ki se skujejo med agentom in okoljem, in ti za agenta predstavljajo pomen nekega predmeta.

Nekateri krogi raziskovalcev poskušajo napraviti še korak dlje, kar se tiče abstraktnosti, in modelirati doživljanje (Schillaci et al., 2016). Ti prav tako trdijo, da doživljanje izhaja iz senzomotoričnih interakcij z okoljem. Agent se uči uporabljanja orodij s poskušanjem ali opazovanjem človeka, izkušnje učenja pa je sposoben tudi podoživljati. Z uporabo nižjih kognitivnih procesov rekreira nekakšen svet, po katerem vznikne doživljanje. To rekreiranje služi tudi učinkovitejšemu delovanju v svetu, npr. uporabljanju orodij.

Področje UUI, ki se ukvarja z modeliranjem višjih kognitivnih procesov, je torej izredno raznoliko in razkropljeno. Skupna točka je nenehen pritisk kritikov, ki so večinoma zagovorniki SDUI. Težko je namreč modelirati višje kognitivne procese brez privzemanja nekega vnaprej podanega znanja; redki so modeli, ki bi bili izključno »spodaj-navzgor«. To krha predpostavke UUI, da kognicija izhaja od »spodaj-navzgor«. S tem se ponovno pojavi vprašanje pomena, ker do tega ne pride od »spodaj-navzgor«, temveč ga, kot v SDUI, določijo programerji. Vprašanje pomena vodi do vprašanja avtonomije, ki je zaradi istih razlogov modelom ne gre pripisovati – ti se namreč ne vedejo glede na notranje cilje, ki bi bili resnično pomembni za njih same, in ne za njihove kreatorje. Tovrstne težave vodijo do kritike, da UUI preprosto ni dovolj utelešena – ali vsaj ne pravilno (Froese, 2007).

5. Filozofija in potreba po novih paradigmah

UUI, ki le težka pojasnjuje višje kognitivne procese le na podlagi svoje paradigme, je zato vplivala na filozofske prakse, ki se ukvarjajo s tovrstnimi vprašanji. Nekateri filozofi so raziskave UUI pripeljale do trditev, da um ne služi za razmišljanje, kot je tradicionalno okarakteriziran, temveč za dejanja, za to, da stori-
mo tisto, kar v resničnem svetu pod časovnim pritiskom pač moramo (Clark, 1997), da je senzomotorika fizično vpletena v konceptualno procesiranje (Rupert, 2009) ali pa to, da je vidno zaznavanje nekaj, kar počnemo (Hickerson, 2007) oz. da sestavlja našo zavest (Gangopadhyay et al., 2010). V splošnem je prišlo do pojava močnih kritik v prid simbolni paradigmi ali pa do radikalizacije pogleda na utelešenost (Chemero, 2009). Nastali sta dve šoli misli: kompatibilistična in puristična (Calvo in Gomila, 2008). Prva trdi, da lahko za razlago različnih kognitivnih procesov uporabljamo različne metode – nekateri procesi bodo razloženi s simbolnimi, drugi pa z utelešenimi. Druga pa popolnoma zavrača simbolno paradigmo in se zavzema za to, da bi bile za raziskovanje kognicije rabljene izključno utelešene metode, brez kognitivističnih teorij (npr. teorija dinamičnih sistemov, konceptualna integracija idr.).

Z začetkom 21. stoletja se zdi, da se je raziskovanje na področju UUI nekoliko ustavilo. Razlogov za to je več. Prvič, glavni akterji so prenehali z udejstvovanjem znotraj področja (Ziemke je nehal s pisanjem člankov, Brooks je ustanovil podjetje iRobot in izdeluje Roombe, Pfeifer, ki je konec 90. prinesel področje v Evropo, se je upokojil itd.). Drugič, samo področje se je močno usmerilo, tj. izoblikovala sta se dva tabora: tehnični (Bongard, Lipson, Floreano) in filozofski (Di Paolo, Froese). Tretjič, področje še vedno ni paradigmatsko poenoteno. In četrtič, pojavile so se močne kritike (in ne s strani zagovornikov SDUI) prevladujočih trendov UUI in UK.

Kritikam Dreyfusa (2007), ki se od 1970-ih niso pretirano spremenile in še vedno opozarjajo, da je UI premalo heideggerska (slednje velja tako za UUI kot za trende UK), so se pridružili predvsem zagovorniki iz filozofskega tabora UUI (Di Paolo, Froese). Gre za kritike, da roboti v UUI še vedno ne ustvarjajo resničnega pomena, pomena, ki bi bil bistveno pomemben za njih same, ki bi izhajal iz njih samih glede na to, kaj potrebujejo za preživetje (Dreyfus, 2007). Heideggerjeve ideje torej niso dovolj realizirane – kljub temu da znanje *da*, vnaprej podano znanje, ni več osrednje, se zdi, da morebitno znanje *kako*, ki naj bi ga zdaj roboti imeli, ni znanje njih samih (temveč prej programerjev). Kot stvari stojijo, je pomen, ki naj bi vzniknil v robotih, pomemben za programerje, ki robotu zadajo to nalogo – s tem tudi same ozemljitve simbolov ni v robotu. Cilje postavijo programerji in ne izvirajo iz robotov.

Ta težava pomena je povezana še z enim konceptom, ki je ključnega pomena: avtonomijo.⁷ Avtonomija je koncept, ki se dandanes v širši UI pojavlja zelo trivialno – če ima robot ločeno telo od svojega kreatorja in če lahko sam izpolnjuje določene naloge v prostoru (kot je seveda vnaprej določen), potem je avtonomen. A avtonomen je lahko le, če cilji izhajajo iz njega samega in so le-ti pomembni zanj – ne za njegovega kreatorja –, šele takrat pa seveda lahko vznikne pomen, ki bo imanenten robotu samemu (Kolenik, 2016). Trenutni trendi UUI tega ne omogočajo niti ne raziskujejo v pravi smeri, saj se raziskovalci ne zavedajo svoje zmote v razmišljanju, da je namreč ustvarjanje pomena lahko ukazano, vsiljeno od njih (Froese, 2007). Zdi se torej, da filozofija še drugič v zgodovini UI zadaja tako močne kritike, da so paradigmatske spremembe nujne (Froese in Ziemke, 2009). Tako Dreyfusova heideggerska kritika iz 1970-ih. še vedno velja: modeli trenutno ne »najdejo pomena v svetu in se ne redigirajo v prepoznavanju za njih relevantnih delov sveta glede na njihovo obliko telesa, način premikanja, osebno in družbeno identiteto itd.« (Dreyfus, 2007: 265).

⁷ Avtonomijo definiram kot določanje lastnih zakonitosti.

6. Zaključek

Morda smo vsi zgrešili neko organizacijsko načelo bioloških sistemov ali pa vsaj neko splošno resnico o njih. Morda obstaja način, ki nam bo preko opazovanja bioloških sistemov omogočil vpogled v to inherentno nujnost interakcij med deli le-teh, ki je do sedaj popolnoma umanjkal iz sistemov umetne inteligence. /.../ Mislim, da trenutno preprosto ne razumemo in da je potrebna bistvena sprememba v našem mišljenju /.../. (Brooks, 1997: 300)

S temi treznimi besedami je Brooks pospremil kritike in očitne težave, s katerimi se bori UUI. Slednja je sicer več kot le močno stopila na prste simbolni paradigmi in SDUI ter pokazala, da delovanje v resničnem svetu ne mora potekati od »zgoraj-navzdol« s predprogramirano knjižnico znanja, ker duševnost pač ni manipulacija z arbitrarnimi simboli, ki ima že vnaprej postavljena pravila. A pristopi brez reprezentacijskega odzivanja na dražljaje od »spodaj-navzgor« ali obravnavanja uma kot simulacije nižjih kognitivnih procesov ali katerega podobnega tretjega trenda, ki se podpisuje pod termin UK, težka razlagajo višje kognitivne procese, vsekakor pa ne rešujejo globljih vprašanj avtonomije in pomena. Vračanje k SDUI ali nekakšna radikalizacija UUI se zdita le še več istega, saj se še vedno ne pripoznava pomembnost netrivialne avtonomije, iz katere šele lahko izhaja pomen. Brooks je, kljub svojim zmotam, imel prav: potrebna je bistvena sprememba v razmišljanju. In ravno to ponujajo nove paradigme (ali »resnične«, »prave« paradigme UK), ki se razvijajo v krogih filozofije in UI. Takšni sta, med drugim, enaktivizem (Varela idr., 1991) in umetno življenje (*artificial life*).

Enaktivizem bi lahko šteli med te »prave« utelešene paradigme. Med prvimi je namreč močno poudaril pomembnost telesa in njegovega odnosa z okoljem, a so se njegove ideje nekoliko izgubile, omilile ali nekako strivializirale, še posebej v UUI (Rohde, 2010). Enaktivizem, ki briše ločnico med spoznavajočim subjektom in spoznavanim objektom, poskuša rešiti težavo pomena prav preko razreševanja vprašanja avtonomije. Vzrok avtonomije je po enaktivizmu *autopoiesis* (Maturana in Varela, 1984). *Autopoiesis* označuje sistemov (biološki) proces samoustvarjanja in samoorganizacije preko neprestanega proizvajanja sposobnosti proizvajanja samega sebe. To povzroči, da je sistem avtonomen. In le avtonomen sistem lahko ustvarja pomen, ki je ozemljen v doživljanju, v *Daseinu*. Umetno življenje kot področje UI veliko pozornost posveča adaptaciji, ki je nekako enačena s kognicijo. Če je torej sistem zmožen avtonomne adaptacije, gre za kognitiven sistem. Ta avtonomija mora biti posledica okoljskih pogojev in ne vnaprej pripravljenega robota (Froese in Ziemke, 2009). Obe paradigmi sicer predstavljata manjšino znotraj svojih področij, zaradi relativne mladosti pa tudi nista tako dobro razdelani kot npr. simbolna paradigma. Kljub temu predstavljata možnost razrešitve nekaterih problemov, ki pestijo filozofijo, UI in kognitivno znanost že dolgo časa (Rohde, 2010).

What Do Heidegger and Roomba Have in Common? Embodied Artificial Intelligence as a Playground for Continental Philosophy

Roomba, the robotic vacuum cleaner, would not exist today without the rich history of philosophy, which presents the groundwork for traditional artificial intelligence. The latter sees cognition as an arbitrary symbol manipulation in its “mind as computer” metaphor. This view led to many problems, such as the “frame” and the symbol grounding problem, which sparked the AI revolution. The ideas of Heidegger, Merleau-Ponty and of other continental philosophers, which were realized by the roboticist Brooks, Roomba’s creator, were among the driving forces of embodied artificial intelligence (EAI). Its basis is the emergence of cognition from the sensorimotor loop between an agent and the environment. However, the immensely dispersed field of EAI is without common paradigmatic grounds. Aware of its problems regarding autonomy and meaning, this paper focuses on the philosophy-EAI relationship. It identifies philosophy as crucial for (E)AI, and offers a review of current trends and real-life examples. The paper concludes with a critique of current trends, and suggests some solutions originating from philosophy.

Keywords: embodied artificial intelligence, embodied cognition, phenomenology, traditional artificial intelligence, enactivism

Literatura

Braitenberg, V. (1984). *Vehicles: Experiments in synthetic psychology*. Cambridge, Mass.: MIT Press.

Bringsjord, S. in Arkoudas, K. (2014). »Philosophical Foundations«. V Ramsey, B. in Frankish, K. (ur.), *The Cambridge Handbook of Artificial Intelligence*, New York: Cambridge University Press, str. 34–63.

Brooks, R. A. (1995). »Intelligence without reason«. V Steels, L. in Brooks, R. A. (ur.), *The artificial life route to artificial intelligence: Building embodied, situated agents*, Hillsdale: Lawrence Erlbaum Associates, Inc., str. 25–81.

Brooks, R. A. (1997). »From earwigs to humans«. *Robotics and Autonomous Systems*, 20(2-4), str. 291–304.

Calvo, P. in Gomila, T. (2008). *Handbook of Cognitive Science: An Embodied Approach*. Oxford: Elsevier Ltd.

Chemero, A. (2009). *Radical Embodied Cognitive Science*. Cambridge, Mass.: MIT Press.

Clark, A. (1997). *Being There: Putting Brain, Body, and World Together Again*. Cambridge, Mass.: MIT Press.

Crevier, D. (1993). *AI: The Tumultuous Search for Artificial Intelligence*. New York: Basic Books.

- Dennett, D. (1997). »Consciousness in human and robot minds.« V Ito, M., Miyashita, Y. in Rolls, E. T. (ur.), *Cognition, computation and consciousness*, Oxford: Oxford University Press, str. 17–30.
- Dreyfus, H. L. (1972). *What Computers Can't Do*. New York: MIT Press.
- Dreyfus, H. L. (2007). »Why Heideggerian AI Failed and How Fixing it Would Require Making it More Heideggerian«. *Philosophical Psychology*, 20(2), str. 247–268.
- Farkaš, I., Malík, T. in Rebrová, K. (2012). »Grounding the meanings in sensorimotor behavior using reinforcement learning«. *Frontiers in Neurobotics*, 6(1).
- Froese, T. (2007). »On the Role of AI in the Ongoing Paradigm Shift within the Cognitive Sciences«. V Lungarella, M., Lida, F., Bongard, J. in Pfeifer, R. (ur.), *50 Years of Artificial Intelligence*, Berlin: Springer-Verlag, str. 63–75.
- Froese, T. in Ziemke, T. (2009). »Enactive artificial intelligence: Investigating the systemic organization of life and mind«. *Artificial Intelligence*, 173(3-4), str. 466–500.
- Gangopadhyay, N., Madary, M. in Spicer, F. (ur.). (2010). *Perception, Action and Consciousness*. New York: Oxford University Press.
- Harnad, S. (1990). »The symbol grounding problem«. *Physica D*, 42, str. 335–346.
- Harvey, I. (2000). »Robotics: Philosophy of Mind Using a Screwdriver«. V Gomi, T. (ur.), *Evolutionary Robotics: From Intelligent Robots to Artificial Life*, Ontario: AAI Books, str. 207–230.
- Havas, D. A., Glenberg, A. M., Gutowski, K. A., Lucarelli, M. J. in Davidson, R. J. (2010). »Cosmetic Use of Botulinum Toxin-A Affects Processing of Emotional Language«. *Psychological Science*, 21(7), str. 895–900.
- Heidegger, M. (1953). *Being and Time*. Tübingen: Max Niemeyer Verlag.
- Hickerson, R. (2007). »Perception as Knowing How to Act: Alva Noë's Action in Perception«. *Philosophical Psychology*, 20(4), str. 505–517.
- Kolenik, T. (2016). »Embodied Cognitive Robotics, the Question of Meaning and the Necessity of Non-Trivial Autonomy«. V Markič, O., Gams, M., Strle, T. in Kordeš, U. (ur.), *Cognitive science: proceedings of the 19th International Multiconference Information Society – IS 2016*, Ljubljana: Institut Jožef Stefan, str. 24–27.
- Lakoff, G. (1987). *Women, fire, and dangerous things: What categories reveal about the mind*. Chicago: University of Chicago Press.

- Liao, J., Beal, D. N. Lauder, G. V. in Triantafyllou, M. S. (2003). »Fish exploiting vortices decrease muscle activity«. *Science*, 302, str. 1566–1569.
- Markič, O. (2010). *Kognitivna znanost: filozofska vprašanja*. Maribor: Aristej.
- Maturana, H. R. in Varela, F. J. (1984). *The Tree of Knowledge: The Biological Roots of Human Understanding*. Boston, Mass.: Shambhala.
- McGeer, T. (1990). »Passive Dynamic Walking«. *The International Journal of Robotics Research*, 9(2), str. 62–82.
- Merleau-Ponty, M. (1962). *Phenomenology of perception*. New York: Humanities Press.
- Merleau-Ponty, M. (1963). *The structure of behaviour*. London: Methuen.
- Mirolli, M. in Parisi, D. (2011). »Towards a Vygotskian cognitive robotics: The role of language as a cognitive tool«. *New Ideas in Psychology*, 29, str. 298–311.
- Noë, A. (2005). *Action in Perception*. Cambridge, Mass.: MIT Press.
- Pfeifer, R., Bongard, J. in Grand, S. (2007). *How the body shapes the way we think: A new view of intelligence*. Cambridge, Mass.: MIT Press.
- Rizzolatti, G. in Arbib, M. (1998). »Language within our grasp«. *Trends in Neurosciences*, 21, str. 188–194.
- Pulvermüller, F. (2005). »Brain mechanisms linking language and action«. *Nature Reviews Neuroscience*, 6, str. 576–582.
- Rohde, M. (2010). *Enaction, Embodiment, Evolutionary Robotics*. Berlin: Springer.
- Russell, S. in Norwig, P. (2003). *Artificial Intelligence: A Modern Approach* (2nd ed.). Upper Saddle River, NJ: Prentice Hall.
- Rupert, R. (2009). *Cognitive Systems and the Extended Mind*. New York: Oxford University Press.
- Schillaci, G., Hafner, V. V. in Lara, B. (2016). »Exploration Behaviors, Body Representations, and Simulation Processes for the Development of Cognition in Artificial Agents«. *Frontiers in Robotics and AI*, 3.
- Searle, J. (1980). »Minds, brains, and programs«. *Behavioral and Brain Sciences*, 3, str. 417–424.
- Steels, L. (2008). »The symbol grounding problem has been solved, so what's next?«. V de Vega, M., Glenberg, A. in Graesser, A. (ur.), *Symbols and Embodiment: Debates on Meaning and Cognition*, Oxford Scholarship Online, str. 223–244.
- Stolorow, R. D. (2006). »Heidegger's investigative method in *Being and Time*«. *Psychoanalytic Psychology*, 23(3), str. 594–602.

Takáč, M. (2008). »Autonomous construction of ecologically and socially relevant semantics«. *Cognitive Systems Research*, 9, str. 293–311.

Varela, F. J., Thompson, E. in Rosch, E. (1991). *The embodied mind: Cognitive science and human experience*. Cambridge, Mass.: MIT Press.

Webb, B. (1995). »Using robots to model animals: a cricket test«. *Robotics and Autonomous Systems*, 16, str. 117–134.

Wilson, A. D. in Golonka, S. (2013). »Embodied cognition is not what you think it is«. *Frontiers in Psychology*, 4(58).

Wilson, M. (2002). »Six views of embodied cognition«. *Psychonomic Bulletin & Review*, 9(4), str. 625–636.

Zebrowski, R. (2010). »In dialogue with the world: Merleau-Ponty, Rodney Brooks and embodied artificial intelligence«. *Journal of Consciousness Studies*, 17(7-8), str. 156–172.

Ziemke, T. (2004). »Embodied AI as science: Models of embodied cognition, embodied models of cognition, or both?«. V Iida, F., Pfeifer, R., Steels, L. in Kuniyoshi, Y. (ur.), *Embodied Artificial Intelligence*, Heidelberg: Springer-Verlag, str. 27–36.

Martina Tomšič

Filozofska fakulteta Univerze v Ljubljani

Reprezentacije pri telesnem zavedanju

Motnje pri doživljanju lastnega telesa kažejo na to, da so reprezentacije telesa konstitutivne za telesno zavedanje. V članku obravnavam razlikovanje med dvema vrstama reprezentacij telesa, med zemljevidi telesa in podobami telesa, kot ga opredeli nevroznanstvenik Damasio. Slednji obe vrsti reprezentacije obravnava kot vzorca, ki se oblikujeta na dveh različnih ravneh – prva na živčni in druga na mentalni ravni. Prva raven je temelj telesnega zavedanja samega sebe (sebstva), šele na drugi ravni pa lahko govorimo o polnem samo-zavedanju, tj. kakor izraz »jaz« uporabljamo v vsakodnevni rabi, ko imamo v mislih sami sebe.

Ključne besede: utelešena kognicija, reprezentacije, telesno zavedanje, shema telesa, zemljevid telesa

1. Uvod

Dandanes si filozofskega preučevanja kognitivnih procesov brez upoštevanja vloge telesa praktično ni več mogoče predstavljati. Če pa ga že nekateri avtorji postavijo v oklepaj, mu vsaj načeloma pripoznajo pomen oziroma vlogo, ki jo kot del organizma ima (možgani so del celote, so del telesa). S pripoznanjem telesa kot aktivnega akterja pri kognitivnih procesih pa se obseg in prepletenost posameznih procesov znotraj celotnega organizma poveča. To ima za posledico tudi različno razumevanje pojmov, ki jih uporabljamo za razlago.

Za lažjo umestitev reprezentacijskih pristopov v utelešeni kogniciji bom v prvem delu članka na kratko osvetlila zgodovinski razvoj kognitivnih znanosti. V osrednjem delu članka se bom osredotočila na preučevanje telesnega zavedanja z reprezentacijskega vidika znotraj utelešene kognicije. Pri tem bom na kratko povzela, zakaj sploh ohraniti koncept reprezentacij pri filozofskem obravnavanju telesnega zavedanja, in predstavila, kako razumeti reprezentacije v tem kontekstu in v kakšnem odnosu so do telesa (reprezentacije telesa).

2. Zgodovinsko ozadje teorije utelešene kognicije

Z razvojem kognitivnih znanosti v zadnjih petdesetih letih se je raziskovanje s kognitivnih procesov razširilo na praktično celoten spekter duševnih procesov, tudi na čustva, zavest in telesne občutke. Kot povzema Markič: »Predmet raziskave so bili v začetnem obdobju predvsem kognitivni procesi v ožjem smislu – zaznavanje, učenje, pomnjenje, mišljenje in jezikovno sporazumevanje, v zadnjem času pa se je raziskovanje razširilo na področje čustev in zavesti, pa tudi socialnih, kulturnih in ekonomskih vidikov« (2010: 14). Vzporedno se je okrepilo tudi filozofsko preučevanje teh pojavov z izkustvenega, fenomenološkega vidika. Tako je uveden dodaten vidik pristopa k obravnavanju filozofskih problemov kognicije.

2.1 Kratka primerjava klasične kognitivne znanosti, teorije dinamičnih sistemov in teorije utelešene kognicije

Za začetke kognitivnih znanosti je značilno, da so znanstveniki preučevali predvsem višje kognitivne procese, zaradi česar je bil uporabljen pristop k obravnavi procesov od zgoraj navzdol. Znotraj tega pristopa je še vedno aktualna in splošno priznana delitev ravni pristopanja, kot jo je na primeru teorije vida in vidnih procesov opredelil nevroznanstvenik Marr (2000). Ločil je tri ravni analize, ki jih po njegovem ima vsak sistem, ki procesira informacije. Te različne ravni analize, ki se medsebojno dopolnjujejo, opredeli kot: □

1. računska raven – kaj je cilj računanja in zakaj je ustrezen;
2. algoritemska in reprezentacijska raven – kako lahko računsko teorijo implementiramo, oziroma, natančneje, kaj so reprezentacije za vhodne in izhodne podatke in kakšni so algoritmi za transformacije;
3. implementacijska/fizična raven – kako so algoritmi in reprezentacije fizično realizirani.

Klasičen kognitivističen model lahko opredelimo z naslednjimi značilnostmi (Markič, 2010):

- znanstveno raziskovanje kognitivnih procesov v ožjem smislu (zaznavanje, pomnjenje, reševanje problemov, ...);
- duševne reprezentacije tvorijo simbolni sistem, jezik misli, ki je neposredno realiziran v fizičnem smislu;
- razlaga poteka »od zgoraj navzdol«;
- funkcionalna raven je tista, ki omogoča razlago in napoved vedênja.

Z razvojem nevroznanosti in zaradi težav, ki so se pokazale pri klasičnem kognitivističnem pristopu, so znanstveniki iskali še alternativne oblike procesiranja. Rezultat tega prizadevanja so konekcioniistični modeli ter nevronske mreže. Oba pristopa, tako konekcionizem kot klasični pristop, še vedno podobno odgovarjata na vprašanje *Kaj je kognicija*, razlikujeta pa se pri odgovoru *Kako kognicija deluje* (Markič, 2010).

Za konekcioniistični pristop je značilno (Markič, 2010):

- procesiranje je podobno procesiranju na ravni bioloških nevronske mreže;
- reprezentacije so posledica učenja in prilagajanja mreže na okolje ter so emergentne lastnosti celotne mreže;
- razlaga poteka »od spodaj navzgor«;
- konekcionizem pokaže, kako lahko okolje sooblikuje kognitiven sistem.

Iz navedenega je razvidno, da konekcionizem in nevronske mreže izpostavijo vpetost kognitivnega sistema v okolje in v telo, s čimer utrejo pot teoriji dinamičnih sistemov ter utelešeni kogniciji. Spremeni se tudi pogled na reprezentacije. Te v klasični kognitivni znanosti tvorijo jezik misli (*language of thought*; Fodor, 1975) in so temeljni gradnik kognitivne teorije, v konekcionizmu in utelešeni kogniciji pa niso podane vnaprej, temveč so posledica učenja in stika organizma (oz. telesa) z okoljem.

Teorija dinamičnih sistemov, ki je že od sredine devetdesetih nepogrešljivo orodje utelešene kognitivne znanosti, daje poudarek razvoju oz. spreminjanju sistemov skozi čas – predmet znanstvenih študij so tako razumevanje, modeliranje in napovedovanje načinov, kako se vedénje sistema spreminja skozi čas. V zadnjih desetletjih so na ta način raziskovali vedénje na področju bioloških, kognitivnih in družbenih sistemov (Richardson in Chemero, 2014).

Znotraj kompleksnih dinamičnih sistemov lahko prepoznamo tri glavne značilnosti (Richardson in Chemero, 2014):

1. sestavljeni so iz številnih interaktivnih komponent ali dejavnikov, ki so lahko homogeni (npr. zbirka kortikalnih področij) ali heterogeni (npr. možgani v telesu, ki je umeščeno v okolje);
2. sistemi kažejo emergentno vedénje (tj. njihovo kolektivno vedénje kaže koherenten vzorec, ki ga ni mogoče predvideti na podlagi obnašanja posameznih komponent ločeno);
3. emergentno vedénje je samo-organizirano (ni rezultat akterja, ki vse skupaj nadzira).

Vse tri značilnosti lahko na primer lepo razberemo pri jati ptic. Te se zbirajo v jate, ki lahko obsegajo od sto do tisoč primerkov in dajejo vtis ogromne, valujoče gmote, poenotenega vedénja (sprememba oblike gmote). Interakcija med posamezniki tako ustvari kolektivne vzorce velikega obsega.

Ena od temeljnih postavk utelešene kognicije je ta, da je za kognicijo pomembno tudi telo in ne le možgani. Če je tradicionalna kognitivna znanost raziskovala predvsem kognicijo v ožjem smislu (osrednje kognitivno procesiranje), ločeno od telesnih mehanizmov, kot je senzorično procesiranje in motorični nadzor, jo utelešena kognitivna znanost obravnava v širšem kontekstu (tudi telesa). (Wilson in Foglia, 2011)

Za pozicijo utelešene kognicije lahko potemtakem rečemo, da zagovarja globoko zakoreninjenost kognitivnih procesov v interakciji telesa s svetom. Kot sem že omenila, je pristop utelešene kognicije precej raznolik. Vseeno se mi zdi, da Wilson (2002: 626) s tem, ko izpostavi šest trditev, dobro povzame glavne značilnosti.

1. Kognicija je umeščena – kognitivna dejavnost se dogaja v kontekstu dejanskega okolja ter inherentno vključuje zaznavo in dejavnost.
2. Kognicija je pod časovnim pritiskom – kognicijo moramo razumeti glede na to, kako v okolju deluje pod časovnim pritiskom dejanskega časa.
3. Kognitivno delo odlagamo v okolje – ker smo v svoji zmogljivosti procesiranja informacij omejeni (npr. delovni spomin, pozornost), izkoriščamo okolje na način, da si zmanjšamo kognitivno delovno obremenitev. Informacij se poslužujemo le, ko jih potrebujemo.
4. Okolje je del kognitivnega sistema – prepletenost med duševnostjo in okoljem je tako obsežna in kontinuirana, da znanstveniki ne obravnavajo duševnosti same kot pomenske enote analize.
5. Kognicija je za dejanje¹ – kognitivni mehanizmi so obravnavani glede na njihov doprinos k ustreznemu vedanju v situaciji.
6. *Offline* kognicija temelji na telesu – dejavnost duševnosti (tj. senzoričnega procesiranja in motoričnega nadzora) temelji na mehanizmih, ki so se razvili za interakcijo z okoljem.

Pri obravnavi zavedanja telesa kot ga predlagam sama, bodo največjo vlogo igrale značilnosti, ki so opredeljene v trditvah ena in pet.

Kljub poudarjanju razlik znotraj posameznih pristopov, med njimi vendarle obstajajo osnovne značilnosti, ki so jim skupne in na podlagi katerih so medsebojno primerljivi. Katere so potemtakem tiste bistvene razlike med klasičnim simbolnim pristopom in utelešeno kognicijo, kot jo sprejemam tudi sama? Povzetek treh glavnih razlik med obema je predstavljen v naslednji razpredelnici.

¹ Angleški izraz *action* slovenim kot *dejanje* in *acting* kot *dejavnost*.

Klasičen kognitivističen vidik	Vidik utelešene kognicije
Kognitivna znanost je znanost o kognitivnih procesih (notranjih stanjih), ki so razumljeni kot manipuliranje s simboli (računanje).	Znotraj modela duševnosti pride do spajanja z zunanjim svetom – kognitivni procesi so omejeni z obliko utelešenja, z okoljem in dejavnostjo organizma ter so usmerjeni v spoznavanje v realnem času.
Kognicija kot pasivni priklic	Kognicija kot aktivna konstrukcija, osnovana na utelešenih, ciljno usmerjenih dejavnostih organizma.
Reprezentacije kot simbolni sistemi	Reprezentacije telesa

Tabela 1: Primerjava vidika klasične kognitivne znanosti in utelešene kognicije

Seveda ima vsak izmed predstavljenih pristopov svoje pomanjkljivosti in prednosti. Na primer, klasični modeli se kažejo kot precej uspešen pristop za modeliranje višjih kognitivnih funkcij, ki zahtevajo uporabo jezika, konekcionistični modeli so uspešni pri prepoznavanju vzorcev, pristop teorije dinamičnih modelov in »utelešene kognicije« pa dokaj uspešno dajejo primat razumevanju bioloških rešitev za preproste naloge.

2.2 Reprezentacijski razlagalni oris v utelešeni kogniciji

Znotraj teorij utelešene kognicije je pristopov k razlagi kognitivnih procesov več, pri čemer nekateri pri svoji razlagi ohranjajo reprezentacije, drugi ne. Clark (1997), na primer, izhajajoč iz vprašanja vloge reprezentacij, opredeli ločnico med kognitivno znanstveno razlago, ki povsem zavrača reprezentacije (t.i. »radikalno različico utelešene kognitivne znanosti«) in med tisto, ki uporablja tako reprezentacijske kot ne-reprezentacijske razlagalne orise, vendar pa zavrača klasične simbolne reprezentacije (gre torej za manj radikalno različico) (Clark, 1997: 148–149). Srž pristopa radikalne različice je v opuščanju starih orodij preučevanja, ki naj jih nadomestijo nova: »Utelešeno kognicijo je najboljšo preučevati s pomočjo nekonceptualnih in nereprezentacijskih idej ter razlagalnih shem, ki jih vključujejo, npr. orodja teorij dinamičnih sistemov« (prav tam: 148). Sam Clark pa je do takega pristopa kritičen.

de Vignemont predlaga, da morda ločnica med radikalnim in neradikalnim utelešenjem, kot jo predlaga Clark, ni najbolj ustrezna, saj je narejena zgolj na osnovi tega, ali posameznik pri svojem pristopu uporablja reprezentacije, ne pa tudi same opredelitve utelešenja. Predlaga torej, da se pri razlikovanju reprezentacijskih pristopov upošteva tudi opredelitev utelešenja samega. Predlagam naslednjo razvrstitev (prilagojeno po Goldman in de Vignemont, 2009 ter Alsmith in de Vignemont, 2012: 3):

1. Razlaga prisotnosti oziroma raba telesa:
 - a) anatomija telesa: deli telesa imajo, zaradi svoje značilne anatomije, pomembno vzročno vlogo v kogniciji.
Primer: Če bi ljudje imeli sistem za eholokacijo, tako kot ga imajo netopirji, bi svet zaznavali drugače. Fizično telo (brez možganov) torej vpliva na nara-vo naše zaznave.
 - b) dejavnosti telesa: dejavnosti in druge značilnosti povezane s telesom (npr. te-lesna drža) imajo pomembno vzročno vlogo pri kogniciji.
Primer: obrazni *feedback* – dejavnost mojih obraznih mišic vpliva na moje razpoloženje ali čustvo (t.i. »*fake it to make it*« učinek).
2. Razlaga reprezentacij telesa, ki so to:
 - a) po vsebini (telesna vsebina): mentalne reprezentacije s telesnimi vsebinami imajo pomembno vzročno vlogo pri kogniciji.
Primer: reprezentacije imajo lahko za vsebino npr. dele telesa, reprezentirajo roko, nogo.
 - b) po obliki oz. kodi (telesna oblika): mentalne reprezentacije v raznih telesnih oblikah ali kodah imajo pomembno vzročno vlogo pri kogniciji.
Primer: vidna reprezentacija, slušna reprezentacija itd.

Vsebina in oblika reprezentacij telesa lahko sovpadajo, ni pa nujno. Različne oblike se namreč lahko delno prekrivajo z vsebino – na primer vidna reprezentacija roke, taktilna reprezentacija roke idr.

Alsmith in de Vignemont glede na to razumevanje utelešene kognicije razlikujeta med močnim in šibkim utelešenjem. Kot močno utelešenje (ali močno utelešeno stališče) razumeta »vsak vidik, ki daje jasno razlagalno vlogo telesu« (Alsmith in de Vignemont, 2012: 3). Tovrstna opredelitev močnega utelešenja sicer ne opredeljuje odnosa do reprezentacij, vendar pa avtorja menita, da ravno to vzdržanje od opredelitve reprezentacije dopušča (saj jih ne izključuje). Nasprotno pa »vidik, ki daje jasno razlagalno vlogo reprezentacijam telesa, medtem ko ne daje tudi jasne razlagalne vloge samemu telesu« (prav tam), poimenujeta kot šibko utelešene-ega (oziroma šibko utelešenje). V skladu s to opredelitvijo sta 1. a) in b) stališči močnega utelešenja, 2. a) in b) pa sta stališči šibkega utelešenja. Avtorja sta pre-pričana, da bo tovrstno razlikovanje v pomoč pri razlagi nekaterih kognitivnih procesov, povezanih s telesom.

Sama se bom v nadaljevanju osredotočila na pristop utelešene kognicije, ki pri višjih kognitivnih procesih ohranja reprezentacijski oris, saj se, kot ugotavlja tudi Markič (2010), z ukinitvijo reprezentacij, ki je prisotna v radikalni različici teorije dinamičnih sistemov in utelešene kognicije, postavlja vprašanje, ali taka teorija sploh še lahko razloži višje kognitivne procese, kot je na primer logično sklepa-nje, zavest idr. Te procese lahko razumemo kot *offline* procese, ki jih Wilson (2002) obravnava v šesti trditvi. Zanje velja, da navkljub svoji utemeljenosti na

telesnih procesih in interakciji z okoljem vseeno omogočajo abstraktno razmišljanje, ki presega trenutno situacijo, v kateri se nahaja organizem.

Tudi številni sodobni nevroznanstveniki (Damasio, 2000a; 2012; Blanke et al., 2004; Gallese, 2003) pri svojih raziskavah in razlagah kognitivnih procesov ohranjajo pojem reprezentacije. Prav tako ga ohranjajo tudi nekateri filozofi kognitivne znanosti (O'Shaughnessy, 2000; de Vignemont, 2010; Shapiro, 2011; Metzinger, 2000). Gre predvsem za tiste avtorje, ki svoje delo sicer utemeljujejo na raziskavah nevroznanstvenikov.

V nadaljevanju članka predstavim reprezentacije telesa in njihovo vlogo pri kogniciji, konkretno pri Damasiju kot avtorju, ki ga lahko umestimo znotraj pristopa šibkega utelešenja.

3. Reprezentacije telesa: sheme, zemljevidi in podobe

Reprezentacijski pristop utelešene kognicije, kot smo videli, predvideva reprezentacije telesa za razlago kognitivnih procesov. Kriterijev za razdelitev in opredelitev je veliko. Avtorji v svojih delih uporabljajo različne izraze, nekateri za reprezentacije telesa uporabljajo izraz shema telesa, drugi podoba telesa, nekateri oboje ter razlikujejo med obema vrstama reprezentacij telesa, tretji spet poizkušajo vpeljati nove izraze, npr. načrt, model idr. Ker splošno sprejeta definicija teh izrazov ne obstaja, je njihova raba dokaj arbitrarna, kar lahko vodi v precejšno zmedo tako na ravni posameznih avtorjev kot znotraj same stroke. To še posebej velja za izraza shema telesa in podoba telesa, ki se jima pripisuje različne pomeni tekom razvoja kognitivnih znanosti. Konceptualne težave, ki se pojavljajo v zvezi s tema dvema izrazoma, je do sedaj obravnavalo več avtorjev, kot v uvodu svojega članka izpostavi že Gallagher (1986), de Vignemont (2010) pa je podala analizo na podlagi nevropsiholoških motenj.

Posledica tega, da je izraz shema konceptualiziran na različne načine, je ta, da avtorji – v izogib nejasnostim – vpeljujejo nove, drugačne izraze, kar pogosto pripelje do dodatne zmede. Na primer, Haggard in Wolpert (2005) namesto sheme uporabljata izraz načrt (ang. *scheme* namesto *schema*), Metzinger (2000) uporablja izraz model, s katerim pravzaprav označuje to, kar Damasio poimenuje zemljevid.

V nadaljevanju bom predstavila Damasijev model reprezentacij telesa, ki mu sledi kratka analiza in izpostavitve nekaterih konceptualnih preprek, s katerimi se znanstveniki soočajo, kot tudi prepletanje samih procesov in posledično tudi pojavov.

3.1 Zemljevidi in podobe telesa pri Damasiju

Antonio Damasio je najbolj poznan po svojih nevroznanstvenih raziskavah, ki so čustvom dodelile pomembno vlogo v družbeni kogniciji in pri sprejemanju odločitev. Njegovo delo se osredotoča na razumevanje živčnih mehanizmov, ki so, med drugim, temelj spomina, jezika in zavesti. Poleg številnih strokovnih člankov, je svoje znanstvene ugotovitve povzel v več knjigah: *Feeling of What Happens* (1994), *Descartes' Error* (2000a), *V iskanju Spinoze* (2008) ter *Self Comes to Mind* (2010), v katerih je, ne glede na različne poudarke (čustva, sebstvo, ...), kot širši okvir prisotna tudi njegova teorija biološke osnove zavesti. V tem članku bodo predstavljene reprezentacije telesa, ki pri Damasiju igrajo temeljno vlogo znotraj teorije zavesti.

Damasijevo izhodišče oziroma teoretski okvir, ki se ponavlja skozi vse njegove knjige, je, da smo ljudje fizična, tj. telesna bitja, ter da je naša primarna skrb (tako kot tudi vseh ostalih organizmov) ohranjanje življenja in homeostaze organizma (Damasio, 2012: 43).

Smo torej primarno telesni organizmi, ki pa imajo obenem tudi zavest in duševnost.² Razlika med človekom in npr. enoceličnim organizmom je v kompleksnosti živčnega sistema. Za amebo sicer lahko rečemo, da je kompleksen organizem, ki ima za glavno nalogo regulacijo življenja. Nima pa možganov. In jih tudi ne potrebuje, saj živi relativno enostavno življenje v relativno enostavnem okolju (približevanje za življenje ugodnim okoliščinam in odmikanje od škodljivih). Ljudje pa živimo v okoljih, ki so bolj raznolika in kompleksna. Zato smo za soočenje z njimi razvili bolj kompleksen živčni sistem. Damasio (2012: 44) je mnenja, da so najbolj kompleksni tisti organizmi, ki lahko tvorijo reprezentacije življenjskega procesa, lastnega telesa in sveta zunaj.

Reprezentacije so narejene v obliki živčnih zemljevidov. Ker pa smo primarno organizmi, ki morajo vzdrževati homeostazo telesa, so zemljevidi telesa temeljni. Možgani z reprezentacijami telesa v vsakem trenutku beležijo stanje oz. dogajanje telesa (samega in v odnosu do okolja). »Da bi lahko možgani uskladili neštete telesne funkcije, od katerih je odvisno življenje, morajo imeti zemljevide, v katerih so trenutek za trenutkom upodobljena stanja različnih telesnih sistemov« (Damasio, 2008: 169). Kljub temu da Damasio kot temeljne postavlja reprezentacije telesa, pa ne obstajajo zgolj te. Živčni vzorci vzniknejo v dveh vrstah bioloških sprememb, v spremembah, ki so povezane s telesnimi stanji, ter v spremembah, ki so povezane s kognitivnimi stanji (Damasio, 2000a: 281). Spremembe, povezane s telesnimi stanji, so dosežene z enim od naslednjih dveh mehanizmov:

² Ang. izraz *mind* je v knjigi *V iskanju Spinoze* slovenjen kot duh (prevajalec je S. Vörös). Jaz bom v tem članku raje uporabljala izraz *duševnost*, saj, saj se mi zdi manj obtežen z dualizmom substanc.

1. Neposredni mehanizem oz. mehanizem »zanke telesa« uporablja tako signale telesnih tekočin (kemična sporočila, ki so prenesena po krvnem toku) kot tudi živčne signale (elektrokemična sporočila, ki so prenesena po živčnih poteh). Obe vrsti signalov se kažeta v spremenjeni telesni pokrajini, ki je posledično reprezentirana v somatosenzornih strukturah centralnega živčnega sistema, vse od možganskega debla navzgor.
2. Posredni mehanizem oz. mehanizem »simuliranega telesa« ustvari reprezentacijo sprememb, povezanih s telesom, neposredno v senzoričnih zemljevidih telesa in sicer pod nadzorom drugih živčnih mest, na primer v prefrontalni skorji. Posebnost tega mehanizma je, da se dejanska sprememba ne zgodi v telesu. Mehanizem telo enostavno zaobide: imamo učinek »kot da bi bilo« telo res spremenjeno, vendar ni.

Damasio reprezentacije telesa iz metodološkega razloga (za lažje preučevanje) deli na zemljevide telesa (*body maps*) in podobe telesa (*body images*), pri čemer pa obe vrsti reprezentacij razume kot vzorce. Razlika med njima je v tem, da telesne zemljevide doživljamo kot podobe, ko imamo zavest, sebstvo. Zemljevidi so torej dinamični živčni vzorci (oziroma vzorci dejavnosti v možganih), podobe pa mentalni vzorci, ki so osnovani na prvih (Damasio, 2012). To osnovanost in hkrati vsebovanost prvih v drugih imenuje tudi ugnezdenje: »Ugnezdenje preprostega v kompleksnem zagotavlja, da se uravnavni namen ohrani tudi pri višjih členih verige« (Damasio, 2008: 49). Princip ugnezdenja se v skladu z Damasijevim evolucijskim pristopom, ki evolucijo obravnava kot nadgradnjo – preprostejše je vsebovano v bolj kompleksnem, staro se ohranja v novem. Tudi s preprosto obliko mehanizmov bi preživeli, vendar je organizem učinkovitejši s kompleksnejšimi.

Damasio v vseh svojih knjigah nenehno izpostavlja, da sta obe vrsti reprezentacij telesa del istega kontinuuma in da je razlika med obema v ravni opisa. Razlogi za to ločevanje na ravni opisa so metodološki. Ker je proces kompleksen, ga za lažje razumevanje in raziskovanje razstavi na manjše komponente, vendar pa na koncu vedno sledi ponovna združitev razstavljenega nazaj v celoto. Gre za »načelno ločevanje med tistim delom procesa, ki je javen, in delom, ki ostane zaseben« (Damasio, 2008: 27–28). Hkrati v vseh svojih delih opozarja, da takšnega razlikovanja ne gre enačiti z dualizmom substanc, saj to, kako se reči kažejo na doživljajski ravni, predstavlja dualizem aspekta (2012: 64–65). V osnovi pa gre za eno substanco, čeprav se morda zdi drugače, ker zaenkrat še ne znamo razložiti vseh mehanizmov.

Ta dvojnost pri reprezentacijah telesa se pojavlja tudi pri drugih procesih (npr. razlikovanje med čustvi in občutki) in v svojem temelju predstavlja dvojnost telo-duševnost:

/.../ glede pojavov, običajno imenovanih »duh« in »telo«, [sem] mnenja, da izhajata iz iste biološke »substance«, a ju kot različna razlikovalna predmeta obravnavam iz istega razloga, kot razlikujem med čustvom in občutkom: gre za raziskovalno strategijo, katere namen je bodisi izboljšati razumevanje ubrane celote, ki jo sestavljata bodisi duh in telo bodisi čustva in občutki. (Damasio, 2008: 28)

Hkrati pa se zaveda, da vprašanja o odnosu med duševnostjo in telesom še ni zadovoljivo pojasnil, saj

čeprav je sodobno sklapljanje možganov in duha zelo dobrodošlo, pa dualističnega razcepa med duhom in telesom ne odpravi, temveč zgolj prestavi mesto razcepa. Po najbolj priljubljenem in aktualnem sodobnem stališču sodijo namreč duh in možgani skupaj na eno stran, telo – celoten organizem brez možganov – pa na drugo stran. Novi razcep loči možgane od »telesa kot takega«, zaradi česar postane razlaga, kako sta duh in telo povezana, veliko težja. (Damasio, 2008: 180)

V svoji zadnji knjigi (2012) se je Damasio – v želji, da bi izpostavil to enotnost in zabilisal mejo (saj gre le za različne ravni opisa) – odločil, da bo zemljevid telesa in podoba telesa uporabljal kot sopomenki. Podobno zabriše tudi mejo med duševnostjo in možgani:

V tej knjigi uporabljam izraze podoba, zemljevid in živčni vzorec skoraj izmenljivo. Občasno zabrišem tudi ločnico med duševnostjo in možgani, namerno, da s tem poudarim dejstvo, da razlika, kljub temu, da je veljavna, lahko zastre pogled na to, kar skušamo razložiti. (Damasio, 2012: 65)

Vprašanje je, ali je tovrsten zabris z vidika preglednosti in razumevanja smiseln. Je pa vsekakor v skladu z Damasijevo načelno predpostavko kontinuuma:

Ker duh nastane v možganih, ki so sestavni del mehanizma, je tudi sam del tega dobro izdelanega aparata. Z drugimi besedami, telo, možgani in duh so različne pojavne oblike istega organizma. Čeprav jih pod mikroskopom lahko razčlenimo, so pod normalnimi delovnimi pogoji za znanstvene namene nerazdružljivi. (Damasio, 2008: 186)

Izraz zemljevid se mu zdi primeren zato, ker se možganski zemljevidi, ki so rezultat trenutne dejavnosti določenih nevronov in nedejavnosti drugih, dobesedno izrisujejo kot vzorci živčne aktivacije v možganih. Še več, za nekatere dejanske predmete se izrišejo točno določeni vzorci aktivacije nevronov (Damasio, 2012: 73–74). Kot so to pokazale raziskave na opicah, se npr. za krog in za kvadrat izrišejo točno določeni vzorci. Žal to ne velja za vse predmete, predvsem pa ne velja

za izkustveno dimenzijo. »Ko so določeni nevroni 'prižgani', v določeni prostorski razporejenosti, se 'zariše' črta, ki je bodisi ravna ali zavita, debela ali tanka, vzorec, ki se razlikuje od ozadja, ki ga ustvarijo nevroni, ki so 'ugasnjeni'« (Damasio, 2012: 66).

Zemljevidi se lahko povezujejo horizontalno in vertikalno. Za lažjo predstavo horizontalne živčne povezave Damasio predlaga, da si predstavljamo tanko rezino možganske skorje (odrezano vzporedno s površino le-te), kot jo pripravimo za preučevanje pod mikroskopom, ki jo ustrezno prepariramo s kemično barvo, tako da se linije obarvajo. Izrisala se bo dvodimenzionalna kockasta mreža nevronov, npr. kot zemljevid Broadwaya in Manhattna (ki nimata poševnih ulic). To horizontalno povezovanje v možganski skorji se dogaja v vseh slojih, tako da je posamezen sloj umeščen med druge sloje, kjer se prav tako zarisujejo zemljevidi. In vsak element tega horizontalnega zemljevida se povezuje tudi vertikalno, v stolpce, ki vsak vsebuje na stotine nevronov. Stolpci možganski skorji zagotovijo vhodne podatke iz drugih predelov možganov, perifernih senzoričnih sond, kot so npr. oči, in iz telesa. Zagotavljajo pa tudi *outpute* za te iste vire ter izvajajo različne integracije in prilagoditve signalov, ki so procesirani na vsaki lokaciji. T.i. stolpične povezave priskrbujejo možganski skorji *inpute* oz. signale, ki prihajajo od drugih živčnih predelov, ter pošiljajo *outpute* iz možganske skorje v ta druga področja.

Istočasno s predpostavko o kartografiranju³ pa Damasio (2008; 2012) priznava, da je način, kako sam proces kartografiranja poteka, težko opredeliti. Zemljevide bi si lahko predstavljali kot preprosto zbirko točk, dejansko pa mora pri njihovem oblikovanju sodelovati ogromno različnih področij, da na primer zgradijo podobe nekega obraza, pokrajine idr. Možgani, ki izgrajujejo bogate zemljevide na več ločenih lokacijah, morajo te med seboj tudi usklajevati – jih postaviti v medsebojni odnos, v koherentno celoto. Tako pri kartografiranju ne gre za to, da bi možgani golo kopirali oz. pasivno preslikovali dogajanje zunaj možganov: »[K]artografiranje ni nujno pasivni proces. Strukture, v katerih nastajajo zemljevidi, imajo besedo pri kartografiranju in so obenem podvržene vplivom drugih možganskih struktur« (Damasio, 2008: 185).

Čeprav Damasio (2012: 73–74) sklepa, da se veliko kartografiranja odvija na kortikalni ravni (najbolj podrobni zemljevidi, ki jih možgani naredijo, vzniknejo tu), hkrati ugotavlja, da imajo tudi nekatere strukture pod možgansko skorjo zmožnost ustvariti grobe zemljevide. Slednje utemeljuje s primerom tetraplegikov (poškodba hrbtenjače), ki še vedno izkazujejo reflekse povezane z bolečino. Na podlagi te ugotovitve je mogoče sklepati na to, da se nekaj kartografiranja dogaja tudi na

³ Ang. glagol *to map*. V opombi prevajalec Vörös pojasni, da »glagol označuje možgansko dejavnost 'risanja' ali 'oblikovanja' živčnih zemljevidov (vzorcev aktivnosti oz. neaktivnosti nevronov v nekem možganskem predelu), ki prikazujejo dogajanje v telesu« (Damasio, 2008: 16).

ravni hrbtenjače, vendar pa informacije o tem ne prehajajo naprej navzgor (in posledično ne pridejo do zavesti). Damasio nadalje predvideva, »da so mentalne reprezentacije somatosenzoričnih dražljajev iz okončin in trupa polno sestavljene šele na ravni jeder gornjega možganskega debla« (Damasio, 2012: 73). Del signalov, ki potujejo iz telesa do možganov, sicer privede do enostavnega kartografiranja (npr. okončin v prostoru), ampak večina signalov je najprej obdelanih v subkortikalnih jedrih, znotraj hrbtenjače in še posebej v možganskem deblu. Damasio je prepričan, da (še posebej) možgansko deblo ni zgolj vmesna postaja telesnih signalov na poti do možganske skorje, temveč da je na tej vmesni stopnji nekaj dodano (Damasio, 2012: 92–93). Strukturi v možganskem deblu (solitarno jedro in parabranchialno jedro) sta po njegovem prvi, ki centralnemu živčnemu sistemu priskrbita zemljevide telesa kot celote. Dejavnosti v teh zemljevidih ustrezajo primordijalnim občutkom (ki bodo natančneje predstavljeni kasneje). Na tej točki naj povem le to, da so ti občutki osnovani neposredno iz signalov iz telesa. So nenadomestljiva komponenta sebstva in tvorijo prvo razkritje duševnosti – da je organizem živ. Ta pomembna jedra v možganskem deblu torej ne proizvedejo le golih virtualnih zemljevidov telesa, ampak proizvedejo tudi občutena stanja telesa – tu vznikneta občutka bolečine in ugodja.

Organizem lahko informacije, vsebovane v nastalih zemljevidih, uporabi na dveh ravneh, na ne-zavedni ali zavestni. Z rabo na ne-zavedni ravni organizem učinkovito usmerja motorično vedenje in posledično izvede pravo dejanje. Na zavestni ravni pa organizem zemljevide doživlja kot podobe, ki jih lahko upravlja in kognitivno obdeluje. Izraz podoba se pri Damasiju ne nanaša zgolj na vidne podobe, pač pa na vse podobe, ne glede na njihov izvor (npr. slušni, visceralni, tipni itd.). »Skratka, možgani kartografirajo svet okoli sebe in svoje lastno početje« (Damasio, 2012: 18). In z izdelavo zemljevidov možgani informirajo sami sebe (Damasio, 2012: 63).

V svoji zadnji knjigi Damasio glede na predmet izvora podatkov razvrsti zemljevide (in posledično podobe) v tri skupine (Damasio, 2012: 76). S tovrstnim razlikovanjem pride do temeljnih gradnikov telesnega samozavedanja, kot bomo videli v nadaljevanju.

Vrste zemljevidov (podob)	Predmet izvora
Zemljevidi notranjih struktur in stanj organizma (interoceptivni zemljevidi).	Funkcionalno stanje telesnih tkiv, kot je npr. stopnja krčenja/raztegnitev gladkega mišičevja; parametri o stanju notranjega okolja (<i>internal milieu</i>).
Zemljevidi drugih vidikov organizma (proprioceptivni zemljevidi).	Podobe specifičnih telesnih komponent, kot so sklepi, progasto mišičevje, nekateri deli viscere (drobovja).
Zemljevidi sveta, zunanjega organizmu (eksteroceptivni zemljevidi).	Vsi predmeti ali dogodki, ki spodbujajo senzorične prevodnike, kot je mrežnica, koščeni polž v ušesu ali mehanoreceptorji kože.

Tabela 2: Vrste zemljevidov in njihov izvor (Damasio, 2012: 76)

Prvi dve vrsti zemljevidov (interoceptivni in proprioceptivni) imata svoj izvor v notranjosti telesa, tretja, eksteroceptivna, pa ima svoj izvor v posebnih delih telesa, preko katerih je telo v odnosu z okolico.

Proizvodni mehanizem je pri obeh vrstah podob – tistih iz notranjosti telesa in tistih iz posebnih senzoričnih sond – enak. Prvič, dejavnost v telesnih strukturah povzroči začasne strukturne spremembe v telesu. Drugič, možgani s pomočjo kemičnih signalov, ki potujejo po krvnem toku, in elektrokemičnih signalov, ki potujejo po živčnih poteh, v številnih ustreznih predelih izdelajo zemljevide teh telesnih sprememb. In končno, ti živčni zemljevidi postanejo duševne podobe. (Damasio, 2008: 186–187)

Podobe iz notranjosti so signalizirane v centralni živčni sistem, tiste iz posebnih senzoričnih sond nastopijo v visoko specializiranih telesnih delih (npr. mrežnici) in so specializirane v specifične predele, ki kartografirajo stanje tega specifičnega telesnega receptorja (npr. pri mrežnici je to del talamusa, možganskega debla in vizualne skorje). Gre torej za prej omenjeni princip horizontalnega in vertikalnega kartografiranja.

Če strnem, Damasijeva hipoteza je, da možgani v vsakem trenutku tvorijo zemljevide telesa, katerih temeljna vloga je ohranjanje homeostaze. (To ohranjanje notranjega ravnovesja telesa se v veliki meri dogaja na ne-zavedni ravni). Najbolj primarni v tem kontekstu so interoceptivni zemljevidi, ki v kombinaciji s proprioceptivnimi in eksteroceptivnimi beležijo druga stanja organizma in so osnova za

specifične telesne občutke.⁴ Hkrati se vsi zemljevidi telesa v možganih arhivirajo (vendar ne kot nevtralni, ampak zabeleženi z neko valenco izkustva). S tem je omogočena vsakokratna primerjava trenutnega zemljevida s preteklimi, ki so zaznamovani z izkustvom (na primer do neznane osebe imamo odpor, pa ne vemo zakaj, ker možgani pri primerjavi zemljevidov prepoznajo, da smo se v podobnem kontekstu v preteklosti počutili ogrožene), tj. v kakšnem stanju je v danem kontekstu bil organizem, in po potrebi se aktivira avtomatski odgovor. Glavni namen tega mehanizma, ki ga Damasio imenuje somatski označevalec (Damasio, 1994: 164), je preživetje organizma in prepoznavanje potencialno ugodnih oziroma neugodnih situacij, v katerih se organizem nahaja.

3.2 Damasio in reprezentacije pri telesnem zavedanju

Zavest in sebstvo⁵ sta pri Damasiju tesno prepletena pojma, ki drug drugega podpirata in omogočata. Sebstvo lahko razumemo tudi kot samo-zavedanje oz. zavest samega sebe, pri čemer je v svoji najbolj osnovni obliki (protosebstvo) to samo-zavedanje telesno in je hkrati pred-zavestno. Telesne zavesti ne gre zamenjevati z zavestjo telesa, ki nastopi šele na višjih stopnjah zavesti (ko se zavedamo, da je telo moje).

Damasio izhaja iz prepričanja, da je homeostaza oziroma ohranjanje življenja organizma tisti temeljni princip, ki mu je vse podvrženo. Regulacija telesa v možganih in telo samo sta tesno povezana (z zunanjim svetom ni take povezave kot je med možgani in mojim telesom).

Urnnavanje življenja torej vodijo občutki. Slednjih ne gre enačiti s čustvi. Občutek je izkustvo stanja telesa in je na nivoju podob. Pri čustvu pa gre za stanje, ki prevzame telo. Občutek čustva je tako le eden od občutkov. Občutki čustev so različice kompleksnih občutkov telesa, ki jih povzroči določen predmet, na katerega se občutki nanašajo. Najbolj temeljni občutki so občutki bolečine in ugodja (primordialni občutki), ki dajejo vsem telesnim reprezentacijam valenco (somatski označevalec). Damasio (2012) meni, da moramo za to, da smo lahko zavestni, najprej občutiti svoje reprezentacije (in na najbolj osnovni ravni so ravno reprezentacije telesa – natančneje, zemljevidi notranjosti telesa – tisti, ki jih najprej lahko občutimo, in sicer ravno v kontekstu bolečine oziroma ugodja). Tako na primer zemljevid rdeče žoge, ki jo vidimo na igrišču, v nas ne more zbuditi občutka bolečine oz. ugodja. Zemljevid, ki ga možgani ustvarijo za rdečo žogo, je sam po sebi nevtralen.

⁴ Ang. *feeling*, kar je v *V iskanju Spinoze* slovenjeno kot *občutek*, s čimer označuje zasebno (zavestno) izkušnjo čustva, ki je telesno stanje. Kljub temu da se mi načeloma zdi bolj ustrezen slovenski izraz *občutje*, saj nam na nek način omogoča distanciranje od tega, kar v običajni rabi besede razumemo z izrazom občutek – tj. senzoričnega občutka, ohranjam izraz.

⁵ Ang. *self* slovenim s *sebstvom*, čeprav ni najbolj ustrezen izraz. V knjigi *V iskanju Spinoze* je *self* preveden tudi kot *jaz*, kar je ustrezno, vendar po mojem mnenju šele pri najvišjih nivojih sebstva.

V skladu z metodo razgradnje, Damasio tudi pri sebstvu razlikuje tri stopnje sebstva (ki so zopet en kontinuum): protosebstvo, jedrno sebstvo in avtobiografsko sebstvo. Protosebstvo je najbolj temeljna reprezentacija sebstva (še pred-zavestni biološki predhodnik), ki orisuje notranja stanja telesa in je referenca za jedrno sebstvo in avtobiografsko sebstvo. »Protosebstvo je koherentna zbirka nevronske vzorcev, ki kartografirajo, trenutek za trenutkom, stanje fizične strukture organizma« (Damasio, 2000a: 154). Iz reprezentacije organizma v možganih vznikne protosebstvo. Ko so protosebstvu dodane informacije (npr. iz okolice, spomin) se razvijeta višji stopnji sebstva, tj. jedrno in avtobiografsko sebstvo. Protosebstvo dobiva informacije iz treh glavnih vrst zemljevidov (Damasio, 2012: 190).

Prvi dve stopnji se odvijata predvsem na ravni možganskega debla, nekaj tudi na ravni korteksa, in sta pred-zavestni. Damasio v svoji zadnji knjigi izpostavi pomembnost možganskega debla (pred tem je poudarek predvsem na možganski skorji, ki je za integracijo zagotovo zelo pomembna, ni pa temeljna za zavest). Zavest in sebstvo torej ne vznikneta zgolj iz možganske skorje. V možganih obstaja struktura, možgansko deblo, ki je evolucijsko stara (plazilski del možganov) in ki igra pomembno vlogo pri zavesti. Na njeno nepogrešljivost pri procesu sebstva Damasio sklepa iz raziskav, ki so pokazale, da ko okvarimo dele možganskega debla, ki imajo opraviti z zavestjo, zavest ni mogoča. Nastopi lahko celo smrt. Primer takšne strukture je tegmentum (zadnji zgornji del možganskega debla) – poškodba na tem področju vodi v komo. Glede na to, da je ta del možganov prisoten tudi pri živalih, Damasio dopušča možnost, da imajo tudi živali sebstvo. Za razliko od možganskega debla, lahko pri korteksu določene dele (seveda ne vseh) poškodujemo, pa zavest ne bo okrnjena. Možganska skorja je tista, ki je pomembna za integracijo – proizvede koordinirane podobe. Za zavest (predvsem na najvišji stopnji) je pomembno, da vse strukture delujejo v povezavi. Najvišje, avtobiografsko sebstvo, ki je osnovano na podlagi predhodnih stopenj sebstva (tj. protosebstva in jedrnega sebstva), obsega dodatno tudi preteklost, ki in kot jo je doživel organizem, in prihodnost, ki jo organizem anticipira (zemljevidi so shranjeni v spominu in jih je mogoče priklicati; s tem so tudi v samem procesu vključeni dodatni predeli v možganih).

Hkrati s tem, ko so možgani (preko katerih spoznavamo vse, kar je okoli njih) lahko informirani le preko telesa, uspejo s kartografiranjem telesa na integrativen način ustvariti bistveno komponento tega, kar bo postalo sebstvo (Damasio, 2012: 20, 190). Na eni strani imamo torej reprezentacije okolice, zunanje organizmu, ki reprezentacije tvori (na podlagi signalov, ki v možgane vstopajo preko telesa), na drugi pa reprezentacije organizma samega. In občutki povezani z integrirano reprezentacijo organizma so to, kar Damasio (2012: 180– 181) imenuje protosebstvo in kar je temelje samo-zavedanja in zavesti širše. »Kartografiranje telesa v svoji prefinjeni obliki podpira tako procese sebstva v zavestnih duševnostih kot reprezentacije sveta, ki je zunanji organizmu. Notranji svet je odprl pot naši

sposobnosti, da spoznamo ne samo taisti notranji svet, pač pa tudi svet okoli nas» (Damasio, 2012: 106).

Zopet torej naletimo na dvojnost, tokrat zavesti same. Damasio (2000b) to dvojnost opredeli kot dva problema zavesti:

- a) kako dobimo »film-v-možganih« (oz. kako živčni vzorci postanejo mentalni vzorci ali podobe);
- b) kako možgani v dejanju spoznanja predmeta skupaj z izgradnjo mentalnih vzorcev zanj ustvarijo tudi občutek za sebstvo (oziroma, kakšna je relacija med tem, da imamo podobe premetov, in občutkom, da smo mi tisti, ki jih imamo).

Damasio je mnenja, da sta oba problema enako pomembna, in predlaga, da naj celostna teorija zavesti obravnava oba. Vendar pa sam za izhodišče vzame sebstvo (oz. samo-zavedanje), saj se mu intuitivno dozdeva, da bi lahko bila osnova za izgradnjo občutka za sebstvo enaka osnovi, ki omogoča spoznanje predmetov, s čemer je občutek za sebstvo nujen pogoj za spoznanje (Damasio, 2000b).

Damasio (2000b) po tem, ko je odkril gradnike, predpostavi izgradnjo zavesti kot neke vrste sestavljanke v treh korakih. Tudi pri izgradnji so temeljni mehanizmi isti (kartografiranje oz. izgradnja opisa), razlika pa je v tem, za kaj so ti mehanizmi uporabljeni.

- 1) Izgradnja zemljevidov predmeta (zemljevidi prvega reda): na tem nivoju so živčni vzorci neposredno povezani s predmetom (predmet pa je lahko organizem sam ali pa mu je zunanji). Možgani za izdelavo živčnih vzorcev oz. zemljevidov predmetov uporabljajo senzorične sisteme (Damasio ta proces imenuje tudi kartografiranje).
- 2) Izgradnja zemljevidov odziva (so še vedno zemljevidi prvega reda): živčni vzorci beležijo čustven odziv na nek predmet. Možgani kartografirajo spremembe v organizmu, ki so rezultat interakcije s predmetom.

Vsi zemljevidi prvega reda so vir mentalnih podob in njihov tok konstituira miselni proces. Vendar to ni zadosti za generiranje zavesti.

Zemljevidi višjega (tudi: drugega) reda nastanejo na podlagi signalov, ki jih preko živčnih projekcij možgani prejmejo iz zemljevidov prvega reda (tj. opisujejo, kar se dogaja v zemljevidih prvega reda). Šele na tej ravni lahko pride do refleksije – spoznanja organizma, da je sam tisti, ki ima podobe in zavest (zavest samega sebe in zavest predmetov).

»... pomemben del tega, kar imenujemo zavest, je konstituirano s podobami, ki jih zemljevidi drugega reda doprinesejo k duševnosti v obliki občutka spoznanja samega sebe« (Damasio, 2000b: 117). Po Damasiju postanemo zavestni, ko pri-poved podob postane del duševnosti.

3.3 Dinamičnost reprezentacij telesa pri Damasiju

V članku je bila do sedaj predstavljena vloga telesnih reprezentacij pri Damasiju, njihovo razlikovanje glede na raven procesiranja ter glede na izvor podatkov, na podlagi katerih so reprezentacije tvorjene.

Reprezentacije telesa Damasio funkcionalno opredeli kot predpogoj za ohranjanje življenja oziroma homeostaze organizma in izgradnjo zavestne duševnosti. Iz metodoloških razlogov razlikuje dve vrsti – zemljevide telesa in podobe telesa. Podlaga te delitve je razlikovanje med živčnimi in mentalnimi vzorci (tj. reprezentacijami). Kot že rečeno, gre za epistemološko razlikovanje, za razlikovanje med zasebnim in javnim, med prvoosebnim in tretjeosebnim dostopom. Prvoosebni dostop je mogoč z introspekcijo oz. na podlagi zasebnih občutkov – podobe (katerekoli vrste) so dostopne neposredno, vendar le organizmu oz. lastniku duševnosti, v kateri se odvijajo. Tretjeosebni dostop pa znanstvenikom omogoča, da opazujejo vedenje in raziskujejo mehanizme, ko, na primer, s pomočjo sodobnih tehnik, kot je MRI, stimulirajo in opazujejo aktivacijo določenih področij v možganih.

Kot je bilo predhodno že predstavljeno, Damasio glede na izvor signalov razlikuje med interoceptivnimi, propioceptivnimi in eksteroceptivnimi zemljevidi in podobami, tj. reprezentacijami. Zemljevidi, ki jih doživljamo, postanejo podobe. Podobe notranjega stanja organizma (interoceptivne) tvorijo primordialne občutke, medtem ko podobe drugih vidikov organizma, kombinirane z interoceptivnimi, tvorijo specifične občutke telesa. Podobe zunanjega sveta običajno spremljajo podobe različic interoceptivnih in propioceptivnih zemljevidov. V normalni duševnosti (pri ljudeh) so prisotne vse tri vrste podob (in zemljevidov).

Kot smo videli v uvodnem delu, je eden od možnih kriterijev razlikovanju med *online* in *offline* reprezentacijami. Wilson (2002) v svoji šesti trditvi tudi *offline* kognicijo opredeli kot temelječo na telesu. V nadaljevanju si bomo pogledali, kako *online* in *offline* reprezentacije obravnava Damasio.

V *Descartes' Error* Damasio razlikuje med dinamičnimi zemljevidi telesa, ki jih poimenuje tudi *online*, ter statičnimi zemljevidi o splošni telesni strukturi (2000a: 152). Slednji imajo izvor predvsem v interoceptivnih in propioceptivnih signalih. Predvideva torej, da *offline* zemljevidi verjetno predstavljajo propiocepcijo in interocepcijo. Za tovrstne zemljevide uporabi tudi izraz *offline* ali dispoziциjski zemljevidi, krajše dispoziциje.

Damasiju ostane odprto vprašanje o izvorni vsebini reprezentacij – od kod vsebina za temeljne zemljevide, ki so osnova in standard za primerjavo z zemljevidi posodobljenih, aktualnih stanj telesa. Na tej točki tako uvede še prirojene dispoziциje.

Glede na spremenljivost (dinamičnost) Damasio torej razločuje stabilne oziroma trajne ter dinamične, nenehno posodablajoče se zemljevide. Najbolj stalni so po njegovem zemljevidi, ki reprezentirajo notranja stanja telesa (npr. viscero) in so

kot taki temelj za ostale reprezentacije telesa ter so lahko točka reference, proto-sebstvo. O notranjosti telesa tvorimo zemljevide v možganih in jih uporabljamo kot referenco za vse ostale zemljevide. Če naj imamo referenco, ki jo poznamo kot sebstvo, jaz, potrebujemo v našem lastnem procesiranju nekaj, kar je stabilno, kar ne odstopa preveč iz dneva v dan. Imamo eno telo (referenčna točka), ki pa ima različne dele (npr. noge rastejo), medtem ko mora biti notranjost vzdrževana znotraj parametrov (zlata sredina homeostaze) iz dneva v dan (drugače nastopi bolezen, smrt). Zato so zemljevidi o notranjosti telesa za Damasija relativno stalni in so kot taki edini lahko točka reference za tvorbo sebstva.

Stabilne so torej reprezentacije, ki reprezentirajo dele in stanja telesa, ki se skozi življenje najmanj spreminjajo (in imajo zato izvor v notranjosti telesa – intero-cepcija, delno priporiocepcija), dinamične pa tiste, ki reprezentirajo spremenljive dele in stanja telesa (in imajo izvor v eksterocepciji, delno propriocepciji). Delitev med stabilnimi in dinamičnimi reprezentacijami Damasio v veliki meri enači z delitvijo med *offline* in *online* reprezentacijami. Gre za razliko med reprezentiranjem oz. kartografiranjem stabilnih in pa dinamičnih aspektov telesa. Dinamične so tiste, ki kažejo telo, kakor je trenutno (npr. stopalo imam velikosti 40, sedim v turškem sedlu itd.), stabilne pa tiste, ki kažejo, kako naj bi telo bilo (običajno v kanonični poziciji). Dinamične vedno spremljajo oziroma so v odnosu do referenčne točke – stabilne reprezentacije sebstva.

Ločnica *online-offline* pa ne pomeni avtomatsko tudi ločnice med zavestnim in nezavednim, saj se obe kartografiranji lahko odvijata bodisi na eni ali drugi ravni. Razlikovanje med zavestnim in nezavednim je druga delitev. Kot povzema Damasio (1994: 228), so reprezentacije, ki jih možgani izgradijo za opis situacije, in gibi, ki jih možgani oblikujejo kot odgovor na situacijo, odvisni od vzajemne interakcije med možgani in telesom. Možgani, s tem ko se spreminjajo, ustvarjajo reprezentacije, ki se stalno nadgrajujejo. Nekatere od teh reprezentacij ostanejo nezavedne, druge pa postanejo zavestne. Istočasno signali iz možganov (tistega dela, katerega delovanje ni nikoli reprezentirano neposredno v zavesti) še naprej potujejo v telo, nekateri hote, drugi avtomatsko. Rezultat tega je, da se telo zopet spremeni in reprezentacija, ki jo dobimo na podlagi te spremembe, se ustrezno spremeni.

Ker *online* zemljevidi praktično sovpadajo s spremenljivimi in *offline* s stabilnimi zemljevidi, v svoji zadnji knjigi ohrani samo poimenovanje spremenljivi in stabilni zemljevidi. Izraza *online* in *offline* pa uporablja v ustaljeni rabi, pri čemer je *online* razumljeno kot nekaj, kar je na razpolago v danem trenutku, *offline* pa kot nekaj, kar je v ozadju. Pravi, na primer, da je jedrno sebstvo vedno *online* (pa četudi zgolj kot pridih), avtobiografsko sebstvo pa živi dvojno življenje – prehaja med *online* in *offline* (slednje opiše kot speče, ki čaka na aktivacijo) (Damasio 2012: 210).

4. Zaključek

V osrednjem delu članka je bila predstavljena Damasijeva teorija telesnega samozavedanja in telesnih reprezentacij znotraj širšega konteksta njegove teorije zavesti. Kot je iz predstavljenega razvidno, smo zagotovo še daleč od teorije, s katero bi razložili zavest v celoti. Jasno pa se kaže, da je reprezentacijski oris, ki ga potrebujemo za opis kognitivnih procesov, nujen ne le na ravni višjih kognitivnih procesov, ampak tudi že na ravni telesnih procesov, ki so osnova prvim. »Vedenje se prične z občutkom, saj je njegov substrat zgrajen iz telesnih signalov« (Damasio, 2012, str. 117).

Glede na izhodišče, ki je blizu monizmu oziroma redukcionizmu, po katerem smo ena substanca in ne dve, lahko po analogiji (dva vidika, nivoja, istega) predpostavimo podobnost (ne enakost) mehanizmov na obeh ravneh (duševni in telesni) znotraj procesa, kontinuuma, kot to stori tudi Damasio. Vrzeli med obema zaenkrat sicer še ne moremo premostiti, zato se ta navidezna dvojnost ohranja za metodološke potrebe:

Duševna raven bioloških pojavov ima dodatne specifikacije, ki niso prisotne na ravni živčnih zemljevidov. Upam, da nam bo redukcionistična raziskovalna strategija sčasoma omogočila razložiti, kako pridemo od ravni »živčnih zemljevidov« do ravni »duševnosti«, čeprav se duševna raven pri tem ne bo »reducirala« na raven živčnih zemljevidov, saj ima emergentne lastnosti, ki so ustvarjene iz ravni živčnih zemljevidov. (Damasio, 2008: 198)

In ravno zaradi te vrzeli tako Damasio kot de Vignemont predlagata, da se k obravnavi višjih kognitivnih procesov pristopa tako od zgoraj navzdol kot od spodaj navzgor. Čeprav vrzel ostaja prisotna, nam ta dvojni pristop omogoča kontrolo konceptov, ki jih postavimo na eni strani, in daje upanje, da lahko na ta način z novimi odkritji oziroma dognanji na eni strani po analogiji pridemo tudi do novih dognanj na drugi.

Pri pristopu utelešene kognicije, ki prosto uporablja tako reprezentacijske kot nerepresentacijske razlagalne orise (neradikalno utelešenje po Clarku), se je na primer pokazal problem kompatibilnosti pojmov (sheme, podobe) in mehanizmov (tvorjenje živčnih, mentalnih vzorcev), ki so pri razlagi uporabljeni, saj jih pristop črpa iz različnih teoretskih okvirjev (npr. senzomotorični in reprezentacijski). Nasprotno pa pri pristopu šibkega utelešenja po predlogu Alsmith in de Vignemont (2012), kamor lahko uvrstimo tudi Damasija, te težave ni, saj je razlikovanje prisotno le na epistemološki ravni (kako te vzorce doživljamo). Kot smo lahko videli, pristopa Damasio k obema ravnema z istimi mehanizmi, to je reprezentacijami oziroma živčnimi vzorci. Še več, na ta način je mogoče vpeljati telo v razlago višjih kognitivnih stanj. Damasio (2012) vidi v reprezentacijah, tj. sposobnosti možganov, da izdelujejo zemljevide, celo možnost rešitve problema duh-telo.

Kompleksni možgani, kot so naši, izdelujejo eksplicitne zemljevide struktur, ki sestavljajo telo, hkrati pa neizogibno kartografirajo tudi funkcionalna stanja, ki jih te komponente telesa imajo. Ker so zemljevidi možganov osnove mentalnih podob, lahko možgani, ki izdelujejo zemljevide, vpeljejo v proces duševnosti telo kot vsebino. S tem postane telo tema duševnosti in posledično postane rešljiv tudi problem telo-duh.

Dodatno, s tem ko ohranjamo podobne oziroma nadgrajene mehanizme na obeh ravneh, omogočamo celo vrsto razlag, povezanih s fizično dejavnostjo in športom. Na primer, v predšolski in šolski vzgoji so nekatere tovrstne teme že vključene v kurikulum (npr. povezovanje športne vzgoje z drugimi področji dejavnosti, tj. gibanje-jezik, gibanje- matematika idr.), oziroma se obravnavajo v okviru specialističnih obravnavanj (npr. avtizem, jecljanje). Pri vrhunskih športih se pogosto uporablja vizualizacija in zavestna aktivacija določenih mišičnih skupin, omenja se uporaba rekvizitov kot podaljškov našega telesa, npr. floret, lopar. Pri starostnikih pa se uvaža proprioceptivna vadba, priporočena je fizična aktivnost kot del preventivnih ukrepov pri različnih boleznih, tudi pri preprečevanju Alzheimerjeve bolezni.

Representations and bodily awareness

Disorders in experiencing our own body suggest that body representations are constitutive to bodily awareness. In this article, I address the distinction between two types of representations, namely body maps and body images, as determined by the neuroscientist Damasio. He considers both types of representations as patterns formed at two different levels – first at the neural level, and then at the mental one. The first level forms the basis of bodily self-awareness (self), whereas only at the second level can we talk of self-consciousness in a full sense, i.e. as the term “I” is used in everyday life, when we mean ourselves.

Keywords: embodied cognition, representations, bodily awareness, body schema, body map

Literatura

Alsmith, A. in de Vignemont, F. (2012). »Embodying the mind and representing the body«. *Review of Philosophy and Psychology*, 3(1), str. 1–13.

Blanke, O., Spinelli L., Landis, T. in Seeck, M. (2004). »Out-of-body experience and autoscapy of neurological origin«. *Brain*, 127, str. 243–258.

Clark, A. (1997). *Being there: Putting brain, body and world together again*. Cambridge: The MIT Press.

Damasio, A. (1994). *Descartes' Error: Emotion, Reason and the Human Brain*. New York: Avon Books.

Damasio, A. (2000a). *The Feeling of What Happens: Body Emotion and the Making of Consciousness*. London: Vintage Random House.

Damasio, A. (2000b). »A Neurobiology for Consciousness«. V Metzinger, T. (ur.), *Neural Correlates of Consciousness – Empirical and Conceptual Questions*, Cambridge: The MIT Press, str. 111–120.

Damasio, A. (2008). *Iskanje Spinoze: Veselje, žalost in čuteči možgani*. Ljubljana: Založba Krtina.

Damasio, A. (2012/2010). *Self Comes to Mind: Constructing the Conscious Brain* (ponatis). London: Vintage.

de Vignemont, F. (2010). »Body schema and body image – pros and cons«. *Neuropsychologia*, 48(3), str. 669–680.

Fodor, J. A. (1975). *The Language of Thought*. New York: Crowell.

Gallagher, S. (1986). »Body image and body schema: A conceptual clarification«. *Journal of Mind and Behaviour*, 7(4), str. 541–554.

Gallese, V. (2003) »A neuroscientific grasp of concepts: From control to representation«. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 358, str. 1231–1240.

Goldman, A. in de Vignemont, F. (2009). »Is social cognition embodied?«. *Trends in Cognitive Sciences*, 13(4), str. 154–159.

Haggard, P. in Wolpert, D. M. (2005). »Disorders of Body Scheme«. V Freund, H.-J. et al. (ur.), *Higher-Order Motor Disorders*, Oxford: Oxford University Press, str. 261–273.

Markič, O. (2010). *Kognitivna znanost: filozofska vprašanja*. Maribor: Aristej.

Richardson, J. M. in Chemero, A. (2014). »Complex Dynamical Systems And Embodiment«. V Shapiro, L. (ur.), *The Routledge Handbook of Embodied Cognition*, New York: Routledge, str. 39–51.

Marr, D. (2000). »Vision«. V Cummins, R. in Dellarosa Cummins, D. (ur.), *Minds, Brains and Computers: The Foundations of Cognitive Science: An Anthology*, Malden, Mass.: Blackwell Publishers, str. 69–196.

Metzinger, T. (2000). »The Subjectivity of Subjective Experience: A Representationalist Analysis of the First-Person Perspective«. V Metzinger, T. (ur.), *Neural Correlates of Consciousness – Empirical and Conceptual Questions*, Cambridge: The MIT Press, str. 287–306.

- O'Shaughnessy, B. (2000). *Consciousness and the World*. New York: Oxford University Press.
- Shapiro, L. A. (2011). *Embodied Cognition*. New York: Routledge.
- Wilson, M. (2002). »Six views of embodied cognition«. *Psychonomic Bulletin & Review*, 9(4), str. 625–636.
- Wilson, R. A. in Foglia, L. (2011). »Embodied Cognition«. V Zalta, E. N. (ur.), *The Stanford Encyclopedia of Philosophy*, (Jesen 2011). Dostopno na: <http://plato.stanford.edu/archives/fall2011/entries/embodied-cognition> [24. 9. 2015].