ANαliZA

časopis za kritično misel

Društvo za analitično filozofijo in filozofijo znanosti (www.daf-drustvo.si)

03
2004 letnik 8.

ANαliZA, časopis za *kritično misel*, številka 3, leto VIII., oktober 2004, cena: 1.200 SIT

Izdajatelj: Društvo za analitično filozofijo in filozofijo znanosti, Aškerčeva 2, 1000 Ljubljana, www.daf-drustvo.si/analiza.htm, e-mail: analiza@uni-lj.si, transakcijski račun: 02010-0013794433

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

Glavni & odgovorni urednik: Božidar Kante

Uredniški odbor: Borut Cerkovnik, Friderik Klampfer, Olga Markič, Danilo Šuster, Miklavž Vospernik

Lektor: Rudi Cerkovnik

Tehnični urednik: Borut Cerkovnik

Oblikovanje: K.reG@

Prelom: B©

Tisk: Biro m, Šmartinska 130, 1000 Ljubljana

Ta številka je izšla z denarno pomočjo Ministrstva za kulturo Republike Slovenije.

ISSN 1408-2969

KAZALO

Boran Berčić
SMRT 3

DUH IN TELO

Tyler Burge
VZROČNOST DUHA IN TELESA TER POJASNJEVALNA PRAKSA 31

Janez Bregant
KRITIKA BURGOVE REŠITVE MENTALNE VZROČNOSTI 51

Paul M. Churchland & Patricia S. Churchland
INTERTEORETSKA REDUKCIJA: PODROČNI VODNIK ZA NEVROZNANSTVENIKE 60

Sebastjan Vörös
NOVI VAL REDUKCIONIZMA 73

VRLINSKA EPISTEMOLOGIJA

Linda Zagzebski
ISKANJE IZVORA SPOZNAVNEGA DOBRA 85

Vojko Strahovnik
UVOD V VRLINSKO EPISTEMOLOGIJO 101

RECENZIJA

ETIŠKI POJMOVNIK 121

Boran Berčić
Filozofska fakulteta Univerze na Reki

SMRT

Avtor v članku kritično pretrese najbolj znane argumente, ki poskušajo dokazati, da smrt za osebo, ki umre, ni nekaj slabega, in pokaže, da nobeden od njih ni veljaven. To pomeni, da je smrt, vsaj v običajnih okoliščinah, za osebo, ki je umrla, zares nekaj slabega. Avtor ugotavlja, da Williamsu ne uspe pokazati, da bi bilo večno življenje nevzdržno in da bi smrt po dovolj dolgem življenju pomenila olajšanje. V nadaljevanju članka je znani Epikurjev argument proti strahu pred smrtjo razdeljen na dva sorodna, pa vendar različna argumenta. Avtor pokaže, da je prvi osnovan na neresnični ali vsaj vprašljivi teoriji vrednosti, drugi pa predstavlja resen intelektualni izziv. Podrobno so prikazani tudi različni odgovori na Lukrečev argument iz asimetrije pretekle in prihodnje neeksistence. Avtor meni, da nobeden od odgovor ni povsem zadovoljiv. Na koncu članka avtor obravnava in zavrne "matematični argument" proti strahu pred smrtjo kot neveljaven.

Ključne besede: življenje, smrt, vrednosti, Epikur, Lukrec, deprivacijska opredelitev, asimetrija pretekle in prihodnje neeksistence, osebna identiteta.

> Aristotel je menil, da je smrt najstrašnejša od vseh stvari, ki se nam lahko pripetijo.[1] Epikur in Lukrec sta mislila, da smrt pravzaprav ne more biti za nas nekaj slabega in da se je zato ni treba bati.[2] Aristotel je mislil, da je njegovo stališče očitno in ga ni posebej dokazoval, medtem ko . sta Epikur in Lukrec svoja stališča podkrepila, z zelo zanimivimi argumenti. Tu bomo predstavili in razpravljali o treh najbolj znanih: o argumentu, da smrt nima za nas nikakršnega pomena, o argumentu iz asimetrije pretekle in prihodnje neeksistence in o tako imenovanem "matematičnem argumentu". Preden preidemo k podrobnejši obravnavi njunih argumentov, bomo poskušali ugotoviti, ali je smrt vselej nekaj slabega za osebo, ki umre, ali samo za osebo, ki umre prezgodaj, to je, ali bi življenje človeškega bitja bilo smiselno tudi tedaj, ko bi bilo večno, ali pa bi si po daljšem ali krajšem času neizbežno želeli konec svoje eksistence.[3] Torej bomo

[1] To trdi v *Nikomahovi etiki*, III, 9, ko razpravlja o pogumu.

[2] Epikur je svoja stališča o smrti izpovedal na več mestih, v *Pismu Menoikeju* in v *Maksimah.* Ta stališča so razložena tudi pri Diogenu Laertskem v *Življenju in mislih znanih filozofov*. Lukrec, privrženec Epikurja, je svoja stališča izpovedal v delu *O naravi stvari*.

[3] Tu se ne bomo lotili problemov, povezanih z evtanazijo ali samomorom, to je razprave okrog tega, kaj narediti

pravzaprav obravnavali tri trditve o vrednosti smrti za osebo, ki je umrla: (1) smrt je *vselej* nekaj slabega; tudi če bi živeli tisoč ali milijon let, bi bila smrt nekaj slabega; (2) smrt je nekaj slabega le, če nastopi *prezgodaj;* za človeka, ki umre v pozni življenjski dobi, se ne more reči, da se mu je zgodilo nekaj slabega in (3) smrt *nikoli* ni nekaj slabega za nas, kadarkoli in kakorkoli zgodaj naj že nastopi.

Ali je smrt za nas vselej slaba?

Vprašanje je, ali je smrt *vselej* slaba za osebo, ki umira, in ali je vselej *enako* slaba. Lahko bi trdili, da je le *prezgodnja* smrt nekaj slabega, da pa smrt v *običajni dobi* pravzaprav to ni. Jasno je, zakaj je smrt dvajsetletnika tragedija; prezgodnja smrt mu je odvzela vse tisto, kar bi preživel, če ne bi nesrečno umrl, v povprečju je to še petdeset let življenja. Toda, *kaj je smrt odvzela osebi, ki je umrla pri osemdesetih ali devetdesetih?* Če- ravno je mogoče vselej preživeti še kako leto, pa preprosto ne moremo pričakovati življenja, ki bi bilo znatno daljše od tega. To je *biološka omejitev vrste,* ki ji pripadamo. Ploščati kačji pastir živi vsega en dan, pes deset let, želva z Galapagosa pa domnevno preko dvesto let, mi pa živimo do osemdesetih ali devetdesetih. Tako kot ne moremo leteti, videti v temi ali dihati pod morjem, *ne moremo* živeti znatno dlje od osemdesetih ali devetdesetih let.[4] Kako potem trditi, da je smrt nekaj slabega, če nastopi v primerni življenjski dobi? Če upoštevamo biološko omejitev vrste, ki jih pripadamo, ne moremo trditi, da je smrt v primerni življenjski dobi za nas nekaj slabega. Vendar vselej hočemo še živeti, vsaj če so zdravstvene in druge ustrezne okoliščine kolikor toliko ugodne. Zamislite si, da iz takega ali drugačnega razloga verjamete, da boste jutri umrli. Pri tem se vam prikaže Bog in vas vpraša, ali želite živeti še leto dni. Kaj bi mu odgovorili? Odgovor bi bil najverjetneje pritrdilen. Po letu dni se vam znova prikaže in vam zastavi enako vprašanje. Odgovor bi bil spet pritrdilen. Tudi po dveh, treh, štirih itn. letih. Ali bi sploh kdaj odgovorili nikalno? Najverjetneje, da ne. Kolikor že živimo, vselej si želimo še živeti.[5] To pa pomeni, da *ne bi hoteli nikoli umreti* in da bi *hoteli pravzaprav živeti večno.* Čeravno zveni celo trivialno, je jasno, zakaj je smrt za nas nekaj slabega celo tedaj, ko nastopi v primerni življenjski dobi: preprosto *ne želimo* umreti. Zavedamo se bioloških omejitev vrste, ki ji pripadamo, vendar si jih ne želimo in jih imamo za nekaj slabega. Sicer bi se tehnološki razvoj civilizacije lahko opredelil ravno kot premagovanje bioloških omejitev vrste. Ne moremo leteti – izdelali smo letalo; ne moremo videti v temi – napravili smo infrardeča očala; ne moremo dihati pod vodo – izumili smo potapljaško obleko in jeklenke

v okoliščinah, v katerih se zdi smrt manjše zlo od nadaljevanja življenja. Razpravljali bomo zgolj o smrti v okoliščinah, v katerih se življenje *tudi sicer* šteje vredno življenja, torej v okoliščinah, v katerih so zadovoljene vse temeljne potrebe, v okoliščinah, v katerih ni nobenih posebnih trpljenj itn.

[4] Nagel kot primer omejitve vrste navaja dejstvo, da je krt slep (Nagel, 2004: 45). Vprašanje je, ali je treba dejstvo, da je krt slep, imeti za nekaj, kar je zanj slabo? Če je smrt v običajni dobi sploh kaj slabega za človeka, potem je lahko slaba zgolj v smislu, v katerem je slepota slaba za krta. Primer s krtovo slepoto sicer uporabi tudi Aristotel, da bi ponazoril razliko med manjkanjem po rodu in manjkanjem samem na sebi; krt je slep v smislu, da nima vida po rodu, ki mu pripada, človek pa je lahko slep v smislu, da nima vida sam na sebi (Aristotel, 1999: V, 22 (1022b, 22)).

[5] Thomas Nagel, *The View From Nowhere.* To Nagelovo stališče je pravzaprav odgovor na Williamsov argument, h kateremu bomo kmalu prešli.

za kisik itn. Če bi se pomirili z omejitvami vrste, bi še vedno živeli v jamah. Povečanje standarda življenja je podaljšalo življenjsko dobo, da ne govorimo o zdravstveni skrbi. Ne zadovoljimo se z omejitvami svoje vrste in jih po skušamo preseči. V tem smislu so biološke omejitve za nas nekaj slabega, v istem smislu pa je za nas nekaj slabega tudi smrt, ki nastopi v primerni dobi. Jasno, v okoliščinah, kakršne so *de facto*, človek, ki doživi osemdeset ali devetdeset let, ne more trditi, da bo s svojo smrtjo prikrajšan. Lahko pa trdi, da je prikrajšan z dejstvom, da so okoliščine take, kakršne so, to je, da ima vrsta, ki ji pripada, omejitve, kakršne ima. Čeravno je res, da ni realno pričakovati, da bomo živeli več kot osemdeset ali devetdeset let, lahko trdimo, da je ravno to dejstvo tisto, s čimer se ne pomirimo in tisto, kar je za nas slabo. Ne toliko in tako slabo kot smrt pri dvajsetih, a vendarle nekaj slabega. Zaradi tega bi bil edini način, da smrt za nas ne bi bila nekaj slabega, to, da ne nastopi *nikoli*, to je, da živimo *večno*. Lahko torej trdimo, da je za nas smrt – kadarkoli naj nastopi in kolikor dolgo smo pred tem že živeli – nekaj, kar je *vselej* slabo. Tudi ko bi živeli sto, dvesto, osemsto ali milijardo let, nam nikoli ne bi bilo dovolj in bi vselej hoteli še več. Zaradi tega je smrt vselej slaba.[6] Četudi se ta sklep zdi verjetno mnogim jasen in resničen, se mu lahko oporeka. Vprašanje je namreč, ali bi večno življenje res bilo nekaj, kar bi bilo za nas dobro in ali bi to lahko sploh bilo nekaj, kar bi si zares želeli. Precejšnje število ljudi ima intuicijo, da bi življenje v takih okoliščinah v določenem trenutku postalo dolgočasno in nesmiselno, v tem primeru pa bi si želeli smrt in ta bi postala za nas nekaj *dobrega*, dobrega v smislu, da bi nas končno oropala nesmiselnega in neznosnega obstajanja. Jasno je, da to ne pomeni, da vsakomur, ki *de facto* živi dovolj dolgo, življenje postane nesmiselno in neznosno. Tako stanje bi lahko nastopilo po petstotih, šeststotih ali milijon letih, ni bistveno kdaj, vendar obstaja intuicija, da bi enkrat moralo nastopiti, da bi bilo neizbežno. To stališče zastopa tudi Bernard Williams in ga podkrepljuje s primerom Eline Makropulos, likom iz Čapkove drame.[7] Elina se je pri dvainštiridesetih letih dokopala do eliksirja večnega življenja in ga potem pila preteklih tristo let, tako da ima zdaj pravzaprav tristodvainštirideset let, vendar je v takem fiziološkem stanju, kot da bi jih imela dvainštirideset. Tako življenje je postalo zanjo nesmiselno in neznosno mučenje, tako da je na koncu eliksir odvrgla in uničila formulo za njegovo proizvodnjo. Williams meni, da njena odločitev ni bila nesmotrna muha, stvar osebne odločitve ali preprosto stvar temperamenta, temveč nekaj takega, do česar je enkrat *moralo* priti, in da bi bilo *vsako* človeško bitje prej ali slej prisiljeno storiti enako. Zanimiv je teoretski okvir, v katerem Williams upravičuje svojo intuicijo: izhaja iz predpostavke, da je tisto, kar nam nudi zadovoljstvo v življenju in ga naredi smiselnega, *zadovoljitev naših želja*.[8] Jasno je, da moramo, da bi lahko zadovoljili svoje želje,

[6] Avtor, ki je znan ravno po takem stališču, je španski filozof Miguel de Unamuno (Unamuno, 1983).

[7] Drama Karla Čapka "Primer Makropulos" iz leta 1922. Nekaj let po objavi jo je uglasbil Leon Janaček in je danes bolj znana kot istoimenska opera. Williams je svoja razmišljanja predstavil v znanem tekstu (glej Williams, 1973: 82-100).

[8] Za nekatere stvari si sicer želimo, da se zgodijo samo pod pogojem, da smo živi, zato Williams uvede dokaj znano razliko med *pogojnimi* in *kategoričnimi* željami. Na primer, nasledno zimo si želimo imeti nov plašč le pod pogojem, da bomo naslednjo zimo živi, kajti če naslednjo zimo ne bomo več živi, ne bomo več potrebovali nobenega plašča. Zaradi tega take *pogojne* želje ne morejo predstavljati razlogov za življenje,

biti živi, zato je smrt za nas nekaj slabega, *ker nam preprečuje zadovoljitev naših želja.* Oseba, ki bi živela dovolj dolgo, bi sčasoma *ali* izpolnila vse svoje želje *ali* pa bi si jih nehala želeti. V stanju, v katerem ne bi več imela nikakršnih želja, bi njeno življenje postalo popolnoma prazno in nesmiselno, ne bi imela več nobenih razlogov za življenje, *smrt pa zanjo ne bi bila več nekaj slabega.* Še več, nastopila bi kot olajšanje, ki bi prekinilo neznosnost nesmiselnega obstajanja. Jasno, možno je, da bi se osebi, ki bi ali izpolnila ali izgubila vse svoje želje, *vselej pojavljale nove želje.* Ali bi v tem primeru oseba še vedno imela razloge, da živi, in ali bi potem smrt bila zanjo še vedno nekaj slabega? Williams meni, da *ne bi* bila, zato ker so želje, ki jih imamo, konstitutivne za našo identiteto, to je, da *nas želje, ki jih imamo, naredijo za osebe, kakršne smo.* Zato bi oseba, ki bi sčasoma spremenila vse svoje želje, preprosto postala neka druga oseba, taka izguba identitete pa bi bila enaka smrti, saj bi predstavljala popolno prenehanje osebe, ki je obstajala. Zaradi tega nam nenehno pojavljanje novih želja ne bi moglo zagotoviti življenja, ki bi bilo večno in smiselno. Če se iztrošene želje torej ne bi zamenjale z novimi, bi naše življenje postalo nesmiselno in neznosno in bi bilo bolje, če bi umrli. Če pa bi se želje zamenjale z novimi, to ne bi bili več mi, temveč bi namesto nas nastal nekdo drug. Zato, misli Williams, smrt nastopi ali prezgodaj ali prepozno, to je, ali prej kot iztrošimo vse svoje želje ali po tistem, ko jih iztrošimo. Na podlagi tega potem sklepa, da bi bilo najbolje, da bi umrli, tik preden iztrošimo vse svoje želje, to je, tik preden naše življenje postane nesmiselno in neznosno.

Čeravno je argument zanimiv, se sooča z dvema očitnima problemoma. Prvi je, zakaj bi želje, ki jih imamo, morale sčasoma izginiti. Mnoge stvari so enako lepe in privlačne tudi po desetih, dvajsetih ali tridesetih letih. Zakaj ne bi bile take za vselej? Užitek dokončanega dela, lepota filozofske argumentacije, privlačnost nasprotnega spola, pivo po nogometu in mnoge druge stvari so vselej enako dobre. Zakaj naj bi želje po teh in mnogih drugih stvareh sploh kdajkoli oslabile? Drug problem v Williamsovem argumentu je, da ni jasno, zakaj bi oseba, ki bi sčasoma spremenila svoje želje, morala izgubiti svojo identiteto in se na koncu spremeniti v neko drugo osebo. Sprememba želja in nalaganje raznih izkustev bi lahko privedla celo do spremembe značaja. Toda zakaj to ne bi bila še naprej ista oseba? Ali psihološka kontunuiteta in enotnost zavesti ne bi bila zadostna za ohranitev identitete? Res je, da ljudje tudi v običajnem trajanju človeškega življenja spreminjajo želje, vendar jih to ne naredi za druge osebe in s tem ni

temveč imajo smisel le pod pogojem, da smo živi. Po drugi strani pa so nekatere stvari smiselne ne glede na to, ali smo živi ali ne. To so stvari, ki si jih želimo, da se zgodijo ne glede na to, ali bomo živi ali ne. Na primer, to, da bi naši otroci končali šolo, si želimo ne glede na to, ali bomo to dočakali ali ne. Ta naša želja je *kategorična* in nam daje razlog za življenje, ker je verjetneje, da bodo naši otroci uspešno končali šolo, ko bomo še živi, kot pa ne. Razlika se lahko jasno vidi na naslednjem primeru. Človek, ki gradi hišo in zve, da bo živel še šest mesecev, lahko reagira na dva načina: lahko pohiti, da jo konča, lahko pa jo popolnoma zanemari, da bi čas, ki mu je ostal, posvetil nečemu drugemu. V prvem primeru je njegova želja kategorična, v drugem pogojna. Zaradi tega so po Williamsu kategorične želje tiste želje, ki da- jejo smisel in vrednost našemu življenju, in tisto, kar določa našo identiteto. Williamsov argument se opira zgolj na kategorične želje. Ta razlika je zelo zanimiva tudi zaradi neodvisnih razlogov. Zdi se namreč, da moramo nujno imeti vsaj eno kategorično željo. Celo asketi, ki zavračajo vse, česar ravno nujno ne potrebujejo za golo preživetje, morajo imeti vsaj eno kategorično željo, sicer ne bi imeli razloga, da si sploh prizadevajo preživeti. Edina izjema ja lahko samomorilec, katerega edina kategorična želja je, da ga ne bi bilo več.

konec njihovega obstajanja. Četudi je mogoče, da bi nam iz tega ali onega razloga večno življenje sorazmerno hitro postalo dolgočasno, se zdi, da Williamsov argument *ne* kaže, da bi se to *nujno* moralo zgoditi in da bi zaradi tega prej ali slej moral nastopiti trenutek, v katerem naša lastna smrt ne bi bila za nas več nekaj slabega in bi pravzaprav postala nekaj dobrega. Ta možnost je sicer še naprej odprta, vendar je ta argument ne dokazuje. Zdaj pa bomo prešli k argumentom, s katerimi si prizadevajo dokazati veliko močnejšo trditev, to pa je, da smrt za nas ni nikoli nekaj slabega.

Smrt nima za nas nikakršnega pomena

Znan je Epikurjev argument, da smrt za nas nima nikakršnega pomena.[9] S tem argumentom si je prizadeval *dokazati,* da naša lastna smrt *ni in ne more biti za nas nič slabega* in da se nam zato *ni treba bati lastne smrti.* Argument ima dve različici, ki ju bom predstavil in o njiju razpravljal ločeno. Prva temelji na zdravorazumski intuiciji, da človek, ki umre, *nič več ne čuti,* tako da ne more imeti nobenih neugodnih izkustev. Druga se opira na bolj spekulativno intuicijo, da človek, ki je umrl, ne obstaja več, tako da stanje smrti za *nekoga* ne more biti več nekaj slabega.

Ko govorimo o Epikurjevih argumentih, je treba imeti v mislih predvsem to, da je bil *hedonist.* To pomeni, da je menil, da so *ugodni občutki tisto, kar je dobro, neugodni pa ono, kar je slabo.* Hedonizem je povsem preprosta in dokaj prepričljiva etična pozicija: *če so ljudje srečni in uživajo, je to dobro, če pa so nesrečni in trpijo, je to slabo* – in to je vse, iz česar sestoji etika! Posamezni postopki, splošna pravila obnašanja, karakterne lastnosti osebe, stanja stvari, dogodki – vse to je dobro ali slabo le v *drugotnem* smislu, torej samo v tisti meri, v kateri privede do ugodnih ali . neugodnih občutkov. Tisto, kar je *primarno* dobro ali slabo, so ugodni ali neugodni občutki. Dalje, dobro in slabo nikoli ni dobro ali preprosto slabo na sebi; dobro ali slabo je dobro ali slabo *za nekoga,* saj mora vedno obstajati *nekdo,* ki ima ugodne ali neugodne občutke. Če v univerzumu ne bi bilo bitij, ki imajo občutke, ne bi bilo niti dobro niti slabo. Da bi katerakoli stvar lahko bila dobra ali slaba, morajo obstajati bitja, pri katerih ta stvar privede do ugodnih ali neugodnih občutkov. Stvari, ki na nas ne morejo vplivati, ne morejo biti za nas niti dobre niti slabe, ker ne morejo v nas vzbuditi nobenih občutkov. Srž prvega Epikurjevega argumenta proti strahu pred smrtjo je torej povsem jasna: *ker je smrt stanje, v katerem nimamo nobenih – niti ugodnih niti neugodnih – občutkov, je to stanje, ki za nas ne more biti niti dobro niti slabo in je za nas pravzaprav nekaj vrednostno nevtralnega.* Zato ni nikakršnega razloga, da se bojimo lastne smrti. Argument lahko rekonstruiramo na naslednji način:

[9] Epikur ga je pojasnil v pismu Menoikeju (Long & Sedley, 1997: 124-127; Bošnjak, 1978; Lucretius Carus, 1959: 523-524). Long in Sedley sta lepo in koncizno sistematizirala in predstavila Epikurjeve in Lukrečeve argumente proti strahu pred smrtjo. Na Epikurjeva in Lukrečeva stališča o smrti se nanaša tudi Sharples (Sharples, 2000: 113-120), vendar njegovi komentarji zadevajo bolj zgodovinsko kulturološki kontekst kot pa samo argumentacijo. *

(P_1) Edino dobro in slabo, ki obstaja, so ugodni ali neugodni občutki.

(P_2) Smrt je stanje, v katerem ni nobenih občutkov.

$\therefore$ Smrt torej ne more biti nekaj slabega.

Argument je zelo preprost in jasen. Toda čeravno ima neko intuitivno prepričljivost, je vprašanje, ali res kaže, da smrt ni nekaj slabega in da se je ne bi smeli bati. Ali nas lahko prepriča, da je stanje lastnega neobstajanja nekaj, kar je za nas pravzaprav *vrednostno nevtralno*, in nekaj, česar se nam ni treba bati? Argument je gotovo ustrezno sredstvo proti praznovernemu strahu pred neugodnimi izkustvi v posmrtnem življenju: peklu, purgatoriju, Hadu itn. Ker takrat, ko umremo, ne bomo nič občutili, gotovo ne bomo doživeli nobene od teh grozot Če gledamo na zgodovinski kontekst, se zdi, da se je ta argument izoblikoval in predstavil predvsem z namero, da odžene strah te vrste, ki je bil v tej dobi domnevno zelo razširjen.[10] Eden od prvenstvenih epikurejskih ciljev je bil namreč osvoboditi ljudi dveh neupravičenih strahov, strahu pred bogovi in strahu pred smrtjo, in jim tako narediti življenje lepše in boljše. Poleg tega je bil Epikur prepričan materialist, tako da v okviru njegove filozofije gotovo ni bilo prostora za verovanje v posmrtno življenje. Toda četudi je argument primerno sredstvo proti strahu od grozot, ki nas domnevno čakajo v posmrtnem življenju, je vprašanje, ali "deluje" tudi proti strahu pred smrtjo, ki je razumljena kot popoln konec obstajanja. Ali je lahko razlog za to, da se nehamo bati popolnega niča lastnega neobstajanja, ki nujno stoji pred vsakim od nas? Zdi se, da ne!

Premisa (P_2) se zdi nevprašljiva; razen redkih, ki so res prepričani v posmrtno življenje[11], se vsi strinjamo, da je smrt popolno nehanje nas kot zavestnih in voljnih bitij in da, potem ko umremo, ne bomo imeli več nobenih izkustev. Sklep je tudi veljaven. Zdi se, da sledi iz (P_1) in (P_2); če bi tudi (P_1) bila resnična, bi moral biti resničen tudi sklep. Toda (P_1) ni resnična. Trdi, da so ugodni in neugodni občutki *edino*, kar je dobro ali slabo, to pa ni točno. Obstajajo še drugi načini, kako so lahko stvari in dogodki za nas dobri ali slabi; za nas so lahko dobre ali slabe tudi tiste stvari in dogodki, ki v nas *ne izzovejo nobenih občutkov*, ne ugodnih in ne neugodnih.

Epikurjeva teorija vrednosti

(P_1) izraža Epikurjev *senzualizem* in *subjektivizem* glede vrednosti, problem pa je v tem, ker se zdi, da taka teorija vrednosti preprosto ni resnična. Grobo govoreč, gre za stališče, da *za nas ne more biti slabo tisto, za kar ne vemo, da se je zgodilo.* Vendar obstaja cel niz primerov, ki kažejo, da to ni res! Videti je, da se

[10] Sharples sumi, da je bil strah pred posmrtnim življenjem res tako razširjen in meni, da so Epikurjevi in Lukrečevi nasprotniki bolj retorični kot dejanski (glej Sharples, 2000: 114-120).

[11] Zelo zanimivo je vprašanje, kdo je zares prepričan v posmrtno življenje. Vojaki, ki so pod artilerijskim napadom molili k Bogu in Devici Mariji, niso molili za boljše mesto v posmrtnem življenju, temveč zato, da preživijo in da jih ne bi zadela granata v tem svetu. Četudi so na deklarativni ravni sprejemali govor o odrešitvi in večnem življenju, so se popolnoma enako bali smrti kot tisti, ki niso bili verniki.

nam lahko zgodijo *stvari, ki so za nas slabe, čeravno mi zanje ne vemo.*[12] Če me žena vara z vodovodnim inštalaterjem, poštarjem ali kolegom v službi, je to nekaj, kar je zame slabo tudi tedaj, ko ne vem, da to počne. Če me sodelavci pri delu opravljajo za mojim hrbtom, je to nekaj, kar je zame slabo tudi tedaj, ko ne vem, da to počnejo. Spodkopavanje mojega ugleda *je zame slabo* tudi tedaj, ko ne vem, da se to dogaja. Ravno tako je tudi krepitev mojega ugleda in reputacije nekaj, kar je *zame dobro* tudi tedaj, ko ne vem, da se to dogaja. Dobre ali slabe so za nas torej tudi stvari, ki jih *ne čutimo* in ki se jih *ne zavedamo,* ne pa, kot je to mislil Epikur, samo tiste, ki v nas vzbujajo ugodne ali neugodne občutke. Morda bi lahko rekli, da so stvari, ki ne vplivajo na naše občutke, lahko za nas dobre ali slabe *zgolj v smislu, da bi v nas vzbudile ugodne ali neugodne občutke, ko bi zanje zvedeli.* Torej da ni treba, da v nas *de facto* izzovejo ugodne ali neugodne občutke, temveč da zadostuje, da imajo *dispozicijo,* da jih izzovejo, ko bi mi zanje vedeli. To stališče pravzaprav izražamo, ko se sprašujemo "Kaj bi rekel, ko bi zvedel?" ali ko rečemo "Zelo rad bi vedel!". Tudi z reklom "Obrača se v grobu", ki je po svoji formulaciji kategorično, ne pa dispozicijsko, v glavnem želimo povedati, da bi se pokojnik odzval z negodovanjem, če bi bil živ in bi vedel, kaj delamo. S tako dispozicijsko interpretacijo bi lahko Epikurjevo merilo dobrega in slabega zadostno razširili, da bi zajelo tudi navedene protiprimere in ga le-ti v tem primeru ne bi več spodkopavali. Toda zdi se, da navedeni primeri kažejo *več;* zdi se, da kažejo, da gre za stvari, ki so za nas dobre ali slabe *tudi tedaj, ko zanje nikoli ne bi zvedeli.* Stvari so za nas dobre ali slabe *objektivno,* neodvisno od občutkov, ki jih v nas izzovejo. Na koncu lahko problem izoblikujemo v okviru Evtifronove dileme in se vprašamo, ali *stvari v nas izzivajo občutke ugodja ali neugodja zato, ker so za nas dobre ali slabe, ali pa so za nas dobre ali slabe zato, ker v nas izzivajo ugodne ali neugodne občutke?* Vprašanje je torej, *kaj kaj določa,* vrednost stvari naše občutke ali naši občutki vrednost stvari? Zdi se, da v velikem številu primerov vrednost stvari predhodi našim občutkom in jih določa, to pa pomeni, da tisto, kar jih naredi slabe, niso neugodni občutki, ki jih izzovejo v nas, temveč nekaj, kar je v njih že prej in neodvisno od dejstva, da v nas vzbudijo te občutke. V velikem številu primerov so naši ugodni ali neugodni občutki *pokazatelji,* ki kažejo na pozitivno ali negativno vrednost stvari, ne pa tisto, kar stvarem *daje* pozitivno ali negativno vrednost. Na primer, siromaštvo pripelje do razočaranja. Vprašanje je, ali je siromaštvo za nas slabo zato, ker privede do razočaranja, ali privede do razočaranja zato, ker je za nas slabo? Zdi se, da je razočaranje, do katerega pripelje siromaštvo, zgolj pokazatelj slabega stanja, v katerem smo, ne pa tisto, kar stanje, v katerem smo, naredi slabo. Siromaštvo je nekaj, kar bi za nas bilo slabo tudi tedaj, ko nad njim ne bi bili razočarani.[13] Poleg tega pa količina ugodnih občutkov ne more biti in pogosto tudi ni merilo vrednosti življenja. Za ekscentrika, ki ves svoj čas in ves svoj denar vlaga v to, da bi zgradil *perpetuum mobile,* bi lahko upravičeno rekli, da trati življenje in da bi *zanj bilo bolje,* če bi

[12] Na to poanto je pokazal Nagel v članku "Smrt", sprejelo pa jo je niz avtorjev.

[13] Jasno je, da lahko trdimo, da je siromaštvo za nas slabo *tudi* zaradi razočaranja, do katerega pripelje, ne pa, kot je to mislil Epikur, *predvsem* zaradi razočaranja, do katerega pripelje.

delal kaj drugega, celo če ga nič drugega ne bi naredilo tako srečnega kot poskusi, da bi zgradil svoj priljubljeni stroj. Ti primeri in ti preudarki kažejo, da je Epikurjeva teorija vrednosti zmotna in da je treba zavrniti (P_1) v navedenem argumentu. Ker neugodni občutki *niso* edino, kar je lahko za nas slabo, je lahko za nas nekaj slabega tudi smrt, četudi je to stanje, v katerem ni nobenih občutkov. Iz dejstva, da je smrt stanje, v katerem ne moremo imeti nobenih občutkov, niti neugodnih, preprosto *ne sledi*, da smrt zaradi tega ne more biti za nas nekaj slabega. Torej je treba Epikurjev argument proti strahu pred smrtjo zavrniti zato, ker je osnovan na zmotni teoriji vrednosti.

Deprivacijska pojasnitev

Največji problem za Epikurjev argument, da za nas smrt nima nikakršnega pomena, je naslednji: tudi tedaj, ko bi sprejeli osnovne postavke njegove teorije vrednosti, senzualizem in subjektivizem, iz njih še vedno ne bi sledilo, da so neugodni občutki *edino*, kar je za nas lahko slabo. *Za nas je lahko slabo tudi tisto, kar nas oropa ugodnih občutkov*, zapravljeno ugodje je lahko za nas večja škoda od pretrpljenega neugodja. Torej obstajata dva načina, na katera je lahko neko stanje za nas slabo: lahko

(1) privede do neugodnih občutkov;

(2) nas oropa ugodnih občutkov.

Dejstvo, da je Epikur zanemaril možnost, da so stvari in dogodki lahko za nas slabi *tudi* zato, ker nas oropajo ugodnih občutkov, ki bi jih sicer imeli, je ključno v razpravi o smrti. Prepričani smo namreč, da je *smrt za nas nekaj slabega ravno zato, ker nas za vselej oropa možnosti ugodnih občutkov in izkustev*, ne pa zato, ker s seboj nosi nekakšna neugodna izkustva.[14] Epikur je imel prav, ko je trdil, da, potem ko umremo, ne moremo imeti več nobenih neugodnih občutkov, vendar je bil v zmoti, ko je iz tega izpeljal sklep, da zaradi tega smrt ne more biti nekaj, kar je za nas slabo. Ta sklep preprosto *ne sledi* zato, ker je za nas slabo tudi tisto, kar nam preprečuje, da imamo ugodne občutke in izkustva, še posebej tisto, kar nam *za vselej* preprečuje, da bi še sploh kdaj imeli *kakršnakoli* izkustva in občutke. Smrt je za nas slaba zato, ker nam *odvzema* nekaj, kar bi imeli, če ne bi umrli.[15] Zato se taka pojasnitev dejstva, da je smrt nekaj, kar je za nas slabo, imenuje *deprivacijska pojasnitev*.[16] Lahko jo oblikujemo takole:

(DP) Smrt je za osebo, ki je umrla, nekaj slabega zato, ker jo je oropala vseh ugodnih občutkov, ki bi jih imela, če ne bi umrla.

[14] Deprivacijska pojasnitev zelo lepo zajema naše intuicije o tem, kaj naredi smrt za nas slabo, bodisi da menimo, da je zgolj *prezgodnja* smrt slaba, bodisi da mislimo, da je smrt *vselej* slaba.

[15] Lahko rečemo, da je smrt *negativno zlo v* smislu, da predstavlja oropanje, deprivacijo nečesa dobrega. Pozitivno dobro je sicer prisotnost nečesa dobrega; negativno dobro je odsotnost nečesa dobrega; pozitivno zlo je prisotnost nekega zla, negativno zlo pa je odsotnost nečesa dobrega.

[16] Stališče se v anglosaški literaturi imenuje "The Deprivation Account of the Badness of Death" in ga zastopa večina avtorjev, ki se ukvarja s problemom smrti: Thomas Nagel, John Martin Fischer, Fred Feldman in drugi.

Torej, ko se bojimo smrti, se ne bojimo tega, da se nam bo, ko nas ne bo več, zgodilo nekaj slabega v smislu, da bomo imela nekakšne neugodne občutke ali izkustva. Tisto, česar se bojimo, ko se bojimo smrti, je ravno to, *da se nam ne bo nič več zgodilo,* niti ugodno niti neugodno. Res je, da se moramo bati tistega, kar se lahko zgodi nam ali komurkoli drugemu, ki ga imamo radi, vendar to ni vse. Povsem razumljiv je tudi strah pred stanjem, v katerem se nam ne more zgoditi nič več. Želimo, da se nam dogajajo dobre stvari, ne pa da se nam ne more nič zgoditi in da ne moremo ničesar storiti. Če torej sprejmemo izjemno plavzibilno deprivacijsko upravičenje prepričanja, da je smrt za nas nekaj slabega, potem je treba zavrniti Epikurjev argument in to zato, ker je njegova prva premisa preprosto neresnična; *neugodni občutki niso edino, kar je slabo, ravno tako slabo je tudi to, da smo takih občutkov oropani.*

V okviru Epikurjeve etike ima sicer odsotnost ugodnih občutkov manjšo težo, kot morda sicer mislimo, da jo ima. Epikur namreč razlikuje *kinetične* in *katastemične'*[17] užitke.[18] Na primer, ko jemo, je užitek v dobrem okusu hrane kinetični užitek, medtem ko je *stanje* sitosti, v katerem nas ne vznemirja lakota, katastemični užitek. Po Epikuru so *katastemični užitki več vredni in boljši od kinetičnih;* odsotnost lakote je več vredno od užitka v dobrem okusu hrane, odsotnost bolečine je več vredno od užitka masaže ali orgazma itn. Tu Epikur deli osnovno intuicijo s tako imenovanim *negativnim utilitarizmom* — s stališčem, da je odsotnost trpljenja veliko pomembnejša od prisotnosti sreče. Naposled, stanje duševne in telesne nevznemirjenosti *(ataraxia),* ki je bilo v svojem antičnem svetu najvišje dobro, je katastemični, ne pa kinetični užitek. Zaradi tega oropanost ugodnih občutkov, na katero se opira deprivacijsko upravičenje, za Epikurja ne predstavlja nekaj, kar bi bilo posebej slabo. Toda deprivacijsko upravičenje zajema vse ugodne občutke, tako tiste, ki jih je Epikur imenoval kinetične, kot one, ki jih je imenoval katastemične. Četudi sprejmemo stališče, da odsotnost kinetičnih užitkov niti ni nekaj res slabega, kaj je potem s katastemičnimi užitki? Ker so ti več vredni in ker predstavljajo najvišje dobro, ali ni potem smrt slaba zato, ker nas oropa tudi njih? Potem ko umremo, gotovo ne bomo mogli biti več v stanju neskaljenega duhovnega zadovoljstva niti v kakršnemkoli drugem stanju, za katerega bi si morali prizadevati za časa življenja. Epikurjev odgovor je zelo zanimiv, vendar ni povsem prepričljiv. On je menil, *da je za katastemične užitke vseeno, kako dolgo trajajo!* Kinematični užitki so podvrženi načelu *čim več, tem bolje,* katastemični pa ne! Nanje *ni mogoče* aplicirati Benthamovega *utilitarističnega računa sreče.* Nekaj minut duševnega spokoja je enako dobro kot nekaj ur ali nekaj let. Stanje neskaljenega duševnega miru je enako dobro ne glede na to, kako dolgo traja, bistveno je, da smo v njem, ne pa kako dolgo. Zato končno dolgo življenje, oropano neugodnih občutkov, ne more biti nič slabše od neskončno dolgega življenja, oropanega neugodnih občutkov.[19] Če je to res, potem je Epikurjevo stališče, da smrt za nas

[17] Lahko rečemo tudi *katastematične,* pogosto se uporablja tudi izraz *statične.*

[18] Lep prikaz Epikurjevih stališč o užitkih lahko najdemo v knjigi J. M. Rista (Rist, 1972: 6. pogl.).

[19] Diogen Laertski, *Život i mišljenje istaknutih filozofa,* točka XIX Epikurjevih naukov. Odlomek navaja tudi Bošnjak (Bošnjak, 1978).

nima nikakršnega pomena, odporno proti deprivacijskemu upravičenju. Če je vseeno, kako dolgo traja spokojno življenje, potem smrt – niti prezgodnja smrt – ne more biti nekaj slabega za tistega, ki umre. Četudi je katastemične užitke težje kvantitativno zaobseči kot kinetične, pa se vendarle zdi, da tudi zanje velja načelo čim več, tem bolje. Ali ni namreč bolje ne biti lačen pet ur kot eno uro? Ali ni bolje ne imeti nobenih bolečin celo življenje kot samo v rani mladosti? Zdi se, da je bil Epikur tu preprosto v zmoti in da je predpostavil tisto, kar bi moral dokazati, to pa je, da nam smrt nič ne pomeni. Deprivacijsko upravičenje prepričanja, da je smrt nekaj slabega, se zdi bolj prepričljivo in močnejše od Epikurjevega stališča, da je vseeno, kako dolgo traja spokojno življenje.

Poleg tega Epikurjev argument, da smrt za nas nima nikakršnega pomena, temelji na predpostavki, da so *neugodni občutki edino, kar je za nas slabo.* Zato smrt ne more biti za nas nekaj slabega, ker je to stanje, v katerem ni nobenih in tako tudi ne neugodnih občutkov. Smrt bi lahko bila nekaj slabega le tedaj, ko bi bila stanje, v katerem bi imeli nekakšne neugodne občutke. Vendar pa je dokaj razširjeno *ravno nasprotno* stališče, po katerem je bolje živeti še tako težko in boleče življenje, kot pa sploh ne živeti. Po tem stališču *je neprimerljivo bolje imeti kakršnekoli občutke, pa naj bodo še tako boleči in slabi, kot pa ne imeti nobenih občutkov.* To stališče ima gotovo precejšnjo težo in ga je treba jemati resno, toda morda ni povsem resnično. Zagovorniki evtanazije se namreč opirajo na intuicijo, da so možne in da dejansko obstajajo skrajne situacije, v katerih je bolje umreti in ne imeti več nikakršnih občutkov, kot pa še nadalje živeti v neznosni bolečini brez kakršnegakoli upa na ozdravitev, in da moramo zato vsakemu, ki se znajde v takem položaju, omogočiti izbor med tema dvema opcijama. Vendar pa nisem prepričan, da zagovorniki evtanazije menijo, da je življenje v neznosni bolečini brez kakršnegakoli upa na ozdravitev nekaj, kar je *vrednostno negativno,* medtem ko je smrt nekaj, kar je *vrednostno nevtralno,* in da je zato evtanazija pravzaprav prehod iz stanja, ki je za to osebo *slabo, v* stanje, ki za to osebo ni *niti dobro niti slabo.* Prej je res, da zagovorniki evtanazije menijo, da je v takih okoliščinah smrt preprosto *majše zlo* od nadaljevanja življenja v neznosni bolečini. To pa pomeni, da je smrt vseeno nekaj slabega za osebo, ki umira, in da je treba pojasniti, zakaj je nekaj slabega. Intuicija, da je bolje imeti kakršnekoli občutke kot pa nobenih, v vsakem primeru kaže, da deprivacijska pojasnitev najverjetneje *ne zajema vsega* tistega, zaradi česar menimo, da je smrt za nas nekaj slabega. Smrt je gotovo slaba zato, ker nas oropa *vseh ugodnih* občutkov, ki bi jih imeli, ko ne bi umrli, vendar se zdi, da ni slaba *zgolj zato,* zdi se, da je slaba *tudi zato,* ker nas oropa *vseh* občutkov, kakršnikoli naj bi že bili – ugodni, neugodni ali nevtralni. Zdi se, da je kakršnokoli izkustvo, pa naj si bo še tako slabo, več vredno od nikakršnega izkustva. Zato bi deprivacijsko upravičenje prepričanja, da je smrt za nas nekaj slabega, *morali razširiti tako, da bi zajelo vse občutke, ne samo ugodne.* Epikur je svoj argument formuliral v okviru prisotnosti in odsotnosti ugodnih in neugodnih občutkov in zato razprava v glavnem še naprej poteka v teh okvirih. Vendar je jasno, da smrt ni slaba zgolj zato, ker nas oropa občutkov, slaba je *tudi* zato, ker nas *oropa*

možnosti delovanja, nam onemogoča, da bi udejanili svoje načrte in da bi storili tisto, kar smo si želeli storiti. Potem ko umremo, *ne bomo mogli nikoli več kaj storiti.* Biti zmožen storiti tisto, kar hočemo storiti, pa je gotovo za nas nekaj dobrega. Tako lahko predstavimo popolnejšo formulacijo deprivacijske pojasnitve:

(DP') Smrt je nekaj slabega za osebo, ki je umrla, zato ker jo je oropala vseh občutkov, ki bi jih ta oseba imela, če ne bi umrla, in ji onemogočila, da stori kaj od tistega, kar si je želela storiti.

Očitno je, da se deprivacijska pojasnitev lahko še naprej dopolnjuje z vsem tistim, kar naredi življenje še vrednejše življenja, vendar je osnovna ideja jasna: *smrt je za nas slaba zato, ker nas oropa vsega dobrega, kar nam omogoča življenje.* Zaradi tega bi morala popolna pojasnitev dejstva, da je smrt nekaj slabega za tistega, ki je umrl, hkrati biti tudi popolna pojasnitev dejstva, da je življenje nekaj dobrega za tistega, ki živi.

Slabo za koga in kdaj?

Druga različica Epikurjevega argumenta, ki pravzaprav dodatno okrepi prvo različico, temelji na preprostem uvidu, *da se nam vse, kar se nam zgodi, lahko zgodi le tedaj, dokler smo živi.* Zato da bi se nam sploh kaj lahko zgodilo, dobro ali slabo, moramo biti živi. *Ko več ne bomo živi, se nam nič več ne more zgoditi, niti dobro niti slabo.* Zato naša lastna smrt ne more biti nekaj, kar bi bilo *za nas* slabo. Naš odhod je lahko nekaj slabega zgolj za tiste, ki nas preživijo, naše bližnje ali naše prijatelje, ne pa tudi za nas same. Smrt bi lahko bila nekaj slabega *le tedaj, ko bi imela koga, za katerega* bi lahko bila nekaj slabega, ker pa nas, potem ko umremo, preprosto ne bo več, je nesmiselno misliti, da bi naša lastna smrt lahko bila za nas nekaj slabega. Smrti bi se pravzaprav morali bati *le tedaj, če bi se nam lahko zgodila, dokler smo še živi,* ker pa naša lastna smrt ni nekaj, kar se nam zgodi, dokler smo še živi, *se je nam ni treba bati,* ker preprosto ni nekaj, kar se lahko zgodi *nam.*[20] Bati se je treba le tistega, kar se nam lahko zgodi, dokler smo živi, ne pa tudi onega, kar se bo morebiti dogajalo po trenutku, ko umremo, saj so vse to stvari, ki se ne bodo mogle več zgoditi nam, ker nas preprosto ne bo več. *Smrti bi se pravzaprav morali bati le tedaj, če bi bila smrt nekaj, kar lahko preživimo,*[21] ker pa po predpostavki ne moremo preživeti svoje lastne smrti, se nimamo ničesar bati. Zato naša lastna smrt za nas ne more biti nekaj slabega, ker se nam *ne more zgoditi niti ko smo živi niti potem, ko umremo;* dokler smo živi, smo živi in smrt ne more biti v tem obdobju za nas nekaj slabega, potem ko umremo, pa nas več ni in tudi takrat ne more biti smrt za nas

[20] Zato je Aristotelova trditev, da je smrt nekaj najstrašnejšega, kar se nam lahko zgodi, pravzaprav zelo presenetljiva, nemara celo paradoksna. Aristotel namreč v nadaljevanju istega stavka trdi, da se mrtvemu človeku ne more več zgoditi nič dobrega niti slabega (Aristotel, 1994: III, 9). Tu je neposredno pod udarom Epikurjevega argumenta.

[21] Lukrec se posmehuje tistim, ki se bojijo, da jih bodo – potem ko umrejo – raztrgale divje živali, da bo njihovo truplo strohnelo v naročju zemlje ali končalo v plamenih itn.; ravno tako tudi onim, ki se pritožujejo, da – potem ko bodo umrli – ne bodo nikoli več videli svojih bližnjih, saj tedaj, ko bodo umrli, ne bodo imeli več nobenih želja (Lukrec, 1959: III, 870-910).

nekaj slabega. Ravno v tem je kleč znanega Epikurjevega izreka, da "dokler smo mi, ni smrti, kadar pa ona pride, tedaj *nas* ni več" (Epikur, 1959b: 524). Ker se nam smrt *nima kdaj* zgoditi, preprosto ne more biti za nas nekaj slabega in se je sploh ni treba bati. Ker nam smrt torej nič ne pomeni niti takrat, ko smo živi, niti takrat, ko smo mrtvi, *smrt za nas nikoli nima nikakršnega pomena in se nas nikoli ne tiče.* Ravno tako bi bilo nesmiselno, niti takrat, ko smo živi, če bi smrt imeli za neko prihodnje zlo, ki nas bo doletelo, zato ker smrt, kot smo videli, preprosto ni neko zlo, ki nas lahko kdajkoli doleti (Epikur, 1959b: 523-524). Če ni zlo, potem ko umremo, kako bi lahko potem bilo zlo, preden umremo? Ker se je iracionalno bati nečesa, kar se nam nikoli ne more zgoditi, je torej vsak strah pred smrtjo preprosto iracionalen. Epikurjev argument se opira na predpostavko, da smrt za nas ne more biti nekaj slabega, ker ni *nikogar, za katerega* bi lahko bila nekaj slabega, in zato, *ker nima kdaj* za nas biti nekaj slabega. Njegovo sklepanje gre takole:

(1) Nekaj je lahko za nas slabo le, če se nam lahko zgodi; tisto, kar se nam ne more zgoditi, ne more biti za nas niti slabo.

(2) Medtem ko smo, ni smrti, ko pa smrt nastopi, nas ni več.

(3) Torej smrt ni nekaj, kar bi nas lahko doletelo.

(4) Torej smrt ne more biti nekaj, kar je za nas slabo.

Ideja je torej v tem, da – dokler smo živi – smrt ne more biti za nas nekaj slabega preprosto zato, ker še nismo umrli. Po drugi strani pa smrt, potem ko umremo, za nas ne more biti več nekaj slabega, zato ker nima več biti za koga nekaj slabega, saj ne obstajamo več. Zato smrt *nima kdaj* biti za nas nekaj slabega in preprosto *nima za koga* biti nekaj slabega.[22] Tako znana Epikurjeva trditev, da "dokler smo mi, ni smrti, kadar pa ona pride, tedaj *nas* ni *več*" postavlja – vsaj *prima facie* - v nezavidljiv položaj vsakogar, ki želi trditi, da je smrt za nas nekaj slabega.

Le-ta mora odgovoriti na vprašanje: *"Kako je smrt lahko nekaj slabega za nekoga, ki ga več ni, in če je lahko nekaj slabega, potem kdaj?"* To je kajpada problem tudi za deprivacijsko stališče, saj namreč ni jasno, *kdo je oropan občutkov*, ker osebe, ki je oropana občutkov, preprosto ni več. Jasno je, kdaj je bolečina slaba za osebo, ki jo boli – *tedaj, ko jo boli.* Bolečina ima lahko kajpada škodljive posledice tudi tedaj, ko mine: lahko privede do strahu, frustracije in neugodnih spominov. Toda bolečina je primarno *slaba, dokler traja.* Očitno je tudi, za koga je bolečina slaba – za osebo, ki jo boli. Kaj pa reči za smrt? *Kdaj je smrt slaba in za koga?* Za osebo, ki je umrla v t_n dejstvo, da po t_n nje ni več, ne more biti nekaj slabega za njo v času pred t_n. To dejstvo bi lahko bilo nekaj slabega za to osebo le tedaj, če bi lahko iz časa po t, vzročno delovala na osebo, ki je obstajala do t_n. To je le v primeru, če bi lahko iz kasnejšega časa vzročno delovala na prejšnji čas, torej le tedaj, če bi lahko *vzročno delovala nazaj v času.* Ker pa je ideja vnazajšnjega vzročnega delovanja nesmiselna, dejstvo, da nekoga

[22] Ta argument pravzaprav kaže, da za osebo, ki je umrla, ne more biti nekaj slabega niti prezgodnja smrt

ni po t_n, ne more biti zanj slabo do t_n. Ravno tako ne more biti zanj slabo niti po t_n, saj po t_n nje preprosto ni več. Zdi se torej, da se soočamo s paradoksno situacijo: *smrt bi lahko bila nekaj slabega za osebo, ki je umrla, le tedaj, če bi ta oseba živela še naprej tudi potem, ko je umrla.* Če tega paradoksa ne rešimo, moramo sprejeti Epikurjev argument in stališče, da smrt za nas ne more biti nekaj slabega preprosto zato, ker za osebo, ki umre, ne more biti nič več dobro niti slabo.[23]

Jasno je, da potem ko umremo, ne bomo nič več občutili niti znali. Vendar se zdi, da se nam v nekem smislu lahko dogajajo stvari, ki so za nas dobre ali slabe, tudi potem ko umremo. Čeravno oseba, ki je umrla, nič več ne čuti, *je zanjo slabo,* če se ne spoštuje njena oporoka ali če se oskruni njen grob.[24] *Zanjo je dobro,* če se po njej imenuje športna dvorana ali če knjige, ki jih je napisala, postanejo prodajna uspešnica. Tudi potem ko umremo, *bo za nas slabo, če* se bodo naši otroci obnašali slabo in sramotno, ravno tako kot bo *za nas dobro,* če se bodo obnašali dobro in častno, ker bo to pokazalo, da smo bili dobri starši in da smo jih dobro vzgojili. Ti primeri kažejo, da se nam v nekem smislu *lahko dogajajo dobre in slabe stvari tudi potem, ko umremo,* ko ti dogodki v nas gotovo ne bodo izzvali več nobenih občutkov. Lahko bi rekli, da je v istem smislu tudi smrt nekaj, kar je za nas slabo. Stanje prihodnjega nebivanja bo za nas slabo, čeravno potem ko bomo umrli, ne bomo ničesar čutili niti znali. To so le analogije, ki nam kažejo, da je smiselno govoriti o stvareh, ki bodo v nekem smislu bile za nas dobre ali slabe tudi potem, ko umremo. Da bi odgovorili na Epikurjev izziv pa moramo natančneje pojasniti, v katerem smislu so lahko stvari za nas dobre ali slabe tudi potem, ko nas več ne bo.

Obstajata dva načina, kako lahko poskusimo rešiti ta problem. Prvi način je poskusiti pokazati, da je *pojem osebe širši in da zajema več kot čisto biološki organizem,* katerega obstoj je omejen v prostoru in času, tako da je smiselno govoriti o tem, kaj je dobro ali slabo za dano osebo tudi potem, ko njen biološki organizem izgine.[25] Res je, pojem osebe je slojevit in zapleten ter ni povsem jasno, kaj vse natanko obsega, zato je to vprašanje zelo zanimivo. Zdi se, da ta pojem, med drugim, ne obsega le vsega tistega, kar ta oseba *je,* temveč tudi vse tisto, kar *bi lahko bila.* Verjetno tudi vse tisto, kar je naredila in celo tudi vse tisto, kar je hotela narediti, vendar ji je spodletelo. Zamislimo si, da se vrhunski strokovnjak v svojem poslu zapije in postane potepuh. Ko bi za takega človeka trdili, da je gromozanska škoda, da je tako končal, bi pijanca in potepuha vrednotili s perspektive tistega, kar bi bil, če se ne bi zapil, to je iz perspektive

[23] Glede tega je zelo zanimiv članek Pallea Yourgraua z naslovom "Can the Dead Really Be Buried", še posebej I. in II. dodatek.

[24] Bogovi Ahilu domnevno niso dovolili, da bi mrtvo Hektorjevo telo pustil nepokopano in ga vrgel psom in mrhovinarjem. Antični bogovi niso menili, da se človeku, ki ga več ni, ne more narediti nič slabega; torej niso bili epikurejci.

[25] To opcijo je predstavil Thomas Nagel v članku z naslovom "Smrt". Zanimivo je, da lahko podobno stališče najdemo tudi pri Heideggru (Heidegger: 1997: §§ 46-53); tudi on *je* menil, da je človeška eksistenca nekaj več kot zgolj biološki organizem. Ko Heidegger govori o smrti, je znano njegovo stališče, da je zavest o lastni končnosti konstitutivna za našo eksistenco. Čeravno je to eksistencialistično stališče zelo zanimivo, je vprašljivo in se ga tu ne bom natančneje lotil.

tistega, kar bi lahko bil, ne pa iz perspektive tistega, kar je *de facto*. Zdi se, da velja enako tudi za situacije, v katerih nam je žal za osebo, ki je umrla. Obžalovanje ni usmerjeno na tisto, kar ta oseba *de facto* je, to je na nič, saj je več ni, temveč na tisto, kar bi ta oseba bila, če ne bi umrla, to je na tisto, kar bi lahko bila. Ker naš pojem osebe zajema tudi vse tisto, kar bi oseba lahko bila, je smiselno trditi, da je smrt za osebo, ki je umrla, nekaj slabega. Po tem odgovoru oseba v nekem smislu živi še naprej tudi potem, ko biološki organizem umre. Jasno je, vprašanje je, ali je ta odgovor dober in ali oseba tudi potem, ko umre, v nekem smislu živi še naprej, zato ker obstajajo možne okoliščine, v katerih bi bila še naprej živa.

Drugi način, kako se poskusiti spoprijeti z Epikurjevim izzivom, je pokazati, da so nekatere stvari za nas dobre ali slabe *nečasovno, atemporalno*, to je, da resničnost teh sodb ni odvisna od časa, v katerem so dane osebe obstajale, in od časa, v katerem so se jim dane stvari dogajale. V tem primeru ne bi bilo ustrezno vprašati, *kdaj* je za neko osebo slabo, da je umrla, temveč bi bila zanjo njena smrt *vselej* nekaj slabega, kadarkoli naj bi se že zgodila in kadarkoli bi mi o tem govorili. Sicer *pa*, kakšno je sploh to vprašanje, kdaj je za nekoga slabo, da je umrl: preden je sploh umrl ali po tistem, ko je umrl? *Če je njegova smrt kdajkoli bila zanj nekaj slabega, potem je bila vselej nekaj slabega!*

Za bolečino je jasno, kdaj se dogaja in komu se dogaja. Toda ali je res jasno, kdaj je bolečina slaba in za koga je slaba? Videli smo, da je bil odgovor, da je slaba za tistega, ki ga boli, kolikor dolgo ga pač boli. Pa je to res dober odgovor? Ali ne velja enako kot za bolečino tudi za smrt? *Če je kdajkoli slaba, potem je vselej slaba!* Trinajstega avgusta 2003 me je bolel zob. Ali ne bi bilo bolje, da me ne bi bolel? Seveda bi bilo! Ali je sploh smiselno vprašati, kdaj? Kdaj bi bilo bolje, da me trinajstega avgusta 2003 ni bolel zob? Zdi se, da je to vprašanje *nesmiselno.*

Če bi bilo bolje, da me trinajstega avgusta 2003 ne bi bolel zob, potem bi *vselej* bilo bolje, da me trinajstega avgusta 2003 ne bi bolel zob. Če je kdo nesrečno umrl pri dvajsetih letih, je to gotovo nekaj slabega. Ali je smiselno vprašati: "Kdaj?" Ne! Edini odgovor je lahko: "Vselej!"[26] Za vselej je slabo, da je umrl pri dvajsetih, lahko pa bi živel osemdeset let. Nesrečna smrt mu je vzela šestdeset let življenja, to pa je nekaj, kar je za vselej slabo.[27] V Italiji je trinajstega avgusta 2003 nek človek dobil 60 milijonov € na lotu. Ta znesek je dobil trinajstega avgusta 2003, vendar: kdaj je bilo zanj dobro, da je trinajstega avgusta 2003 dobil ta znesek? Vselej! Kajti vselej je bolje dobiti 60 milijonov €, kot jih ne dobiti. Zdi se, da je pravzaprav neprimerno in nemogoče določiti, *kdaj* je nekaj za nekoga dobro ali slabo. Epikur je iz tega dejstva izpeljal sklep, da smrt *nikoli* ni slaba za človeka, ki umre, vendar je videti, da se je motil – smrt je *vselej* slaba za človeka, ki umre.[28]

[26] Nagel to intuicijo podpre z analogijo z višino; Abraham Lincoln je bil višji od Louisa XIV. Vprašanje je: "Kdaj?"

[27] To opcijo zastopa Fred Feldman (Feldman, 1991: KI. del); prisotna je tudi pri Nagelu v že omenjenem članku, zastopa pa jo tudi Harry Silverstein (Silverstein, 1980) v članku "The Evil of Death Revisited".

[28] Zanimivo je, da se obe opciji pravzaprav opirata na primerjavo dejanskega stanja z možnim stanjem stvari ali

Ta dva odgovora sta zelo zanimiva in imata precejšnjo intuitivno moč, vendar je vprašanje, ali sta sprejemljiva in ali jima res uspe pokazati, da je smrt za osebo, ki umre, nekaj slabega. Epikur bi odgovoril, da v navedenem primeru smrti dvajseletnika njegova prezgodnja smrt do njegovega dvajsetega leta ni mogla biti zanj nekaj slabega, saj je bil živ, po dvajsetem letu pa ni mogla biti zanj nekaj slabega, saj preprosto ni imela biti za koga kaj slabega, ker on ni več obstajal. Če so predstavljeni odgovori točni, potem druga različica Epikurjevega argumenta odpade, če niso, potem argument ostaja nepobit.

Simetrija preteklega in prihodnjega nebivanja

Drug znan argument, ki bi moral pokazati, da smrt pravzaprav ni nič slabega in da se je sploh ni treba bati, je predstavil Lukrec, najbolj znan Epikurjev privrženec, v delu *O naravi sveta* (Lucretius, 1959: III, 830- 870). Argument je zelo lep in zanimiv in se opira na *simetrijo med preteklim in prihodnjim nebivanjem*. Simetrija ali dozdevna simetrija sestoji iz naslednjega: *pred rojstvom nismo obstajali, ravno tako kot ne bomo obstajali niti potem, ko umremo.* Edina razlika med tema dvema vrstama nebivanja je v tem, da je eno bilo v preteklosti, drugo pa bo v prihodnosti, eno se je že končalo, drugo pa se še ni začelo. Naj bo točka *R* na časovni premici trenutek rojstva, točka S pa trenutek smrti:



Torej nismo obstajali pred *R*, ravno tako kot ne bomo obstajali po S; ni nas bilo pred rojstvom, ravno tako kakor nas ne bo tudi po smrti. Vprašanje je, ali je sploh kaka razlika med tema dvema nebivanjima, razen tega, da je bilo eno v preteklosti, drugo pa bo v prihodnosti? Zdi se, da – razen razlike v času – *ni* nobene druge razlike. Ali imamo dejstvo, da nismo obstajali, preden smo se rodili, za nekaj, kar je za nas slabo? Ali se bojimo tega, da nas ni bilo, preden smo se rodili? Ali se bojimo tega, da je pred našim rojstvom vladal ravno takšen nič, kot bo vladal potem, ko umremo? Ne! Nihče ne meni, da je to, da ni obstajal, preden se je rodil, zanj nekaj slabega, in nihče se ne boji niča, ki je obstajal pred njegovo eksistenco.[29] Bilo bi preprosto smešno in nesmiselno misliti, da smo bili s tem, da nismo obstajali, preden smo se rodili, oropani. Srž argumenta je v naslednjem: *če menimo, da preteklo nebivanje za nas ni nekaj slabega, zakaj potem mislimo, da je prihodnje nebivanje nekaj slabega?* Če je strah pred preteklim nebivanjem smešen, zakaj ni potem smešen tudi strah pred prihodnjim nebivanjem? *Ker ni nikakršne razlike med preteklim in prihodnjim nebivanjem in ker se sploh ne bojimo preteklega nebivanja, se nam* v *enaki meri ni treba bati*

– strokovno rečeno – na primerjavo dveh možnih svetov.

[29] Zdi se, da sicer otroci, ki so stari približno 3-5 let, z nekakšno bojaznijo razmišljajo o letih, ki so neposredno pred njihovim rojstvom in v katerih še niso obstajali. Poskušajte se spomniti, kako vam je bilo v teh letih, ko so se odrasli pogovarjali o času, v katerem vas še ni bilo, in ko ste razmišljali o tem, da vas ni bilo. Morda se ljudje bolj bojijo tistega nebivanja, ki jim je bliže na časovni premici – otroci preteklega, odrasli pa bodočega. Če bi se izkazalo, da je tako stališče med otroci res zelo razširjeno in celo običajno, bi to dejstvo gotovo vrglo novo in zelo zanimivo luč na Lukrečev argument in na naravo življenja in smrti nasploh.

niti prihodnjega nebivanja. Če naše preteklo nebivanje ne more biti za nas nekaj slabega, potem ravno tako ne more biti za nas nekaj slabega tudi naše prihodnje nebivanje. Če nimamo razloga, da bi se bali preteklega nebivanja, potem nimamo niti razloga, da bi se bali tudi prihodnjega nebivanja. Potemtakem smrt za nas ne more biti nekaj slabega, strah pred smrtjo pa je popolnoma neutemeljen in je torej enako smešen, kot bi bil strah pred dejstvom, da nismo obstajali, preden smo se rodili. Torej se smrti ni treba bati! Argument je videti takšen:

(P_1) V dejstvu, da nismo obstajali, preden smo se rodili, ni nič slabega.

(P_2) Ni nobene pomembne razlike med preteklim in prihodnjim nebivanjem.

$\therefore$ Torej ni nič slabega niti v dejstvu, da ne bomo obstajali, potem ko umremo.

Lukrečev argument nas privede v situacijo, da *ali* (1) sprejmemo stališče, da smrt ni nekaj slabega, *ali* da (2) spremenimo stališče do nebivanja pred rojstvom in začnemo trditi, da je to enako slabo kot tudi prihodnje nebivanje, *ali* da (3) poskušamo pokazati na pomembno razliko med preteklim in prihodnjim nebivanjem in tako upravičimo asimetrično stališče, ki ga imamo. Četudi je argument lep in zanimiv, se zanj ne zdi, da ravno "deluje". Tudi potem ko ga preberemo in razumemo, imamo še naprej asimetrično stališče do svojega preteklega in prihodnjega nebivanja. Morda nam argument omogoča, da svoj obstoj ugledamo z nekoliko drugačne perspektive, vendar pa nam ne pomaga, da se nehamo bati smrti: še naprej menimo, da ni nič slabega v dejstvu, da nismo obstajali, preden smo se rodili, vendar da je še kako slabo, da ne bomo obstajali, potem ko umremo. Zato je osnovno vprašanje v razpravi okrog tega argumenta naslednje: *"Kako upravičiti asimetrično stališče do preteklega in prihodnjega nebivanja?"*[30]

Zgodnejše rojstvo in kasnejša smrt

Ko smo razpravljali o Epikurjevem argumentu, da nam smrt nič ne pomeni, smo videli, da *deprivacijska pojasnitev* najbolje zajema intuicije o tem, zakaj je smrt nekaj slabega, namreč zato, ker nas oropa izkustev, ki bi jih imeli, če ne bi umrli. Toda Lukrečev argument iz simetrije na zelo zanimiv način podvomi o tej pojasnitvi. *Če je namreč zgodnejša smrt za nas slaba zato, ker nas oropa izkustev, ki bi jih imeli, če ne bi umrli, potem bi moralo biti za nas ravno tako slabo tudi kasnejše rojstvo, zato ker nas oropa izkustev, ki bi jih imeli, če bi se prej rodili.* Če bi torej sama dolžina življenja bila edino, do česar nam je, potem bi zgodnejša smrt morala biti enako slaba kot tudi kasnejše rojstvo. Toda vsi imajo zgodnejšo smrt za nekaj slabega, kasnejšega rojstva pa nihče nima za nekaj slabega. Nihče ne meni, da je s tem, ko se je rodil takrat, ko se je rodil, ne pa 10 let prej, oškodovan, po drugi strani pa bi vsi menili, da smo oškodovani, ko bi vedeli, da bomo umrli 10 let prej, kot bi sicer umrli. Ko torej upoštevamo asimetrično stališče do preteklosti in prihodnosti, vidimo, da smrt ni za nas nekaj

[30] Najboljši prikaz problemov povezanih s splošno asimetrijo časa je nemara knjiga Paula Horwicha (Horwich, 1987).

slabega zgolj zato, ker nam zmanjšuje *skupno količino* občutkov in izkustev, temveč prej zato, ker nam zmanjšuje *prihodnjo količino* občutkov in izkustev. Slaba je zato, ker nas oropa nečesa, kar je *šele pred nami*, ne pa *nečesa, kar je že za nami*. Zato mora zadovoljivo upravičenje prepričanja, da je smrt za nas nekaj slabega, zajeti tudi to asimetrijo ali pa pokazati, da je le-ta pravzaprav iracionalna. Pokazati mora torej, zakaj ni pomembno zgolj *koliko*, temveč *koliko še*. *Deprivacijska pojasnitev sama na sebi ne more pojasniti, zakaj je bolje imeti 20 let v življenju, ki traja 60 let, kot imeti 79 let v življenju, ki traja 80 let.* Jasno, če je to sploh bolje. Morda je tako stališče zgolj zadeva temperamenta in nemara se ne bi vsi strinjali. Morda bi nekateri menili, da je 80 let bolje od 60 in pika, ne glede na to, koliko so trenutno stari, 20 ali 79. V vsakem primeru pa večina ljudi daje prednost prihodnjemu življenju pred preteklim življenjem. Načelo *čim več, tem bolje* je imuno na razliko med preteklostjo in prihodnostjo, naša stališča pa niso. Res je, drago nam je tudi preteklo življenje, vendar veliko manj od prihodnjega. Na primer, če bi bili gotovi, da bomo jutri umrli, ali bi raje imeli 20 ali 80 let? V takih okoliščinah bi raje imeli 80 let. Bolje je imeti 80 let v preteklosti in en dan v prihodnosti, kot 20 let v preteklosti in en dan v prihodnosti. Tudi preteklo življenje je vredno, vendar ne enako kot prihodnje. *Če bi dolžina življenja bila edino, za kar nam gre, potem bi nam moralo biti vseeno, ali bomo živeli dlje v prihodnosti ali dlje v preteklosti; moralo nam bi biti vseeno, ali bomo kasneje umrli ali se prej rodili.* Jasno, to nam *ni* vseeno! Hoteli bi živeti dlje v prihodnosti, ne pa dlje v preteklosti. Hoteli bi imeti še nekaj *pred seboj*, ne pa *za seboj!*

Zamislimo si, da smo naredili neko dobro delo in da nas za nagrado Bog vpraša, ali hočemo, da nam pokloni še eno leto življenja, da živimo leto dlje, kot bi sicer živeli. Rečemo, da hočemo, on pa nas naredi za leto *starejše*, naredi, da smo se rodili leto dni *prej*. Umrli bomo torej takrat, kot bi sicer umrli, zdaj pa smo rojeni leto dni prej. Ali bi se čutili prevarani? Bi! Vsekakor! Zakaj pa, saj je naredil ravno to, kar je obljubil: podaljšal nam je življenje za leto dni, le da nam to leto ni dodal k prihodnosti, temveč k preteklosti. To leto je naše, preživeli smo ga. Preteklo ali ne, naše je! Naše življenje je res postalo leto dni daljše. To leto pa bi vendarle hoteli imeti v *prihodnosti*, hoteli bi, da imamo leto več *pred seboj*, ne pa za *seboj*. Zamislimo si, da Bog pri svoji odločitvi omahuje in da nam ponudi izbiro: ali hočemo dobro leto v preteklosti ali negotovo leto v prihodnosti?[31] Kaj bi izbrali? Zakaj? Kaj pa, če nam ponudi dobro leto v preteklosti in povprečno v prihodnosti? Morda bodo nekateri sprejeli Lukrečev argument in izenačili svoja stališča do preteklosti in prihodnosti. Taki bi potem morali imeti raje dobro leto v preteklosti kot povprečno leto v prihodnosti; ni bistveno, ali smo ga preživeli ali ga bomo šele preživeli, bistveno je, da je boljše! Četudi sprejmemo simetrično stališče do preteklega in prihodnjega življenja, še vedno ni jasno, ali moramo isto stališče zavzeti tudi do preteklega in prihodnjega nebivanja. Ali se bomo nehali bati prihodnjega nebivanja ali bomo strah pred nebivanjem enako porazdelili na

[31] Nekaj manj kot tretjina študentov, ki so bili anketirani med predavanjem, bi raje vzela dobro leto v preteklosti kot negotovo leto v prihodnosti.

preteklo in prihodnje nebivanje? Zdi se, da asimetrično stališče do preteklega in prihodnjega nebivanja zrcali naše asimetrično stališče do preteklosti in prihodnosti nasploh; *več nam je do prihodnosti kot do preteklosti!* Zakaj? Kako upravičiti to razliko? Zakaj se bolj veselimo prihodnjih kot preteklih uspehov? Zakaj se bolj bojimo prihodnje kot pretekle bolečine? Obstajajo kajpada ljudje, ki "živijo v preteklosti". To so v glavnem starejši ljudje, ki so vložili ogromno truda in čustev v nekaj, česar več ni, in so še naprej ostali navezani na te načrte in projekte, zasebne ali javne. Toda celo taki ljudje so v zelo pomembnem in preprostem smislu obrnjeni v prihodnost. Ali bi si jugonostalgik, star fašist, propadli aristokrat ali vdova, ki so ji tudi otroci že umrli od starosti, bolj želeli, da bi si *zob, ki si ga morajo jutri izdreti, izdrli že pretekli teden?* Mislim, da je odogovor očiten. *Bi!* Ravno tako bi si bolj želeli, da se jim *zgodi nekaj lepega v* prihodnjem mesecu dni kot pa v *preteklem* mesecu. Celo *politiki, ki so obsedeni* z zgodovino, trdijo, da je poznavanje *nacionalne zgodovine* pomembno predvsem zato, ke nam pomaga *pri sprejemanju pravilnih političnih odločitev* v sedanjosti in prihodnosti.

Vprašanje *je* torej, *ali lahko* navedemo kakšen racionalni razlog, zaradi katerega bi nam moralo biti več do prihodnosti kot do preteklosti? Ali lahko nekako pojasnimo, zakaj nam je prihodnost vredna več od preteklosti in *zakaj bi hoteli umreti kolikor je le mogoče pozneje, vendar pa se ne bi hoteli roditi karseda prej?* Po eni strani je jasno, da lahko na razne načine odložimo lastno smrt: lahko se zdravo hranimo, nehamo kaditi, se redno rekreiramo, smo pazljivi v prometu itn. Ti ukrepi res dajejo določene rezultate in dokaj verjetno je, da tako lahko res *pomaknemo svojo smrt dlje v prihodnost.* Po drugi strani pa ni *nobenega načina, kako bi lahko svoje rojstvo pomaknili dlje* v *preteklost.* Na preteklost ne moremo vplivati! Rodili smo se, ko smo se rodili in tu ni možna *več* nobena sprememba.[32] Zakaj se sploh truditi za nekaj, na kar ne moremo vplivati? *Edini način, da podaljšamo svoje življenje, je, da ga podaljšamo v prihodnost; ne moremo ga podaljšati* v *preteklost.* Jasno nam je torej, zakaj nam je *de facto* bolj pri srcu prihodnost kot preteklost. To dejstvo je povsem dober razlog, da nam bo več do prihodnosti kot do preteklosti. Vendar pa se zdi, da to *ni edini* razlog in da nam je do prihodnosti bolj kot do preteklosti *še zaradi nečesa drugega.* Zdi se, da bi nam bilo v okoliščinah, v katerih bi lahko po mili volji premikali datum lastnega rojstva, več do prihodnosti kot do preteklosti. Ravno to kaže zamišljeni primer z božjo ponudbo. Razen tega, tudi ko bi lahko vplivali na preteklost, bi si še naprej bolj želeli, da smo si zob, ki ga moramo izdreti naslednji teden, že izdrli pretekli teden. Dejstvo, da ne moremo vplivati na preteklost, da pa lahko na prihodnost, ne more pojasniti olajšanja, ki nastopi, ko so slabe stvari za nami. "Hvala Bogu, končano je!" bi veljalo za mnoge stvari tudi tedaj, ki bi obstajali načini, kako spremeniti preteklost. Zato je vprašanje, zakaj bi tudi v okoliščinah, v katerih bi lahko to naredili, preferirali prihodnost pred preteklostjo, ravno tako kot jo preferiramo tudi v okoliščinah, v katerih *de facto* smo. Vprašanje pravzaprav

[32] Dummett v članku "Bringing about Past" mimogrede omeni, kako je za Jude zelo resen greh, če molijo k Bogu, da bi spremenili preteklost. To sodobnim Judom, celo tistim, ki poznajo tradicijo, ni znano.

predstavlja miselni eksperiment, ki bi nam moral pomagati, da izostrimo intuicije o tem, zakaj nam je in ali nam je res več do življenja v prihodnosti kot do življenja v preteklosti in zakaj imamo dejstvo, da nas – potem ko umremo – ne bo več, za nekaj slabega, dejstvo, da nas ni bilo, preden smo se rodili, pa imamo za nekaj, kar je vrednostno nevtralno. S tem miselnim eksperimentom poskušamo natančneje pokazati, da premisa (P_2) – da ni nobene pomembne razlike med preteklim in prihodnjim nebivanjem – v Lukrečevem argumentu ni resnična in da je strah pred prihodnjim nebivanjem upravičen, medtem ko strah pred preteklim nebivanjem ni upravičen. Ravno tako poskušamo hkrati tudi dopolniti deprivacijsko pojasnitev dejstva, da je smrt za nas nekaj slabega, in pokazati, da nam je predvsem pomembna dolžina življenja, ki je pred nami. Ne pa zgolj skupna dolžina življenja. Ogledali si bomo nekaj možnih odgovorov.

Ti premisleki bi nas lahko napeljali k stališču, da je asimetrija pravzaprav *iracionalna* in da je potemtakem *povsem vseeno, ali se bomo rodili prej ali bomo umrli kasneje.* Če so sicer vse druge stvari enake, je daljše življenje boljše od krajšega in pika. Vseeno, koliko je česa ostalo, več je boljše in konec zgodbe! Vseeno je, koliko je izza nas in koliko pred nami. Bolje je imeti 79 let v življenju, ki traja 80 let, kot 40 let v življenju, ki traja 60 let. Preteklo in prihodnje življenje je enako vredno. Racionalno je preferirati večje dobro pred manjšim, kakdarkoli naj se to dobro že zgodi. Zaradi tega je v primeru božje ponudbe racionalno preferirati dobro leto v preteklosti pred negotovim letom v prihodnosti. Ravno tako kot je racionalno preferirati večji dobitek v daljnji prihodnosti pred manjšim dobitkom v bližnji prihodnosti. Ko bi lahko izbirali, kaj boste vzeli, 8.000 € jutri ali 10.000 € po šestih mesecih? Racionalno je izbrati večjo vsoto – zato ker je večja ne glede na to, kdaj nam bo izplačana.[33] Zato je iracionalno preferirati kasnejšo smrt pred zgodnejšim rojstvom, bistvena je skupna dolžina življenja, vendar ne v primeru, ko je življenje umeščeno v čas.[34] Jasno, pod predpostavko, da je leto enega življenja vredno enako kot leto drugega.

Možno je tudi stališče, da je asimetrija do preteklosti in prihodnosti preprosto *primitivno dejstvo,* ki ga ni mogoče in ga ni treba upravičiti s sklicevanjem na neko še bolj temeljno značilnost, povezano z našim obstojem v času. Tako smo in pika. Ker gre za eno od temeljnih dejstev, bi to – v skrajni liniji – lahko bilo celo točno. Vendar se zdi, kot bomo videli, da je tudi to asimetrijo mogoče raziskati in da lahko najdemo vsaj nekakšne razloge, ki jo upravičujejo. Poleg tega pa ima to, da neko dejstvo sprejmemo kot primitivno in zavračamo, da bi se ga poskušalo pojasniti, prizvok intelektualne kapitulacije. Če ga je mogoče nekako pojasniti,

[33] Ljudje so *de facto* bolj naklonjeni manjši koristi v bližnji prihodnosti kot večji koristi v daljnji prihodnosti. Ali so zaradi tega iracionalni? Obstaja niz razlogov, s katerimi bi lahko upravičili take preference. Najbolj pogost odgovor je, da verjetnost, da bo do izplačila res prišlo, pada z oddaljenostjo v času; bolje vrabec v roki kot golob na strehi; bolje je imeti gotovih 8.000 € kot manj gotovih 10.000 €. Poleg tega racionalni osebi tudi ni niti malo vseeno, kdaj bo prišla do neke koristi. Kdaj je bolje zadeti sedmico na lotu, pri petindvajsetih ali pri petinšestdesetih? V izreku "Kdor hitro da, dvakrat da!" je gotovo neka modrost. Poleg tega, čas ne igra enake vloge pri nekom, ki umira od lakote, in pri nekom, ki ima vse, kar potrebuje. Zato sploh ni jasno, zakaj bi bili ljudje, ki preferirajo manjšo korist v bližnji prihodnosti pred večjo v daljši prihodnosti, iracionalni?

[34] To opcijo zastopa Derek Parfit, ki meni, da je to racionalno stališče. Toda ker smo evolucijsko ustrojeni tako, da smo obrnjeni v prihodnost, nas evolucija pravzaprav onemogoča, da bi imeli naj boljše stališče do lastnega življenja in smrti (Parfit, 1984: 77).

potem je to vsekakor potrebno poskusiti.

Prvi možni odgovor je, da nam gre predvsem za tisto, kar se dogaja *zdaj, v sedanjosti,* ne pa za ono, kar je bilo, ali ono, kar bo. Zato je racionalno, če damo največjo težo dogodkom v sedanjosti. Toda prihodnost bo enkrat postala sedanjost, preteklost pa nikoli. Zato bi nam v sedanjosti moralo biti tudi do prihodnosti, vendar pa ne tudi do preteklosti; čim bližji nam je dogodek v prihodnosti, večjo vrednost ima za nas, čim bolj je oddaljen, tem manjšo vrednost ima. To stališče odlično pojasnjuje in upravičuje preferiranje manjše koristi v bližnji prihodnosti pred večjo koristjo v daljnji prihodnosti. Kaže, da je bolj racionalno vzeti 8.000 € jutri kot pa 10.000 € v roku šestih mesecev.[35] To stališče ravno tako upravičuje tudi asimetrijo v stališčih do prihodnjega in preteklega nebivanja. Naše preteklo nebivanje ni nekaj, kar bi lahko bilo za nas slabo preprosto zato, ker ne bo nikoli več v sedanjosti, medtem ko je po drugi strani naše prihodnje nebivanje za nas nekaj slabega zato, ker bo enkrat prišlo v sedanjost. Še več, čim bliže nam je, tem slabše je za nas. Jasno, vprašanje je, v kolikšni meri je ta odgovor obranljiv, vendar se zdi, da samo takšen ali podoben odgovor lahko pojasni asimetrijo v stališčih.

Drugi možni odgovor je, da ustrezna razlika med preteklim in prihodnjim nebivanjem sestoji iz dejstva, da smisel, v katerem nismo obstajali, preden smo se rodili, *ni enak* smislu, v katerem ne bomo obstajali, ko bomo umrli.[36] Preden se je namreč rodil človek, ni obstajal *nekdo,* ki se še ni rodil, vendar pa je potem, ko je človek umrl, obstajal *nekdo,* ki je umrl.[37] Nesmiselno je govoriti o *ljudeh, ki se še niso rodili;* oni preprosto nikoli niso obstajali, vendar pa je povsem smiselno govoriti o ljudeh, *ki so umrli.* O tistih, ki so obstajali, lahko govorimo: o tem, kakšen značaj so imeli, kaj so radi počeli itn. O onih, ki še ne obstajajo, ne moremo govoriti, in to ne le zato, ker še *ne vemo,* kakšni bodo, temveč zato, ker preprosto *še ne obstajajo.* Res je, da Lukrec ni obstajal, preden se je rodil. Vendar pa je to resnica *zdaj,* potem ko je obstajal. V času *pred njegovim rojstvom* to ni bila resnica. Preden se je rodil, ni bilo *nikogar,* ki ne bi obstajal, vendar pa je potem, ko je umrl, bil *nekdo,* ki ni obstajal.

Lukrec ni čakal – preden se je rodil – na vrsto, da bi se rodil in da bi začel obstajati, temveč ga preprosto ni bilo. Le če bi bil fatalizem resničen, bi bila v tem primeru trditev, da Lukrec ni obstajal, preden se je rodil, lahko resnična, tako preden se je rodil, kot potem, ko se je rodil. Torej le v primeru, da je bilo *od vekomaj* resnično, kdo vse bo obstajal, kdaj se bo rodil, kako bo živel in kdaj bo

[35] Tudi to je ena od opcij, ki jo obravnava Parfit (Parfit, 1984: II. del).

[36] Lukrec seveda meni, da je ta smisel enak. Pri Longu in Sedleyju – v angleškem prevodu – in pri Sovretu – v slovenskem prevodu – stoji, da ni razlike med nebivanjem tistega, ki je umrl, in onega, ki *nikoli ni bil rojen* (Lukretius, 1959, III, 868). Ne vem, kaj je zapisano v izvirniku, toda če je angleški in slovenski prevod točen, potem Lukrečevo stališče ne more biti resnično v nobenem od obeh smislov.

[37] To je razlika, na katero se je mogoče opreti v razpravi okrog dopustnosti splava. Tisti, ki namreč menijo, da je splav uboj, bi lahko trdili, da razlika med med kontracepcijo in splavom sestoji iz tega, da v primeru kontracepcije nikomur ni bil odtegnjen obstoj, ker *nihče še ni začel obstajati,* medtem ko smo v primeru splava nekomu vendarle odvzeli življenje, ker je *nekdo že začel obstajati.* Jasno, tisti, ki menijo, da splav ni uboj, bodo trdili, da iz tega sestoji razlika med splavom in detomorom, ne pa razlika med kontracepcijo in splavom.

umrl. Če fatalizem ni resničen in če prihodnost ni določena v enakem smislu kot preteklost, potem v času, preden smo se rodili, preprosto ni bilo res, da še ne obstajamo. Stavek "Lukrec ne obstaja" je bil nesmiseln, preden se je rodil, neresničen, medtem ko je živel, in resničen, potem ko je umrl. Ta asimetrija v času pojasnjuje tudi asimetrijo v naših stališčih do preteklega in prihodnjega neobstajanja; ker v času, preden smo se rodili, *ni bilo nikogar,* ki ne bi obstajal, nam za ničemer ne more biti žal, saj to pravzaprav niti ni bilo *naše* neobstajanje. Po drugi strani pa je jasno – potem ko umremo —, *kdo* je tisti, ki ne bo več obstajal: *mi* smo tisti, ki ne bomo več obstajali, in to je tudi tisto, kar obžalujemo in česar se bojimo. Naše preteklo neobstajanje za nas ne more biti nekaj slabega preprosto zato, ker ni imelo biti za koga slabo, medtem ko je naše prihodnje neobstajanje za nas nekaj slabega zato, ker ima biti za koga slabo. Od tod asimetrija v stališču.

Tretji možni odgovor, soroden drugemu, je, da asimetrija izhaja iz dejstva, da bi življenje, ki se nadaljuje v prihodnost, še naprej bilo *naše* življenje, medtem ko življenje, raztegnjeno v preteklost, *ne bi bilo* več *naše* življenje.[38] Po tej intuiciji preprosto ni smiselno govoriti o tem, ali bi bilo za nas bolje ali ne, da smo se rodili 5, 10 ali 20 let prej. Jaz se nisem mogel roditi pet let prej, kajti če bi se rodil pet let prej, *to ne bi bil več jaz, temveč nekdo drug.* Če bi se rodil pet let prej, bi imel druge prijatelje, drugačno družbo iz vrtca, šole in fakultete, bral bi druge knjige, morda bi opravljal drugo delo. Če be se rodil pet let prej, bi se morda rodil v drugem mestu ali v drugi državi itn. Moja biografija bi bi la drugačna, imel bi drugačna izkustva in drugačne spomine. Ker pa so ravno izkustva in spomini tisti, ki me naredijo za osebo, kakršna sem, ne pa za neko drugo osebo, bi bil – če bi se rodil pet let prej – neka druga oseba, ne pa ta, katera sem. Tisto, kar me naredi za osebo, kakršna sem, je *moja biografija* in *psihološka kontinuiteta,* ki jo imam.[39] Če bi se rodil pet let prej, bi imel drugačno biografijo in drugačno psihološko kontinuiteto in zaradi tega to ne bi bil več jaz, temveč nekdo drug. Zato se *nisem mogel* roditi prej, kot sem se rodil. Nemožnost, o kateri tukaj teče beseda, je *logična nemožnost,* saj bi moji starši *fiziološko* lahko imeli otroka pet let prej, kot sem se rodil jaz.[40] Vendar je *logično nemogoče,* da bi ta oseba bila jaz, to bi bil nekdo drug, verjetno podoben meni, vendar vseeno nekdo drug. V skladu s to intuicijo je želja, da bi se rodili pet ali deset let prej, *logično protislovna* v istem smislu kot želja, da bi kvadrat bil okrogel ali da bi bilo železo leseno. Psihološko je morda mogoče razmišljati o tem, kako bi bile videti stvari, če bi bili pet ali deset let starejši – to je si je morda mogoče tudi želeti, toda v strogem smislu je to protislovje, nekaj, kar je nujno nemogoče.[41]

[38] To trdita Thomas Nagel v članku "Smrt" in Frederik Kaufman v članku "Pre-Vital and Post-Mortem Non-Existence".

[39] Psihološka kontinuiteta zavzema osrednje mesto v razpravah okrog osebne identitete. Idejo je prvi sistematično predstavil Locke (Locke, 1924: II. knjiga, XXVII poglavje).

[40] Zanimiva so tudi Kripkejeva stališča o osebni identiteti; on na primer meni, da ne bi mogli biti ista oseba, kakršni smo, če bi imeli druge starše (glej Kripke, 2000: 99).

[41] Vprašanje je, ali *je* to dober argument. Morda gre prej za nejasno intuicijo kot za odločilni argument. Ali si lahko konsistentno zamislite, da ste pravzaprav alkemik, ki je v dvanajstem stoletju odkril eliksir večne mladosti, tako da imate sedaj okrog 850 let? Ali bi to še naprej bili vi? Zdi se mi, da psihološka kontinuiteta in enotnost zavesti ne izključujeta take možnosti. Vendarle, ali lahko odgovorite na to vprašanje z da ali ne?

Ker so protislovna stanja stvari nujno nemogoča, jih racionalen človek ne bo želel[42] Zamislimo si, da se je nekdo rodil leta 1900 in da je umrl leta 1980. Oglejmo si naslednja protidejstvena pogojnika:

(1) Tudi če ne bi umrl leta 1938, ne bi dočakal konca druge svetovne vojne.

(2) Tudi če se ne bi rodil leta 1880, bi se rodil leta 1890.

Za prvega ni nujno, da je, vendar je *lahko* resničen. Zamisliti si je mogoče okoliščine, v katerih nekdo iz takega ali drugačnega razloga ne bi dočakal leta 1945, celo če ne bi umrl leta 1938. Oseba, za katero se tr- di, da tako in tako ne bi dočakala konca vojne, *je ista oseba*, za katero se trdi, da je umrla leta 1938. Drugi pogojnik pa *ne more* biti resničen: če se oseba, o kateri gre beseda, ne bi rodila leta 1880, preprosto ne bi nikoli obstajala. Oseba pa, ki bi se rodila leta 1890, *bi bila druga oseba*, drugačna od osebe, ki se je *de facto* rodila 1880.[43] Zdi se, da ta asimetrija ne velja le za ljudi, temveč tudi za druga bitja. Možno je, da bi stol še obstajal, če se Ben ne bi usedel nanj, vendar pa ni mogoče, da bi ga – če ga ni izdelal Rogina v Zagrebu leta 1958 – napravil Perman v Kastvu leta 1972. Mogoče je, da bi Fanuko še vozil starega Golfa, če ga bi pravočasno dal kleparju, vendar pa ni mogoče, da bi bil ta avto – če ne bi bil napravljen v TAS-u leta 1979 – napravljen v Nemčiji leta 1982 itn. Vse to so primeri, ki kažejo, da čas nastanka neke stvari sodi v merilo njene identitete, čas njenega izginitve pa ne.[44] Če bi stvar pozneje zginila, bi bila to še vedno ista stvar, če pa bi kasneje nastala, to ne bi bila več ista stvar. Enako velja tudi za naše obstajanje: lahko umremo tudi kasneje, kot bomo umrli, vendar se logično ne moremo roditi bistveno prej ali kasneje, kot smo se rodili. *Zato je jasno, zakaj si želimo kasneje umreti, vendar pa se ne želimo prej roditi* in zakaj zgodnejšo smrt štejemo za zlo, kasnejše rojstvo pa ne. Ko bi živel pet let dlje, bi to še naprej bil jaz, toda če bi se rodil pet let prej, to ne bi bil več jaz, temveč nekdo drug. Tisto, kar je v tem kontekstu pomembno, je, da si jaz želim, da bi *jaz* bil ta, ki bo živel dlje, ne pa da bi to bil kdo drug. Zaradi tega je edini način, da živim dlje, da živim dlje v prihodnosti. Celo ko bi lahko spremenil preteklost, ne bi (logično) mogel narediti, da bi se rodil prej.

Ti trije odgovori bi morali pojasniti, zakaj štejemo zgodnejšo smrt za izgubo, kasnejše rojstvo pa ne, in zakaj bi, če bi lahko izbirali, hoteli umreti kasneje, ne bi se pa hoteli prej roditi. Če so pri tem uspešni, potem je jasno, zakaj zavrniti Lukrečev argument: zato, ker premisa (P_2), ki trdi, da ni nobene pomembne razlike med preteklim in prihodnjim neobstajanjem, preprosto ni resnična. Četudi

Parfit na primer obžaluje, da se ni rodil 300 let prej, da bi lahko užival v družbi Huma, Byrona, Čehova, Nietzscheja in Sidgwicka (glej Parfit, 1984: 176).

[42] Vprašanje je, ali nam lahko racionalnost prepreči, da si želimo, da protislovna stanja stvari postanejo resnična. Morda bi si tudi racionalna oseba lahko želela, da vsota kvadratov nad katetama ne bi bila enaka kvadratu nad hipotenuzo ali da zatrjevanje konsekvensa ne bi bila logična napaka, temveč veljaven sklep ali da bi bilo parnih števil dvakrat manj kot parnih in neparnih itn. Čeravno

[43] je ta argument dokaj prepričljiv, ga ne sprejemajo vsi. Na primer, Alex Byrne dvomi, da je sodba "če se oseba ne bi rodila 1890, se sploh ne bi rodila" nujno resnična, in meni, da gre morda prej za izjemno nizko verjetnost kot pa za nujno neresničnost.

[44] V merilo identitete za fizične predmete sodi poleg časa nastanka verjetno tudi kraj nastanka, material, iz katerega je predmet izdelan, v primeru artefaktov pa tudi ljudje ali podjetja, ki so jih izdelala itn.

so ti odgovori sprejemljivi in imajo precejšnjo težo, je vprašanje, ali so točni in – kar je še pomembnejše – ali so zadostni, da upravičijo asimetrijo v naših stališčih do preteklosti in prihodnosti. Drugi in tretji odgovor namreč kažeta, da neobstajanje pred rojstvom pravzaprav sploh ni bilo *naše* neobstajanje, tako da to za nas ne more biti nekaj slabega. Toda ti odgovori ne upravi- čujejo naše splošne preference prihodnosti pred preteklostjo; na primer dejstva, da bi raje videli, da bi si zob, ki ga moramo izdreti naslednji teden, izdrli že pretekli teden. Samo prvi odgovor (da nam je več do sedanjosti in tistega, kar bo šele prišlo v sedanjost, kot do onega, kar več nikoli ne bo v sedanjosti) lahko upravičuje to splošno asimetrijo v stališčih do preteklosti in prihodnosti. Toda vprašanje je, ali je ta odgovor točen in – če je – ali je zadosten, da upraviči asimetrijo v naših stališčih. Zato se zdi, da ostaja vprašanje, ali še obstajajo razlogi, zaradi katerih nam je več in zaradi katerih bi nam moralo biti več do prihodnosti kot do preteklosti, še nadalje odprto.

"Matematični argument"

Lukrec je predstavil tudi tako imenovani *"matematični argument"* proti strahu pred smrtjo (Lucretius, 1959: III, 1070-1099). Argument se opira na *razmerje* časa, ki ga bomo živeli, in časa, ki ga ne bomo živeli: kakorkoli dolgo naj že živimo, je čas, ki ga bomo preživeli, *končen*, pa naj si bo to ena sekunda ali milijarda let. Po drugi strani pa bomo, potem ko bomo umrli, *neskončno* dolgo mrtvi. Torej: kakorkoli dolgo naj že živimo, je naše življenje pravzaprav neskončno majhen del skupnega časa. Kolikorkoli naj je že dolgo naše življenje, je čas, ko bomo mrtvi, vselej neskončnokrat daljši od časa, ko bomo živeli. Ko na svoje življenje in smrt pogledamo s te perspektive, nam bi moralo biti *vseeno, kako dolgo bomo živeli.* Z drugimi besedami, ne smemo se obremenjevati s tem, ali bomo umrli prej ali kasneje. *Ker bomo neskončno dolgo mrtvi, je vseeno, kako dolgo bomo živeli.*[45] Jasno je, da je vprašljivo, ali "matematični argument" res odstranjuje ali vsaj zmanjšuje strah pred smrtjo in ali res kaže, da je vseeno, kako dolgo bomo živeli. Zdi se, da ne. Res je, da je čas našega obstajanja, kolikor dolgo naj že bo, neskončnokrat krajši od časa našega neobstajanja. Vendar pa ni jasno, kako bi to lahko bil razlog za to, da nam postane vseeno, kako dolgo bomo živeli, in da se nehamo bati prihodnjega neobstajanja. *Čeravno vemo, da bomo neskončno dolgo mrtvi, si mi še naprej želimo živeti čim dlje.* Ne zanima nas, kako dolgo bomo mrtvi, ne zanima nas niti razmerje med trajanjem življenja in trajanjem smrti, *zanima nas, kako dolgo bomo živi!* To je tisto, za kar nam gre. Če je eden umrl pri dvajsetih, drugi pa pri osemdesetih, se je prvemu zgodila tragedija, drugemu pa ne, četudi bosta tako en kot drug neskončno dolgo mrtva. Iz dejstva, da bomo neskončno dolgo mrtvi, preprosto *ne sledi*, da nam mora biti vseeno, kako dolgo bomo živeli. Še več, zdi se, da vztrajanje pri dejstvu, da "bomo neskončno dolgo mrtvi", to je, da neskončno dolgo ne bomo obstajali, ne

[45] Zdi se, da ima argument implikacijo, da bi bilo smiselno hoteti živeti dlje, ko bi bilo mogoče biti krajši čas mrtev, to je, ko bi bil čas, ko bomo mrtvi, končen. Na to poanto je opozoril Bernard Williams (Williams, 1973: 84).

zmanjšuje strahu pred smrtjo, temveč ga ravno povečuje.

Perspektiva večnosti nas v nekem smislu res postavlja v položaj, da svoje lastno obstajanje in neobstajanje vidimo kot nekaj neskončno malega in nepomembnega. Toda v tem primeru smrt neha biti za nas nekaj pomembnega le v meri, v kateri tudi naše življenje postane za nas nepomembno: vseeno je, kdaj bomo umrli le zato, ker je vseeno, ali smo kdajkoli živeli. V tem primeru smrt neha biti za nas nekaj slabega le zato, ker življenje neha biti nekaj dobrega. Res ni jasno, na kaj bi se lahko naslonili v argumentaciji z osebo, ki ji ni do lastnega obstajanja, in kako bi lahko taki osebi pokazali, da je konec njenega obstajanja zanjo nekaj slabega. Toda argument bi moral tudi ljudem, ki jim je do lastnega obstajanja, pokazati, da njihova smrt, četudi prezgodnja, zanje ne more biti nekaj slabega zato, ker bodo neskončno dolgo mrtvi. Toda zdi se, da argument tega ne more pokazati osebam, ki jim je do lastnega obstajanja. Zato argument ni dober, ker ima lahko v najboljšem primeru nekakšno težo le za tiste, ki tako in tako verujejo v tisto, kar z njim poskušamo dokazati.

Mislim, da je analiza argumentov in stališč o vrednosti lastne smrti pokazala, da Epikurjevi in Lukrečevi argumenti, četudi zelo zanimivi in kvalitetni, vseeno niso pokazali, da smrt ne more biti za nas nekaj slabega, *kadarkoli naj že nastopi*. Ravno tako mislim, da niti Williamsov argument ne pokaže, da je smrt za nas lahko nekaj slabega le v primeru, če nastopi *predčasno*, temveč da je smrt – vsaj v običajnih okoliščinah – *vselej* nekaj slabega za osebo, ki umira, navzlic biološki omejitvi vrste. *prevedel Božidar Kante*

DEATH

In this article author critically examines well known arguments that purport to show that death is not something bad for the person that dies, and tries to show that these arguments are not sound, that is, author tries to show that death really is something bad for person that died. Author believes that Williams did not show that eternal live would be unbearable and that death after sufficiently long life would be a relief. Further, famous Epicurus' argument against the fear of death author splits into two related but nevertheless separated arguments, and tries to show that the first one is based on false or at least questionable theory of value, while the other one represents a serious intellectual challenge. Different answers to Lucretius' argument from the asymmetry of post and future nonexistence are examined, author argues that none of these answers is completely satisfactory. Finally, author examines and rejects "mathematical argument" against the fear of

death as invalid.

Keywords: life, death, values, Epicurus, Lucretius, deprivational account, asymmetry of post and future nonexistence, personal identity.

Literatura

Aristoteles (1994). *Nikomahova etika.* Ljubljana: Slovenska matica.

Aristoteles (1999). *Metafizika.* Ljubljana: ZRC.

Bošnjak, B. (1978). *Filozofija od Aristotela do renesanse.* Zagreb: Naknadni zavod Matice hrvatske.

Diogen Laertski (1979). *Životi i mišljenja istaknutih filozofa.* Beograd: Beogradski grafičko-izdavački zavod.

Epikur (1959a). "Maksime". V T. Lucretius Carus, *O naravi stvari,* Ljubljana: Slovenska matica, str. 501-504.

Epikur (1959b). "Pismo Menoikeu". V T. Lucretius Carus, *O naravi stvari,* Ljubljana: Slovenska matica, str. 523-526.

Feldman, F. (1991). *Confrontations with the Reaper.* New York: Oxford University Press.

Heidegger, M. (1997). *Bit in čas.* Ljubljana: Slovenska matica.

Horwich, P. (1987). *Asymmetries in Time.* Cambridge, Mass.: MIT Press.

Kripke, S. (2000). *Imenovanje in nujnost.* Ljubljana: Krtina.

Locke, J. (1924). *O človeškem razumu.* Ljubljana: Slovenska Šolska matica.

Long, A. A. in Sedley, D. N. (1997). *The Hellenistic philosophers,* zv. 1. Cambridge: Cambridge University Press.

Nagel, T. (1986). *The View from Nowhere.* Oxford: Oxford University Press.

Nagel, T. (2004). "Smrt". *Analiza,* 8, št. 1-2, str. 39-45.

Parfit, D. (1984). *Reasons and Persons.* Oxford: Oxford University Press.

Rist, J. M. (1972). *Epicurus: An Introduction.* Cambridge: Cambridge University Press.

Sharples, R. W. (2000). *Stoiki, epikurejci in skeptiki.* Šentilj: Aristej.

Silverstein, H. (1980). "The Evil of Death". *Journal of Philosophy, 77,* št. 7 (julij 1980), str. 401-424.

Unamuno, M. de (1983). *Tragični občutek življenja.* Ljubljana: Slovenska matica.

Williams, B. (1973). “The Makropulos Case: Reflections on the Tedium of Immortality”. V B. Williams, *Problems of the Self,* Cambridge: Cambridge University Press, str. 82-100.

Duh in telo

Tyler Burge[†]
prevedel Marko Cerkvenik

VZROČNOST DUHA IN TELESA TER POJASNJEVALNA PRAKSA[*]

V zadnjem času se številni filozofi sprašujejo, ali lahko upravičeno verjamemo, da so mentalne lastnosti vzročno učinkovite. Ukvarjajo se s tem, ali so intencionalne oziroma kvalitativne lastnosti mentalnih dogodkov epifenomenalne, tj., da nimajo vzročne moči in so nepomembne za vzročne procese. Domnevajo sicer, da mentalni dogodki so vzroki, kar pa je možno zgolj, če sprejmemo predhodno predpostavko, da so mentalni dogodki fizični dogodki. Na posamezne mentalne dogodke gledajo kot na uprimeritve ali primerke dogodkov fizične vrste. Od tod pa izhaja njihov dvom v to, da imajo mentalne lastnosti teh stanj in dogodkov v vzročnih procesih kakršnokoli vlogo.

Mislim, da se lahko tega dvoma znebimo, ne da bi pri tem morali zapustiti materialistični metafizični okvir, znotraj katerega se je pojavil. Sicer bom povedal nekaj o tem, zakaj mislim, da so nekateri izmed vidnejših argumentov, ki vodijo v epifenomenalizem, napačni, drugače pa me tukaj bolj zanima, od kod omenjeni dvom izvira. Mislim, da je posledica napačnega izbora filozofskih prioritet. Materialistični metafiziki je namenjene preveč pozornosti, razmišljanju o pojasnjevalni praksi pa premalo. Metafizična osnova, iz katere ta dvom izhaja, je veliko manj prepričljiva kot pa vsakdanja praksa, s pomočjo katere se ga lahko znebimo.

Najprej bom opisal skrbi zaradi grozečega epifenomenalizma in izpostavil nekatere šibke točke v argumentih, ki ga podpirajo. Potem bom pojasnil, zakaj menim, da je že sama osnova, iz katere te skrbi izvirajo, sporna. Na koncu bom naredil nekaj splošnih zaključkov v zvezi z (vsaj po mojem) napačnimi prioritetami, iz katerih omenjene skrbi izhajajo.

1

Slika, ki v nas vzbudi strah pred epifenomenalizmom, temelji na plavzibilni ideji, tj., da določena neintencionalno opisana stanja in dogodki, ki "se nahajajo pod" mentalnimi stanji in dogodki, sodelujejo v vzročnih procesih, ki so uprimeritve fizičnih zakonov, slednji pa ne omenjajo mentalnosti. V to, kaj tukaj pomeni "nahajati se pod", se ne bom spuščal. Predpostavljal bom, da se mentalna stanja in dogodki ne bi pojavili, če se ne bi pojavila neka "spodaj ležeča" fizična stanja

[†] Nedu Blocku se zahvaljujem za vrsto predlogov.

[*] Prevedeno po "Mind-Body Causation and Explanatory Practice", v Heil, J. in Mele, A. (ur.). *Mental Causation*, Oxford: Oxford University Press, 1995, str. 97-120. – op. prev.

in dogodki. V verigah, sestavljenih iz fizičnih dogodkov, ni vrzeli. Tako npr. obstajajo spodaj ležeči, tesno povezani živčni procesi, ki so uprimeritve zakonov nevropsihologije, in mentalni dogodki se ne bi pojavili, če se ne bi pojavili nekateri izmed teh procesov. Glede tega nimam posebnih pomislekov.

To, da se ta fizična stanja in dogodki "nahajajo pod" mentalnimi stanji in dogodki, se običajno razume v bolj specifičnem smislu. Trdi se, da so primerki prvih identični s primerki drugih. Sam te materialistične trditve ne sprejemam, čeprav je na splošno zelo priljubljena, in to predvsem zaradi svojih domnevnih prednosti pri pojasnjevanju mentalne vzročnosti. Že od Descartesa se zdi, da obstaja nekakšna skrivnost glede tega, kako lahko mentalni dogodki vplivajo na fizične in obratno. Omenjeno materialistično stališče, pa tudi druga, poskušajo to uganko rešiti s predpostavljanjem zgolj ene vrste vzročnosti, tj. vzročnosti med fizičnimi dogodki. V tem poglavju bom to ontologijo sicer sprejel, vendar zgolj zavoljo argumenta.

Strah pred epifenomenalizmom izvira iz domneve, da so nekatere lastnosti dogodkov (ali stanj) relevantne za njihova vzročna razmerja, druge pa ne. Lastnost biti tretja velika eksplozija v dani državi v zadnjih desetih letih res ni relevantna za višino škode, ki jo je ta eksplozija povzročila, medtem ko vročina, ki se je ob njej sprostila, gotovo je. Tako se pojavlja vprašanje, ali so mentalne lastnosti vzročno učinkovite oziroma ali so mentalni opisi lastnosti relevantni za razumevanje vzročnosti. Mnogo filozofov bolj zaupa temu, da je vzročnost odvisna od "spodaj ležečih" živčnih lastnosti, kot pa temu, da so zanjo odgovorne intencionalne lastnosti mentalnih dogodkov. Če se ti živčni procesi dogajajo, je gibanje telesa in zato vse, kar mi imenujemo obnašanje, odvisno od lastnosti teh procesov. Intencionalne oziroma fenomenalne lastnosti mentalnih dogodkov pri povzročanju ne bi imele nobene vloge, bile bi epifenomenalne; njihova vzročna moč bi bila v najboljšem primeru izpeljiva iz dejanskih "spodaj ležečih" vzročnih procesov.

Običajna je primerjava mentalnih lastnosti z lastnostmi, kot so fenotipi v biologiji. Obstajajo predpisana razmerja med fenotipi staršev in fenotipi njihovih neposrednih potomcev, ki so ob izpolnjenih dodatnih pogojih lahko celo nomološka. Ampak fenotipi staršev vzročno ne vplivajo na nastajanje fenotipov njihovih potomcev. Vzroke za nastanek fenotipov potomcev gre iskati izključno v genotipih njihovih staršev. Nekatere filozofe močno skrbi, da so intencionalna stanja podobno kot fenotipi del nomološko opisljivega sistema, vendar sama niso vzročno učinkovita.

Kaj je motivacija za to? V grobem je to ideja, da so vsi "spodaj ležeči" procesi tako ali tako prisotni in da mentalni dogodki v resnici svojo vzročno moč dobijo od fizičnih lastnosti, ki so pravi akterji vzročnosti. To prepričanje vodi do sklepa, da so mentalne lastnosti odveč. Četudi mentalne lastnosti ne bi bile ločene od fizičnih (ker nad njimi nujno supervenirajo), bi bile v nekem smislu brez pomena, saj se primarni mehanizmi vzročnosti nahajajo v "spodaj ležečih" fizičnih

lastnostih.

Poskusimo to prikazati z argumentom. Predpostavili bi lahko naslednje: (A) da so primerki mentalnih dogodkov identični s primerki fizičnih dogodkov; (B) da vzročno moč fizičnega dogodka določajo zgolj njegove fizične lastnosti in (C) da mentalne lastnosti niso reducibilne na fizične lastnosti. Zdi se, da iz tega sledi sklep, da mentalne lastnosti mentalnega dogodka nimajo pri določanju njegove vzročne moči nobene vloge.

Tukaj je še en argument, ki podpira podoben sklep. Predpostavimo, (a) da je svet fizičnih dogodkov in lastnosti popoln in zaprt sistem v smislu, da lahko fizične dogodke povzročajo samo fizične lastnosti drugih fizičnih dogodkov in (*b*) da mentalne lastnosti niso reducibilne na fizične lastnosti. Spet se zdi, da iz tega sledi, da ne more biti noben fizični dogodek povzročen zaradi mentalnih lastnosti, pa čeprav so le-te lastnosti fizičnih dogodkov. Ta argument sicer dopušča možnost, da mentalne lastnosti mentalnih dogodkov povzročajo mentalne lastnosti drugih mentalnih dogodkov, vendar gre v tem primeru, ker mentalna vzročnost ne vpliva na pojav fizičnih dogodkov, za posebno omejeno vrsto vzročnosti.[1]

Oglejmo si najprej prvi argument. Njegove glavne fraze, kar je ena izmed značilnosti argumentov za epifenomenalizem, niso dovolj jasne. "Je določen z" je ena izmed njih. To bi lahko pomenilo "supervenira nad", "je razložen z" ali "je individualiziran z". Videli bomo, da je trditev (B) problematična ne glede na to, katerega izmed omenjenih pomenov uporabimo, tako da si lahko ugotavljanje, kateri je pravi, prihranimo.

Prav tako ni jasno, kaj vse naj bi pojem fizične lastnosti vseboval. Še posebej ni jasno, ali spadajo med fizične "lastnosti" tudi različne relacije do okolja. Če ne, potem je premiso (B) težko braniti. Ker se zdi, da je bil namen avtorjev argumenta upoštevati tudi relacije do okolja, bom pojem fizične lastnosti uporabljal v širokem smislu, tj., kot da vključuje tudi relacije do okolja, ki so opisane v nepsiholoških, neintencionalnih pojmih.

Ključna nejasnost tiči v pojmu vzročne moči.[2] Vzročno moč dogodka neke vrste moramo razumeti v pojmih vzorcev vzročnosti, v katerih dogodki te vrste nastopajo. Takšne vzorce potem v vzročnih pojasnitvah obravnavamo kot pojasnjevalne. Lastnosti, ki "določajo" vzročno moč dogodka, pa so zgolj tiste, ki nastopajo v vzročnih razlagah. Druga premisa je plavzibilna le, dokler vzročna moč fizičnega dogodka izhaja ali iz vzorcev lastnosti, ki so opisani v pojmih naravoslovnih znanostih, ali pa iz drugih zdravorazumskih razlag, v katerih nastopajo fizični pojmi. Idejo, da fizične lastnosti določajo vzročno moč fizičnega dogodka, moramo razumeti v smislu, da so fizične lastnosti znotraj

[1] Za razpravo o takšnih argumentih, s katerimi se do neke mere strinjam, glej Robert Van Gulick, "Who*s in Charge Herc? And Who's Doing Ali the Work?" (ta knjiga, 13. poglavje). Moja kritika omenjenih argumentov se od njegove sicer razlikuje, vendar sta v splošnem združljivi.

[2] Glej moj članek iz 1989a za podroben opis razumevanja vzročne moči.

vzorcev vzročnosti, ki so opisani v naravoslovnih znanostih in zdravorazumskem fizikalističnem diskurzu, dovolj za obstoj in razumevanje vzročne moči fizičnih dogodkov – v takšnih vzorcih ni prostora za nobene druge lastnosti. Niti vzročne verige niti pojasnjevalni vzorci ne potrebujejo dopolnitve s področja, ki je zunaj fizikalističnega okvira, in kjer ne najdemo fizičnih lastnosti. To je mamljivo in plavzibilno tolmačenje druge premise, iz katere pa sklep prvega argumenta ne sledi.

Če fizični dogodki vsebujejo mentalne lastnosti, potem stališče, da samo fizične lastnosti (lastnosti opisane v naravoslovnih znanostih ali v običajnem fizikalističnem diskurzu) določajo vso vzročno moč fizičnega dogodka (ali da je vsa vzročna moč povezana s fizikalističnimi razlagami fizičnega dogodka), ni prepričljivo, razen če lahko pokažemo, da je psihološka pojasnitev ali nevzročna ali pa se ne nanaša na vzorce, v katerih ne nastopajo vzročne lastnosti. Vzročna moč fizičnega dogodka, ki je mentalen, bi namreč lahko vključevala učinek, ki je opisan v psihološki pojasnitvi. Nihče še ni pokazal, da bi bila psihološka pojasnitev ali nevzročna ali neopisna. Kaj takega tudi ne bi bilo plavzibilno.

Običajno obravnavamo vzročno moč neke entitete glede na pojme, v katerih je opisana vrsta, ki ji pripada. Na primer, če se sprašujemo o vzročni moči srca, pričakujemo, da bodo kot vzroki navedeni fiziološki vzorci lastnosti. Ne pričakujemo navajanja moči, ki bi jih preučevala fizika. Črpanje krvi navadno štejemo za relevantno, medtem ko zmečkanje hrošča, če pademo z lestve, ne. Če bi srce opisovali kot fizični objekt, ki ima takšen in takšen fizični obseg, bi se nam slednja lastnost mogoče zdela relevantna. Omenjeno tolmačenje druge premise (B) prvega argumenta je plavzibilno, vendar nezadostno, saj so mentalni dogodki opisani kot fizični. Glede na to se sicer res zdi, da mentalne lastnosti nimajo nobenega vzročnega vpliva, toda če vzročno moč opišemo v mentalnih pojmih (“kaj je vzročna moč misli, da dežuje?”), je ideja, da so samo lastnosti, opisane v naravoslovnih znanostih, relevantne za določanje vzročne moči, kratko malo nenavadna. Zato druge premise tega argumenta ne moremo preprosto vzeti kot splošen, samoumeven metafizični princip.[3] Če jo tolmačimo tako, da ostane argument veljaven, potem je ali napačna ali pa se izogne vprašanju.

Oglejmo si sedaj še drugi argument. Tukaj je problematična prva premisa, tj. (a). Trditev, da lahko fizične dogodke povzročajo samo fizične lastnosti drugih

[3] Včasih se druga premisa glasi “vzročna moč fizičnega dogodka je v *celoti* določena z njegovimi fizičnimi lastnostmi”. Tudi to premiso doletijo iste težave kot tisto, ki jo navajam sam. Je pa šibkejša od navedene, saj dopušča mentalnim lastnostim – čeprav so fizične lastnosti tiste, ki v celoti določajo vzročno moč fizičnega dogodka -, da imajo pri določanju vzročne moči fizičnega dogodka prav tako neko vlogo. Za dokončanje argumenta potrebujemo premiso, ki izključi “prekomerno določitev”. Vprašanja o “prekomerni določitvi” so precej podobna tistim, ki jih obravnavam sam. Kljub vsemu pa moram dodati, da se mi zdi ta pojem zelo zavajajoč, saj izenačuje vzročno interakcijo med različnimi ravnmi ali vrstami z vzročno interakcijo v povsem običajnih primerih, ko se (recimo) hkrati pojavi več fizičnih vzrokov istega učinka, bil bi pa že vsak sam zadosti za njegov pojav. Vendar so med tema dvema primeroma velike razlike. Dodam lahko, da mora imeti materialist v vseh teh argumentih neko razlago za to, da obravnava mentalne lastnosti kot lastnosti z materialnim značajem. Njihovega obstoja ne smemo obravnavati kot nezdružljivega z materializmom. Menim, da to ni lahka naloga, kljub temu pa so na voljo različne rešitve, za katere se zdi, da uspešno zadovoljijo potrebe večine materialistov. In ker v tem poglavju zavoljo argumenta materializem sprejemam, te zadeve ne bom naprej raziskoval.

fizičnih dogodkov, naleti na podobne težave kot druga premisa prvega argumenta. Obstoj zaprtega sistema odraža vzorec vzročnih razmerij in vzročnih pojasnitev, ki ne rabi nobene dopolnitve od zunaj. Ta sistem ne vsebuje nobenih vrzeli. Vendar iz tega ne sledi sklep, da takšen sistem izključuje ali razveljavlja vzročne relacije ali vzročne pojasnitve, opisane v lastnostih zunaj njega. Če bi takšen sklep, kot je bilo že večkrat rečeno, res sledil, potem ne bi bilo prostora za vzročno učinkovitost niti v posebnih znanostih niti v naravoslovnih, kot sta kemija in psihologija. Kajti v vzročnih relacijah, pojasnjenih v pojmih lastnosti fizike, ne obstajajo nobene vrzeli (razen morda kvantne). Toda le redkim se zdi sprejemljiva ideja, da fizični dogodki ne morejo biti povzročeni s fiziološkimi lastnosti fizičnih dogodkov.

Je vzročna učinkovitost lastnosti, ki nastopajo v kemiji in fiziologiji, odvisna od njihove reducibilosti na lastnosti, ki nastopajo v fiziki? Tudi to se zdi podobno nesprejemljivo. Ali so lastnosti teh posebnih znanosti reducibilne na lastnosti, ki se omenjajo v fiziki, je namreč široko odprto empirično vprašanje. V bistvu se zdi splošno redukcijo skoraj nemogoče izvesti. Vprašanje o vzročni relevantnosti lastnosti teh posebnih znanosti se zdi neodvisno od vprašanja o reducibilosti.

Slabosti omenjenih argumentov so znane. Kljub temu pa občutek nelagodnosti ostaja. Mnoge moti ideja o tem, da mentalne lastnosti supervenirajo nad fizičnimi in tako pri povzročanju dogodkov nimajo nobene vloge. Zato so se pojavili številni poskusi avtorjev, da bi ugotovili, katera vrsta supervenientne relacije bi mentalnim lastnostim “dovoljevala”, da bi bile vzročno učinkovite, čeprav bi supervenirale nad fizičnimi (Kim, 1984, Sosa, 1984).

Ti poskusi znajo biti koristni. Vendar sam mislim, da je skrb zaradi epifenomenalizma odveč. Mnogo bolj verjetno je namreč, da je napačen epifenomenalizem, kot pa, da so resnične različne domneve (vključno z materialistično), ki naj bi vodile do njega. Prav tako je mnogo bolj verjetno, da je napačen epifenomenalizem, kot pa, da so resnični različni poskusi, da bi ga prikazali kot napačnega ali kot nekaj, čemur se lahko izognemo s sklicevanjem na protidejstvenike, določila zakonov ali supervenienco. Epifenomenalizem se pogosto omenja kot resna metafizična možnost, vendar ga lahko v najboljšem primeru jemljemo zgolj kot vir zahtev po tem, da bi razjasnili naše vsakodnevne pojme. Izgleda bolj kot ena izmed manj verjetnih oblik skepticizma, ki jo lahko uporabimo kot orodje za filozofsko razjasnitev in ki ima malo resnične prepričevalne moči. Mislim, da bi bil drugačen, manj metafizičen pristop k temu problemu stvarnejši in plodnješi.

Ironično je, da so se številni materialisti, medtem ko so poskušali razjasniti mentalno vzročnost, dokopali do prepričanja, da je to, ali so mentalni opisi lahko vzročno relevantni, resno metafizično vprašanje. Mislim, da je to enako, kot če bi priznali, da materializmu ni uspelo pojasniti mentalne vzročnosti. Prepričanje, da materializem lahko pojasni vzročnost, pa je bilo vedno njegova glavna privlačnost. Čeprav materializem ne rabi sprejeti epifenomenalizma, pa samo

dejstvo, da so ga jemali kot resno metafizično možnost, nakazuje, da se je pri iskanju pojasnitve mentalne vzročnosti zgodila napaka.

Mentalne vzročnosti (vzročnosti, ki vključuje mentalne ali intencionalne lastnosti) in vzročne moči mentalnega ne moremo pojasniti tako, da se osredotočimo na lastnosti, ki jih opisujejo naravoslovne znanosti. Naše razumevanje mentalne vzročnosti izhaja predvsem iz našega razumevanja psiholoških pojasnitev, in to neodvisno od naše vednosti – ali, bolje, ne glede na našo nevednost – o procesih, ki se dogajajo na spodaj ležeči ravni. Materialistične razlage ne znajo zapolniti vrzeli med metafiziko mentalne vzročnosti in tistim, kar dejansko vemo o naravi in obstoju mentalne vzročnosti in kar skoraj v celoti izhaja iz psiholoških razlag in opažanj.

2

Do sedaj še nisem podvomil v materialistično metafiziko, ki med drugim predstavlja osnovo za strah pred epifenomenalizmom. V tem poglavju želim predstaviti razloge za dvom v najbolj splošno obliko materializma in v njegovo osrednjo vlogo pri razumevanju mentalne vzročnosti.

Vsekakor je razumno verjeti, da se pod našimi mentalnimi stanji in procesi nahajajo fizični, kemični, biološki in živčni procesi, ki delujejo skladno z zakoni, ki tukaj veljajo. Nekateri izmed njih so verjetno nujni za to, da bi bili intencionalni (ali fenomenalni) mentalni dogodki vzroki za obnašanje. Zdi se, da so nujni celo za sam obstoj mentalnih dogodkov. Kakorkoli, sam mislim, da je narava razmerja med mentalnimi dogodki in temi fizičnimi procesi še vedno povsem nejasna. Najbolj priljubljeno stališče v zvezi s tem razmerjem pa je materialistična teorija identitete primerkov.

Pred leti sem razvil argument proti njeni pomembnejši različici (Burge, 1979, str. 109-113). Tista, ki sem jo imel v mislih, pravi, da sta vsaka uprimeritev mentalnega stanja in primerek mentalnega dogodka identična z uprimeritvijo fizičnega stanja ali primerkom dogodka, ki uprimerja naravno fizično vrsto, opisano v pojmih naravnih znanosti, kot jih poznamo danes. Različni primerki mentalnega dogodka iste vrste so lahko primerki različnih dogodkov fizičnih naravnih vrst.

Zahteva, da je primerek fizičnega dogodka instanca fizične vrste, ki je opisljiva v pojmih naravoslovnih znanosti (fizike, kemije, biologije, nevropsihologije itd.), je pomembna zato, ker zagotavlja materialistom identifikacijo mentalnega s fizičnim, poleg tega pa je nesporno fizična in ima določeno vlogo pri pojasnjevanju fizične vzročnosti. Mislim, da kljub vsemu obstajajo materialistične teorije, ki od nas zahtevajo manj od omenjenega. Moj argument ne bo pokazal, da so le-te napačne. Nekaterim izmed njih bom malo več pozornosti namenil kasneje.

Argument proti teoriji identitete primerkov, katere element je omenjena zahteva,

delno temelji na miselnih eksperimentih o Zemlji dvojčici, za katere predpostavljam, da so znani.[4] Skladno z njimi je mogoče, da je telo neke osebe, če ga obravnavamo izoliranega od okolja, fiziološko in molekularno isto, pa čeprav ima ta oseba misli, ki imajo različno intencionalno vsebino. Razlika v vsebini je odvisna od razlik v zgodovini relacij med posameznikom in njegovim okoljem. Mogoče je npr., da ima oseba, ki je določen odnos do aluminija ali artritisa v svoji zgodovini ponotranjila (npr. s pomočjo videnja ali pogovora) zgolj na osnovi tega, da gre za aluminij ali artritis, prepričanje, da je aluminij lahka kovina ali da je artritis huda bolezen, pa čeprav ni zmožna ločiti aluminija oziroma artritisa od drugih dejanskih ali možnih kovin oziroma bolezni. Možno pa je, da obstajajo protidejstvena okolja, v katerih ima pri ustvarjanju in nastajanju misli enako vlogo, kot jo imata aluminij in artritis, ena izmed teh drugih "podobnih" kovin ali bolezni. V takšnih protidejstvenih okoljih bi lahko imela oseba za izpolnitev svojih namenov in ciljev telo, ki je kemična in fiziološka kopija dejanskega telesa neke osebe. Vendar bi kopija kljub vsemu imela misli z drugačno vsebino. Kopija ne bi mislila, da je aluminij lahka kovina ali da je artritis huda bolezen.

Te miselne eksperimente bom jemal kot veljavne, na njih pa temelji (iz njih ni izpeljana) tudi prva premisa mojega argumenta zoper teorijo o identiteti primerkov:

(1) Možno je, da ima oseba misli z različno vsebino, pa čeprav so vsi primerki dogodkov, ki se pojavijo v posameznikovem telesu in so verjetni kandidati za identifikacijo z mentalnimi dogodki ter opisljivi v pojmih naravoslovnih znanosti, kot so fizika, kemija in nevropsihologija, enaki.

Druga premisa se na omenjene miselne eksperimente navezuje manj:

(2) Noben pojav misli ne more imeti drugačne intencionalne vsebine in biti primerek istega dogodka ali biti isti dogodek.

Vzemimo sedaj katerikoli primerek fizičnega dogodka *b* v posameznikovem telesu, ki je primeren kandidat za identifikacijo s pojavom posameznikove misli (primerek mentalnega dogodka) *a,* tj., da je aluminij lahka kovina (ali da je artritis huda bolezen). Po (1) so možne situacije, v katerih se pojavi isti primerek *b* in edino misli (primerki mentalnih dogodkov) z različno intencionalno vsebino. Po (2) noben izmed teh pojavov misli ni primerek istega dogodka kot *a.* Ker se torej *b* lahko pojavi, ne da bi se pojavil *a, b* ne more biti identičen z *a.*

Ta argument je kritiziral Donald Davidson. Sprejema drugo premiso, njeno zanikanje pa imenuje "ne le neverjetno, ampak absurdno. Če imata dva mentalna dogodka različno vsebino, potem sta gotovo dva različna dogodka." S prvo premiso pa se odkrito ne spopade. Izgleda sicer, kot da jo zavrača. Meni, da ustrezni miselni eksperimenti kažejo, da "se ljudje, ki so si podobni v vseh

[4] Burge 1979, 1982, 1986a, 1986b, 1989b. Miselni eksperimenti temeljijo na metodologiji, ki jo je razvil Putnam, 1975.

relevantnih pogledih /.../, lahko razlikujejo v tem, kar menijo ali mislijo /.../. Seveda pa obstajajo med njimi *neke* razlike tudi v fizičnih pogledih; njihove vzročne zgodovine so si različne."[5]

Ta odgovor se mi ne zdi prepričljiv. V relevantnih miselnih eksperimentih gotovo obstajajo fizične razlike med dejanskimi in protidejstvenimi situacijami. Vprašanje je, ali so vedno fizično različne entitete tiste, ki so plavzibilne kandidatke za identifikacijo z različnimi mentalnimi dogodki ali uprimeritvami mentalnih stanj. Ene izmed njih – različne fizične vzročne zgodovine – to gotovo niso. Le-te nimajo enakih vzrokov in učinkov kot relevantni mentalni dogodki (stanja). Poleg tega pa je vprašljivo, ali relevantno opisane vzročne zgodovine sploh uprimerjajo kakšno pojasnjevalno naravno vrsto znotraj naravoslovnih znanosti.

Glede na to, da obstajajo v protidejstveni situaciji razlike vsaj v daljni vzročni zgodovini vsakega relevantnega fizičnega dogodka, ki se pojavi v posameznikovem telesu, bi lahko pomislili, da je vsak tak dogodek primerek dogodka, ki se razlikuje od kateregakoli drugega primerka dogodka v dejanski situaciji. Skladno s takšnim stališčem se dogodki razlikujejo le na podlagi tega, da imajo različne vzročne zgodovine.[6] Toda to izgleda izredno neplavzibilno stališče o identiteti dogodkov. V protidejstveni situaciji se mi zdi možno zamisliti isti dogodek v posameznikovem telesu, ne da bi pri tem moral sprejeti stališče, da za vsak dogodek v njegovi vzročni zgodovini, pa če je še tako oddaljen, velja identiteta primerkov (ali vrst). Protidejstveno razmišljanje o dogodkih temelji v veliki meri na naši fleksibilnosti. Pogosto govorimo o določenem dogodku v protidejstvenih situacijah, v katerih pa ne predpostavljamo, da je vsak predhodni dogodek v zgodovini sveta, ki je s tem dogodkom vzročno povezan, isti.[7]

Predpostavimo, da sem dobil svoje predstave o aluminiju s pomočjo prebiranja knjig. V protidejstveni situaciji vidim zapis, ki izgleda enako. Bistvena razlika leži na koncu dolge vzročne verige: protidejstvena veriga gre nazaj do neke druge

[5] Citati so iz Davidson, 1986, str. 452. Davidson je v zasebnem razgovoru potrdil, da zavrača premiso (1).

[6] To bi lahko bila posledica Davidsonovega kriterija za individualizacijo dogodkov; glej tudi Davidson, 1969. Ker pa kriterij ni oblikovan modalno, ni jasno, ali ga namerava Davidson uporabiti v vseh protidejstvenih situacijah. Res pa je, da v razpravi o njem upošteva vsaj eno protidejstveno situacijo.

[7] Davidson (1986, str. 451-453) priznava, da lahko dogodki v posameznikovem telesu pripadajo isti vrsti tako v relevantnih dejanskih kot tudi v protidejstvenih situacijah. Ne navaja pa razlogov za to, da za dogodke v posameznikovem telesu v dejanskih in protidejstvenih situacijah ne bi mogla veljati identiteta primerkov, tj., da bi bili isti dogodki. Na str. 453 je njegov nepojasnjen namig, da naj bi v mojem argumentu določeno vlogo imele 'esencialistične' predpostavke. Ne moti me, če mi pripisujejo neke vrste esencializem. Toda moj argument na njem ne temelji. Vse modalne zahteve (vključno s tisto v drugi premisi) je mogoče tolmačiti s tem, da so podprte z najboljšimi razpoložljivimi metodami za individualizacijo. V prvi premisi trditve o esenci ali nujnosti sploh nisem postavil, postavil sem le eno o možnosti. Če hoče Davidson tej trditvi ugovarjati, se mora sklicevati na določeno nemožnost. Zanikati bi moral možnost, da bi lahko bil nek dogodek, če bi bila njegova vzročna zgodovina glede načinov, na katere tvorimo različne misli, drugačna, isti primerek dogodka. Zdi se mi, da bi takšno zanikanje namigovalo na skrajno obliko esencializma o posameznih dogodkih. Namiguje na to, da je celotna vzročna zgodovina nekega dogodka vedno del njegovega bistva. Zanikanje bi lahko ponovno tolmačili v pojmih metod za individualizacijo (glej zgoraj), vendar se mi zdi, da naše dejanske pojasnjevalne prakse tega ne podpirajo. Davidson v bistvu misli, da so dogodki isti, če in samo če imajo iste vzroke in učinke (glej op. 9). Toda na tej točki razprave bi bilo dovolj, če bi trdil, da dogodek ne more biti isti. v kolikor ima drugačno vzročno zgodovino. Kolikor jaz vem, te modalne teze ni branil. Mislim pa, da ni ubranljiva.

kovine, medtem ko gre dejanska veriga nazaj do aluminija. Stališče, ki ga obravnavamo, pravi, da mora biti vsak dogodek v omenjenih verigah drugačen, biti mora drugačna posameznost. Tako fiziološko opisani dogodki v mojih možganih, ki jih povzroči branje zapisa, v tem primeru ne morejo biti identični. Sporno pri tem stališču je, da individualizira možganske dogodke glede na stvari, ki za fiziologijo možganov niso relevantne. Ne spomnim se ničesar iz naše pojasnjevalne prakse, kar bi podpiralo takšno metafizično pozicijo.

Zdi se mi, da je v številnih protidejstvenih situacijah za vse fizične dogodke (ali primerke dogodkov), ki so kandidati za identifikacijo z intencionalnimi mentalnimi dogodki, mogoče, da ostanejo isti, medtem ko se relevantni mentalni dogodki, misli, razlikujejo. Ideja, da vedno obstajajo fizični dogodki, ki so opisljivi v pojmih naravoslovnih znanostih in ki jih lahko identificiramo z mentalnimi, njihova identiteta pa se bo v vseh možnih protidejstvenih situacijah spreminjala natanko takrat, ko se bodo spreminjali mentalni dogodki, se mi ne zdi prav nič verjetna. Vsekakor pa bi morali ugovori Davidsonove vrste v luči miselnih eksperimentov pokazati, katere razlike v fizičnih dogodkih bi lahko bile identične z različnimi primerki mentalnih dogodkov.

Davidsonov ugovor mojemu argumentu pa vsebuje še nekaj drugega. Trdi, da upoštevanje zunanjih dejavnikov, ki sooblikujejo naše vsakdanje načine identificiranja mentalnih stanj, ne škodi teoriji identitete primerkov. Kot primer stanja, katerega identifikacija je odvisna od okolja, navaja sončne opekline. Davidson pravilno misli, da je primerek sončne opekline identičen s primerkom stanja kože. Meni tudi, da so mentalna stanja in dogodki v splošnem podobni sončnim opeklinam.

Mislim, da primer s sončnimi opeklinami za moj argument ni relevanten. Le-ta ne trdi, da nobeno neindividualistično opisano stanje ali dogodek ne more biti identično s stanji ali dogodki v posameznikovem telesu. Trdi le, da nekatera mentalna stanja in dogodki ne morejo biti.[8] Strinjam se, da je primerek stanja biti opečen identičen s primerkom fiziološko opisljivega stanja kože. Vendar stanje biti opečen ni podobno stanju imeti prepričanje ali dogodku mišljenja. Prvič, stanja opeklin kože znamo identificirati na načine, ki so sistematični in pojasnjevalni ter popolnoma neodvisni od dejstva, da so sončne opekline. To je del tistega, kar naredi identifikacijo sončnih opeklin s fiziološkim stanjem kože verodostojno. Učinek, ki ga ima sonce na kožo, lahko vidimo ter podamo kemično ali fiziološko razlago zanj. Pomembno je, da pri tej razlagi ne rabimo

[8] Davidson (1986, str. 452) pravi, da mogoče delam napako, ko mislim, da so miselni eksperimenti sami po sebi nezdružljivi s teorijo identitete primerkov. Nadaljevanje njegove razprave namiguje na to, da res delam to napako. Tega nisem nikoli mislil, niti nisem nikoli napisal nič takega, iz česar bi to sledilo. V kasnejšem članku (Davidson, 1989) pravi, da argument za moje zanikanje teorije identitete primerkov "predpostavja, da če je neko stanje ali dogodek identificirano (mogoče je to nujno, če gre za mentalno stanje ali dogodek) s pomočjo sklicevanja na stvari zunaj telesa, mora biti tudi to stanje ali dogodek zunaj telesa oziroma vsaj ne sme biti identično z nobenim dogodkom v telesu". Resnično mislim, da določeni mentalni dogodki, ki so nujno individualizirani s pomočjo sklicevanja na stvari zunaj telesa, niso identični z nobenim dogodkom v telesu, ki je opisljiv v pojmih vrst naravoslovnih znanosti. Toda tega stališča nisem nikoli jemal kot predpostavko. Dokazoval sem ga. Davidson se tako ne dotakne bistva najinega nesoglasja. Kar sledi, bi moralo zadevo razjasniti.

nobenih predpostavk o tem, ali je ta učinek morda povzročilo kakšno nebesno telo. Pri identifikaciji stanj telesa, ki so (domnevno) prepričanja, pa takšnih načinov identificiranja, ki bi bili neodvisni od predpostavk o teh prepričanjih, nimamo. Menim, da ta razlika odraža našo nevednost o razmerju med prepričanji in telesnimi stanji. To ni nezdružljivo s teorijo identitete, vendar postavlja pod vprašaj enostavno analogijo s primeri, kjer so identitete očitne.

Drugič, moj argument za sončne opekline ne bi deloval na isti način, kot deluje za misli. Zdi se, da je prva premisa resnična tudi za sončne opekline: isto stanje (primerek stanja) iste kože, opisljivo v fizioloških pojmih, bi lahko bilo (metafizično) povzročeno tudi na kakšen drug način, ne le s sončnimi žarki. To se zdi resnično, saj obstajajo nesporni in očitni načini identifikacije sončnih opeklin s fiziološko opisljivimi stanji (primerki stanj) kože. Ti opisi imajo veliko pojasnjevalno moč in se zdijo individualno relevantni. Kljub temu pa njihove predpostavke o stanjih kože ne izhajajo iz znanja o sončnih žarkih.[9]

Kakorkoli, zdi se, da druga premisa (2) za sončne opekline ne drži. Mislim, da ni razlogov za naslednjo rditev: če fiziološko opisljivo stanje (primerek stanja) kože ne bi bilo sončna opeklina, ne bi bilo mogoče, da bi bilo isto stanje (primerek stanja) kot zdaj. To trditev podpira naša sposobnost identificirati stanja kože s sončnimi opeklinami in jih uvrstiti v sistematično, individualno relevantno pojasnjevalno shemo, ne da bi pri tem upoštevali karkoli v zvezi s soncem. Vrsta *sončna opeklina* pa, nasprotno, spada v mnogo manj sistematično, informativno in pomembno pojasnjevalno shemo (če sploh lahko rečemo, da spada v kakšno). Njena trditev o individualizaciji primerkov stanj preko protidejstvenih situacij je veliko šibkejša kot trditev fiziološke sheme. Podobno je tudi njena trditev o zagotavljanju temeljne oblike individualizacije (kar za intencionalne pojme pravi druga premisa) veliko šibkejša od trditve o mentalnih vrstah. Primer s sončnimi opeklinami zaradi neuspeha analogije med njimi in mišljenjem tako ne more spodkopati mojega argumenta.

Medtem ko Davidson dvomi v prvo premiso mojega argumenta, pa nekatere teorije identitete primerkov zanikajo drugo. Običajno se trdi, da lahko vrste intencionalnih stanj in dogodkov identificiramo z vrstami funkcionalnih stanj, ta pa so lahko realizirana s pomočjo različnih fizičnih stanj ali dogodkov. Primerki realizirajočih fizičnih stanj in dogodkov naj bi bili identični s primerki mentalnih stanj ali dogodkov. Iz tega ne sledi zanikanje (2). Toda ponavadi se domneva

[9] Po mojem mnenju tako primer s sončnimi opeklinami ni ravno združljiv z Davidsonovo kritiko mojega argumenta, saj želi Davidson zanikati premiso (1). Davidson najverjetneje zanika, da analogija iz (1) velja za sončne opekline. Glej njegov članek iz leta 1986, str. 451-455, kjer priznava le to, da bi lahko za stanje kože še vedno držala identiteta vrste. Ne navaja pa nobenih razlogov za prepričanje, da posamezne uprimeritve stanja ne bi moglo povzročiti nič drugega razen sonca. Morda meni, da so stanja podobna dogodkom v smislu, da so, tako kot to nujno velja za dogodke, njihove uprimeritve identične le, če imajo enake vzroke (glej op. 7). Vendar ne vidim nobene metafizične ali celo fiziološke ovire za prepričanje, da je bila dana uprimeritev stanja proizvedena s pomočjo svetlobnih žarkov s popolnoma enakimi fizičnimi lastnostmi in vektorji, kot jih imajo sončni žarki, ki pa niso izvirali od sonca. Ni mi jasno, ali Davidson sprejema analogijo s premiso (2). Izgleda pa, da bi verjetno sprejel trditev, da nobena uprimeritve stanja imeti sončne opekline ne bi mogla biti ista uprimeritev tega stanja, če bi jo povzročilo kaj drugega namesto sonca. Kot pravim v nadaljevanju, se mi to ne zdi verjetno.

tudi, vsaj implicitno, da je identiteta spodaj ležečih fizičnih stanj in dogodkov neodvisna od njihove točne funkcionalne vloge v danem sistemu. Če pa predpostavimo to, potem ta teorija zavrača (2).

Takšna teorija izgleda zelo neobetavno. V naši pojasnjevalni praksi je malo stvari, ki bi podpirale identifikacijo intencionalnih mentalnih vrst z neintencionalno opisanimi funkcionalnimi vrstami.[10] Identifikacija intencionalnih mentalnih primerkov s primerki, opisanimi v pojmih naravoslovnih znanosti, pa ima po mojem mnenju v pojasnjevalni praksi še manj podpore. Vendar je tisto, na kar se želim osredotočiti, nekaj drugega, tj., kaj so razlogi za sprejetje (2).

Sistem pripisovanja intencionalne vsebine je temeljno sredstvo za identifikacijo intencionalnih mentalnih stanj in dogodkov v psiholoških razlagah in pri samopripisovanju. Pravzaprav nimamo nobenega drugega sistematičnega načina za identifikacijo takšnih stanj in dogodkov. Davidson trdi, da je (2) intuitivno očitna in da so njene zavrnitve absurd. S tem se strinjam. Toda menim, da bi to lahko upravičili s sklicevanjem na to, da je to naš edini način individualiziranja intencionalnih mentalnih dogodkov, ki zagotavlja sistematično razumevanje, opis in pojasnitev mentalnih dogodkov in intencionalne aktivnosti. Takšne kognitivne prakse so naš najbolj zanesljiv način poznavanja narave tega, kar obstaja.

Ne poznam nobenega plavzibilnega ali celo resnega argumenta proti (2).[11] Napaka, da smo kljub vsemu prepričani, da takšen argument obstaja, pa je glavni razlog za našo zgrešeno skrb zaradi epifenomenalizma. Dokler mislimo, da so intencionalne vsebine le naključno povezane z dogodkom misli, smo v položaju, ko si lahko predstavljamo, da je na eni strani živčni dogodek verjetno identičen z mentalnim, na drugi pa, da ima ta mentalno/fizični dogodek različno vsebino. Ob predpostavki, da so vzročni zakoni, ki povezujejo živčne procese, zadostni za pojav njihovih učinkov, se lahko vprašamo, kako so lahko vzročno relevantne intencionalne lastnosti. To vprašanje je posledica stališča, ki sem ga omenil že v prvem odstavku tega članka in ki pravi, da je intencionalna vsebina mentalnih stanj in dogodkov zgolj njihova epifenomenalna značilnost.

Toda takšna vsebina je v resnici pojasnjevalno in identifikacijsko središče teh dogodkov. Le malo stvari je še, na katere se lahko v razpravi o vzročnih močeh in ontologiji mentalnega sklicujemo. Sistematične, informativne in pomembne pojasnjevalne sheme, kot je naša psihološka, so najboljše zagotovilo za ustrezne individualizirane opise, ki kažejo na to, kaj je bistveno za identiteto posameznikov, delov ali primerkov.[12] Zdi se mi, da je vsaka metafizična teorija,

[10] Za nedavno razpravo o razlogih za to glej Putnam, 1988.

[11] Slišal sem za predlog, da čeprav (2) ne drži, pa drži naslednje: ni mogoče, da bi lahko obstajala dva dogodka, ki bi bila mentalno enaka, imela pa bi različni vsebini. (2) je še vedno zanikana, ker bi lahko bil isti fizičen dogodek baje identičen z mentalnim dogodkom, ki ima v dejanski situaciji takšno vsebino, ali z mentalnim dogodkom, ki ima v protidejstveni situaciji drugačno vsebino. Ne gre za isto vsebino misli, ampak za isti miselni/živčni dogodek. Zdi se mi, da to ne pripomore mnogo k ubranitvi zanikanja (2). Izgleda, da trditve iz prejšnjega odstavka še vedno držijo.

[12] Spet bi lahko kdo pogovor o esenci preoblikoval v pogovor o najbolj temeljnih metodah za identifikacijo posameznikov in njihovo ponovno identifikacijo skladno s protidejstvenimi ugibanji; glej op. 8.

ki želi pojasniti mentalno vzročnost ali ontologijo mentalnih dogodkov, pri tem pa zanika (2), na napačni poti.

Kot sem že omenil, so sistematične, informativne in pomembne pojasnjevalne sheme dogodkov in stanj hkrati tudi naši najboljši kazalci vzročne relevantnosti. Ideja, da so vzročna razmerja, opisana v psihologiji, v resnici ali v bistvu vzročna zgolj, če so opisana v nepsiholoških pojmih, se zdi zelo vprašljiva. Ne vidim razloga za prepričanje, ki je sedaj skupno večini filozofov, da se "pravo" vzročno delo odvija na nižji ravni. Prav tako ne vidim razloga za mišljenje, da lahko razumemo mentalno vzročnost s pomočjo analize supervenience (čeprav mislim, da je to, da ugotovimo, na kakšen način mentalni dogodki supervenirajo nad fizičnimi, zelo pomembno). Naše razumevanje mentalne vzročnosti ne izhaja prvenstveno iz ponovnih opisov v fizičnih pojmih.[13] Prvenstveno izhaja iz našega razumevanja mentalnih pojasnitev. To razumevanje pa je v veliki meri neodvisno od sklicevanja na spodaj ležeče procese.

Zdi se mi, da je pomanjkljivost Davidsonovega zelo vzpodbudnega argumenta za teorijo identitete primerkov to, da podcenjuje osrednjo vlogo, ki jo ima kognitivna praksa pri razumevanju mentalne vzročnosti in ontologije (Davidson, 1970). Njegov argument se delno opira na premiso, da morajo biti vzročne relacije med dogodki podprte z zakoni (ali biti njihova posledica) popolnega in zaprtega sistema pojasnitev, kjer predikati teh zakonov veljajo za dogodke, ki so vzročno povezani. Davidson to premiso združi z domnevo, da obstajajo vzročne relacije med mentalnimi in fizičnimi dogodki, ter domnevo (ki jo tudi zagovarja), da psihološke pojasnitve same ne morejo biti del popolnega in zaprtega sistema pojasnitev. Iz teh treh premis sklepa, da morajo primerki mentalnih dogodkov pasti v okvir predikatov fizičnih zakonov, ki pa tvorijo del popolnega in zaprtega pojasnjevalnega sistema. Zdi se mi, da morajo to biti zakoni fizike.

Mislim, da ne vemo in ne moremo vedeti *a priori*, da je posledica vzročnih stavkov obstoj zakonov ali pojasnjevalnega sistema, ki vsebuje takšne posebne lastnosti. Toliko *a priori* o obliki zakonov narave ne moremo vedeti, pa naj bodo opisani v pojmih katerekoli znanosti. Prav tako ne moremo *a priori* poznati relacij med ontologijami in vzročnimi shemami različnih znanosti. To pomeni, da ne moremo vedeti *a priori*, da zakoni ali nomološke posplošitve psihologije opisujejo dogodke, ki imajo kakšno specifično zvezo z dogodki, ki padejo v okvir zakonov kakšne druge znanosti. Mislim, da ni razloga, še manj apriornega, za sprejetje Davidsonove prve premise. Zato ne moremo vedeti *a priori*, da padejo

[13] Ali še slabše v sintaktičnih pojmih. Priznamo lahko, da mnogo psiholoških stanj in dogodkov vključuje operacije, ki temeljijo na sintaktičnih entitetah. Te notranje mentalne sintaktične elemente lahko razumemo kot izražanje intencionalnih vsebin. Toda ideja, da so možgani občutljivi zgolj na sintaktične oblike, na vsebino pa ne, se mi zdi zavajajoča metafora. Možgani so občutljivi na živčne oblike. Predpostavljamo sicer lahko, da med sintaktičnimi entitetami obstajajo vzročna relacije, vendar s tem vzročnost že opisujemo na višji abstraktni ravni, kot je tista, kjer se nahaja možganska fiziologija. Če pa lahko vzročnost opišemo na tej ravni, potem jo lahko opišemo tudi na ravni intencionalne vsebine. Ne vidim razloga za prepričanje, da so sintaktične lastnosti glede vzročne relevance važnejše od intencionalnih lastnostih. Sintakso uporabljamo v pojasnitvah, ki so v osnovi intencionalne. Po mojem mnenju ni sintaksa glede ontologije ali vzročnosti v nobenem smislu v prednosti.

mentalni dogodki, ki so vzroki, v okvir kakšne druge (nepsihološke) vrste zakonov.

Najboljša pot do razumevanja mentalne vzročnosti in ontologije vodi preko razumevanja najbolj zanesljivih načinov pojasnjevanja in opisovanja mentalnih dogodkov. Naša dejanska kognitivna praksa ne govori v prid ideji, da padejo intencionalni mentalni dogodki v okvir fizičnih zakonov, ki ne dopuščajo izjem, ali da uprimerjajo fizične opise najbolj temeljne naravoslovne znanosti.

Vrnimo se zdaj k argumentu zoper teorijo identitete primerkov, ki sem jo opisal na začetku (in ki naj bi jo podpiral Davidsonov argument). Ker se premisa (1) in premisa (2) zdita zelo verjetni in varni pred ugovori, se mi argument zdi močan in zdrav. Izgleda, da najbolj splošni metafizični razlog za skrb glede epifenomenalizma stoji na šibkih temeljih.

Obstajajo pa tudi manj zavezujoče oblike materializma v zvezi z mentalnimi dogodki. Vse obravnavajo primerke mentalnih dogodkov kot primerke fizičnih dogodkov, ki niso uprimeritve nobene vrste, opisane v pojmih naravoslovnih znanostih. Morda pa to ni intrinzično neplavzibilna trditev. Posamezne vojne, plazovi, nevihte in obroki hrane mogoče ne padejo v okvir nobenega dogodka naravne vrste, ki je opisljiv v pojmih kakšne naravoslovne znanosti. Dvomim, da vedno obstaja en določen primerek dogodka, ki uprimerja vrste naravoslovne znanosti (mislim tudi na fiziko), s katerim ga lahko identificiramo. Kljub temu pa sta oba dogodka nesporno fizična. Morda so takšni tudi mentalni dogodki.

Po drugi strani pa se te analogije ne zdijo posebno obetavne. Relevantni opisi so nesporno fizični. Težava pri njihovi identifikaciji z dogodki, opisanimi v fiziki, je zgolj v tem, da so kompleksni in da padejo v okvir vzorcev, ki so značilni za makro opise, niso pa posebno uporabni za sistematične pojasnitve. V nasprotju s tem pa psihološka stanja padejo v okvir sistematičnih vzorcev, ki so uporabni za pojasnitve. Vendar ti vzorci, tako se zdi, ne vključujejo običajnih fizičnih lastnosti (npr. maso, energijo, zgradbo itd.), ki pa so uporabne za fizične pojasnitve in opise.[14] Zato se zdi ideja, da so ti vzorci fizični vzorci v običajnem smislu, vprašljiva.

Vsekakor pa takšne oblike materializma, dokler relevantni primerki fizični dogodkov, s katerimi naj bi identificirali primerke mentalnih dogodkov, ne bodo opisani v nepsiholoških pojmih, ne pripomorejo veliko k pojasnitvi mentalne vzročnosti. V naši kognitivni praksi je malo stvari, ki bi namigovale na to, da bomo v bližnji prihodnosti takšne opise uprimeritev osrednjih intencionalnih stanj ali primerkov dogodkov tudi dobili. Pa četudi bi bila takšna liberalizirana različica materializma resnična, bi k razumevanju mentalne vzročnosti pripomogla bore malo.

[14] Razmerje sestave je verjetno ključnega pomena. Obstajajo vrste materializma, ki trdijo, da so vsi predmeti razstavljivi na neorganske fizične delce. Če odmislimo dejstvo, da se te vrste ne izkažejo za najbolj primerne, ko govorimo o lastnostih in dogodkih (da ne govorimo o predmetih, kot so številke, intencionalne vsebine, metode), pa kljub temu predstavljajo osnovo za trditev, da je za naše razumevanje mentalnih entitet pomembna njihova fizična sestava, trditev, za katero se mi (za zdaj) zdi, da je ne podpira nič, kar vemo.

3

Težava glede mentalne vzročnosti, ki navadno služi kot motivacija za sprejetje materializma, je pojasniti mehanizem, ki bi omogočal vzročno interakcijo med dvema tako različnima stvarema, kot sta fizični dogodek (ali substanca) in mentalni dogodek (ali substanca). Materialisti trdijo, da težava izgine, če na mentalne dogodke gledamo kot na fizične. Omenjeno težavo je izpostavil Descartes in priznal, da je imel z njeno rešitvijo nemalo problemov. Zanj je bila, ker je obravnaval mentalne in fizične entitete v zastarelem smislu – kot entitete, ki zaradi svoje narave ali obstoja niso bile odvisne od ničesar -, ta težava tipična. Toda mentalni dogodki in druge mentalne entitete, kot npr. misli, niso substance v takšnem smislu. Tako za nas priznavanje različnih vrst strukturalne in eksistencialne odvisnosti mentalnega od fizičnega ne predstavlja nobenega problema. Kljub temu pa ni očitno, da bi morala ta odvisnost vključevati materialno zgradbo, tj., ni očitno, da bi morali biti mentalni dogodki identični s fizičnimi ali iz njih sestavljeni.

Sam nimam zadovoljivega odgovora na vprašanje, kako pojasniti omenjeni mehanizem. Prepričan pa sem, da pravih razlogov za njegovo sprejetje, ki bi tehtnico prevesili na stran materializma, ni v izobilju. Najprej ni jasno, ali je vprašanje o mehanizmu sploh relevantno. Ni jasno, ali je to sploh primerno vprašanje. Zahtevati obstoj mehanizma, s katerim bi lahko pojasnili vzročnost duha in telesa, je isto kot zahtevati obstoj fizičnega modela za razumevanje takšne vzročnosti. Še zdaleč pa ni očitno, da je takšen model ustrezen. Očitno ni niti, zakaj bi kakršenkoli model sploh potrebovali. Argument, ki sem ga navedel zgoraj, tako ne vključuje nobenega jasno formuliranega problema o mentalni vzročnosti, ki bi nas silil v to, da kot rešitev sprejmemo kakršenkoli materializem (vključno z njegovo komputacionalistično različico). Zahteva po mehanizmu je isto kot implicitna zahteva po materialistični rešitvi.

Bolj sistematičen argument za materializem, ki temelji na podobnih razlogih, je naslednji. Fizične učinke povzročajo predhodna fizična stanja ali dogodki skladno s povsem deterministčnimi fizikalnimi zakoni. Mentalni vzroki povzročijo fizične premike naših teles. Če takšna vzročnost ni del fizičnih procesov, potem bi obstajala odstopanja od povsem deterministčnih fizičnih vzorcev, opisanih s fizičnimi zakoni. To bi nasprotovalo fizičnim iznosom, jih spreminjalo ali na njih kakorkoli “drugače vplivalo”. Vendar razloga za mišljenje, da se to dogaja, ni. Predhodna fizična stanja zadostujejo za pojav fizičnih učinkov. Sklicevanje na mentalno vzročnost, ki ni del fizične vzročnosti, bi tako od nas zahtevalo, da dvomimo v ustreznosti fizikalnih pojasnitev celo znotraj fizike. Sprejetje takšne vzročnosti moramo zato zavrniti.

To sklepanje se mi zdi dokaj prepričljivo, toda kljub vsemu ne tako, kot morda izgleda. Zakaj bi mentalni vzroki, če niso del fizičnih procesov, fizični sistem spremenili ali si prišli z njim navzkriž? Mišljenje, da se to zgodi, temelji nedvomno na ocenjevanju mentalnih vzrokov po fizičnem modelu – le-ti učinku

nekaj dodajo. Zdi se, da je ideja v tem, da vzrok na učinek prenese del energije ali nanj s silo vpliva. V takšnem modelu bi potem mentalni vzroki fizični učinek speljali s poti, katero mu je začrtal fizični vzrok. Primeri, kjer bi mentalni in fizični vzrok poskrbela za isti fizični iznos – primeri "prekomerne določitve" -, bi po takšnem modelu izgledali kot protinaravna naključja. Tako se zdi sklicevanje na prekomerno določitev učinka, kot posledico pojava dveh vzrokov zanj, mentalnega in fizičnega, neprivlačno, hkrati pa meče še slabšo luč na idejo, da lahko mentalni vzroki fizične iznose, pojav katerih pričakujemo izključno na fizični osnovi, spreminjajo ali na njih vplivajo.

Vendar, ponavljam, vprašanje, ali je fizični model mentalne vzročnosti ustrezen, je del tega, s čimer se šele ukvarjamo. Kot smo videli, lahko najdemo različne načine, na katere mentalni vzroki "vplivajo na spremembe" in ki niso v nasprotju s fizičnimi pojasnitvami. Spremembe, za katere so odgovorni, so opisane s psihološkimi vzročnimi pojasnitvami in z njimi povezane s protidejstveniki. Takšne "spremembe", za katere poskrbijo psihološki vzroki, ne zahtevajo, da morajo v fizičnih verigah vzročnosti ostati vrzeli. Zdi se, da niso odvisne od nobenih posebnih domnev o fizičnih dogodkih, ki se nahajajo pod mentalnimi vzroki.

Mislim, da dobimo dovolj dober razlog za prepričanje, da si psihološko in fizikalistično stališče o vzročnih procesih ne nasprotujeta že po analizi pojasnjevalnih ciljev in prakse, kar je veliko prej, preden se sploh začnemo ukvarjati z ontologijo. Del poante sklicevanja na mentalna stanja tiči v pojasnjevanju intencionalne aktivnosti, ki vključuje (ali je identična s) fizičnim gibanjem. Tek moškega v lekarno pojasnimo z njegovim prepričanjem, da bo njegov otrok, če ne bo dobil zdravila, trpel, in njegovo odločitvijo, da ne bo počakal na zdravnika. Mislimo, da sta tek moškega povzročila oblikovanje prepričanja, ki ga ima, in njegova odločitev.

Bilo bi protinaravno misliti, da morajo takšni mentalni dogodki spreminjati oziroma vplivati na verigo fizioloških dogodkov, ki vključuje gibanje mišic med tekom, ali celo v njej zapolniti neko vrzel. Bilo bi protinaravno misliti, da psihološka pojasnitev izključuje ali nasprotuje neintencionalnim pojasnitvam fizičnega gibanja. Menim, da se te ideje ne zdijo protinaravne zato, ker vemo, da so mentalni dogodki materialni. Protinaravne se nam zdijo zato, ker vemo, da ti dve vzročni pojasnitvi razlagata isti fizični učinek kot posledico dveh zelo različnih vzorcev dogodkov. Pojasnitvi teh dveh vzorcev pa odgovarjata na dve zelo različni vrsti vprašanj. Nobena izmed teh dveh pojasnitev o drugi ne predpostavlja ničesar bistvenega, specifičnega. Tako relacij med entitetami, na katere se v teh dveh pojasnitvah sklicujemo, ne moremo razbrati iz vzročnih posledic, ki ju imata. Protinaravnost mišljenja, da morajo mentalni vzroki zapolniti vrzeli v fizičnih verigah dogodkov, verjetno izvira iz tradicionalnega dualizma ali liberalističnih pomislekov glede svobodne volje. Toda protinaravnost ostaja ne glede na to, od kod izvira.

Poanta tega razmišljanja je, da nimamo razlogov za domnevo, da mora biti neuspeh mentalnih vzrokov, da bi se vključili v fizično verigo dogodkov, pojasnjen v pojmih mentalnih vzrokov kot delov fizičnih dogodkov. Njihova vključitev bi bila glede na predhodne domneve o mentalnih in fizičnih pojasnitvah presenetljiva. Zato nevključevanja ni treba pojasnjevati v ontoloških pojmih.

Prav gotovo obstajajo nekakšne sistematične, celo nujne relacije med mentalnimi dogodki in spodaj ležečimi fizičnimi procesi. Za prepričanje, da so mentalni procesi odvisni od spodaj ležečih fizičnih procesov, imamo na voljo dobre razloge. S sondiranjem ali poškodovanjem dela možganov lahko povzročimo mentalna stanja in procese, na njih vplivamo ali pa jih celo uničimo. Toda relaciji identitete in fizične sestave sta relaciji, ki sta znanstveno uporabni. Obnašanje molekule npr. razlagamo v pojmih obnašanja njenih sestavnih delov. Še zdaleč pa ni jasno, da so te strukturalne relacije sistematično znanstveno uporabne za povezovanje psihologije in nevrofiziologije (glej op. 16). To so zgolj ugibanja o tem, kakšne relacije lahko obstajajo. Meni se zdijo le ena izmed možnosti za zagovarjanje določenih relacij med entitetami, na katere se – znotraj teh zelo različnih pojasnjevalnih praks – nanašajo. Kakšno obliko naj stališče, ki pojasnjuje relacije med dvema ravnema, zavzame, pa se mi zdi še vedno odprto.

Zdi se mi plavzibilno verjeti v neke vrste široko supervenienco: ni sprememb v mentalnih stanjih brez neke vrste spremembe v fizičnih stanjih. Toda sklepati iz tega na materializem, je po mojem mnenju metafizična spekulacija, za katero se zmotno domneva, da je postala relativno očitna znanstvena ali zdravorazumska puhlost.

Dokler psihološka pojasnitev zagotavlja vednost in razumevanje in dokler je ta pojasnitev (včasih) vzročna, lahko trdno verjamemo, da je vzročnost duha in telesa del našega sveta. Osnoven način razumevanja takšne vzročnosti je s pomočjo razumevanja psiholoških vzročnih in pojasnjevalnih stavkov, pri čemer gre za “razumevanje” v navadnem in nefilozofskem smislu. Koliko lahko k razumevanju mentalne vzročnosti doprineseta filozofija in nevrofiziologija, bomo šele videli. Na vsak način pa psihološka pojasnitev in mentalna vzročnost od materialistične metafizike ne potrebujeta priznanja o tem, da sta veljavni.

Zdi se mi, da bi se morali filozofi vprašanja, ali je določena oblika materializma resnična, lotiti bolj sproščeno. Mislim, da je vprašanje, ali so mentalni dogodki v kakršnemkoli očitnem smislu fizični, popolnoma odprto in nepomembno. Morda znanost nikoli ne bo mogla pojasniti ničesar drugega razen omejenih relacij med nižjimi, bolj avtomatičnimi deli kognitivnega sistema. Morda identitete in relacije del-celota nikoli ne bodo sistematizirane. Morda bo tradicionalna ideja o kategorični razliki v znanstveni praksi ostala še naprej prisotna. Tisto, kar je pomembno, je, da naše psihološke pojasnitve delujejo in da ne nasprotujejo fizikalističnim. Kot filozofi želimo dobro razlago tega, v kakšnem medsebojnem odnosu se te pojasnitve in njihovi predmeti spoznave nahajajo. Nima pa smisla

sporu med različnimi ontološkimi pristopi ali zaslugam materialističnega stališča pripisovati prevelikega pomena.

Za poplavo predlogov v zadnjih dveh desetletjih, ki poskušajo mentalno vzročnost ali mentalno ontologijo umestiti v "naturalistično sliko sveta", se mi zdi, da imajo le-ti več skupnega s politično ali religiozno ideologijo kot pa s filozofijo, ki naj bi ohranila pregled nad tem, kar vemo, in tistim, o čemer špekuliramo. Znanost materializma niti ne dokazuje niti ga jasno ne podpira. Metafizika bi morala tvegati in se podati na raziskovanje področij onkraj znanosti, pri čemer pa bi morala ohraniti svoj prefinjen občutek o odgovornosti.

4

Dokazoval sem, da epifenomenalizma ne rabimo obravnavati kot resne metafizične možnosti, pa četudi materializem drži. Jasnost psihološke vzročne pojasnitve je večja od resnosti metafizičnih pomislekov, ki jo postavljajo pod vprašaj. Prav tako sem dokazoval, da materialistična metafizika ni najbolj plavzibilno izhodišče za razmišljanje o vzročnosti duha in telesa, medtem ko pojasnjevalna praksa je.

Mislim, da je bolj naravno in plodno začeti s predpostavko, morda ovrgljivo, vendar trdno, da ima pripisovanje intencionalnih mentalnih dogodkov osrednji pomen tako v psiholoških pojasnitvah kot tudi v vsakodnevnem življenju in v različnih delih psihologije. Prav tako lahko predpostavimo, da so intencionalni mentalni dogodki pogosto vzroki in da ima psihološka pojasnitev pogosto obliko vzročne pojasnitve. Glede na ti predpostavki se zdi "skrb" zaradi nevarnosti epifenomenalizma zelo oddaljena. Kajti če so intencionalni mentalni dogodki individualizirani po vrstah v pojmih svojih intencionalnih lastnosti, tj., če so te lastnosti temeljne pojasnjevalne lastnosti in če ti dogodki vstopajo v vzročna razmerja in so (v pojmih teh lastnosti) citirani v pojasnitvah, potem se zdi, da imamo dober razlog za sklep, da so te lastnosti vzročno učinkovite. Metafizični pomisleki, ki sem jih predstavil v tej razpravi, se mi še zdaleč ne zdijo dovolj močni, da bi ta sklep ogrozili.[15]

Zaključil bom z omembo še enega razloga v prid ideji, da so intencionalni dogodki kot takšni učinkoviti. Zakaj so filozofi, ki se ukvarjajo z epifenomenalizmom, zaradi njega zaskrbljeni? Zakaj se ne držijo bolj nepristransko?

Del odgovora se skriva v dejstvu, da mnogi filozofi instinktivno verjamejo, da če je posledica njihove metafizike epifenomenalizem, potem je le-ta

[15] Že samo dejstvo, da se v pojasnitvah sklicujemo na intencionalne mentalne dogodke kot takšne, le-te razlikuje od fenotipov. Fenotipi so tisto, s čimer razlagamo v biologiji, intencionalni mentalni dogodki pa so del vzročnih razlag v psihologiji in vsakodnevnem življenju – ne le tisto, s čimer razlagamo razlage. Značilnost večine razprav s tega področja je, da so takšne majhne razlike v pojasnjevalnem statusu zgolj omenjene. O metafiziki situacije se pogosto razpravlja v navidezni izolaciji od pojasnjevalne prakse.

problematična.[16] To je s filozofskega vidika smiselna drža. Vsi namreč vemo, da zunaj naših filozofskih razprav epifenomenalizem ne drži.

Obstaja pa tudi globlji razlog, zakaj naj bi se nam epifenomenalizem zdel intelektualno neprivlačen. Cilj psiholoških pojasnitev, tako v psihologiji kot v vsakodnevnih razpravah, je pomagati ljudem, da bi sebe razumeli kot akterje. Vzročne implikacije so v naše intencionalne pojme in intencionalne načine pojasnjevanja že vgrajene. Prepričani smo, da, ker sprejemamo odločitve ali želimo nekaj narediti, sami povzročamo stvari. Mislimo, da podajamo trditve, oblikujemo teorije in ustvarjamo kulture, zato ker imamo določene misli in določene cilje, ki jih izražamo in izpolnjujemo. V tem kontekstu se v osnovi identificiramo v pojmih naših intencionalnih mentalnih značilnosti – naših želja, misli, vrednot. Naša dejavnost temelji na tem, da so naše potrebe, želje, misli in vrednote kot takšne v našem svetu v nekem smislu učinkovite. V posameznih primerih našo dejavnost definiramo v pojmih intencionalnega značaja, ki ga imajo naši mentalni dogodki.

Večina naših intelektualnih in praktičnih pravil ter ocen predpostavlja, da smo akterji. Če je naša želja ali odločitev botrovala temu, da se je nekaj zgodilo, vendar ta želja ali odločitev ni bila vzročno učinkovita (torej je učinkovitost izvirala iz neke spodaj ležeče živčne lastnosti), potem ta dejavnost ni bila naša. Če naših teoretičnih razmišljanj ne moremo nadzorovati, potem se ne moremo šteti niti za avtorje teorij niti se ocenjevati kot misleci. Tiste ocene naših intelektualnih in praktičnih aktivnosti, ki so najbolj normativne, bi bile v tem primeru prazne.

Če bi intencionalne psihološke razlage “osmislile” naše početje, ga “racionalizirale”, ne bi pa ponudile vpogleda v naravo vzročne učinkovitosti, bi izgubile veliko teže. Ne bi ponudile vpogleda v različne oblike dejavnosti, ki dajejo življenju njegov pomen in cilj, psihologom pa vzbujajo posebno zanimanje.

Na trdim, da ni možno, da bi psihološka razlaga odpovedala ali da bi bila vsaj veliko bolj omejena, kot si predstavljamo. Mislim, da je vprašanje zamisljivosti precej rafinirano in zapleteno. Poanta je, da obravnavana oblika psihološke razlage niti približno ni odpovedala. V okviru znanih omejitev zelo dobro deluje v vsakodnevnem življenju, v velikem obsegu se uporablja v psihologiji in družboslovju, prav tako pa je potrebna tudi za razumevanje naravoslovnih znanosti – pravzaprav bi lahko rekli, da je potrebna za kakršnokoli obliko razumskega udejstvovanja. Imamo razloge za prepričanje, da nam zagotavlja večje razumevanje naše dejavnosti. Hkrati pa, kot sem poudaril, obstaja tudi močna povezava med intencionalno psihološko pojasnitvijo in stališčem, ki nas na eni strani postavlja za akterje, na drugi pa mentalnim stanjem in dogodkom kot takšnim pripisuje neke vrste vzročno učinkovitost.

[16] Glej Block, 1990 za jasen opis tega stališča.

Mislim, da to niso edini razlogi za sprejetje te oblike pojasnitve. A tudi če bi bili, bi bila le-ta že tako močno zakoreninjena v nas, da ni nobenega pravega upanja za to, da bi lahko epifenomenalizem kdaj postal plavzibilna teorija. Na epifenomenalizem je podobno kot na skepticizem bolje gledati kot na instrument za razjasnitev naših najbolj globokih prepričanj. Zdi pa se mi, da je tradicionalni skepticizem glede dejavnosti, želja in odgovornosti za ta namen boljše orodje.

Literatura*

Block, N. (1990). "Can the Mind Change the World". V Boolos, G. (ur.), *Meaning and Method: Essays in Honor of Hilary Putnam*, Cambridge: Cambridge University Press, str. 137-170.

Burge, T. (1979). "Individualism and the Mental". *Midwest Studies in Philosophy*, 4, str. 73-121.

Burge, T. (1982). "Other Bodies". V Woodfield, A. (ur.), *Thought and Object: Essays on Intentionality*, Oxford: Clarendon Press, str. 97-120.

Burge, T. (1986a). "Cartesian Error and the Objectivity of Perception". V Pettit, P. in McDowell, J., *Subject, Thought and Context.* Oxford: Clarendon Press, str. 117-136.

Burge, T. (1986b). "Intellectual Norms and the Foundations of Mind". *Journal of Philosophy*, 83, str. 697-720.

Burge, T. (1989a). "Individualism and Causation in Psychology". *Pacific Philosophical Quaterly*, 70, str. 303-322.

Burge, T. (1989b). "Wherein is Language Social". V George, A. (ur.), *Reflections on Chomsky*, Oxford: Basil Blackwell, str.175-191.

Davidson, D. (1969). "The Individuation of Events". V Rescher, N. (ur.), *Essays in Honor of Carl G. Hempel*, Dordrecht: Reidl, str. 216-234.

Davidson, D. (1986). "Knowing One's Own Mind". *Proceedings of the American Philosophical Association*, 60, str. 441-458.

Davidson, D. (1989). "The Myth of the Subjective". V Krausz, M. (ur.), *Relativism: Interpretation and Confrontation*, Notre Dame: University of Notre Dame Press, str. 159-172.

Davidson, D. (1970). "Mental Events". V Foster, L. in Swanson, J., *Experience and Theory*, Amherst: University of Massachusetts Press, str. 79- 101.

Kim, J. (1984). "Epiphenomenal, and Supervenient Causation". *Midwest Studies in Philosophy*, 9, str. 257-270.

Putnam, H. (1975). "The Meaning of 'Meaning'". V Gunderson, K. (ur.),

* dodatek prevajalca

Language, Mind and. Knowledge (Minnesotta Studies in the Philosophy of Science, 7), Minneapolis: University of Minnesotta Press, str. 131-193.

Putnam, H. (1988). *Representation and Reality*. Cambridge: MIT Press.

Sosa, E. (1984). "Mind-Body Interaction and Supervenient Causation". *Midwest Studies in Philosophy*, 9, str. 271-281.

Van Gulick, R. (1995). "Who's in Charge Here? And Who's Doing All the Work?". V Heil, J. in Mele, A. (ur.), *Mental Causation*, Oxford: Oxford University Press, str. 97-120.

Janez Bregant
Pedagoška fakulteta Univerze v Mariboru

KRITIKA BURGOVE REŠITVE MENTALNE VZROČNOSTI

Nekateri antiredukcionisti so se pri iskanju odgovorov na to, kako so lahko mentalne lastnosti vzročne učinkovite, če pa – ob predpostavki supervenience kot metafizične relacije med mentalnim in fizičnim – spodaj ležeče fizične lastnosti oropajo mentalne kakršnekoli vzročne moči, zaradi česar slednje izpadejo iz vzročnih pojasnjevalnih teorij o našem fizičnem obnašanju, oprli na idejo, da ta ti. *problem vzročne izključitve* izhaja iz našega napačnega prepričanja, da morajo biti zgoraj ležeče mentalne lastnosti na nek način odvisne od fizičnih, saj naj bi jim le takšna metafizična zveza v nekem smislu zagotavljala vzročni status. Mentalno vzročnost želijo rešiti zgolj s sklicevanjem na pojasnjevalno prakso, v kateri mentalne lastnosti vsak dan opisujemo kot vzroke naših fizičnih dejanj, zaradi česar naj bi le-ta predstavljala pravi temelj za to, da lahko mentalnemu pripišemo vzročno vlogo, ne da bi se za to morali sklicevati na kakršnokoli metafizično relacijo med duhom in telesom. V članku je predstavljena kritika Burgove rešitve problema vzročne izključitve, ki spada v opisani razred rešitev, namiguje pa na to, da se s sklicevanjem na pojasnjevalno prakso omenjenemu problemu sploh ne izognemo.

Ključne besede: mentalna vzročnost, problem vzročne izključitve, epifenomenalizem, fizikalizem primerkov, supervenienca, pojasnjevalna praksa.

Uvod

Pri poskusu ubranitve mentalne vzročnosti, tj. izognitvi problemu vzročne izključitve, so se *nekateri* antiredukcionisti zatekli k ideji, da v zvezi z vzročno učinkovitostjo mentalnega ne obstaja nič, kar bi bilo problematično. Vse težave, ki izhajajo iz razprave o vzročni relevantnosti psihičnega, predvsem *epifenomenalizem,* so posledica napačnega prepričanja, da morajo argumenti za mentalno vzročnost temeljiti na metafizičnih predpostavkah (Burge, Baker, Van

Gulick).[1]

Vir nepotrebnih skrbi, ki jih imamo v zvezi s tem, ali ima mentalno vzročno moč, leži tako v napačnem razumevanju tega, kaj je pogoj za vzročno učinkovitost psihičnega. Po njihovem mnenju moramo razloge za to, kako je lahko mentalno v svetu, ki je v osnovi fizičen svet, vzročno relevantno, iskati v *pojasnjevalni moči* psihičnega, tj. v njegovi sposobnosti, da nastopa v razlagah dogodkov, ne pa v *metafizičnih relacijah* med duhom in telesom. Tako bi ugotovili, da so vsi problemi, ki so posledica prevelike zaverovanosti v metafiziko in podcenjevanja epistemologije, v resnici le navidezni. Burge pravi takole:

> "Sicer bom povedal nekaj o tem, zakaj mislim, da so nekateri izmed vidnejših argumentov, ki vodijo v epifenomenalizem, napačni, drugače pa me tukaj bolj zanima, od kod omenjeni dvom izvira. Mislim, da je posledica napačnega izbora filozofskih prioritet. Materialistični metafiziki je namenjene preveč pozornosti, razmišljanju o pojasnjevalni praksi pa premalo. Metafizična osnova, iz katere ta dvom izhaja, je veliko manj prepričljiva kot pa vsakdanja praksa, s pomočjo katere se ga lahko znebimo." (Burge, 2004: 31)

Vendar se zdi prepričanje, da gre pri mentalni vzročnosti le za vprašanje izbire med metafizičnim principom na eni strani in epistemološko prakso na drugi, sporno. Zdi se, da nas glede mentalne vzročnosti ne skrbi vprašanje, *ali* lahko npr. naše želje, prepričanja in namere vplivajo na premikanje našega telesa, saj je vsaj v vsakodnevnem življenju vsakomur jasno, da je odšel v kuhinjo po kozarec mleka zaradi želje po njem, ampak, *kako* je lahko mentalna vzročnost ob predpostavljanju fizikalizma, ki zahteva ustrezno povezavo fizičnega in psihičnega, možna. Vprašanje v zvezi z mentalno vzročnostjo na prvi pogled ni epistemološko, ampak metafizično, res pa je, da odgovor na to, *kako* je možna, vpliva na odgovor, *ali* je možna.

Kljub temu pa je v središču nekaterih ponujenih rešitev problema vzročne izključitve ideja, da za razumevanje mentalne vzročnosti utemeljitev metafizične relacije ni nujna in da jo lahko – zato da bi bila rešitev problema vzročne izključitve uspešna – mirne duše pozabimo. V prvem delu članka bom opisal epifenomenalizem, v drugem bom povzel Burgov odgovor na problem vzročne izključitve, ki spada v omenjeni razred rešitev, v tretjem pa bom predstavil njegovo kritiko.

I Epifenomenalizem

Oglejmo si najprej, kaj teorija, ki zavrača obstoj mentalne vzročnosti, zaradi česar je Burgu trn v peti, sploh pravi. Predpostavimo, da žeja povzroči željo po pitju, bolečina pa željo po jemanju aspirina. Takšno povzročanje mentalnega z mentalnim je dober izhodiščni primer za prikaz tega, kar pravijo nasprotniki mentalne vzročnosti. Motivacijo za njihove trditve lahko najdemo v ugotovitvah

[1] *Glej* npr. Kim (2001: 37-38) ali Van Gulick (1995: 233-237).

nevrofiziologije. Nevrofiziologi predpostavljajo, da so mentalni pojavi le "sence" nevrofizioloških procesov, ki se dogajajo v možganih, in ti fizični procesi so v končni fazi tisto, kar zagotavlja biološkemu organizmu, da normalno deluje. Tako tisti, ki zavračajo obstoj mentalne vzročnosti, pravijo, da željo po pitju in željo po jemanju aspirina povzročajo možganski procesi, ne pa žeja in bolečina kot mentalna pojava, na kar bi prisegli na prvi pogled in kar se nam zdi v vsakdanjem življenju več kot očitno. Žeja in bolečina sta lahko le "senci" fizičnih procesov, ki so pravi vzroki za želji po pitju in jemanju aspirina. Zato je naše prepričanje, da sta žeja in bolečina povzročili želji po pitju in jemanju aspirina, napačno.

Kot vzrok in učinek sta lahko povezana samo fizični in mentalni dogodek (dva mentalna dogodka ne moreta biti *nikoli* povezana kot vzrok in učinek), pa še v tem primeru se mora mentalni dogodek *vedno* nahajati na koncu vzročne verige (gre v bistvu za povzročanje mentalnega s fizičnim). Vzročne zveze med mentalnimi dogodki, ki se nam zdijo na prvi pogled očitne, ne predstavljajo pristnih vzročnih relacij, ampak so le odsev tega, kar se dogaja na bolj temeljnih ravneh. Tako nasprotniki mentalne vzročnosti trdijo, da so vsi mentalni dogodki povzročeni s fizičnimi, zaradi česar so mentalni dogodki "epifenomeni", tj. dogodki brez moči, da bi povzročili druge dogodke. Mentalni dogodki so učinki fizičnih procesov, sami ne povzročajo ničesar, niti fizičnih dogodkov niti drugih mentalnih, in vedno predstavljajo konec vsake vzročne verige. *(glej* Kim, 1996: 129) Takšno stališče se, ker predpostavlja, da mentalni dogodki niso vzročno relevantni oziroma učinkoviti (mentalno nima v vzročni verigi dogodkov nobene vloge), imenuje *epifenomenalizem.* Opozoriti pa velja, da epifenomenalizem ne dvomi v realnost mentalnega, le vzročne moči mu zaradi opisanih razlogov ne pripisuje.

V sodobni razpravi o mentalni vzročnosti epifenomenalizem ni priljubljeno stališče in skoraj nihče ga neposredno ne sprejema (vsi se mu v bistvu poskušajo izogniti, zdi pa se, da vsem to ne uspe in da obstajajo filozofi, ki jih njihovo stališče o mentalni vzročnosti vodi v epifenomenalizem). Verjetno bi veliko epifenomenalistov našli v nevrofiziologiji, ki, kot smo že omenili, obravnava mentalno, še posebej zavest, le kot senco ali odsev zapletenih živčnih procesov, ki se dogajajo v možganih. Zakaj pa je epifenomenalizem, čeprav sprejema celo neko vrsto mentalne vzročnosti (povzročanje mentalnega s fizičnim), sploh pa ne zanika obstoja mentalnih pojavov, v filozofiji tako nepriljubljen? Standardni odgovor na to vprašanje je, da zato, ker je njegova predpostavka o vzročni neučinkovitosti mentalnega nekaj, kar je tako močno v nasprotju z našimi intuicijami in s tem, kar čutimo globoko v nas, da se nam zdi napačna in da je nikakor nismo pripravljeni sprejeti.

Nekateri avtorji pa se bojijo tudi daljnosežnih posledic, do katerih lahko pripelje epifenomenalizem. "Ker predpostavlja, da v naravi obstaja nekaj, kar ne služi nobenemu namenu, nekaj imenitnega, kar je odvisno od dela tistih, ki so mu podrejeni /.../ nevarno pa je, da bi, in nedvomno sčasoma tudi bo, to izginilo." (Alexander, 1920: 8) Mentalno v svetu ne počne ničesar in je zato povsem nekoristno. Tako se zdi nesmiselno, da bi ga obravnavali kot nekaj realnega.

Naša prepričanja in želje ne vplivajo niti na sprejemanje naših odločitev niti na izvrševanje naših dejanj, zato niti naša zaznava niti naše znanje nimata nobenega opravka z našimi umetniškimi stvaritvami in tehnološkimi izumi. Imeti vzročno moč tako pomeni biti realen. Prikrajšati mentalno za vzročno učinkovitost zato pomeni prikrajšati ga za realnost. *(glej* Kim, 1996: 129-130) Alexandrov citat nas tako na prepričljiv način opozarja na to, kam nas sprejetje epifenomenalizma *lahko* vodi.

II Burgova rešitev mentalne vzročnosti

Po njegovem težave, ki pestijo mentalno vzročnost, izhajajo iz prepričanja, da so ene lastnosti dogodkov vzročno relevantne, druge pa ne, zaradi česar se poraja vprašanje, ali so višje ležeče mentalne lastnosti resnično vzročno učinkovite ali pa moramo morda vso vzročno moč pripisati nižje ležečim fizičnim lastnostim. Številni so prepričani, da je premikanje udov, kot je npr. dvigovanje in spuščanje roke, v tolikšni meri odvisno od fizičnih procesov, ki potekajo na nižjih ravneh, da so fizične lastnosti tiste, ki so vzrok za naše obnašanje, in tako mentalne lastnosti izpadejo iz kroga tistih, ki jim pripisujemo vzročno moč (epifenomenalizem).

Motivacija za takšen *skrajni* odgovor se skriva v ideji, da zaradi sprejemanja fizikalizma primerkov mentalni dogodki črpajo vzročno moč iz svojih fizičnih lastnosti, ki so edine sposobne poskrbeti za pravo vzročnost, zaradi česar so mentalne lastnosti odveč. (Tudi supervenienca mentalnih lastnosti nad fizičnimi naj ne bi bila rešitev, saj imamo po epifenomenalistovem mnenju še vedno dva vzroka za isti dogodek, pri čemer fizična lastnost, v kateri se skriva pravi vzročni mehanizem, mentalno izključuje iz vzročnih relacij.)

V obliki argumenta lahko navedeno zapišemo takole:

(1) Primerki mentalnih dogodkov so identični s primerki fizičnih dogodkov.

(2) Vzročno moč fizičnega dogodka določajo zgolj njegove fizične lastnosti.

(3) Mentalne lastnosti niso reducibilne na fizične lastnosti.

∴ Mentalne lastnosti niso vzročno učinkovite. *(glej* Burge, 2004: 33)

Burge meni, da je ključnega pomena za sprejetje tako skrajnega stališča, kot je epifenomenalizem, ki mentalnemu odreka kakršnokoli vzročno moč, prva premisa, ki izraža metafizično relacijo med duhom in telesom ter pripelje mentalno tako blizu fizičnemu, da izgubi ves vzročni vpliv. Prepričan pa je, da je poudarjanje fizikalizma primerkov neupravičeno, saj je takšna relacija med mentalnimi in fizičnimi dogodki vse prej kot jasna, zaradi česar je osnova za tako skrajni sklep, ki nas sili v neučinkovitost mentalnega, vprašljiva. Mentalno bi prav lahko imelo vzročno moč, če se ne bi tako osredotočili na iskanje metafizičnih relacij med duhom in telesom, ki potem, pa čeprav niso niti najmanj jasne, predstavljajo temelje za naše sklepe o psihičnem, popolnoma pa pozabljamo na pojasnjevalno prakso, h kateri bi se najprej morali napotiti po odgovor.

Njegov argument zoper fizikalizem primerkov v klasičnem smislu, tj. da so lahko različni primerki iste vrste mentalnega stanja primerki različnih vrst fizičnih stanj, temelji v glavnem na miselnem poskusu z Zemljo dvojčico. Spomnimo se, da Putnam z njim dokazuje naslednje: vsebina misli ni v naših možganih, ampak se nahaja v svetu zunaj nas. Razlog za to je v tem, da imata Peter in njegov dvojnik isto misel, oba želita popiti požirek vode, vendar se vsebina njunih misli razlikuje. Peter se namreč nahaja na Zemlji, kjer je voda sestavljena iz dveh atomov vodika in enega atoma kisika, njegov dvojnik pa na Zemlji dvojčici, kjer je voda, čeprav je vse ostalo isto, sestavljena iz po enega atoma X, Y in Z. Tako se Petrova misel nanaša na H_2O, misel njegovega dvojnika pa na XYZ, pa čeprav sta oba natančna fizična duplikata. To pomeni, da imata lahko njuni misli, čeprav se oba nahajata v istem fizičnem stanju, različno vsebino.

Burge zaključi, da je lahko posameznik v istem fizičnem stanju (Petra in njegovo kopijo lahko obravnavamo kot eno osebo), vendar ima zaradi različne vzročne zgodovine misli z različno vsebino, iz česar izhaja prva premisa njegovega argumenta za to, da je fizikalizem primerkov napačen.[2]

Druga premisa argumenta pa je posledica splošnega prepričanja, da sta dva dogodka, ki imata različno vsebino, tudi različna dogodka, zanikanje česar Davidson imenuje ne le neverjetno, ampak absurdno – saj če imata dva mentalna dogodka različno vsebino, potem sta gotovo dva različna dogodka. *(glej* Burge, 2004: 37)

Burgov argument za to, da je fizikalizem primerkov napačen, potem v standardnem zapisu izgleda takole:

(1) Možno je, da ima oseba misli z različno vsebino, pa čeprav so vsi primerki dogodkov, ki se pojavijo v posameznikovem telesu in so verjetni kandidati za identifikacijo z mentalnimi dogodki ter opisljivi v pojmih naravoslovnih znanosti, kot so fizika, kemija in nevropsihologija, enaki.

(2) Noben pojav misli ne more imeti drugačne intencionalne vsebine in biti primerek istega dogodka ali biti isti dogodek.

∴ Fizikalizem primerkov je napačen.

Burge zaključi, da metafizična relacija identitete primerkov, iz katere v skrajni točki izhaja ideja o vzročni neučinkovitosti mentalnega, ne drži, zaradi česar je nesmiselno, da se ji pripisuje tako velik pomen. "/.../ ne moremo vedeti *a priori*, da zakoni ali nomološke posplošitve psihologije opisujejo dogodke, ki imajo kakšno specifično zvezo z dogodki, ki padejo v okvir zakonov kakšne druge znanosti." (Burge, 2004: 43)

Čemu pa moramo po njegovem nameniti največ pozornosti oziroma, kaj je tisto, s čimer bi lahko glede na to, da so relacije med duhom in telesom, na katerih temeljijo vse materialistične rešitve mentalne vzročnosti, sporne, upravičili vzročno relevantnost psihičnega?

[2] Davidson se s tem ne strinja *(glej* Burge, 2004: 37-38).

> "Sistematične, informativne in pomembne pojasnjevalne sheme, kot je naša psihološka, so najboljše zagotovilo za ustrezne individualizirane opise, ki kažejo na to, kaj je bistveno za identiteto posameznikov, delov ali primerkov. /.../ Najboljša pot do razumevanja mentalne vzročnosti in ontologije vodi preko razumevanja najbolj zanesljivih načinov pojasnjevanja in opisovanja mentalnih dogodkov. /.../ Kot sem že omenil, so sistematične, informativne in pomembne pojasnjevalne sheme dogodkov in stanj hkrati tudi naši najboljši kazalci vzročne relevantnosti." (Burge, 2004: 41-43)

Omeniti moramo, da Burge ni kartezijanski dualist, saj meni, da "prav gotovo obstajajo nekakšne sistematične, celo nujne relacije med mentalnimi dogodki in spodaj ležečimi fizičnimi procesi. Za prepričanje, da so mentalni procesi odvisni od spodaj ležečih fizičnih procesov, imamo na voljo dobre razloge". (Burge, 2004: 46) Opozoriti velja, da je za Burga dovolj izključitev metafizične relacije iz kroga tistih, ki vplivajo na to, ali imajo mentalne lastnosti vzročno moč ali ne, zaradi česar sprejetje njenega obstoja, kakršnakoli naj bi ta zveza že bila, ne minira njegove ideje o tem, da se moramo pri pojasnitvi mentalne vzročnosti ozirati zgolj na pojasnjevalno prakso, metafizično relacijo med duhom in telesom pa zanemariti. Pravi celo naslednje:

> "Zdi se mi plavzibilno verjeti v neke vrste široko supervenienco: ni sprememb v mentalnih stanjih brez neke vrste spremembe v fizičnih stanjih. Toda sklepati iz tega na materializem, je po mojem mnenju metafizična spekulacija, za katero se zmotno domneva, da je postala relativno očitna znanstvena ali zdravorazumska puhlost." (Burge, 2004: 46)

Ker je metafizični model mentalne vzročnosti vprašljiv, se moramo ozreti po nečem drugem in ni boljšega od vzročnih psiholoških pojasnitev, ki predpostavljajo, da ima mentalno vzročno moč, poleg tega pa niso odvisne od nobene določene predpostavke o nižje ležečih fizičnih procesih. Burge trdi, da "bi bilo protinaravno misliti, da psihološka pojasnitev izključuje ali nasprotuje neintencionalnim pojasnitvam fizičnega gibanja. /.../ Protinaravne se nam zdijo zato, ker vemo, da ti dve vzročni pojasnitvi razlagata isti fizični učinek kot posledico dveh zelo različnih vzorcev dogodkov. Pojasnitvi teh dveh vzorcev pa odgovarjata na dve zelo različni vrsti vprašanj. Nobena izmed teh dveh pojasnitev o drugi ne predpostavlja ničesar bistvenega, specifičnega." (Burge, 2004: 45-46)

Zaključimo lahko, da je po njegovem mentalistična razlaga, dokler nam omogoča znanje in razumevanje, dovolj za predpostavko, da je mentalna vzročnost del našega sveta. Osnovo za takšno vzročnost predstavljajo psihološke pojasnjevalne sheme oziroma vsakodnevna pojasnjevalna praksa. Ali lahko nevrofiziologija pripomore k večji jasnosti takšne vzročnosti, bomo šele videli, nikakor pa niti mentalistična razlaga niti mentalna vzročnost ne potrebujeta odobritve s strani materialistične metafizike.

III Kritika Burgovega predloga

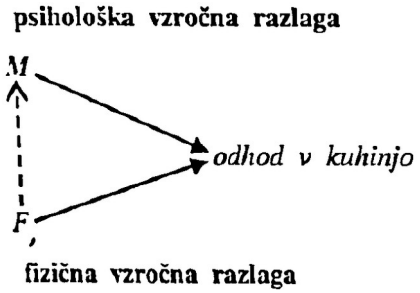
Vprašanje je, kot smo že omenili, ali se pri pojasnjevanju mentalne vzročnosti metafiziki resnično lahko meni nič tebi nič odrečemo in začnemo iskati napotke za rešitev vzročne učinkovitosti v epistemologiji. Burge trdi, da "je bolj naravno in plodno začeti s predpostavko, morda ovrgljivo, vendar trdno, da ima pripisovanje intencionalnih mentalnih dogodkov osrednji pomen tako v psiholoških pojasnitvah kot tudi v vsakodnevnem življenju in v različnih delih psihologije. Prav tako lahko predpostavimo, da so intencionalni mentalni dogodki pogosto vzroki in da ima psihološka pojasnitev pogosto obliko vzročne pojasnitve. Glede na ti predpostavki se zdi "skrb" zaradi nevarnosti epifenomenalizma zelo oddaljena. Kajti če so intencionalni mentalni dogodki individualizirani po vrstah v pojmih svojih intencionalnih lastnosti, tj. če so te lastnosti temeljne pojasnjevalne lastnosti, in če ti dogodki vstopajo v vzročna razmerja in so (v pojmih teh lastnosti) citirani v pojasnitvah, potem se zdi, da imamo dober razlog za sklep, da so te lastnosti vzročno učinkovite. Metafizični pomisleki, ki sem jih predstavil v tej razpravi, se mi še zdaleč ne zdijo dovolj močni, da bi ta sklep ogrozili." (Burge, 2004: 47)

Kim se s takšnim odgovorom ne strinja in pravi, da je goreča razprava o tem, ali so intencionalne razlage dejanj vzročne razlage, ki je potekala v petdesetih in šestdesetih letih prejšnjega stoletja, nazorno pokazala, da pojasnjevalne prakse ne moremo očistiti kakršnegakoli metafizičnega vpliva. Davidsonova rešitev, ki je psihološkim razlagam zagotovila vzročni status, je vključevala precejšnjo mero metafizike, npr. idejo o dogodkih in njihovih opisih, cel argument zanjo pa verjetno ne bi mogel mimo njegove sporne doktrine o anomalnosti mentalnega. Tako mora Burge pri sklicevanju na vzročno moč mentalističnih razlag upoštevati, da je vzročno učinkovitost psihičnim lastnostim zagotovil Davidsonov filozofski argument in velika količina metafizike v njem, "ne pa prepričljivost pojasnjevalne prakse, ki racionalizira dejanje v pojmih prepričanj in želja. Ni šlo za nesoglasja v zvezi s pojasnjevalno prakso, razprava se je nanašala na njeno naravo in principe." (Kim, 2001: 63) To bi nas moralo prepričati, da metafizike iz razprav o mentalni vzročnosti ne moremo kar tako izpustiti, saj se zdi, da razlagalne zveze temeljijo na ontoloških, zaradi česar Burge, ki prisega na neko vrsto metafizične relacije med duhom in telesom (še več, sprejema celo neko vrsto supervenience), problema vzročne izključitve ne more rešiti brez njenega upoštevanja.

Pa vzemimo kot primer željo po mleku *(M),* ki povzroči premikanje mojega telesa, npr. odhod v kuhinjo, in si oglejmo, kako bi se z njim spopadel Burge. Na voljo imamo dve razlagi, nevrofiziološko, ki pojasnjuje moje gibanje v pojmih prenašanja živčnih signalov in krčenja ustreznih mišic in se, gledano od zadaj naprej, konča z vibriranjem določenih živčnih vlaken v osrednjem živčnem sistemu mojih možganov, in psihološko, ki pojasnjuje moje gibanje v pojmih prepričanj in želja. Prva vključuje kot vzrok za moj odhod v kuhinjo fizično lastnost *F* (vzburjenost živčnih vlaken), druga pa *M* (željo po mleku). Vprašanje je, kakšno je razmerje med njima, saj je očitno, da imamo dve vzročni verigi, ki

ponujata dva različna vzroka za učinek, zaradi česar nam grozi problem vzročne izključitve, saj je fizična vedno sama zadostna za to, da proizvede učinek, ki nas sili v to, da ju *povežemo*. Burge bi glede na to, da pravi, da bi bilo "protinaravno misliti, da psihološka pojasnitev izključuje ali nasprotuje neintencionalnim pojasnitvam fizičnega gibanja" (Burge, 2004: 45), verjetno odgovoril, da nam tega ni treba storiti, saj si fizične in psihološke razlage pri pojasnjevanju fizičnega obnašanja niso konkurenčne.

Ampak, ali nas to ne vodi spet v vzročno izključitev mentalnega? Kim pravi, da "je zanimivo dejstvo v zvezi z razlagami, ki ga je Burge spregledal, to, da sta si lahko dve *(ali* več) razlagi, če pojasnjujeta (še zlasti, če vzročno pojasnjujeta) isti eksplanandum, konkurentki ne glede na to, da so njune pojasnjevalne premise konsistentne in resnične". (Kim, 2001: 65) To razločno kaže naslednja slika:



(Prekinjena puščica označuje relacijo supervenience, polni puščici pa vzročnosti.)

Tako se zdi, da problem vzročne izključitve, čeprav razlage pripadajo različnim ravnem opisa oziroma različnim področjem raziskovanja, *ne izgine*, in da sklicevanje na epistemologijo pri njegovem reševanju ne razveljavi potrebe po metafizičnem modelu mentalne vzročnosti. Kim zaključi takole:

> "Obstajata kratki in ravni poti od mentalistične razlage do mentalne vzročnosti in od dopuščanja dveh razlag istega pojava fizičnega [dogodka] do grozečega problema vzročno pojasnjevalne izključitve." (Kim, 2001: 67)

A REVIEW OF BURGE'S SOLUTION OF MENTAL CAUSATION

Some anti-reductionists try to find a plausible answer to the question how can mental properties be causally efficacious under a supposition of supervenience as a metaphysical relation between the mental and the physical which makes underlaying physical properties be the only causes of physical events kicking thereby mental properties out of explanatory theories about our physical behaviour

in a peculiar manner. They appeal to the idea that this problem, called the problem of causal exclusion, follows from our false assumption that upperlaying mental properties must be in some way dependent on physical properties since it is only under this condition that they have, at least in some sense, a causal status. They try to save mental causation solely by appealing to explanatory practice in which mental properties are everyday described as causes of our physical behaviour. Therefore, explanatory practice is supposed to be a true fundament upon which mental properties have causal roles no matter what the metaphysical relation between mind and body is. This article introduces a critique of Burge's solution of the exclusion problem, which belongs to this class of proposals, suggesting that we do not avoid the mentioned problem by appealing to explanatory practice.

Keywords: mental causation, exclusion problem, epiphenomenalism, token physicalism, supervenience, explanatory practice.

Literatura

Alexander, S. (1920). *Space, Time, and Deity.* London: Macmillan.

Baker, L. R. (1995). "Metaphysics and Mental Causation". V Heil, J. in Mele, A. (ur.), *Mental Causation,* Oxford: Clarendon.

Burge, T. (2004). "Vzročnost duha in telesa ter pojasnjevalna praksa". *Analiza,* št. 3, str. 31-50.

Davidson, D. (1995a). "Mental Events". V Moser P. K. in Trout, J. D. (ur.), *Contemporary Materialism,* London: Routledge.

Heil, J. in Mele, A. (ur.) (1995). *Mental Causation,* Oxford: Clarendon.

Horgan, T. (1997). "Kim on Mental Causation and Causal Exclusion". *Philosophical Perspectives,* 11, str. 165–184.

Kim, J. (1993). *Supervenience and Mind.* Cambridge: Cambridge University Press.

Kim, J. (1996). *Philosophy of Mind.* Boulder: Westview.

Kim, J. (2001): *Mind in a Physical World.* Cambridge: MIT Press.

Van Gulick, R. (1995). "Who's in Charge Here? And Who's Doing All the Work?". V Heil, J. in Mele, A. (ur.), *Mental Causation,* Oxford: Clarendon.

Paul M. Churchland in Patricia S. Churchland
prevedel Sebastjan Vörös

INTERTEORETSKA REDUKCIJA: PODROČNI VODNIK ZA NEVROZNANSTVENIKE*

"Redukcionizem" je znotraj akademskih krogov sporen termin. Nekaterim se zdi, da označuje pravilen pristop k vsem znanstvenim področjem, saj stremi k interteoretski enotnosti in pravi sistematičnosti opisa pojavov. Šlo naj bi torej za pristop, ki ga velja aktivno izvajati in vneto zagovarjati.

Drugim pa se zdi, da označuje napačno naravnan pristop, ki je ozkogleden in slep za bogastvo pojavov. Je neumen primer "ničkotstva"****, ki je slepo za emergentno kompleksnost in organizacije višjega reda. Gre za pristop, ki se ga velja ogibati.

Slednji ugovor največkrat najdemo pri različnih družbenih znanostih, kot so, na primer, antropologija, sociologija in psihologija. Prvo stališče pa je največkrat prisotno znotraj naravoslovnih znanosti, kot so, recimo, fizika, kemija in molekularna biologija. Povsem pričakovano je torej dejstvo, da je problem redukcionizma najbolj pereč tam, kjer se ta dva intelektualna tokova srečata, torej v moderni nevroznanosti.

Sporno vprašanje je, če lahko upravičeno pričakujemo in stremimo k redukciji vseh psiholoških pojavov na nevrobiološke in nevroračunske [nevro*computational]* pojave. Velik in še vedno ugleden del akademske skupnosti ostaja pri zavračanju te možnosti. Njihovo zavračanje je načelno. Nekateri opozarjajo na obstoj tega, čemur filozofi pravijo takšnosti [qualia], tj. raznovrstne subjektivne kakovosti, ki se nam kažejo v naših čutih: pomislimo na bolečino, vonj rože, občutek rdečosti itd. Za te takšnosti velja, da presegajo vsakršno zmožnost materialistične obrazlage ali redukcije (Jackson, 1982; Nagel, 1974). Drugi opozarjajo na semantično vsebino oziroma *intencionalnost* naših misli, ki naj bi ravno tako bila izven dosega redukcije (Popper in Eccles, 1978; Searle, 1980, 1990). Spet nekateri trdijo, da lahko najbolj pomembne aspekte človeškega obnašanja razložimo zgolj s pomočjo *emergentnih lastnosti* in njim lastnih zakonitosti, tj. s pomočjo lastnosti, ki ireducibilno vključujejo socialni nivo, kot so, recimo, vdanost kaki moralni ideji, dojemanje nekega političnega dejstva ali pre-poznava osebne izdaje (Taylor, 1970, 1987). Spet drugi vidijo

* Prevedeno po Paul M. Churchland in Patricia S. Churchland "Intertheoretic Reduction: A Neuroscientist's Field Guide," V *Philosophy and the Neurosciences* (ur. Bechtel, W., Mandik, P., Mundalc, J. Stufflebcam. R.), 2001, Oxford: Blackwell, str. 419-430. – op. prev.

** Skovanko iz "nothing but" – tj. "nič (drugega) kot" – "nothing-butery" prevajam s slovensko skovanko "ničkotstvo". – op. prev.

navzkrižje s pomembno in globoko utrjeno idejo o *človeški svobodi* (Popper in Eccles, 1978). Končno, nekateri materialisti izpostavljajo problem *večvrstne realizacije*. Opozarjajo namreč na dozdevno dejstvo, da lahko zavestno inteligenco vzdržujejo fizični sistemi, ki so drugačni od človeku lastne biokemije, npr. sistem tranzistorjev, ravno tako kot lahko nacionalno finančno ekonomijo vzdržujejo žetoni, ki so drugačni od trenutno uporabljanih kovancev in bankovcev. Vendar nihče med nami ne verjame, da lahko makroekonomijo reduciramo na kemijo kovin in papirja. Kako lahko torej predpostavljamo, da je mogoče psihologijo zreducirati na nevrobiologijo človeka? (Fodor, 1975)

Namen tega članka je trojen. Prvič, skušala bova predstaviti uporaben pregled splošne narave interteoretske redukcije, pri čemer se bova opirala na številne primere iz zgodovine znanosti. Tukaj je bistvenega pomena, da razširimo svoja obzorja, kajti namesto, da nepretrgoma zremo v problem, ki nam trenutno buri duhove (možnost redukcije psiholoških fenomenov na nevralne fenomene), si raje oglejmo primere, kjer se je prah spora že polegel, saj nas taki primeri lahko marsikaj naučijo. Drugič, pri teh primerih bova dognala, kaj so dejanske vrednosti, ki jo ti primeri kažejo, in kaj slabosti, ki se jih velja ogibati. Na koncu bova skušala uporabiti te zgodovinske lekcije na našem primeru – kognitivna nevroznanost – skušala z njihovo pomočjo spodbiti prej omenjene ugovore in pomisleke.

1. Interteoretska redukcija: nekaj prototipskih primerov

Ker so posamični zgledi najboljši učitelji, predlagava, da si jih na tem mestu nekaj tudi ogledamo. Eden najzgodnejših primerov obsežne interteoretske redukcije je bila redukcija Keplerjevih treh zakonov o gibanju astronomskih teles na Newtonovo mehaniko. Keplerjeva teorija je bila omejena zgolj na gibanje planetov znotraj našega osončja, medtem ko se je pri Newtonovi teoriji zdelo, da gre za pravilen opis gibanja v splošnem. Zato je bil nedvomno sijajen uspeh, ko je Newtonu uspelo pokazati, da se lahko, ob predpostavki, da je masa katerega koli planeta zanemarljivo majhna v primerjavi z maso sonca, vse tri Keplerjeve zakone izpelje iz njegove lastne teorije.

Keplerjevi trije zakoni za gibanje planetov so:

1. Vsi planeti se gibljejo po elipsah okoli Sonca, pri čemer se le-to nahaja v gorišču elipse.
2. Planet pokrije v enakih časih enake ploščine.
3. Za vsak planet velja, da je kvadrat planetovega obhodnega časa sorazmeren kubu planetove velike polosi.

Newtonovi trije zakoni so:

1. Inercijsko gibanje je konstantno in premočrtno.
2. Pospešek = sila/masa.
3. Medsebojna učinka dveh teles sta si vedno enaka in nasprotno usmerjena.

Izkazalo se je torej, da gre pri Keplerjevi razlagi zgolj za poseben primer oz. posebno aplikacijo bolj splošne Newtonove razlage. Prav tako se je izkazalo, da gre pri gibanju astronomskih teles zgolj za poseben primer splošnih inercijskih gibanj masivnih teles, ki jim gibanje usmerjajo sile. Božanski oz. nadnaravni značaj nebes je bil tako izgubljen za vekomaj. Sublunarna in superlunarna domena sta bili združeni v eno samo domeno, znotraj katere je en sam zbir zakonov urejal obnašanje vseh objektov.

Newtonova mehanika je prav tako zaslužna tudi za drugi veliki primer interteoretske redukcije, do katere pa je prišlo šele v 19. stoletju. Porodilo se je namreč vprašanje, ali lahko taista mehanika, glede na to, da ji je doslej uspelo razložiti gibanje tako na astronomski ravni kot na ravni človeške velikosti, ponudi razlago tudi za gibanje na mikroskopskem nivoju.

Poskusi izgradnje tovrstne razlage so ustvarili še eno združitev, katere rezultati so vključevali tudi nepričakovan dodatek, ki se tiče teorije toplote. Če predpostavimo, da se katera koli plinasta snov sestoji iz velikega števila mrgolečih submikroskopskih delcev, ki znotraj dane posode poskakujejo sem ter tja v skladu z Newtonovimi tremi zakoni, potem lahko izpeljemo zakon za pritisk, ki ga bodo ti delci kolektivno povzročali na stenah posode. Ta "kinetični" zakon ima sledečo obliko:

$PV = 2N/3 \cdot mv^2/2$

Ta zakon je imel enako obliko kot takrat že znani "zakon idealnega plina", namreč

$PV = nRT$

(Pri obeh velja, da P = pritisk, V = prostornina.) Čeprav se zakona razlikujeta v zapisu, pa tako "2N/3" kot "nR" označujeta količino plina v posodi (N – število delcev plina v posodi; n – število molov plina). Edina preostala razlika je ta, da ima "kinetični" zakon izraz za *kinetično energijo povprečnega delca* ($mv^2/2$) na mestu, kjer ima "zakon idealnega plina" izraz za *temperaturo* (T). Mar gre pri fenomenu, ki ga poznamo pod imenom "temperatura", pravzaprav za povprečno kinetično energijo na molekularnem nivoju? Ta presenetljiva konvergenca načel je pripravila Bernoullija, Joula, Kelvina in Boltzmanna, da so na to vprašanje odgovorili pritrdilno. Dodatne raziskave so pokazale, da ima povprečna molekulama kinetična energija *vse* vzročne lastnosti, ki jih je klasična teorija pripisovala temperaturi. Skratka, izkazalo se je, da temperatura ni nič drugega kot povprečna molekularna kinetična energija. Newtonska mehanika je tako slavila še eno redukcijsko zmago. Gibanje na vseh treh ravneh je bilo opisano znotraj ene same teorije, dobro znana pojavna lastnost, *temperatura*, pa je bila razložena in predstavljena na nov ter nekoliko nenavaden način.

Ni nepomembno poudariti, kako je pri tovrstni redukciji šlo za enačenje dobro poznane *fenomenalne* lastnosti vsakdanjih predmetov z zelo nevsakdanjo mikrofizično lastnostjo. (Z izrazom "fenomenalno" mislimo tisto lastnost, ki jo lahko zanesljivo izkustveno prepoznamo, vendar pa s sklicevanjem na preprostejše izkustveno prepoznavne elemente ne moremo razložiti, na kakšen

način pravzaprav prepoznavamo to lastnost.) Od tod sledi, da redukcija ni omejena zgolj na pojmovne sisteme, skrite v višavah teoretične stratosfere. Včasih se izkaže, da gre pri pojmovnem sistemu, ki ga vključimo v neko širšo vizijo, pravzaprav za del našega zdravorazumskega sistema, za del, katerega pojme redno uporabljamo v vzročnih opazovanjih, ki temeljijo na naših vrojenih zaznavnih sistemih. Oglejmo si še nekaj drugih primerov: že pred Newtonom so vedeli, da gre pri *zvoku* pravzaprav za kompresijske valove v atmosferi, pri *višini tona* pa za valovno dolžino; to vedenje je bilo rezultat obsežnejše redukcije zdravorazumskega pojmovanja zvoka in glasbene teorije na mehanično akustiko. Stoletje in pol po Newtonu pa so ugotovili, da so tudi *svetloba* in njene mnogovrstne *barve* enake elektromagnetnim valovom in njihovim različnim valovnim dolžinam; ta ugotovitev je bila rezultat obsežnejše redukcije geometrične optike na elektromagnetno teorijo, kot jo je predstavil Maxwell leta 1864. Tudi *žarečo toploto*, torej še en zdravorazumski pojav, so pričeli pojmovati kot dolgovalovno elektromagnetno valovanje, ko so jo nekoliko kasneje vključili v prej omenjeno teorijo. Očitno je torej, kako dejstvo, da je neka lastnost oz. neko stanje izpostavljeno kateri od naših čutnih zaznav, nikakor ne pomeni, da te lastnosti oz. stanja ne moremo pojmovati drugače, kadar ga vključimo v pojmovni sistem teorije z boljšo opisno močjo.

K temu nadvse pomembnemu dejstvu se bomo vrnili nekoliko kasneje. Za zdaj pa si oglejmo še nekaj primerov interteoretske redukcije. V 20. stoletju je redukcija klasične (valenčne) kemije na atomsko in subatomsko (kvantno) fiziko v 20. stoletju še en nazoren primer pojmovnega poenotenja. Tukaj so struktura atomovih zaporednih elektronskih lupin in lastnosti stabilnih regij pri porazdelitvi elektronov v medatomskih vezeh omogočile sistematično in poučno rekonstrukcijo elektronske strukture mnogih atomskih elementov, klasičnega zakona valenčne porazdelitve in celotne strukture periodnega sistema. Kot se to zelo pogosto pripeti v primeru interteoretske redukcije, je nova teorija lahko razložila prenekatero vprašanje, na katerega stara teorija ni imela ustreznega odgovora, kot, na primer, specifične toplotne kapacitete različnih substanc in interakcijo kemijskih spojin s svetlobo.

Redukcija kemije na fiziko je pomembna tudi zato, ker sam proces še ni zaključen in zelo verjetno tudi nikoli docela ne bo. Eden od razlogov, ki botruje temu dejstvu, se skriva v neskončnem številu možnih kemičnih spojin in njim značilnih kemičnih, mehaničnih, optičnih in termičnih lastnostih. Drugi razlog pa je ta, da je računanje teh raznolikih lastnosti, izhajajoč iz osnovnih principov kvantne mehanike, računsko grozljivo zamudno, tudi če se omejimo zgolj na približne rezultate, k čemur pa smo iz čisto matematičnih razlogov navadno primorani. Potemtakem ni res, da je bilo vso kemijsko znanje uspešno rekonstruirano v pojmih kvantne mehanike. Rekonstruirane so bile le osnove, pa še te zgolj v približkih. Vendar pa nam tukaj naše izkustvo trdno kaže, da je kvantni mehaniki res uspelo uspešno zajeti osnovne elemente kemijske resničnosti. Zato pričakujemo, da bi lahko vsak del kemije približno rekonstruirali s pomočjo kvantnomehanske terminologije, v kolikor bi kaj takega kdaj res bilo potrebno.

Predhodni primeri lepo ponazarjajo, da gre pri interteoretski redukciji v bistvu za odnos med dvema različnima pojmovnima sistemoma, ki opisujeta pojave, ne pa za dve različni domeni pojavov. Končni smisel redukcije je namreč pokazati, da je tisto, za kar smo mislili, da sta dve različni domeni, pravzaprav ena domena, ki pa jo lahko opišemo z dvema (ali večimi) besednjaki.

Brž̌čas najbolj znana redukcija pa je Einsteinova redukcija Newtonovih treh zakonov gibanja na bistveno drugačno mehaniko posebne teorije relativnosti (PTR). PTR je zajela Newtonove zakone na sledeč način: Če (napačno) predpostavimo, da se vsa telesa gibljejo s hitrostjo, ki je bistveno manjša od hitrosti svetlobe, potem PTR vsebuje zbir zakonov za gibanje, ki ga eksperimentalno ne moremo ločiti od Newtonovega starega zbira. Če upoštevamo relativno omejenost človeškega izkustva, tako ni nerazumljivo, da so se zdeli stari Newtonovi zakoni resnični.

Vendar pa so zakoni PTR semantično in logično bistveno drugačni od Newtonovih zakonov, saj svetu pripisujejo znatno drugačne lastnosti, čeprav so v zgoraj omenjenem posebnem primeru eksperimentalno nerazločljivi od Newtonovih zakonov. Tako v vsaki situaciji, kjer Newtonovi zakoni pripišejo telesu neko notranjo lastnost (npr. maso, dolžino, gibalno količino itd.), PTR pripiše *relacijo*, tj. dvomestno lastnost (npr. *x* ima masoglede-na-inercijski-sistem-F itd.), kajti predstavitev vesolja in tega, kar to vesolje vsebuje (enoten 4-D prostor-čas kontinuum s 4-D svetovnimi osmi), je bistveno drugačna kot pri Newtonu.

Tu gre za primer redukcije, kjer posebni primeri in deduktivne posledice nove in bolj splošne teorije niso identični, marveč zgolj podobni stari in bolj ozki teoriji, ki naj bi jo prva reducirala. Rekonstrukcija posebnega primera znotraj nove teorije namreč ustreza stari teoriji v zadostni meri, da lahko pokaže, čemu je stara teorija veljala za ustrezno v okviru določene domene, in kako je možno staro teorijo nadomestiti z novo, ne da bi pri tem izgubili razlagalno in napovedno moč znotraj domene stare teorije; vendar pa nova rekonstrukcija ni docela izomorfna s staro teorijo, kajti izkaže se ne le, da je stara teorija preozka, marveč da je v nekaterih ozirih tudi napačna. Čas in prostor nista različna, kot je predpostavljal Newton, ravno tako kot ne obstajajo neke od inercijskega okvirja neodvisne notranje lastnosti, kot sta npr. masa ali dolžina.

Na koncu omenimo še nekaj primerov, kjer nova in bolj splošna teorija v ničemer ne prevzame in podpre prikaza realnosti, kot jo slika stara teorija. Takšen primer bi bil neposredna zamenjava (brez redukcije) stare flogistonske teorije z novo kisikovo teorijo izgorevanja, kot jo je predstavil Lavoisier. Stara teorija je trdila, da pri izgorevanju nekega telesa pride do *izgube* duhu podobne substance, flogistona, katerega funkcija pred izgorevanjem je priskrbeti snovi njeno specifično kakovost (npr. kovinskost, lesenost); flogiston naj bi se vezal na žgano apno kot osnovo, ki ostane prisotna tudi po izgorevanju (torej po izgubi flogistona). Flogiston je bil “duh”, ki naj bi kovinam dajal njihovo ustrezno obliko. Po sprejetju Lavoisierove trditve, da med izgorevanjem pride do absorbcije neke popolnoma materialne substance, kisika, je bil flogiston

preprosto odstranjen iz našega opisa sveta.

Drugi primeri teoretičnih entitet, ki so jih eliminirali iz resne znanosti, so kaloričen fluid, rotirajoče kristalne sfere ptolemejske astronomije, štirje temperamenti srednjeveške medicine, vitalni duh predmoderne biologije in svetlobo oddajajoč eter predeinsteinske mehanike. V vseh teh primerih nove teorije niso imele zadostnih virov, da bi rekonstruirale strukturo starejših teorij ali zakone, ki naj bi veljali znotraj teh teorij; vendar so se nove teorije izkazale za bistveno boljše od starih, zato so jih v končni fazi tudi izpodrinile.

Če si na tem mestu ogledamo pot, ki smo jo prehodili, ugotovimo, da imamo na eni strani spektra pare teorij, kjer je stara teorija gladko reducirana na novo, pri čemer je ontologija stare teorije (tj. zbir stvari in lastnosti, ki jih le-ta postulira) preživela, čeprav je bila na novo opisana v novem in bistveno bolj poglobljenem besednjaku. Tukaj navadno srečamo trditve o medteoretični identiteti, kot npr. "Toplota je enaka povprečni molekularni kinetični energiji" in "Svetloba je identična elektromagnetnemu valovanju". V sredini spektra najdemo pare teorij, kjer v viziji nove teorije najdemo zgolj šibak odsev stare ontologije, ki "preživi" le v znatno izpremenjeni obliki. In končno, na drugem koncu spektra naletimo na pare teorij, kjer bistveno bolj uporabna ontologija in znatno bolj uspešni zakoni nove teorije v celoti izpodrinejo staro teorijo in njeno stare ontologijo.

Preden zaključimo ta kratek pregled, bi si veljalo ogledati še primere kjer starejša teorija ni bila niti vključena v novejšo in bojda bolj splošno teorijo niti ni bila iz nje eliminirana. V takih primerih se starejša teorija uspešno upira poskusom prevzema in se kaže kot poseben primer trenutno veljavne splošne teorije. Maxwellova elektromagnetna teorija (od tod naprej bomo jo imenovali EM teorija) je tipičen primer te vrste. Od leta 1864 do 1905 je večina znanstvenikov pričakovala, da bo EM teorija gotovo našla primerno redukcijo v mehaničnih lastnostih etra, tj. elastične substance, v kateri naj bi se EM valovi širili. Sproženih je bilo kar nekaj pomembnih poskusov rekonstrukcije EM fenomenov znotraj mehanske terminologije, ki pa niso nikoli bili izvedeni do konca. Na veliko presenečenje raziskovalcev se je izkazalo, da je obstoj absolutnega medija, v katerem naj bi se širili EM valovi, očitno nekonsistenten z lastnostmi časa in prostora, kot ga je opisala Einsteinova posebna teorija relativnosti iz leta 1905. EM teorija se je tako izkazala za temeljno teorijo in ne zgolj za poseben primer mehanike. Poskus vključevanja je klavrno propadel.

Drug primer je povezan s teorijo obnašanja zvezd, ki jo je zagovarjala klasična astronomija ob koncu devetnajstega stoletja. V tistem času so mnogi verjeli, da bosta mehanika ali fizika lahko v zadostni meri opisali vzorce oddajanja žarkov zvezde. Vendar je sčasoma postalo očitno, da so možne zaloge kemične ali mehanske energije, ki jih vsaka zvezda potrebuje za vzdrževanje enormnih izbruhov termalne in svetlobne energije, omejene na zgolj nekaj deset milijonov let. Ta omejeni obseg je bil v nasprotju z geološkimi dokazi, ki so zgodovino merili v *milijardah* let. Vendar pa se je Lord Kelvin odločil, da bo tvegal, in je geološkim ugotovitvam navkljub določil starost zvezd na nekaj deset milijonov let. Spor se je razrešil šele po odkritju velikanskih količin energije znotraj

atomskega jedra. Zvezdna astronomija je tako dejansko bila reducirana, vendar ne na kemijo ali mehaniko, marveč na kvantno fiziko. Tako je spodletel še en poskus redukcije, le-temu pa je sledil tak, ki je uspel.

2. Nauki za nevroznanost

Po tem pregledu vseh zgoraj navedenih primerov in pa spektra zgledov, ki ga ti primeri definirajo, se lahko vprašamo, česa se lahko iz tega nauči nevroznanstvenik. Eden od naukov je gotovo ta, da gre pri interteoretski redukciji za normalen in razmeroma pogost pojav v zgodovini znanosti. Drugi pa je ta, da je pristna redukcija, če do nje lahko pridemo, nedvomno nekaj pozitivnega. Dobra je iz več razlogov, le-ti pa so znantno močnejši, če jih navedemo v medsebojni povezani.

Prvič, staro teorijo s tem, ko jo prikažemo kot poseben primer znotraj (domnevno resnične) nove teorije, *potrdimo,* tj., dokažemo njeno veljavnost vsaj v grobem orisu ali znotraj neke ustrezno omejene domene. Drugič, ker gre pri rekonstrukciji stare teorije znotraj nove zgolj redko za popolnoma zrcalno podobo, je navadno potrebna *korekcija* določenih delov stare teorije, ti popravki pa kažejo na napredek v našem znanju. Tretjič, redukcija nam omogoča znatno bolj *poglobljeni vpogled* in bistveno bolj *učinkovit nadzor* nad fenomeni v okviru domene stare teorije. Četrtič, s pomočjo redukcije se dokopljemo do *preprostejšega* splošnega pojmovanja narave, saj so očitno raznovrstni pojavi združeni pod isto razlagalno streho. In petič, nova in splošnejša teorija *postane dedič vseh zakonov,* ki so se nakopičili znotraj reducirane stare teorije, saj navsezadnje nudi razlago za iste podatke.

Narobe je, če skušamo nadvse zadovoljivo staro teorijo položiti na Prokrustovo posteljo, tj. izvršiti redukcijo, kadar teorija, ki naj bi nadomestila staro teorijo, ne poseduje zadostnih sredstev, ki bi upravičevala tovrstno nadomestitev. Ali dana sredstva zadoščajo za redukcijo ali ne, pa je kljub človeškim intuitivnim prepričanjem le redko takoj razvidno. Tudi če je redukcija kakšne teorije nemogoča, lahko to priča o napačnosti te teorije in ne o njeni fundamentalni resničnosti. V takem primeru se lahko zgodi, da nova teorija docela izpodrine (eliminira) staro, namesto da bi jo zgolj gladko reducirala. Mogoče bodo človeška pojmovanja, kot so npr. “prepričanja” in “volja”, eliminirana in zamenjana z nekimi docela drugačnimi zgodbami o shranjevanju podatkov in povzročanju obnašanja.

Dejstvo je, da se mnenja o upravičenosti redukcije na področju nevroznanosti in psihologije zelo razlikujejo. Po eni strani se namreč zdi, da je upravičenost redukcije (ali eliminacije) na tem področju bistveno večja kot v vseh zgoraj navedenih zgodovinskih primerih. Kajti, za razliko od svetlobe, toplote ali gibanja nebesnih teles, je za psihološke fenomene v splošnem že znano, da nastanejo kot posledica evolucijskega in ontogenetskega razvoja snovi, natančneje razvoja biološke organizacije. Zato *pričakujemo,* da bomo zmožni razumeti psihološke fenomene s pomočjo biološke osnove.

Po drugi strani pa se nam zdi upravičenost redukcije vprašljiva, kar je lepo izraženo v petih skupinah ugovorov, ki smo jih predstavili na začetku te razprave. Nam lahko zgodovinski primeri, s katerimi smo se seznanili, kakor koli pomagajo pri oblikovanju odgovorov na te ugovore? Predlagava, da si ogledamo vsako skupino posebej.

Prva skupina ugovorov se ukvarja z možnostjo, kako razložiti subjektivni značaj čutnih takšnosti. Vsi negativni argumenti, ki sodijo v to skupino, imajo neko skupno osnovo, namreč to, da si ne moremo predstavljati, kako bi nam lahko kakršen koli opis delovanja nevronov razložil neizrekljive subjektivne fenomene. Očitno je, da ta ugovor postavlja našo zmožnost oziroma nezmožnost, da si kaj predstavljamo, kot merilo za to, da nekaj je ali ni mogoče. Ta zahteva očitno presega to, kar običajno zahtevamo od poskusa. Namreč, kdo pred Jamesom C. Maxwellon si je lahko predstavljal, da bo teorija nabitih energijskih krogel in nihajočih igel pri kompasu zadostovala za razlago fenomena svetlobe? Kdo pred Descartesom, Bernoullijem in Joulom si je mogel zamisliti, da bo lahko mehanika biljardnih krogel razložila na prvi pogled bistveno drugačen fenomen – toploto? Komu bi se zdelo sprejemljivo, da je višina zvoka pravzaprav frekvenca, ali sploh, da sam zvok sestoji iz niza tlačnih valov v atmosferi?

Moramo se spomniti, da je uspešna interteoretska redukcija navadno kompleksen postopek, ki vključuje sistematično rekonstrukcijo vseh ali večine starih pojmov s pomočjo sredstev novega pojmovanja. Rekonstrukcija navadno ni le zelo zapletena, marveč je pogosto tudi silno presenetljiva. Pri tem ne gre za nekaj, kar bi lahko pričakovali, da se bo porodilo v domišljiji kogar koli v obliki retorične zahteve, kot, recimo, vprašanje: "Le *kako* bi lahko bili *Aji* kaj drugega kot *Bji?*".

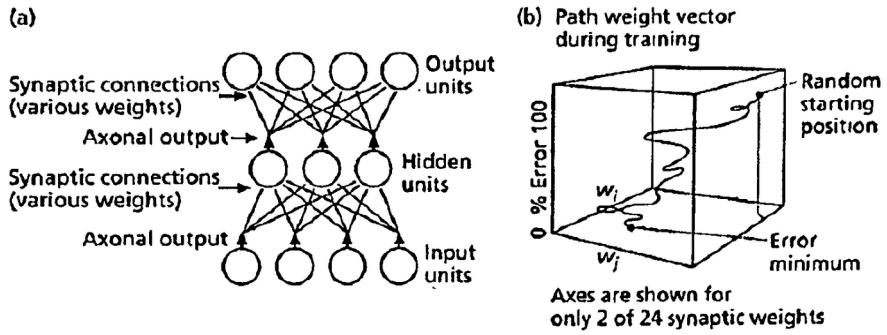
Nadalje, tovrstno retorično vprašanje nas ne more zbegati, če je naša domišljija obveščena o najnovejših teorijah o zaznavnem kodiranju. Ideja, da so vtisi okusa vkodirani kot štiri-dimenzionalni vektorji s koničastimi frekvencami, nam omogoči reprezentacijo prostora vseh človeških okusov, v katerem so vsi sorodni okusi združeni glede na njihove podobnosti, razlike in podobne relacije, kot je npr. vmesnost (Bartoshuk, 1978). Landova retinetalna teorija barvnega vida (Land, 1977) govori o podobni razporeditvi naših barvnih vtisov. Ta teorija vsebuje tudi napoved osnovnih oblik barvne slepote, do katerih pride, kadar je zaradi izgube ene od treh vrst čepkov posameznikov tridimenzionalni barvni prostor reduciran na dve dimenziji.

Pri tem že gre za rekonstrukcijo nekaterih lastnostni pojavov, ki jih preiskujemo, v luči nove teorije. Zdaj je potrebno le še izpeljati rekonstrukcijo do konca, kot je bilo to strjeno v primeru zgodovinskih razlag prej omenjenih objektivnih fenomenalnih lastnosti (toplota, svetloba, višina tona). Nekatere stvari so lahko po svoji naravi res neizrekljivo fenomenalne, saj jih zaznavamo s pomočjo naših osnovnih čutov. Toda zaradi tega niso nič bolj varne pred poučno interteoretsko redukcijo. Zgodovina nas uči ravno o nasprotnem.

Druga skupina ugovorov opozarja na pomen, semantično vsebino oz. intencionalnost naših misli in ostalih mentalnih stanj. Tovrstni antiredukcionistični argumenti so zelo podobni tistim iz prejšnje skupine.

Sklicujejo se predvsem na našo nezmožnost predstavljati si, kako bi lahko pomen bil zgolj posledica interakcije med signali ali procesiranja inertnih simbolov (Searle, 1980, 1990; za zavrnitev Searlovih izpeljevanj glej Churchland in Churchland, 1990. Searle, strogo gledano, nasprotuje le popolnoma računski redukciji; ker pa je ta izrednega pomena za nevroznanost, ga bomo vključili med ostale antiredukcioniste). Pri tovrstnih dokazih gre pravzaprav za argumente iz nevednosti. Vsi imajo sledečo obliko: "Ne morem si *predstavljati*, kako bi deloval nevrokomputacijski proces s smiselnimi reprezentacijami; torej tak proces ni možen." Za kratkoročno spodbijanje tovrstnih izpeljevanj je dovolj, če izpostavimo samo to napako.

Za bolj dolgoročno zavrnitev pa je potrebno nekaj več truda. Potrebno je dejansko predočiti opis, kako možgani reprezentirajo zunanji svet in zakonitosti, ki v njem veljajo. Ravno s tem pa se ukvarjajo moderne teorije o nevronskih mrežah. Te teorije trdijo, da so realnočasovne informacije o svetu kodirane v visokodimenzionalnih aktivacijskih vektorjih, splošne informacije o svetu pa so kodirane v konfiguraciji ozadja mrežnih sinaptičnih uteži. Aktivacijski vektorji so procesirani s konfiguracijami uteži, mimo katerih potujejo, učenje pa ni nič drugega kot prilagoditev posameznikove globalne konfiguracije uteži (glej sliko 22.1) S pomočjo tovrstnih razlag lahko do neke mere že opišemo mnogo različnih stvari, kot so, recimo, prepoznavanje kompleksnih predmetov kljub le pomanjkljivim čutnim podatkom, hiter priklic pomembnih informacij iz obsežnega vsebinsko naslovljivega spomina, zmožnost prepoznavanja majhnih in neubesedljivih razlik in uporaba kompleksne senzomotorične koordinacije (P. M. Churchland, 1989). Še zmeraj vemo znatno premalo, da bi mogli z gotovostjo trditi, da bodo tovrstne hipoteze v zadostni meri razložile reprezentacijske zmožnosti razuma. Ravno tako pa je napačno tudi nasprotno prepričanje, namreč, da bodo take hipoteze nujno neustrezne. Gre za empirično vprašanje in potrebno je počakati na končno sodbo.



Slika 22.1*

Slika 22.1 (a) Preprosta neposredna umetna "nevronska mreža". Vnosi so

* (a) *Synaptic connections (various weights)* – sinaptične povezave (različne uteži); *Axonal outputs* – aksonalni izhodi; *Output units* – izhodne enote; *Hidden units* – skrite enote; *Input units* – vhodne enote. (b) *Path weight vector during training* – vektor poti uteži med treningom; *Random starting position* – naključna začetna točka; *Error minimum* – minimum napake; *Axes are shown for only 2 of 24 synaptic weights* – Prikazani sta osi za le 2 od 24 sinaptičnih uteži. – op. prev.

kodirani kot vzorec ali vektor aktivacijskih vrednosti na vhodnih enotah. Vzorec se prenese na skrite enote, a se spremeni, ko gre preko vmesnih sinaptičnih povezav z različnimi utežmi. Vsaka skrita enota nato sešteje prejete obtežene vnose in na podlagi rezultata privzame ustrezno aktivacijsko vrednost. To privede do novega vzorca ali vektorja aktivacijskih vrednosti na skritih enotah. Zgodba se ponovi na zadnjem sloju izhodnih enot, ki prevzamejo tretji vzorec aktivacij. Mreža tako skrbi za transformacije med vektorji. Za kakšno vrsto transformacije gre, pa je odvisno od specifične konfiguracije sinaptičnih uteži. S pomočjo sigmoidnih izhodnih funkcij na prvih dveh ravneh lahko mreže izračunajo približno vrednost vsakršne izračunljive transformacije. (b) Shematski prikaz abstraktnega "prostora uteži" za preprosto mrežo. Učenje v taki mreži poteka kot zaporedno prilagajanje njene konfiguracije uteži, tako da bi postopno zmanjšali njene napake pri izvedbi. Ta postopek se nadaljuje, dokler mreža končno ne izvede transformacije med vhodnimi in izhodnimi podatki, ki je bila implicitno prisotna v mnogih primerih, na katerih smo to mrežo urili (povzeto po P. M. Churchland, 1998).

Tretja skupina ugovorov poudarja, da pri tem, kar mi razumemo kot človeška zavest, ne gre le za neko intrinzično lastnost določenega bitja, marveč za bogato matrico posameznikovih odnosov do soljudi, običajev in institucij njegove kulture. Ker je redukcionistični pristop k zavesti in obnašanju omejen na mikroskopske dejavnosti znotraj posameznikovih možganov, mu ne bo nikoli uspelo zajeti več kot le droben delež tega, kar je razlagalno pomembnega.

Na tovrstne ugovore najbolje odgovoriva tako, da jih sprejmeva. Človeško obnašanje je nedvomno funkcija prej omenjenih faktorjev. Kultura, v kateri nek posameznik prebiva, bo imela velik vpliv na razvoj posameznikove osebnosti. To pa pomeni, da mora vsaka ustrezna nevrokomputacijska razlaga človeške duševnosti vsebovati tudi opis, kako možgani reprezentirajo ne le grobe lastnosti fizičnega sveta, marveč tudi značaj kognitivnih bitij, s katerimi je posameznik v interakciji, ter značilnosti družbenega, moralnega in političnega življenja, v katerem ta posameznik živi. Možgani družbenih živali se navsezadnje naučijo igrati vlogo interaktivnega elementa znotraj skupnosti možganov, in ta vloga je izredno pomembna kognitivna prednost. Moramo spoznati, kako to počnejo.

To je nedvomno izredno velik izziv, ki pa mu nevroznanstveniki zaenkrat ne posvečajo preveč pozornosti, kaj šele da bi se ga resneje lotili. Vendar to ni nič presenetljivega. Če je izredno težavno opisati že samo to, kako navadna muha določi svoj položaj v prostoru, gre pri opisu sramu ljubljene osebe, značaja kakega politika ali poznavanja skrivnih načrtov našega sovražnika za toliko težjo nalogo. Vemo pa, da umetne nevronske mreže, ki jih izurimo s pomočjo primerov, lahko prepoznavajo neverjetno subtilne vzorce in podobnosti v naravi ter nanje tudi reagirajo. Če to velja za fizične vzorce, čemu ne bi isto veljalo tudi za družbene vzorce? Tu ne gre za kakršno koli principielno težavo, marveč zgolj za izredno velik izziv.

Morebiti se zdi nerealno pričakovati, da nam bo kdaj uspelo podati izčrpen splošni opis nevralne in vedenjske poti nekega posameznika. Povsem možno je,

da zaradi izredne kompleksnosti živčnega sistema ne bomo mogli nikoli priti dlje od zgolj uporabnih približkov za želeni idealni opis. Vzemimo primer kemije in njeno razmerje do kvantne fizike. Tudi tukaj matematika kompleksnih dinamičnih sistemov narekuje določene omejitve glede na to, s kolikšno mero natančnosti lahko rekonstruiramo kemična dejstva s pomočjo fizikalnih principov. To pomeni, da naša redukcija nikoli ne bo popolna, vendar pa smo vseeno upravičeno prepričani, da so kemični pojavi le refleksija kvantnih fizičnih pojavov na makro ravni. Podobno kot za kemične pojave velja tudi za psihološke.

To nas pa pripelje še do četrtega sklopa ugovorov, ki opozarjajo na nevarnost, ki naj bi jo redukcija predstavljala človeški svobodi. Tukaj bova kratka. Ali obstaja kakršna koli svoboda, ki se razlikuje od relativne avtonomnosti, značilne za vse kompleksne dinamične sisteme, ki so delno izolirani od preostalega sveta, je docela empirično vprašanje. Namesto da skušamo pokazati konsistentnost nevroznanosti s tem ali onim vnaprej ustvarjenim pojmovanjem svobode, dovolimo raje znanstvenim raziskavam, da nas *poučijo*, na kakšen način in v kolikšni meri so človeška bitja res "svobodna". To bo nedvomno povzročilo spremembe ali celo popolno zavračanje nekaterih trenutno veljavnih pojmovanj svobode. Toda to bo bistveno bolje, kot pa da postavimo trenutno zmedo za standard, s katerim se bodo bodoče teorije primorane uskladiti.

Peta in zadnja skupina ugovorov trdi, da za psihologijo velja ireducibilni abstraktni status, kar utemeljujejo s tem, da lahko zelo različni fizični sistemi enako dobro realizirajo abstraktno organizacijo, ki tvori kognitivno ekonomijo nekega sistema. Le kako bi lahko reducirali psihološke fenomene na nevrobiologijo, ko pa lahko povsem drugačni fizični substrati služijo enakemu namenu?

S premiso tega sklopa ugovorov se bomo verjetno strinjali v glavnem vsi. Toda iz nje ne sledi antiredukcionistični sklep. To je razvidno, če si pobližje ogledamo enega od primerov iz naše zgodovine znanosti. Temperatura je, kot smo že omenili, enaka povprečni molekularni kinetični energiji. Vendar pa, strogo gledano, to velja zgolj za pline, kjer se lahko molekule prosto gibljejo v balistični maniri. V trdnem agregatnem stanju, kjer delci oscilirajo sem ter tja, pa energija konstantno niha med kinetično in potencialno energijo. V plazmi, segreti do visokih temperatur, sploh ne moremo govoriti o molekulah, saj je vse razstavljeno na subatomske delce. Tukaj je temperatura kompleksna mešanica različnih energij. V vakuumu, kjer sploh ni mase, pa temperatura sestoji iz distribucije valovnih dolžin – "krivulja črnega telesa" – EM valov, ki potujejo skozenj.

Ti primeri lepo kažejo, kako je lahko redukcija vezana na specifično domeno: v primeru plina velja temperatura za eno stvar; v primeru trdne snovi za nekaj drugega; v primeru plazme spet za nekaj tretjega; v primeru vakuuma za nekaj četrtega itd. (Pri vseh gre za "temperaturo", saj so medsebojno povezane in so podvržene enakim zakonom ravnotežja.) Vendar pa zaradi tega ne bomo trdili, da je klasična termodinamika avtonomna, ireducibilna znanost, ki je za vedno izvzeta iz zgodbe spodaj ležeče mikrofizične domene. Ravno nasprotno, zgoraj

navedeno nas uči, da obstaja več kot le en način, na katerega se energije manifestira na mikrofizični ravni.

Podobno je z vizualnim doživetjem – le-to je lahko ena stvar pri sesalcih, rahlo drugačna stvar pri hobotnici in bistveno drugačna stvar pri potencialnem androidu, zgrajenem iz kovine in polprevodnikov. Vendar gre pri vseh za vizualna doživetja, saj so deležni nekih skupnih abstraktnih lastnosti na višji opisni ravni. Tega, da je nevrobiologija zmožna razlož vse psihološke pojave pri ljudeh, tako ne ogroža možnost, da lahko neka *druga* teorija, recimo elektronika polprevodnikov, ponudi razlago za psih loške pojave pri *robotih.* Taki redukciji si ne bi nasprotovali, marveč se medsebojno dopolnjevali,

Drugod sva navedla bolj natančne in izčrpne opise, kako sodobna dejavnost v nevroznanosti pojasnjuje dejstva v psihologiji in kognitivni teoriji (P. S. Churchland, 1988; P. M. Churchland, 1989). Tukaj bova zaključila z dvema svarilnima opazkama. Prvič, čeprav sva bila v tem članku izredno optimistična glede možnosti redukcije psihologije na nevroznanost, se zavedava, da bo dejanska izvedba tovrstne redukcije izredno dolgotrajna in silno zahtevna naloga. V pričujočem članku nama je šlo predvsem za to, da ovrževa zagovornike nemožnosti in da umestiva redukcijske teženje nevroznanosti v ustrezen zgodovinski kontekst.

Drugič, ne smemo domnevati, da bo psihologija kot znanost v postopku redukcije izginila ali da bo njena vloga omejena na pasivno tarčo nevralnih razlag. Nasprotno, kemija ni izginila po kvantno-mehanski razlagi njenih osnov; ravno tako ni izginila biologija, kljub kemijskim eksplikacijam njenih osnov. Še več, tako kemija kot biologija sta bistveno pripomogli k oblikovanju in razvoju spodaj ležečih znanosti. Nedvomno taka prihodnost čaka tudi psihologijo in nevroznanost. Na tej ravni kompleksnosti se interteoretska redukcija ne zdi kot nenaden prevzem ene discipline s strani druge, marveč je bolj podobna dolgemu in počasi zorečemu zakonu.

Literatura

Bartoshuk, L. M. (1978). “Gustatory system”. V Masterton, R. B. (ur.), *Handbook of Behavioral Neurobiology*, 1. zv., *Sensory Integration.* New York: Plenum Press, str. 503-567.

Churchland, P. M. (1989). *A Neurocomputational Perspective: The Nature of Mind and the Structure of Science.* Cambridge, MA: MIT Press.

Churchland, P. M. in Churchland, P. S. (1990). “Could a machine think?”. *Scientific American,* 262, str. 32-37.

Churchland, P. S. (1986). *Neurophilosophy: Toward a Unified Understanding of the Mind/Brain.* Cambridge, MA: MIT Press.

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

Jackson, F. (1982). “Epiphenomenal qualia”. *Philosophical Quaterly,* 32, str.

127-136.

Land, E. (1997). “The retinex theory of color vision”. *Scientific American, 237,*

Nagel, T. (1974). “What is it like to be a bat?”. *Philosophical Review,* 83, str. 435-450.

Popper, K. and Eccles, J. (1978). *The Self and the Brain.* New York: Springer.

Searle, J. (1980). “Minds, brains, and programs”. *Behavioral and Brain Sciences,* 3, str. 417-457.

Searle, J. (1990). “Is the brain’s mind a computer program?”. *Scientific American,* 262, str. 26-31.

Taylor, C. (1970). “Mind-body identity: A side issue?”. V Boršt, C. V. (ur.), *The Mind/Brain Identity Theory*, Toronto: Macmillian, str. 231-241.

Taylor, C. (1987). “Overcoming epistemology”. V Baynes, K., Uohman, J. in McCarthy, T. (ur.), *After Philosophy: End or Transformation?,* Cambridge, MA: MIT Press, str. 464-488.

Sebastjan Vörös
Filozofska fakulteta Univerze v Ljubljani

NOVI VAL REDUKCIONIZMA

V članku obravnavam nekatere temeljne pristope novega redukcionizma v kognitivni znanosti. V prvem delu sestavka na kratko orišem zgodovino redukcionizma in predstavim najpomembnejše teoretične modele interteoretske redukcije (Nagel, Hooker in Bickle), v drugem delu po prikažem Bicklov poskus empirične potrditve interteoretske redukcije na primeru redukcije asociativnega učenja (t.i. "kombinatorična" redukcija). Novi val redukcionizma označim kot zelo obetaven in nadvse zanimiv pristop k razrešitvi vprašanja fizičnega in psihičnega.

Ključne besede: komputacijski funkcionalizem, modeli interteoretske redukcije, novi val redukcionizma, "kombinatorična" redukcija.

1. Uvod

Še do nedavna si je bila velika večina filozofov duha edina v mnenju, da je redukcionizem zastarel in zgrešen način obravnave vprašanja telesa in duha. Takemu gledišču je botroval predvsem vsesplošni sprejem osnovnih predpostavk komputacijskega funkcionalizma, namreč, da so v psiholoških razlagah ključnega pomena mentalne reprezentacije in komputacije nad njihovo vsebino. Ker pa teorij, ki uporabljajo pojme reprezentacije in komputacije, tako zagovorniki funkcionalizma in nereduktivnega fizikalizma, ne gre reducirati na sklop teorij, ki se ukvarjajo zgolj z nevrofiziološkimi stanji in procesi, lahko sklepamo, da redukcija psihologije na nevroznanost ni možna.

V zagovor nereducibilnosti vseh komputacijsko-reprezentacijskih pristopov je bilo navedenih več argumentov. Najbolj znan in najpogosteje navajan je nedvomno *ugovor večnačinovne realizacije*, temu pa sledijo še Davidsonovi argumenti o anomalnosti mentalnega, LePore-Loewerjevi argumenti o normativnosti vsebinskih lastnosti ipd. Bistvo ugovora večnačinovne realizacije, ki ga je razvil Putnam in kasneje dodelal Fodor, je v tem, da je lahko vsako mentalno stanje realizirano v zelo raznovrstnih fizičnih substratih, raznolikost in mnogovrstnost teh substratov pa naj bi onemogočala postavitev primernega fizikalnega opisa taistega stanja. Če povemo isto z besedami LePora in Loewerja:

> "Razmerje med psihološkimi in nevrofiziološkimi lastnostmi je takšno, da slednje realizirajo prve. Še več, vsaka psihološka lastnost je lahko realizirana v velikem, morebiti celo v neskončno mnogem številu različnih fizičnih in celo nefizičnih lastnosti." (navajam po Kim, 1992: 310)

Čeprav se mnogim še dandanes zdi, da zgoraj navedeni ugovor dodobra opravi z redukcionističnimi pristopi znotraj kognitivne znanosti, pa se je v poslednjih letih nabralo kar nekaj literature, v kateri skušajo različni avtorji pokazati bistvene pomanjkljivosti argumenta o nezmožnosti redukcije zaradi večvrstne realizacije in prikazati tako teoretične kot empirične razloge za sprejem redukcionizma. Preden pa se spopademo z opisom enega bolj sodobnih pristopov v nakazani smeri, tj. s "kombinatoričnim" redukcionizmom Johna Bickla, predlagam, da si sprva ogledamo še kratko zgodovino razvoja modelov teoretične redukcije.

2. Kratka predstavitev razvoja različnih modelov interteoretske redukcije

2.1. Nagelov model

U. T. Place in J. J. C. Smart, znana kot začetnika redukcionizma, nista zasnovala kakšnega posebnega modela interteoretske redukcije. To nalogo je prevzel Ernst Nagel, katerega model je kaj kmalu postal splošen vodič za vse pristaše redukcionizma. Nagel je predpostavljal, da redukcija teorije RT (reducirana teorija) na teorijo OT (reducirajoča ali osnovna teorija) temelji na logični izpeljavi zakonov oz. principov RT iz zakonov oz. principov OT. V primerih, kadar opisno besedišče OT nima ustreznih izrazov za pojme RT, je potrebno vpeljati ti. "principe premostitve" (PP). Nagel je opredelil RT, OT in PP sintaktično, in sicer kot niz propozicij. Formalni zapis tovrstnega modela redukcije je sledeč:

OT ∧ PP (po potrebi) → RT

Nagelov pristop je znatno prispeval k zamenjavi do takrat najpogosteje uporabljanega, a v tančico nejasnosti zavitega pojma "ontološke redukcije" z dobro preučenim, znanstveno utemeljenim pojmom interteoretske redukcije. S tem je ponudil stabilen konceptualni okvir vsem zagovornikom redukcije psihologije na nevrofiziologijo, ki je slednjim omogočil sklicevanje na že uspele primere redukcije iz zgodovine znanosti.

Vendar so se kaj kmalu pokazale tudi temeljne pomanjkljivosti Nagelovega modela, in sicer se je za eno največjih slabosti izkazal *problem modus tollensa.* Podrobne zgodovinske raziskave so namreč pokazale, da je v nekaterih primerih redukcij v znanosti RT napačna. Iz zgoraj podanega formalnega zapisa pa je očitno, da mora v Nagelovem modelu v primeru napačnosti RT biti napačna tudi OT. To pa je nesmiselno, saj Einsteinova fizika, ki je reducirala (napačno) Newtonovo fiziko, ni nepravilna zaradi nepravilnosti slednje.

Čeprav je Nagelu uspelo razložiti manjša odstopanja v veljavnosti OT in RT z vpeljavo dodatnih premis, ki naj bi poskrbele za potencialno neresnične t.i. mejne pogoje (MP) in omejujoče predpostavke (OP), pa njegov model v celoti odpove pri razlagi redukcij, v katerih nastopa popolnoma napačna RT. Primeri, kot so npr. flogiston, klasično (newtonsko) pojmovana masa ipd., niso udejanjeni nikjer v fizični realnosti, zato je nemogoče določiti PP, MC ali RT, ki bi bili zmožni povezati Lavoisierovo kisikovo teorijo izgorevanja in Einsteinovo teorijo

relativnosti s prej omenjenimi entitetami.

2.2. Hookerjev model

Clifford Hooker je nadaljevalec Nagelove misli, ki pa je slednjo odel v nekoliko drugačno opravo. Hooker se strinja z Nagelom v tem, da mora interteoretska redukcija temeljiti na dedukciji, v kateri kot premise nastopajo OT in potencialno neresnični/e MP/OP. Bistvena razlika med modeloma pa je ta, da pri Hookerju sklep izpeljave ni RT, marveč ekvipotentna izomorfna slika RT (EIS), ki je določena znotraj konceptualnega okvirja in besedišča OT. Posplošitve, ki sestavljajo EIS, se skladajo s sintaktično strukturo posplošitev znotraj RT in vsebujejo razlage za domeno uporabe RT. Formalna struktura Hookerjevega modela je sledeča:

OT ∧ MP/OP (po potrebi) → EIS, ki je relevantno izomorfna (analogna) RT[1]

Hookerjev model ustvari specifičen spekter interteoretske redukcije, ki se razteza od "razmeroma gladkih" do "ekstremno zavitih/zapletenih" interteoretskih redukcij. Primeri, v katerih OT razmeroma gladko reducira RT, padejo na "razmeroma gladki" pol (redukcija svetlobe na elektromagnetno valovanje), medtem ko primeri, v katerih je RT napačna, padejo na "ekstremno zaviti" pol (redukcija oz. eliminacija flogistonske teorije na Lavoisierovo kisikovo teorijo izgorevanja). "Mešani" primeri, tj. primeri, ki vsebujejo nekaj lastnosti iz obeh ekstremov, pa najdejo svoje mesto na sredini spektra (redukcija Newtonove mehanike na Einsteinovo posebno teorijo relativnosti).

Vendar pa se zgodba tukaj ne konča, kajti ne samo, da se zgodovinski primeri razvrstijo po pravkar opisanem spektru, ampak se podobno zgodi tudi z ontološkimi posledicami, do katerih pride znotraj posamičnih primerov. Te posledice segajo od identitete lastnosti in entitet, preko znatne konceptualne revizije, do popolne eliminacije. Lokacija nekega primera iz zgodovine znanosti na spektru interteoretske redukcije je v zelo ozki vzajemni povezavi z njegovo lokacijo na spektru ontoloških posledic. Na ta način lahko predvidimo ontološke posledice neke redukcije na podlagi lociranja preiskovanega primera na spektru interteoretske redukcije.

Vendar pa so kritiki tudi pri Hookerjevem modelu našli določene slabosti. Prva pomembnejša hiba je ta, da Hooker nikoli ne ponudi jasno razvidnih in s kvantitativnimi podrobnostmi podprtih primerov za "analogno strukturo" EIS. Druga, še bolj problematična pomanjkljivost pa se skriva v vse prej kot zadovoljivi formulaciji osrednjega koncepta, tj. "analognega razmerja" med EIS in RT. Ker na slednjem temelji nivo korekcije za neko RT in posledično celoten spekter interteoretske redukcije, je razumljiva Hookerjeva zadrega, ko prizna, da ni zmožen ponuditi zbir formalnih pogojev za razlikovanje med EIS in RT.

2.3. Bicklov model

[1] Očitno je, da tak model ne potrebuje PP, saj je EIS že določena znotraj OT. Premise in sklep namreč ne vsebujejo nekih nesoizmerljivih besednjakov, ki bi jih bilo potrebno premostiti s posebnimi principi.

John Bickle je v osnovi prevzel temeljno podobo Hookerjevega modela, vendar pa je, da bi se ognil težavam, na katere je naletel sam Hooker, za svoj model uporabil značilno drugačno formalno strukturo interteoretske redukcije. Tako Bickle, zgledujoč se po "strukturalističnih" programih v filozofiji znanosti, ne pojmuje strukturo teorije kot množico propozicij, marveč kot množico modelov, ki ustrezajo specifičnim pogojem teorije množic. Namesto da bi opredelil interteoretsko redukcijo v smislu sintaktičnih (stavčnih) izpeljav, Bicklov pristop opredeli razmerje med teorijama kot potencialno izgradljivost slike (EIS) RT znotraj OT. Ker bi natančna predstavitev Bicklovega modela v znatni meri presegla okvir pričujočega sestavka, se v podrobnejšo analizo ne bom spuščal, vedoželjnemu bralcu pa v branje predlagam na koncu članka navedeno literaturo, zlasti Bickla (1998: 57-103) in Uleta (1992).

Bicklov model tako bolj precizno opredeli razmerje med EIS in RT ter ponudi jasen in kvantitativno podkrepljen primer EIS na primeru redukcije splošne plinske enačbe na statistično mehaniko in mikrofiziko. Bickle je, kot eden vodilnih zagovornikov novega vala redukcionizma, prepričan, da bo tekom let pogoj "izgradljivosti slike" uresničen tudi v primeru psihologije in nevroznanosti. Njegova sugestija naj bi temeljila na domnevnem dejstvu, da tovrstna povezava v sodobnih kognitivno psiholoških in nevroznanstvenih teorijah *že obstaja.* V nadaljevanju bom predstavil primer, za katerega je Bickle prepričan, da potrjuje domnevo o reducibilnosti psihološkega na fizično.

3. Bicklov primer redukcije asociativnega učenja

Bickle se loti prikaza neresničnosti trditve, da komputacijsko-reprezentacijskih teorij mentalnega ni mogoče reducirati na nevrofiziologijo, s sklicevanjem na najnovejše izsledke empiričnih znanosti. Odloči se namreč, da bo namesto zatekanja k analizi in spodbijanju argumentov, na katerih je zgoraj navedena trditev utemeljena, skušal predstaviti najnovejša dognanja na področju asociativnega učenja, ki naj bi pričala ne le o tem, da je redukcija psihičnega na fizično možna, temveč tudi, da je slednja znotraj znanstvenih krogov že udejanjeno in splošno sprejeto dejstvo. Shematizirana oblika njegovega argumenta je sledeča:

(i) Sodobne znanstvene teorije, ki preučujejo asociativno učenje, prepoznavajo potrebo po vpeljavi osnovnih kognitivnih pojmov (reprezentacije in komputacije nad njihovimi vsebinami) za razlago celo najbolj osnovnih oblik asociativnega učenja (Pavlovo pogojevanje).

(ii) Trenutno najbolj ustrezna teorija asociativnega učenja, ki se poslužuje kognitivnih virov, je že bila uspešno reducirana na povsem nevrofiziološki opis vseh vedenjskih podatkov.

(∴) Torej je redukcija teorije s področja kognitivne psihologije že uveljavljeno znanstveno dejstvo.

Kot dokaz resničnosti prve premise navaja Bickle dognanja nekaterih novejših raziskav s področja asociativnega učenja, še posebej se sklicuje na delo Leona

Kamina, ki je bilo kasneje nadgrajeno in dopolnjeno s strani Roberta Rescorla in Alana Wagnerja. Osnovni viri za podporo druge premise pa so mu predvsem rezultati raziskav nevrobiologov Richarda Hawkinsa in Erica Kandela.

Kamin je bil eden izmed prvih kognitivnih psihologov, ki je preko preučevanja "zaviralnega učinka" *[blocking effect]* začutil potrebo po povezovanju klasičnih behavioristističnih modelov asociativnega učenja s kognitivističnim pojmovnim besednjakom. Pa si na kratko oglejmo, kako so potekale Kaminove raziskave. V prvi fazi poskusa so neko žival (podgano) z uporabo averzivnega brezpogojnega dražljaja (BD) (blagi elektrošok) naučili, da se je odzivala na pogojni dražljaj (PD_1) (blisk svetilke), in sicer do te mere, da je sama prisotnost PD_1 v celoti zavrla nek prej naučeni odgovor (pritiskanje droga za hrano). V drugi fazi poskusa so uporabili BD (blagi elektrošok) v povezavi z mešanim PD, ki je bil sestavljen iz PD_1 iz prve faze (blisk svetilke) in novega PD (zvok; PD_2). Po nekajkrat ponovljenem predvajanju kombinacije mešanega PD in BD so eksperimentalni živali predvajali le novo komponento, tj. PD_2. Kaminovi rezultati so pokazali, da se živali na PD_2 ne odzivajo v niti približno tolikšni meri kot na PD_1 v prvi fazi (tj. blokiranje prej naučenega odgovora je znatno manjše kot v prvi fazi; gl. tabelo 1). Od tod izvira tudi ime tega efekta: predhodna priučitev PD_1 *zavira* učenje o vsaki novi PD_2-BP povezavi znotraj kombinacije mešanega PD in BD.

	faza 1	faza 2	predvajanje →	rezultat
eksperimentalna skupina	PD_1 – BD	sestavljeni PD (PD_1 in PD_2) – BD	PD_2	slabo zaviranj
kontrolna skupina	(nič)	sestavljeni PD (PD_1 in PD_2) – BD	PD_2	znatno zaviranje

Tabela 1 (povzeto po Bickle, 1995: 270)

Dobljene podatke je bilo mogoče razložiti na vsaj dva različna načina: prvič, predhodno pogojevanje živalim onemogoči *opaziti, zaznati, prepoznati* PD_2 med drugo eksperimentalno fazo; drugič, na PD_1 predhodno priučene živali PD_2 sicer zaznajo, vendar pa se nanj ne odzovejo, ker je *odvečen:* umanjkala naj bi nujna komponenta pogojevanja, tj. "presenečenje", saj PD_1 povsem zadovoljivo napove nastop averzivnega BD. Čeprav je bil Kamin – kot resen behavioristični znanstvenik iz šestdesetih let – trdno prepričan v ustreznost prve interpretacije, kajti ta se je zdela kompatibilna z osnovno behavioristično terminologijo, pa so ga dodatne raziskave primorale v to, da je priznal pravilnost *druge,* zelo kognitivno navdahnjene razlage.

Kaminovi objavljeni zaključki so več kot zgovoren primer potrebe po privzemu kognitivnih konceptov in izrazoslovja:

> "Mogoče pa je nujno, da za dosego povečane asociativne moči BD spodbudi pri živali določeno *mentalno delovanje.* To mentalno delovanje pa se pojavi zgolj v primeru, ko je BD *nepričakovan,* in tako na nek način

prestraši žival." (navajam po Bickle, 1995: 272)

Vendar pa Kamin tovrstne razlage ne omeji zgolj na asociativne fenomene "višjega reda" (kot je npr. zgoraj omenjeni zaviralni efekt), marveč jo uporabi za splošni opis asociativnega učenja. Za Pavlovo pogojevanje tako pravi:

> "V zgodnjih fazah normalnega eksperimenta pogojevanja je BD *nepričakovan*, a motivacijsko pomemben dogodek, zato pride do vzpostavitve povezave med BD in PD." (navajam po Bickle, 1995: 273)

Robert Rescorla in Alan Wagner sta v svoji teoriji Pavlovega pogojevanja Kaminove pojme "presenečenja", "pričakovanja" itd. vgradila v matematične formule in jim s tem podelila večjo preglednost in jasnost. Intuicije, ki sta jih skušala kvantificirati, sta razložila na sledeči način:

> "Organizmi se česa naučijo samo v primeru, kadar dogodki *niso v skladu* z *njihovimi pričakovanji.* Nekatera pričakovanja so zgrajena z ozirom na dogodke, ki sledijo kompleksu dražljajev: *pričakovanja*, ki jih sproži ta kompleks in njegovi sestavni dražljaji, so potem modificirana samo, kadar nadaljnji dogodki niso v skladu s *sestavljenim pričakovanjem."* (navajam po Bickle, 1995: 273)

Bickle poudarja, da velja Rescorla-Wagnerjeva teorija za trenutno najbolj ustrezno in splošno sprejeto teorijo asociativnega učenja in hkrati opozarja, da je slednja več kot očitno prežeta s klasičnimi kognitivističnimi koncepti, kar naj bi potrjevalo prvo premiso njegovega argumenta.

V iskanju podpore za drugo premiso se Bickle obrača k delu Richarda Hawkinsa in Erica Kandela, ki sta preučevala refleks krčenja škrg in sifona pri morskem polžu *Aplysia californica* in podala osnovno nevrofiziološko sliko asociativnega učenja. V svojem eksperimentalnem delu sta se Hawkins in Kandel osredotočila predvsem na spremembe v predsinaptičnem nevronu, ki povzročijo spremembe v količini nevrotransmiterja, sproščenega v medsinaptični prostor in posledično v povečanem ali zmanjšanem odzivu posinaptičnih nevronov. Refleksivno pogojevanje temelji na utrjevanju specifične sinaptične moči (le-ta je določena s količino nevrotransmiterja, ki ga izločijo predsinaptične celice) vseh parov nevronskih celic, ki se nahajajo med čutnimi (senzornimi) nevroni, ki jih vzdraži BD, in gibalnimi (motornimi) nevroni, ki povzročijo refleks krčenja škrg in sifona. Najbolj pomembna v tem procesu je aktivnost tistih internevronov, ležečih med čutno in gibalno periferijo, ki jih vzdraži prisotnost BD in so preko medsebojnih sinaps vzvratno povezani s čutnimi nevroni, ki jih vzdraži PD. Aktivnost teh internevronov povzroči širjenje aktivacijskih potencialov v čutnih nevronih, vzdraženih s strani PD, kar ima za posledico porast števila kalcijevih ionov znotraj čutne celice. Povečana koncentracija intracelularnih kalcijevih ionov sproži številne encimsko vodene procese, katerih končni rezultat je povečana količina nevrotransmiterja, ki se sprosti ob prisotnosti PD in tako tudi povečan posinaptični odgovor gibalnih nevronov.

Hawkins in Kandel trdita, da lahko kognitivne fenomene višjega reda pri asociativnem učenju (npr. zaviralni efekt) razložimo nevrofiziološko kot

zaporedje in kombinacijo pravkar opisanih intracelularnih procesov, pri čemer pa je potrebno upoštevati še dva dodatna pogoja. Prvi (anatomski) pogoj je ta, da so čutni nevroni povezani z internevroni, katerih dejavnost je bistvenega pomena za vzpostavitev pogojnega odziva, drugi (fiziološki) pogoj pa je ta, da se morajo izhodni dražljaji *[output]* internevronov zmanjšati v primeru trajne (nenehne) stimulacije. Ker pa sta za vse višje organizme značilna tako gosta medsebojna prepletenost nevronskih slojev kot celični mehanizmi akomodacije in povratne inhibicije, sta za te organizme oba pogoja izpolnjena.

Zanimivo je, da lahko vse vedenjske podatke, ki jih zajema Rescorla-Wagnerjeva teorija, načeloma razložimo s pomočjo Hawkins-Kandelovega modela. Vedenjska dinamika, ki jo napoveduje matematična izpeljava Rescorla-Kaminovega modela, je zajeta v moči mednevronskih povezav in vplivu slednjih na pogostnost proženja gibalnih nevronov in posledično na samo obnašanje. Bickle zaključi, da pravkar povedano nudi lep primer udejanjene konstrukcije ekvipotentne izomorfne slike RT (kognitivna psihologija) znotraj OT (nevrofiziologija), ki tako izpričuje verodostojnost Bicklovega modela interteoretske redukcije kot tudi neveljavnost antiredukcionističnih argumentov.

Na morebitne ugovore, da druga premisa njegovega argumenta ni zadovoljivo dokazana zaradi neustreznosti izbranega anatomskega in fiziološkega modela (morski polž), Bickle ponudi več odgovorov. Prvič, treba se je zavedati, da morski polž služi res samo kot nevrobiološki *model* za Rescorla-Kaminovo teorijo. Zaradi preproste anatomske in fiziološke zgradbe so morski polži zelo prikladni za raziskovanje osnovnih znotrajceličnih procesov, tj. "črk" celične "abecede". Kombinacije in zaporedja teh osnovnih oblik celične plastičnosti in procesov, ki jih najdemo tudi v preprostejših organizmih, lahko uporabimo za razlago višjih oblik učenja in mišljenja. V podporo tovrstni "kombinatorični" redukciji Bickle navaja tudi evolucijo, namreč znano je, da višje razviti organizmi gradijo na strukturah, ki so že prisotne v nižjih organizacijah, zato se zdi izgradnja višjih ravni učenja/mišljenja na osnovi elementarnih celičnih struktur plastičnosti in dinamičnosti toliko bolj sprejemljiva.

Drugič, višje ravni asociativnega učenja so za človeka izrednega pomena. Bickle opozarja, da so mehanizmi višjih ravni asociativnega učenja bistveni pri zaznavi medsebojne povezanosti preteklih, sedanjih in prihodnjih dogodkov, ta naloga pa je nujna komponenta v za človeka zelo pomembni zmožnosti, namreč zmožnosti zaznavanja vzročnih razmerij. Tako lahko osnovne strukture, ki jih preučujemo pri morskemu polžu pripomorejo k razumevanju znatno kompleksnejših in za nas izredno pomembnih procesov, kar zopet potrjuje izjemno uporabnost "kombinatorične" redukcije.

Tretjič, navedeni primer "kombinatorične" redukcije nikakor ni edini primer tovrstne redukcije. Rezultati najnovejših preučevanj možganske skorje in hipokampusa, spomina in učenja pri višjih živalskih vrstah pričajo o številnih znotrajceličnih procesih, ki se zdijo izjemno obetavni kandidati za nevrofiziološko osnovo višjih kognitivnih funkcij (učenja in spomina). Najbolj znan in tudi najbolj raziskan med temi procesi je nedvomno proces *dolgoročne*

potencialnosti [LTP, *long-term potentiation]*, pri katerem ravno tako pride do utrjevanja medsebojnih povezav med posamičnimi nevroni. Odlike, ki ga delajo še posebej prikladne za nevrofiziološko osnovo spomina, so *selektivnost* (do utrjevanja povezav pride samo v tistih sinapsah, pri katerih sta tako predkot posinaptični nevron ustrezno vzdražena), *hitrost* (do sprememb v ustreznih sinapsah pride v manj kot minuti, stabilizirajo pa se v nekaj minutah), *trajnost* (sinaptične spremembe, ki jih sproži LTP, lahko vzdržijo več tednov) in *asociativnost* (potrebna je skorajda istočasna aktivnost določenega minimalnega števila vhodnih enot, tj. predsinaptičnih nevronov). Zaradi pravkar naštetih lastnosti in dejstva, da so biokemijski mehanizmi znotraj LTP-ja že zelo dobro raziskani, je Bickle prepričan, da nam nudi LTP tisto osnovo celične plastičnosti, ki lahko ravno tako pripomore pri "kombinatorični" razlagi višjih procesov/mehanizmov kognicije[2].

4. Zaključek

Pričujoči članek je želel bralca seznaniti z nekaterimi modernejšimi redukcionističnimi pristopi znotraj kognitivne znanosti in ga opozoriti na dejstvo, da redukcionizem nikakor ni preživeta ali zastarela smer v reševanju problema psihičnega in fizičnega. V prvem delu članka sem na kratko predstavil razvoj različnih interteoretskih modelov, v drugem pa sem bralca seznanil s "kombinatorično" redukcijo asociativnega učenja na nevrofiziologijo Johna Bickla. Z omenjeno redukcijo skuša Bickle bralca prepričati, da je redukcija (vsaj dela) psihologije na nevroznanost že uveljavljeno dejstvo in da so tako "intuicije, nastale kot rezultat premišljevanja v naslonjaču, znova podlegle dejanskim napredkom v znanosti". Ena od odlik Bicklovega pristopa je nedvomno ta, da skuša združiti funkcionalizem z redukcionističnimi pristopi, tj. pokazati neprotislovnost med obema pristopoma. Redukcionizem naj bi bil tako zgolj razširitev funkcionalizma in nikakor ne njegov nasprotnik. Kompatibilnost obeh pristopov pa Bickle ne skuša prikazati le kot konceptualno možnost, temveč kot splošno sprejeto in potrjeno znanstveno dejstvo. V marsikaterem od svojih esejev Bickle poudarja, da redukcija ne pomeni eliminacije in da psihologija z redukcijo na nevroznanost ne bo izgubila svoje avtonomije, marveč bo kvečjemu svojo avtonomnost utrdila, saj bo zasnovana na jasni in trdni znanstveni podlagi (podobno kot dvojica kemija – kvantna fizika). Bickle v svojih izvajanjih nekoliko spominja na Alana Turinga (seveda gre samo za formalno podobnost), kajti tako kot slednji ponuja način reševanja problema, ki bi se izognil prekomernemu teoretiziranju in pustil spregovoriti empiriji. Jasno je, da Bicklu s svojim pristopom še ni uspelo odgovoriti na marsikatero pereče vprašanje znotraj filozofije kognitivnih znanosti, ne gre mu pa odrekati drznosti in inovativnosti, odlik torej, ki se lahko izkažeta za več kot potrebni pri reševanju nekaterih teoretičnih vozlov, na katere je naletela moderna filozofija duha.

NEW WAVE REDUCTIONISM

[2] O redukciji dolgoročnega spomina na LTP glej posebej Bickle (2002) in Bickle (2003).

The present article deals with some of the basic approaches to the new wave reductionism in the cognitive sciences. In the first part of the article, a short history of reductionism is introduced and some of the most important theoretical models of intertheoretic reduction (Nagel, Hooker and Bickle) are presented. The second part deals with Bickle's attempt to empirically ground intertheoretic reduction on the example of the reduction of the associative learning (the so-called "combinatorical" reduction). The new wave reductionism is characterized as an extremely interesting and promising approach to the mind-body problem.

Keywords: computational functionalism, models of intertheoretic reduction, new wave reductionism, "combinatorical" reduction.

Literatura

Bickle, J. "Concepts of Intertheoretic Reduction". http://host.uniroma3.it/progetti/kant/field/cir.htm.

Bickle, J. (1995). "Psychoneural reduction of the genuinely cognitive: some accomplished facts". *Philosophical Psychology*, 3, str. 265-285.

Bickle, J. (1998). *Psychoneural Reduction.* Cambridge, Ma.: MIT Press.

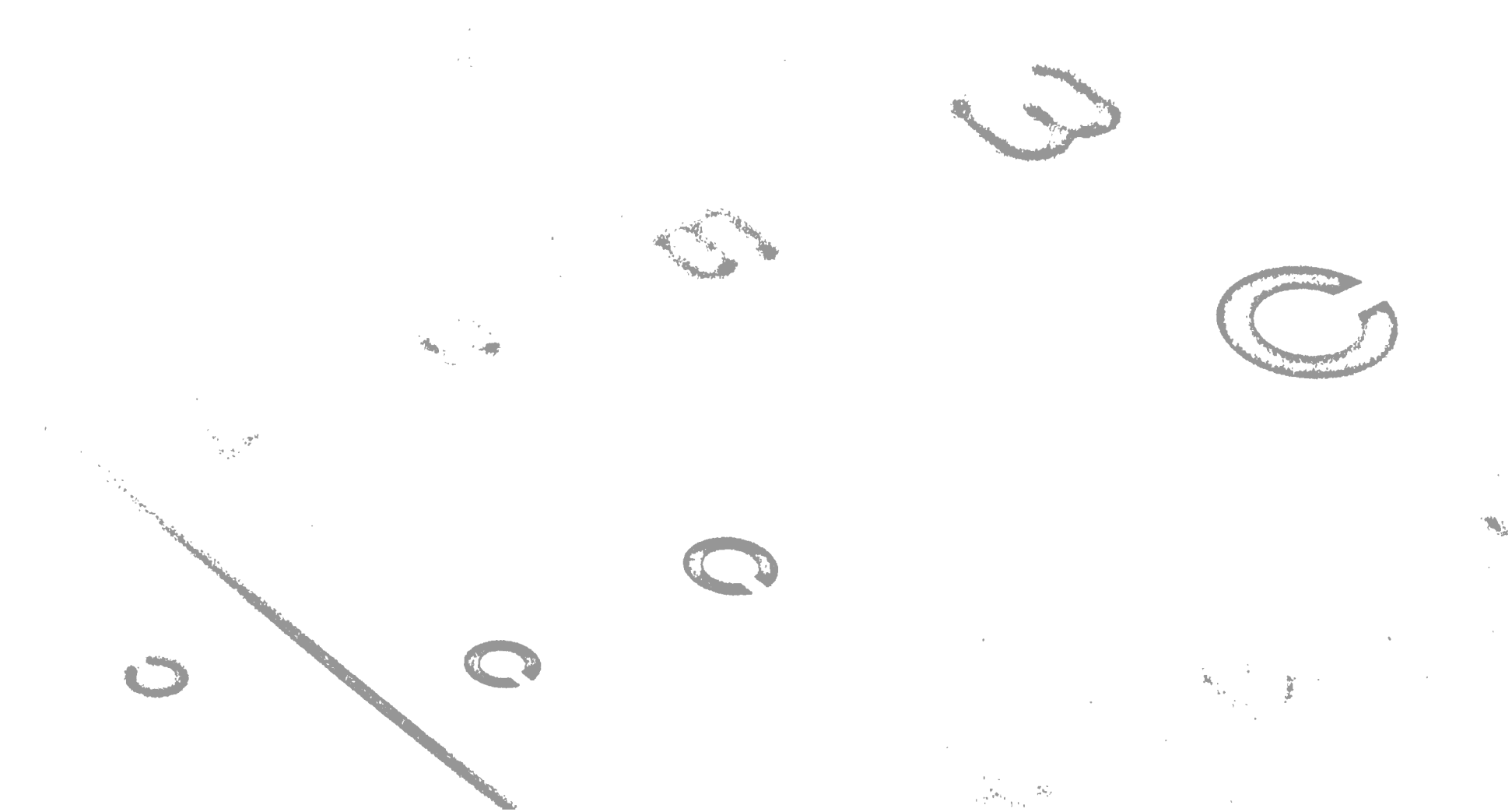
Bickle, J. (2002). "Concepts Structured Through Reduction: A Structuralist Resource Illuminates the Consolidation-Long-Term Potentiation Link". *Synthese*, 130.

Bickle, J. (2003). "Philosophy of Mind and the Neurosciences". V Stich, S. P., Warfield, T. A. (ur.), *The Blackwell Guide to Philosophy of Mind*, Oxford: Blackwell Publishing, str. 322-351.

Kim, J. (1992). "Multiple Realization and the Metaphysics of Reduction". *Philosophy and Phenomenological Research*, 52 (ponatisnjeno v *Supervenience and Mind* (1993), Cambridge: Cambridge University Press, str. 309- 335).

Markič, O. (1997). "Klasična kognitivna znanost in simbolni modeli". *Analiza*, 1, str. 38-52.

Ule, A. (1992). *Sodobne teorije znanosti*. Ljubljana: Znanstveno in publicistično središče.



Vrlinska epistemologija

Linda Zagzebski
prevedel Vojko Strahovnik

ISKANJE IZVORA SPOZNAVNEGA DOBRA*

Običajno filozofi védenje obravnavajo kot visoko vredno spoznavno stanje, morda celo kot eno izmed velikih dober v življenju. Védenje naj bi imelo vsaj večjo vrednost kot resnično prepričanje. Toda sodobni pogledi na naravo védenja nam otežujejo razumevanje njegove vrednosti, kar je v zgodovini filozofije pritegnilo tolikšno pozornost. Nekatere izmed najbolj razširjenih teorij ne morejo pojasniti niti tega, zakaj ima védenje večjo vrednost kot resnično prepričanje. Iskanje izvora spoznavne vrednosti nam lahko razkrije nekatere napetosti in omejitve v načinu opredelitve védenja. Hkrati menim, da bo razkrilo tudi iluzoričnost običajnega stališča o neodvisnosti spoznavnega dobra od moralnega dobra.

1. Kaj naredi védenje boljše od resničnega prepričanja?

Skoraj vedno je kot dognano privzeto to, da je védenje dobro, boljše ko resnično prepričanje *simpliciter*, a izredno težko je pojasniti, kaj je tisto kar ga naredi boljšega. Temu pravim "problem vrednosti".[1] V preteklost sem zagovarjala, da se večina oblik reliabilizma še posebej težko spoprime s tem problemom.[2] Glede na standardne reliabilistične modele *je* védenje resnično upravičeno prepričanje, ki je rezultat zanesljivih kognitivnih *[belief-forming]* procesov in sposobnosti. Toda zanesljivost izvora prepričanja ne more pojasniti razlike v vrednosti med védenjem in resničnim prepričanjem. Eden izmed razlogov je v tem, da zanesljivost *per se* nima vrednosti ali nevrednosti *[disvalue]*. Zanesljiv aparat za espresso je dober, ker je espresso dober. Zanesljivo puščajoča pipa ni dobra, ker puščanje vode ni dobro. Dobrost produkta naredi zanesljivost izvor tega produkta dobro, a zanesljivost izvora ne prenese na produkt nobene dodatne vrednosti. Tekočina v tej skodelici ni nič boljša zaradi dejstva, da prihaja iz zanesljivega aparata za espresso. Če je espresso okusen, je vseeno, če prihaja iz nezanesljivega aparata. Če je rožni vrt lep, je vseeno, če ga je zasadil nezanesljiv vrtnar. Če nas knjiga prevzame, je vseeno, če jo je napisal nezanesljiv avtor. Če je prepričanje resnično, je vseeno, če izhaja iz nezanesljivega kognitivnega izvira.

---

* Prevedeno po Linda Zagzebski, "The Search for the Source of Epistemic Good", *Metaphilosophy*, (2003) let. 34, štev. 1-2, str. 12-28. – op. prev.

[1] Kot izjemo skoraj občemu naziranju o večji vrednosti védenja kot resničnega prepričanja, glej Sartwell 1992. Tak korak zgolj premesti problem k problemu pojasnitve vrednosti resničnega prepričanja, kar bo obravnavano v naslednjem razdelku.

[2] O problemu vrednosti na kratko spregovorim v Zagzebski 1996 in ga podrobneje obravnavam v Zagzebski 2000. Še eno različico problema vrednosti predlaga DePaul 2001.

Slednje velja za kateri koli izvor prepričanja, pa naj gre za proces, sposobnost, vrlino ali veščino – za vsak vzrok prepričanja, katerega vrednost naj bi se prenesla na prepričanje kot na njegov učinek in katerega vrednost naj bi se prenesla zaradi zanesljivosti. Če je védenje resnično prepričanje, ki izvira iz uporabljanja dobrih značilnosti in veščin, potem zanesljivost teh značilnosti in veščin ne more biti tisto, kar mu doda vrednost. Te značilnosti in veščine morajo biti dobre iz nekega razloga, ki svoje moči ne črpa popolnoma iz dobrega svojega produkta, tj. resničnega prepričanja. Ko se je reliabilizem izpopolnil, se je umestitev zanesljivosti premaknila od procesov k sposobnostim in k spoznavnim akterjem.[3] Ta razvoj vsekakor ima svoje prednosti, toda če je značilnost, ki naredi kognitivne procese in sposobnosti dobre, zgolj v zanesljivosti le teh, potem te različice reliabilizma delijo skupno težavo; biti produkt zanesljive sposobnosti ali akterja ne doda vrednosti temu produktu.[4] Če torej védenje izvira iz nečesa, kot je intelektualna vrlina *[intellectual virtue]* ali iz intelektualno vrlih dejanj, to, kar naredi intelektualno veščino dobro, preprosto ne more biti to, da zanesljivo vodi do resničnega prepričanja. To je tudi prvi nauk problema vrednosti: *Resničnost plus zanesljivi izvor resničnosti ne moreta pojasniti vrednosti védenja.*

Iz tega sledi, da mora biti v vzroku resničnega prepričanja vrednost, ki je neodvisna od zanesljivosti ali sposobnosti prenašanja resničnosti [truth *conductiveness*], pa naj jo imenujemo vrlina ali kako drugače. Predpostavimo, da nam to vrednost uspe najti. Ali bo to dovolj za razrešitev problema vrednosti? Na žalost ne, vsaj dokler menimo, da je védenje nek zunanji produkt dobrega vzroka. Skodelica espressa ni boljša zaradi tega, ker bi bila iz aparata, ki bi imel neko vrednost, tudi če bi ta vrednost bila neodvisna od vrednosti okusnega espressa. Analogija z espressom ne pokaže le-tega, da zanesljiv vzrok ne prenese svoje vrednosti na svoj učinek, ampak tudi to, da obstaja splošnejši problem pri pripisovanju vrednosti učinkom zaradi vrednosti njihovih vzrokov, tudi če je vrednost vzrokov neodvisna od vrednosti učinkov. Nočem trditi, da vzrok nikoli ne prenese vrednosti na svoj učinek. Včasih sta vzrok in učinek notranje povezana, kot npr. motiv in dejanje, kar bom obravnavala v nadaljevanju. Kar želim povedati, je to, da se vrednost vzroka ne prenese na učinek samodejno, vsekakor ne pri modelu učinka kot nečesa, kar je z vzrokom proizvedeno. Torej, tudi če ima vzrok resničnega prepričanja neodvisno vrednost, nam to še vedno ne da odgovora na vprašanje, kaj naredi védenje boljše od resničnega prepričanja, če je védenje kot resnično prepričanje dobro na način različen od njegove resničnosti. Drugi nauk problema vrednosti je tako: *Resničnost plus neodvisno vredni izvor (resničnosti) ne moreta pojasniti vrednosti védenja.*

Iz tega drugega nauka sledi, da za rešitev problema vrednosti ni dovolj zgolj najti

[3] Zgodnjo teorijo Sose tako imenujem sposobnostni reliabilizem *[faculty reliabilism]*. Greco svojo teorijo imenuje akterski reliabilizem *[agent reliabilism]*. V Greco 1999 uporablja to ime za razred teorij tudi onstran svoje, ki vključuje Sosovo, Plantingino in tudi mojo zgodnjo teorijo.

[4] Po drugi strani pa imajo reliabilisti v mislih posamezne sposobnosti in lastnosti akterjev, ki jih imenujejo vrline, npr. dober spomin, oster vid in dobro razvite sposobnosti mišljenja. Dobrota teh vrlin ni omejena na njihovo zanesljivost in dokler imamo v mislih to, potem imamo tudi rešitev problema vrednosti. Toda iz istega razloga je te teorije sporno poimenovati kot oblike reliabilizma.

neke druge vrednosti pri analizi védenja, ampak jo je potrebno najti na pravem mestu. Poglejmo teorijo garantov oz. zagotovil kot pravih funkcij [spoznavnih mehanizmov, op. prev.] *[theory of warrant as proper function]* Alvina Plantinge. Pravilno delujoč aparat, nič bolj kakor zanesljiv, ne prenese svoje vrednosti na svoj produkt. Težava ni v tem, da bi pravilna funkcija ne bila dobra, ampak v tem, da ni tisto vredno v samem stanju védenja. Prva dva nauka problema vrednosti tako razkrivata še globlji problem. Če vztrajamo pri aparat – produkt[*] sliki prepričanj, ki je v spoznavni teoriji tako pogosta, potem ne bomo mogli pojasniti, kaj naredi védenje vrednejše od resničnega prepričanja.[5] Védenja ne moremo istovetiti s stanjem resničnega prepričanja, ki je rezultat vrednega vzroka, ne glede na to, ali ta vrednost je ali ni neodvisna od vrednosti resničnega prepričanja.[6]

V nekem mojem drugem članku sem predlagala, da v stanju védenja akter dospe do resnice zaradi vrlin svoje kognitivne dejavnosti.[7] Odgovor Waynea Riggsa in Johna Greca na problem vrednosti je, da je dodatna vrednost védenja glede na resnično prepričanje v stanju stvari, kjer spoznavnemu akterju lahko pripišemo zasluge za resničnost njegovega prepričanja.[8] Tudi odgovor Ernesta Sose je podoben, pravi namreč, da je v stanju védenja resničnost pripisana akterjevi lastni dejavnosti.[9] Podobnost teh odgovorov je očitna, toda problem vrednosti razrešijo le, če opustimo aparat – produkt kot model védenja.[10] Iz istega razloga kot skodelica espressa ni boljša zaradi dejstva, da je bila narejena z zanesljivim ali pravilno delujočim aparatom za espresso, ni nič boljša tudi zaradi tega, če lahko aparatu pripišemo zasluge za njen nastanek. To se pravi, kava v skodelici ni nič okusnejša.

Zaključimo lahko, da resnično prepričanje, ki izhaja iz spoznavnih dejavnosti, ne more biti kot espresso iz aparata. Ne samo, da je zanesljivost aparata nezadostna za izboljšanje okusa kave; nič v zvezi z aparatom ne more izboljšati produkta. Torej, če je védenje resnično prepričanje, ki ga nekaj izboljša, potem ne more biti zunanji produkt nosilca prepričanja na način, kot je skodelica espressa zunanji produkt aparata.

Oglejmo si zamisel o védenju kot nečem, kar je povezano z akterjem, ki mu lahko pripišemo zasluge za resničnost zaradi nečesa, kar je na njem kot spoznavajočem – njegovih vrlin in vrlih dejanj. Za to zamisel obstajajo teoretični

[*] Tukaj bi lahko ta model prevajali tudi z bolj ustaljenima izrazoma v slovenščini, ko gre za reliabilizem, in sicer mehanizem – rezultat, vendar skuša Zagzebskijeva poudariti analogijo s stroji in na tem mestu uporablja dvojico *machine – product*, medtem ko na drugih mestih govori tudi o *mechanism – output*, kar potem prevajam z mehanizem – rezultat. – op. prev.

[5] Aparat – produkt model uporabljajo Alston, Plantinga, Sosa, Goldman in drugi. Pogosto je uporabljen pojem *rezultat [output]*, nekateri pa tudi ponazoritve preko analogij z aparati in njihovimi produkti.

[6] Moj kolega Wayne Riggs je razmišljal o vprašanju umestitve (vrednosti) kot rešitvi problema vrednosti. Gl. Riggs 2002.

[7] Zagzebski 1996, 3. del.

[8] Gl. Riggs (1998) in Greco (v pripravi).

[9] Gl. Sosa (v pripravi).

[10] Kolikor mi je znano, ga Greco in Riggs zavrneta, Sosa pa ga večkrat uporablja, tudi v članku, v katerem predstavi svojo rešitev problema vrednosti, Sosa (v pripravi).

razlogi, ki niso povezani s problemom vrednosti, kot je npr. predlog, da se s tem izognemo gettierovskim problemom[11], torej je podprta tudi s strani ostalih napetosti v pojmovanju védenja. Toda v tem članku je moje zanimanje osredotočeno na to zamisel kot rešitev problema vrednosti. Če imam prav, da védenje ni rezultat akterja, potem mora biti njegovo stanje. Ne domnevam, da je to edina alternativa modelu aparat – produkt[12], toda če razmišljamo o prepričanju kot delu akterja, lahko prepričanje dobi evaluativne lastnosti od značilnosti akterja na isti način kot dejanja dobijo evaluativne lastnosti od akterja. Ideja o stanju védenja kot nečem, za resničnost česa lahko pripišemo zasluge za resničnost akterju, dejansko napeljuje na to, da so spoznavna stanja povezana z njim na isti način kot tudi njegova dejanja. Dejanje ni produkt akterja, ampak njegov del, in akterju lahko pripišemo zasluge za dejanje zaradi akterjevih značilnosti. Podrobneje, akterju lahko pripišemo zasluge za nekatere dobre značilnosti dejanja, npr. dobre posledice ali ujemanje z moralnim načelom, ker značilnosti dejanja izhajajo neposredno iz akterja, npr. iz njegovega namena ali motiva. Če je biti prepričan (verjetje) isto kot delovati (delovanje), potem imamo model, po katerem lahko akterju pripišemo zasluge za resničnost prepričanja zaradi značilnosti prepričanja, ki izvirajo iz akterja. Predlagam, da je to torej tretji nauk problema vrednosti: *Védenje ni povezano s nosilcem védenja kot produkt z aparatom, ampak kot dejanje z akterjem.*[13]

Problem vrednosti se pojavi tudi pri skupini teorij, ki je širša od zgolj reliabilističnih ali celo eksternalističnih. Internalisti v splošnem ne pojmujejo resničnega prepričanja kot produkta tega, kar ga upravičuje, in torej sprejemajo prvi del tretjega nauka. Ne glede na to so nekateri ranljivi za prvi nauk problema vrednosti, ker analizirajo upravičenja na način, po katerem je njegova vrednost pojasnjena z njegovo sposobnostjo prenašanja resničnosti. Laurence BonJour to izrecno pravi v naslednjem odlomku:

> "Poglavitna vloga upravičenja je vloga sredstva za doseganje resnice, bolj neposredno dosegljiv vmesni člen med našim subjektivnim izhodiščem in našim objektivnim ciljem. [...] Če spoznavno upravičenje ne bi vodilo do resnice, če odkrivanje spoznavno upravičenih prepričanj ne bi znatno povečalo možnost odkrivanja resničnih prepričanj, potem bi bilo spoznavno upravičenje nepomembno za naš glavni kognitivni cilj in bi bilo dvomljive vrednosti. Samo če imamo kak razlog za to, da nas spoznavno upravičenje vodi do resnice, lahko imamo kot kognitivna bitja tudi motiv za dajanje prednosti spoznavno upravičenimi prepričanjem pred

[11] To sem zagovarjala v Zagzebski 1996. Gl. tudi Riggs (1998) in Greco (v pripravi). DePaul 2001, opom. 7, trdi, da gettierovski primeri ustvarjajo drugačno obliko problema vrednosti, ker mislimo, da je vrednost akterjevega spoznavnega stanja v gettierovskem primeru manj vredna od stanja vednosti.

[12] Druga možnost bi bila, da védenje istovetimo s celotnim procesom pridobitve prepričanja in da vrednost dobi iz vrednosti tega procesa kot tudi od resničnosti rezultata tega procesa. Dejala sem že. da bi namenu Sosove opredelitve spoznavne vrednosti ustrezalo takšno pojmovanje védenja kot organske celote v smislu, kot sta ga uporabljala F. Brentano in G. E. Moore. To bi dovoljevalo, da bi vrednost celote presegala vrednost posameznih delov. Gl. Zagzebski (v pripravi-b). Tudi DePaul (2001, razdelek 6) obravnava to možnost védenja kot organske celote.

[13] Zahtevo delovanja pri védenju proučujem v Zagzebski 2001.

> neupravičenimi. Spoznavno upravičenje ima torej konec koncev zgolj instrumentalno vrednost, ne pa intrinzične." (BonJour 1985, 7-8)[14]

V tem odlomku BonJour razume vrednost upravičenja na enak način kot reliabilist, namreč kot nekaj, kar je dobro, ker vodi do resnice. Internost upravičenja nima za BonJourja nobene zveze z njegovo vrednostjo. Toda kakor smo videli, če je značilnost, ki spremeni resnično prepričanje v védenje, dobra zgolj zaradi tega, ker vodi do resničnosti, ostanemo brez razlage, zakaj je vedeti, da p, boljše od zgolj resničnega prepričanja, da p. In temu je tako ne glede na to, ali je ta značilnost dostopna zavesti ali refleksiji nosilca prepričanja ali ne. BonJour se ne opira na model aparat – produkt in zato je problem v njegovem primeru bolj subtilen kot pri reliabilizmu. Ne glede na to pa je problem prisoten, ker resnično prepričanje od upravičenja ne pridobi nobene dodatne vredne lastnosti.

Za razliko od tega se tradicionalna opredelitev védenja kot resnične upravičenega prepričanja ne sreča s problemom vrednosti, ker upravičena prepričanja ne proizvedejo ali ne vodijo enostavno do prepričanja, ki je kandidat za védenje. Namesto tega mu dajo neko lastnost, upravičenost. Naredijo *ga* upravičenega. Zaključek je, da če je vedeti, da p, boljše od biti resnično prepričan, da p, potem mora biti nekaj drugega kot resničnost p-ja, kar *naredi prepričanje, da p, boljše.* Moj predlog je, da je prepričanost kot dejavnost – izboljšajo jo lahko nekatere lastnosti akterja.

Zamislimo si nekaj načinov, na katere lahko dejanje pridobi lastnosti zaradi značilnosti akterja. Množica dejanj, ki jih lahko moralno ovrednotimo, je dobila tradicionalno ime prostovoljna oz. svobodna *[voluntary]*. Prostovoljno dejanje je dejanje, za katerega lahko akterju pripišemo zasluge ali grajo. Prostovoljna dejanja vključujejo nekatera dejanja, ki so namerna *[intentional)*, in nekatera, ki so nenamerna. Prostovoljna, a nenamerna dejanja so lahko motivirana in morda vedno tudi so. Prepričana sem, da dejanja prepričanosti v splošnem spadajo v to kategorijo nenamernih prostovoljnih dejanj, četudi to za namene tega članka ni nekaj, kar bi morali nujno sprejemati. Pomembna je ideja, da so prepričanja tipično motivirana in da motiv lahko vpliva na ovrednotenje prepričanja na način, analogen vplivu motiva na ovrednotenje premišljenega [overt] dejanja.

Z motivom imam v mislih čustveno oz. emocionalno stanje *[affective state]*, ki sproža in usmerja delovanje. V moji teoriji čustev je motiv čustvo, ki je na delu ob sprožitvi delovanja. Vrednotenje *[appreciation for a value]* je čustvo, ki lahko sproža in usmerja delovanje. Ko je temu tako, je to v mojem pomenu motiv. Dejanja, motivirana z vrednotenjem, niso nujno namerna, tudi ko so prostovoljna. Moja trditev je, da so dejanja motivirana z ljubeznijo do neke vrednosti *ceteris paribus* sama visoko vredna.[15]

[14] DePaul (1993, pogl. 2) razkriva problem BonJourja in ostalih pri pojasnjevanju vrednosti védenja. Zahvaljujem se mu, da me je napotil na ta odlomek.

[15] Prav tako menim, da so dejanja, motivirana z ljubeznijo, do neke vrednosti vrednejša kot tista, ki merijo k neki

Kakor jaz razčlenjujem vrlino, je motivna dispozicija *[motive disposition]* sestavni del vrline. Vrlo dejanje je dejanje, motivirano z motivom neke vrline *V* in je značilno za dejanja, motivirana z *V* v danih okoliščinah.[16] Dejanje je lahko sočutno, pogumno ali plemenito, ali pa nepošteno, kruto itd. Ime vrline ali hibe [vice], zaradi katere je dejanje storjeno, je ponavadi vzeto iz imena motiva, iz katerega je storjeno, in motiv je značilnost akterja, ki dejanje stori. Če je prepričanost podobna dejavnosti, potem je lahko vrla ali slaba. Lastnosti resničnega prepričanja, ki ga naredijo boljšega od zgolj resničnega prepričanja, so lastnosti, ki jih pridobi od akterja na isti način, kot dobra dejanja pridobe evaluativne lastnosti od akterja. Podrobneje, prepričanje lahko pridobi vrednost iz njegovega motiva kot dodatno vrednosti poleg tiste, ki jo že ima zaradi svoje resničnosti.

Ideja o védenju kot delovanju te dni ni najbolj pogosta, četudi ima v zgodovini filozofije kar nekaj opore.[17] Včasih se uporablja beseda *presoditi oz. sodba [judge]* za razlikovanje med tem, kar lahko pretvorimo v védenje, in prepričanjem, ki je običajno razumljeno kot dispozicija ali pasivno stanje in ne kot dejanje. Sama bom še naprej uporabljala besedo *prepričanje* kot nanašujočo se ne dejanje, ker se mi zdi ta termin sprejemljiv, nekateri bralci pa bodo nadaljevanje lažje razumeli, če jo zamenjajo s *presojo oz. sodbo.*

Kateri motivi akterja bi lahko naredili prepričanje boljše? Zagovarjala sem že, da so to motivi, ki so oblike temeljnega motiva ljubezni do resnice.[18] Motivacijski elementi posameznih intelektualnih vrlin kot npr. odprtosti, intelektualne poštenosti, intelektualne temeljitosti *[intellectual thoroughness}* ali pazljivosti se medsebojno razlikujejo, a vsi temeljijo na splošni ljubezni do resnice ali vrednotenju le-te in nevrednotenju zmote.[19] Motivacijski elementi posameznih intelektualnih vrlin so najverjetneje bolj kompleksni, ker lahko npr. intelektualna poštenost temelji deloma na spoštovanju do drugih kakor tudi na ljubezni do resnice.[20] Toda ljubezen do resnice je najverjetneje primarni motiv, ki leži v temelju širokega spektra intelektualnih vrlin.[21] Če je ljubezen do resnice dober

vrednosti brez motiva ljubezni do ali cenjenja neke vrednosti. Nekatere nenamerna dejanja imajo moralno vrednost, ker izvirajo iz dobrega motiva. Za razliko od tega lahko nekatera namerna dejanja merijo k dobremu cilju, a imajo manj vrednosti, ker ne izvirajo iz dobrega motiva. O tem več razpravljam v Zagzebski (v pripravi-a).

[16] V Zagzebski 1996 razlikujem med vrlim dejanjem *[virtuous act]* in dejanjem vrline *[act of virtue]*. Za razliko od slednjega vrlo dejanje ni nujno uspešno v svojem cilju/namenu. *Dejanje vrline* uporabljam kot strokovni pojem *[temi of art]* za označitev v vseh ozirih dobrega dejanja. Gre za dejanje, ki izvira iz vrlega motiva, ki bi ga značilno storil vrl akter v teh okoliščinah, je uspešno pri doseganju cilja vrlega motiva in je uspešno zaradi preostalih vrlih značilnosti dejanja.

[17] Zdi se, da so Akvinski in drugi srednjeveški filozofi razmišljali o védenju kot vključujočem neko dejanje razuma *[intelect]*. Tudi nekateri odlomki iz Platona nakazujejo isto. Gl. Riggs 2000. pogl. 9.

[18] To zagovarjam v Zagzebski 1996 in bolj podrobno v Zagzebski (v pripravi-a).

[19] V Zagzebski (v pripravi-a) zagovarjam, da ljubezen do resnice ni isto kot sovraštvo do neresnice oz. zmote, a ta razlika ni pomembna za bistvo tega članka.

[20] Spoštovanje, ljubezen in cenjenje *[appreciation]* so v večini kontekstov dokaj različni, toda zdi se mi, da te razlike, ko gre za kontekst čustvene naravnanosti do resnice, niso pomembne. Ker večina spoznavnih teoretikov ne misli, da lahko ima katere koli čustvena naravnanost do resnice kakršne koli posledice za spoznavni status, je dovolj, če nasprotno pokažemo zgolj za eno od teh.

[21] Nekatere intelektualne vrline lahko merijo na razumevanje in ne na resnico. Zdi se mi, da so spoznavni teoretiki v splošnem zanemarjali vrednost razumevanja. Gl. Zagzebski 2001b. GL tudi Riggs (v pripravi).

motiv, potem lahko prispeva vrednost intelektualnim dejanjem, ki jih motivira.

Kakšne vrste vrednost ima ljubezen do resnice? Če predpostavimo, da če je nekaj vredno, je to vredno tudi ceniti ali ljubiti, potem ima ljubezen do resničnega prepričanja vrednost, ker ima vrednost resnično prepričanje. Toda motiv ljubezni do resnice črpa svojo moč tudi iz povsem ločenih moralnih motivov. Temu je tako, ker moralna dopustnost, pohvala in graja temeljijo na spoznavni dopustnosti, pohvali in graji.[22]

Naj predlagam kriterij za nedopustnost. Ko je v igri nekaj, kar je moralno pomembno pri izvršitvi dejanja *S*; in če je S primer delovanja glede na prepričanje *B*, potem je moralno pomembno, da je *B* resnično. Tako je nedopustno za akterja biti prepričan na način, ki mu ne uspe spoštovati pomembnosti resničnosti *B.* To pomeni, da mora biti akter prepričan iz nekih motivov. Podrobneje, predlagam, da morajo biti njegovi motivi takšni, da vključujejo vrednotenje resnice, ali vsaj, da ne vključujejo njenega nevrednotenja ali zanemarjanja.[23]

Če moralna grajevrednost *[blameworthiness]* temelji na spoznavni grajevrednosti, potem isto razmišljanje vodi do sklepa, da moralna hvalevrednost *[praiseworthiness]* ali zasluge *[credit]* temeljijo na spoznavni hvalevrednosti ali zaslugah.[24] Predpostavimo sedaj, da je dejanje *S* primer delovanja glede na prepričanje *B* in da je primerek tipa moralno hvalevrednega dejanja v ustreznih okoliščinah. Predlagam, da so zasluge za dejanje *S* pripisana akterju zgolj v primeru, ko mu lahko pripišemo tudi zasluge za resničnost prepričanja *B.* Če je torej vedeti, da *B*, nekaj podobnega kot biti resnično prepričan v primeru, ko je za resničnost *B* zaslužen ak- ter, sledi, da le-ta prejme zasluge za dejanje S temelječe na prepričanju *B* samo če ve, da *B*.[25]

Predpostavimo tudi, da imam prav o motivacijski zahtevi pri pripisu zaslug za dosego resnice, ki vsebuje ljubezen do resnice. Sledi, da je motiv iz ljubezni do resnice pogoj za ljubezen do moralnih dober [goods], ali vsaj pogoj za ljubezen do tistih moralnih dober v dejanjih, za katera akterju pripišemo pohvalo ali grajo. Hvalevrednost ljubezni do resnice je pogoj za moralno hvalevrednost. Torej

[22] *Locus classicus* razprave o povezanosti moralne dopustnosti dejanj in dopustnosti prepričanj je Cliffordov članek "The Ethics if Belief". W. K. Clifford zaključi, da je neupravičeno prepričanje moralno nedopustno. Gl. tudi Montmarquet 1993 za dobro razpravo o razmerju med dopustnostjo dejanj in prepričanj.

[23] Spoznavna dopustnost je težavna zaradi načela "iz *naj bi* sledi *lahko bi*" *[ought implies can]*. Toda če nismo pripravljeni zatrditi, da ni nedopustnih prepričanj, potem so neke stvari, ki jih smemo ali ne smemo verjeti, in tako nam "iz *naj bi* sledi *lahko bi"* načelo ne preprečuje govoriti o spoznavni dopustnosti. Ne bom se spuščala v razpravo o tem, do kakšen mere smo sploh sposobni nadzorovati naša prepričanja. Želim poudariti le, da v kolikor verjamemo, da obstajajo dejanja verjetja oz. prepričanosti, ki so nedopustna, sledi, da ali imamo moč nad prepričanostjo, ki je predvidena v "iz *naj bi* sledi *lahko bi"* načelu, ali pa to načelo preprosto za ta prepričanja ne velja. Povedano drugače, mislim, da je intuicija o nedopustnosti prepričanja močnejša kot "iz *naj bi* sledi *lahko bi"* načelo.

[24] Hvalevrednost se na določen način razlikuje od pripisa zaslug – glede na poglede večine ljudi – v tem, da je biti vreden hvale močnejša pohvala kakor pa biti zaslužen za nekaj. Mislim, da gre zgolj za razliko v stopnji in da nima večjega pomena.

[25] Brez dvoma moramo tukaj dodati kopico pogojev. Na primer, v splošnem lahko akterju pripišemo določene zasluge za *S*, tudi če je *B* neresnično, v kolikor njegovi intelektualni motivi v zadostni meri spoštujejo pomembnost resničnosti *B*, ko ravna, kakor bi značilno ravnala intelektualno vrla oseba v teh okoliščinah in ko je neresničnost *B* zgolj posledica nesrečnega slučaja.

imamo moralni motiv za védenje. Vrednost, ki spremeni resnično prepričanje v védenje, je pogoj za moralno vrednost dejanj, ki so odvisna od prepričanj.

Navzlic moralni pomembnosti resničnih prepričanj ponavadi mislimo, da ima resnično prepričanje vrednost samo po sebi. Vrednost resničnega prepričanja je posebna spoznavna vrednost, ki dopušča spoznavnim teoretikom obravnavo področja prepričanj in védenja kot nekaj neodvisnega od dejanj, ki so predmet moralnega vrednotenja. To nas pripelje do globljega problema vrednosti védenja: V kakšnem smislu, če sploh katerem, je resnično prepričanje dobro? Če resnično prepričanje ni dobro, potem imamo pred seboj mnogo težji problem kot zgolj iskanje vrednosti, ki naredi védenje boljše od resnične prepričanosti.

2. Vrednost resničnega prepričanja

Védenje sem obravnavala kot nekaj, kar si vedoči zasluži oz. pridobi. Je stanje, v katerem so mu pripisane zasluge za resnico, morda si zasluži celo pohvalo. Toda zakaj bi mislili tako? Omenila sem že, da je bila ta ideja razvita zato, da bi se s tem izognili gettierovskim problemom, toda ta cilj je vsekakor zgolj manjši del naloge opredelitve védenja. O védenju je vredno razpravljati, ker ga je vredno imeti. Toda dejstvo, da je védenje vredno, nas ne sili v to, da bi ga sprejeli kot nekaj, kar si pridobimo ali zaslužimo in smo zanj odgovorni ali pa vredni pohvale, ne glede na to, da takšen način gledanja sledi iz analogij s športom, ki so uporabljene pri razpravi o problemu vrednosti pri Sosi, Grecu in Riggsu, in iz analogije z zmago v bitki, ki jo uporabi Michael DePaul.[26] Vsi smatrajo védenje kot dosežek ali kot v igri priigrane točke in ne kot naklonjenost srečne usode. Mislim, da imajo glede tega prav, toda pomembno je poudariti, da nas dejstvo o védenju kot vrednem stanju ne sili, da bi o njem razmišljali na tak način. Nekatera dobra so dobra, tudi če se nam za njih ni potrebno truditi – npr. zdravje ali varno okolje, in nekatera so lahko celo boljša, če je temu tako – npr. ljubezen in prijateljstvo. Zdravje, varnost, ljubezen in prijateljstvo so vsa dobra v smislu, da so zaželena *[desirable]*. Smisel dobrega, ki ga pridobimo ali zaslužimo, je dobro kot vredno občudovanja *[admirable]*. Trdila sem, da če razmišljamo o védenju kot o delovanju, potem je to stvar, ki je lahko vrla ali slaba (hibava), kar se pravi, vredna občudovanja ali vredna graje *[reprehensible]*. Védenje je vredno občudovanja, gotovo pa je tudi zaželeno, ker je njegov osnovni sestavni del, resnično prepričanje, zaželen. Kar se pravi, da menimo, da je resnično prepričanje dobro *za* nas.

Resnično prepričanje je lahko zaželeno, gotovo pa ni vredno občudovanja. Ni nekaj, za kar bi bili zaslužni ali vredni pohvale. To pomeni, da resnično prepričanje s sabo ne nosi zaslužnosti, čeprav sem rekla, da smo v primeru védenja zaslužni za resnico zaradi drugih značilnosti prepričanja. Vrednost,

[26] DePaul 2001, 179. DePaul uporablja tudi primer oglasa za finančno institucijo, v katerem uglajen gospod naznani: "Mi delamo denar na staromoden način: služimo ga." [We *make money the old-fashioned way: we earn it.*] To pomeni, da je bolje pridobiti denar z delom zanj, kot pa po sreči ali z dedovanjem. Kot poudari DePaul, to pomeni, da je poleg denarja samega še nekaj vrednega.

zaradi katere je védenje boljše od resničnega prepričanja, je vrednost občudovanja, medtem ko je vrednost, ki jo ima resnično prepričanje, zaželenost. Toda sedaj naletimo na problem, ker vsa resnična prepričanja gotovo niso zaželena. V prvi vrsti mnogi opozarjajo na nekatere resnice, ki so trivialne. To predstavlja problem za vrednost védenja, kajti tudi če je védenje o neki trivialni resnici boljše od zgolj resničnega prepričanja o njej, koliko boljše je lahko? Zgolj toliko, kolikor dobrega lahko to védenje ima. Če je resnično prepričanje o tem, kolikokrat je besedica *"the"* uporabljena v reklami za McDonaldi, v temelju brez vrednosti, potem tudi védenje o tem nima vrednosti. Torej, tudi če smo v trivialne resnice prepričani na najbolj vrl, vešč, razumski in upravičen na čin, trivialnost le teh naredi tudi védenje o njih samo trivialno. Neizogiben zaključek je, da nekaj védenja za nas ni dobrega. Nekaj je lahko c< lo slabega. Lahko je slabo za akterja ali pa za druge – npr. védenje tem, kaj natančno kirurg počne z mojo nogo, medtem ko mi odstranjuje kožnega raka; védenje o sosedovem privatnem življenju. Sledi, da ali ni vse védenje zaželeno ali pa da nekaterih resničnih prepričanj ni mogoče spremeniti v védenje.

Pogost odgovor na ta problem je ta, da je resnica pogojno vredna. Ni vredno resnično prepričanje *per se*, temveč to, da imamo odgovore na naša vprašanja. Naši interesi določajo razliko med vrednimi resničnimi prepričanji in nevrednimi *[disvaluable]* ali brezvrednimi *[nonvaluable]*. Sosa poda primer štetja zrnc peska na plaži in pravi, da prepričanje o tem številu nima vrednosti, ker ne služi nobenemu našemu interesu.[27] Seveda bi koga to število *lahko* zanimalo, a vseeno se mi zdi, da védenje o tem ne bi postalo nič boljše. Če je resnica trivialna, prepričanje o njej ni nič boljše zaradi tega, da ima spoznavni akter neke posebne ali perverzne interese. Še več, takšni interesi ga lahko narede celo slabšega, ker tako združimo perverznost interesov s trivialnostjo resnic.[28]

Morda se lahko obrnemo k ideji pomembnosti, da bi ohranili intuicije o naših interesih in ciljih kot povezanih z resnicami, ki so za nas vredne, s tem ko bi povečali pomen vrednosti.[29] Morda so nekatere stvari preprosto pomembne *simpliciter*, kar pomeni, da obstajajo resnice, katerih pomembnost ni zvedljiva na to, kar je pomembno za to ali ono. Morda v pojmu pomembnosti obstajajo stopnje oddaljitve od posameznika, pri čemer so nekatere stvari pomembne za ljudi v določeni vlogi ali določeni družbi, in nekatere za vsakogar. Toda mislim, da nam ta korak ne more pomagati. Ni pomembnih "resnic", če so to resnične propozicije, ker propozicije same na sebi niso pomembne. Če je resničnost lastnost propozicij, potem resničnost ni to, kar je pomembno. Pomembno je stanje resnične prepričanosti v propozicijo. Ko torej pravimo, da so nekatere resnice pomembne in druge ne, s tem v bistvu mislimo, da se nekatera resnična *prepričanja* pomembna in druga ne. In to ne pomeni nič drugega kako to, da se

[27] Gl. Sosa 2001.

[28] Poleg Sose tudi Christopher Stephens uporabi interes(e) kot rešitev problema dveh vrednosti – doseganja resnice in izogibanja zmoti. Goldman 2001 istoveti interes(e) z vrednostjo, ki poenoti spoznavne vrline.

[29] Ta zamisel je na kratko obravnavana v Riggs (v pripravi).

vrednost resničnih prepričanj spreminja. A to smo že vedeli. Kar si želimo vedeti, je to, kaj povzroča te spremembe. Zamisel o pomembnih resničnih prepričanjih je zgolj drug način zastavitve problema, ne pa njegova rešitev.

Druga oblika pogojne vrednosti je instrumentalna vrednost. Zatrjeno je bilo, da naš razum stremi k zadovoljitvi naših želja in doseganju naših ciljev. Resnično prepričanje je vsekakor sredstvo za doseganje naših ciljev, večina katerih je nespoznavnih *[non-epistemic]*. Dober primer tega je stališče Richarda Foleya, ki zagovarja, da je spoznavni cilj resnice instrumentalno vreden kot sredstvo za dosega drugih ciljev, katerih vrednost pa je nedoločena.[30] Jasno je, da imajo mnoga resnična prepričanja instrumentalno vrednost, toda le ta je oblika pogojne vrednosti, ker je vrednost cilja pogoj za vrednost sredstev. Če cilj nima vrednosti, potem je nimajo tudi sredstva.[31] Pogojna vrednost je kakor domnevni terorist; nekdo, ki to je, ni nujno terorist, in prepričanje z instrumentalno vrednostjo ni nujno vredno. Nobena oblika pogojne vrednosti resničnega prepričanja nima za posledico tega, da bi bila vsa resnična prepričanja vredna.

Še vedno preostane možnost, da bi imela resnična prepričanja intrinzično vrednost. Morda ima vsako resnično prepričanje neko intrinzično vrednost preprosto zaradi tega, ker je resnično, ne glede na to, ali je za nas dobro ali ne. Temu je res lahko tako, toda ne vidim, da bi to imelo za posledico, da bi bilo vsako resnično prepričanje, gledano v celoti, vredno, ker intrinzičnost ni vezana na stopnjevitost. Intrinzičnost se nanaša na izvor prepričanja, ne pa na njegovo količino [amount]. Torej, tudi če ima vsako resnično prepričanje določeno intrinzično vrednost, je malo verjetno, da bo le ta pri vsakem prepričanju zadostna, da bo odtehtala njegovo nezaželenost ali drugo negativno vrednost, ki pri nekaterih resničnih prepričanjih izvira od drugod.

Če upoštevamo vse to, je sklep, da vsako resnično prepričanje ni dobro, neizogiben. Ne glede na to, ali upoštevamo občudovanost ali zaželenost, intrinzičen ali ekstrinzičen vir vrednosti, je gledano kot celota verjetno, da obstajajo nekatera resnična prepričanja, ki nimajo vrednosti, in nekatera, ki imajo negativno vrednost.

Sedaj premislimo, kaj to pomeni za vrednost védenja. V prvem razdelku sem zaključila, da je védenje boljše od resnične prepričanosti, samo če gre za resnično prepričanost, za katere resničnost gredo zasluge akterju. Toda če določeno resnično prepričanje nima vrednosti, kako lahko gredo zasluge za njegovo resničnost akterju, če ta v tem primeru ni nekaj, za kar bi pripisovali zasluge? Dokler so nekatera resnična prepričanja brez vrednosti, lahko za njih enako smiselno pripišemo tudi krivdo. Če predpostavimo, da je vsako resnično prepričanje intrinzično dobro, je dobro, da akterju pripišemo zasluge za

[30] Gl. Foley 2001. Foley je pripravljen dopustiti, da vrednost cilja določa akter.

[31] Določena sredstva lahko služijo tudi večim ciljem. Menim, da je v posamičnem primeru vrednost sredstev določena z njihovim ciljem v tem primeru. To je združljivo s tem, da lahko imajo sredstva tega tipa vrednost, ko služijo nekemu drugemu, vrednemu cilju.

resničnost zaradi tega, kar je vredno občudovanja v njegovem spoznavnem védenju – njegovim intelektualno vrlin motivom in dejanjem. Toda resnica, za katero mu pripišemo zasluge, ni ravno visoka nagrada.

Premislimo tudi, kako je z mojim predlogom, da je védenje boljše od resničnega prepričanja, ker gre za stanje, v katerem je resnica dosežena z intelektualno vrlimi motivi in dejanji, katerih vrednost lahko izvedemo iz vrednosti motiva vrednotenja resnice. Toda če v nekaterih primerih gledano v celoti resnica nima vrednosti, zakaj bi bili motivirani, da jo vrednotimo. Seveda predpostavljamo, da ima resnično prepričanje neko intrinzično vrednost. Prav tako lahko predpostavimo, da je običajno za nas dobro, in v tem primeru je smiselno misliti, da ga je dobro *vrednotiti kot* nekaj, kar ima intrinzično vrednost, kakor koli neznatna že je, kakor tudi kot nekaj, kar je za nas običajno dobro. Toda če iščemo vrednost, ki ima možnost, da bi bila pomembno dobro, potem je še vedno nismo našli.

Še več, dokler nekatera resnična prepričanja niso zaželena, lahko akterju za doseganje resnice pripišemo zasluge za to, ne glede na to, da mu ne moremo pripisati zaslug za doseganje zaželene resnice. In tudi ko resnica *je* zaželena, je morda zgolj stvar gole sreče, da je akter prispel do nje, ne pa do neke nezaželene resnice. Po mojem nas to privede k vzporednicam gettierovskim problemom. Gettierovski primeri vzniknejo, ko gre za naključno zvezo med občudovanostjo prepričanja in njegovo resničnostjo. Podobno je možno, da obstaja naključna zveza med občudovanostjo prepričanja in njegovo zaželenostjo. Mislim, da bi bilo preveč, če bi v takšnih primerih zanikali védenje, ne glede na to, da niso tako dobri, kot bi lahko bili. Niso najboljši primerki védenja; niso tisti, ki so pomembna dobra. Rešitev za gettierovske primere je zaprtje vrzeli med občudovanostjo prepričanja in njegovo resničnostjo. Rešitev novega problema vrednosti je zaprtje vrzeli med občudovanostjo resničnega prepričanja in njegovo zaželenostjo. Da bi dobili res zanimivo vrednost védenja, se mora izkazati ne samo, da lahko pripišemo akterju zasluge za resničnost prepričanja, temveč tudi za njegovo zaželenost. To je splošna formula, ki jo lahko natančneje izrišemo na mnogo različnih načinov, tako kot lahko storimo tudi s formulo opredelitve védenja glede na teoretska izhodišča o zaslužnosti, ta pa so odvisna od splošne teorije ovrednotenja akterjev. V naslednjem razdelku bom orisala takšno vrlinsko teoretično opredelitev védenja, ki zadovolji zahteve izpostavljene v prvih dveh razdelkih članka.

3. Védenje, motivi in *eudaimonia*

Trdila sem, da dobri motivi dodajo vrednost dejanjem, ki jih motivirajo, in ta vključujejo spoznavna dejanja. Motivi so kompleksni in v tem članku jih podrobneje še nisem raziskovala, toda njihova značilnost, ki je za nas pomembna, je, da so sami pogosto motivirani z motivi višjega reda. Motivi višjega reda so pomembni, ker povezujejo celoto motivacijske strukture in nam. omogočajo

konsistentnost motivov prvega reda. Če lahko dobri motivi prenašajo vrednost, ne dejanja, katerih motivi so, sledi, da lahko motivi višjega reda prenašajo vrednost na motive nižjega reda, za katere motivirajo akterja, da jih sprejme. Ker iščemo dodaten izvor vrednosti v nekaterih primerih védenja, je smotrno iskati izvor vrednosti motiva resničnega prepričanja v posamičnih primerih védenja, ki je vrednejše od običajnega védenja.

Do sedaj smo srečali najmanj dva načina, pri katerih je vrednotenje resnice v posamičnih primerih zahtevano od drugimi stvari, ki so vredne. To je, imamo motiv za motiv za resnico zaradi drugih dobrih motivov. Prvič, če je v igri nekaj moralno pomembnega, ko izvršimo neko dejanje in je to odvisno od resničnosti nekega prepričanja, potem je resničnost tega prepričanja moralno pomembna. Motiv za resničnost prepričanja je v takšnih primerih motiviran s strani motiva višjega reda, da bi delovali moralno in živeli dobro življenje. Drugič, ker je resničnost prepričanja sredstvo za večino praktičnih ciljev, je motiv za vrednotenje resničnosti na nekem področju motiviran s strani vrednotenja teh ciljev, kar je naprej motivirano s strani želje po dobrem življenju. Predlagam, da motiv višjega reda po dobrem življenju zaobsega tudi motiv, da imamo neke druge motive, vključno z motivom za vrednotenje resnice na nekaterih področjih. Motiv višjega reda motivira akterja, da ima motive, ki so konstitutivni za moralne in spoznavne vrline, in oboje na ta način tudi povezuje. Če je védenje resnično prepričanje, za katero lahko zasluge pripišemo akterju zaradi njegove umestitve v akterjevo motivacijsko strukturo, pridobi vrednost ne samo od motiva za resnico, temveč tudi od motiva višjega reda, ki motivira akterja, da vrednoti resnico na nekaterih področjih in ob nekaterih priložnostih. In ta motiv nima nič neposredno opraviti s spoznavno vrednostjo; je sestavni del motiva po dobrem življenju.

Moj predlog je torej tak. Spoznavnemu akterju lahko pripišemo zasluge za pridobitev resničnega prepričanja, ko do le tega dospe zaradi svojih vrlih intelektualnih dejanj motiviranih z ljubeznijo do resnice. Zasluge za pridobitev zaželenega resničnega prepričanja mu lahko pripišemo, ko do le-tega dospe zaradi dejanj motiviranih z ljubeznijo do resničnih prepričanj, ki so sestavni del dobrega življenja. Motiv za zaželena resnična prepričanja ni popolna razlaga za pripis zaslug akterju, ki je takšno prepričanje pridobil, iz istega razloga kot motiv za resnično prepričanje ni popolna razlaga za pripis zaslug akterju, ki je takšno prepričanje pridobil. Moje stališče je, da so motivi v prvi vrsti vzroki preostalih vrednih značilnosti naše kognitivne aktivnosti. Ko akterju uspe pridobiti zaželeno resnično prepričanje zaradi njegovih občudovanja vrednih intelektualnih motivov, obstaja ne-naključna povezava med občudovanostjo prepričanja in njegovo zaželenostjo. Ta povezava se izogne zgoraj omenjenim gettierovskim problemom in rezultira v tem, da so nekateri primeri védenja veliko dobro *[great good]*.

Preglejmo sedaj nekaj različnih načinov, na katera so lahko prepričanja dobra.

Vsa resnična prepričanja imajo verjetno neko intrinzično vrednost preprosto

zaradi njihove resničnosti, ne glede na to, ali so za nas dobra ali ne. Ko zasluge za resničnost (lahko) pripišemo akterju, je prepričanje vredno tudi občudovanja. To je védenje.

Nekatera resnična prepričanja so za nas dobra, ta so zaželena. Lahko so zaželena, ne glede na to, ali so občudovanja vredna ali ne. Toda nekatera resnična prepričanja so nezaželena. Možno je tudi, da so nekatera neresnična prepričanja zaželena, toda o njih v tem članku nisem govorila.

Občudovanja vredna prepričanja so tista, ki so vrla. Občudovanja vredna prepričanja so lahko neresnična.

Nekatera resnična prepričanja so oboje, zaželena in občudovanja vredna. Najbolj zanimivi so primeri, pri katerih obstaja povezava med njihovo občudovanostjo in njihovo zaželenostjo. Določeno prepričanje je vredno občudovanja, in glede na to dejstvo, ni naključje, da ima akter zaželeno resnično prepričanje. To so najbolj vredni primeri védenja.

Problemi, ki smo jih srečali pri vrednosti resničnega prepričanja, kažejo na neustreznost standardnega pristopa k določanju vrednosti védenja Vprašanje se ne bi smelo glasiti, kaj dodati resničnemu prepričanju, da bi imelo vrednost, temveč, kaj dodati vrlemu prepričanju, da bi dobili védenje. Seveda je odgovor na to vprašanja kot na dlani: Biti mora resnično. Ko pristopimo k problemu vrednosti na ta način, je težje vprašanje odgovorjeno prej kot pa lažje. To ni običajem vrstni red, toda mislim, da je pravi. Če pričnemo, kot je običajno, z resničnim prepričanjem, lahko pričnemo z nečim, kar je lahko brez vrednosti, niti zaželenosti niti občudovanosti. Nadalje, če pričnemo z vrlo prepričanostjo, lahko pojasnimo, zakaj je celo napačno, a po motivaciji vrlo prepričanje, občudovano.

Naj sklenem s kratkim razmislekom o tem, kaj dela vrlino v splošnem dobro stvar. Recimo, da ima Aristotel prav, ko vrla dejanja vidi kot del *eudaimonie*, srečnega življenja. Če imam tudi jaz prav, da je prepričanost vrsta dejanja, sledi, da so vrle prepričanosti del *eudaimonie*. *Eudaimonia* je iz več razlogov za razlago zelo izzivalen pojem, toda vidik, ki se zdi sodobnim komentatorjem zelo težaven, je Aristotelova zamisel, da *eudaimonia* spaja občudovanost z zaželenostjo. Nihče ne oporeka pojmovanju *eudaimonie* kot zaželenega življenja, v bistvu je to njena najpogostejša opredelitev. Pokazati je potrebno, da je vrla (tj. občudovanja vredna) dejavnost sestavni del zaželenega življenja. In to je seveda še kako sporno. Enak problem se pojavi pri vrednosti védenja. Nihče ne bo oporekal trditvi, da so nekatera resnična prepričanja zaželena. Čemur lahko oporekamo, je to, ali so intelektualno vrla prepričanja – ali na način, ki sem ga opisala jaz, ali pa na kak drug način – tudi sestavni del zaželenega življenja. Vprašanje "Zakaj bi želeli imeti občudovanja vredna prepričanja?" se v resnici ne razlikuje od vprašanja "Zakaj bi želeli izvrševati občudovanja vredna dejanja?". Če so vrla dejanja zaželena, je to zato, ker je zaželeno delovati v občudovanja vredni smeri. Podobno, če je védenje propozicije bolj zaželeno kot zgolj resnično prepričanje, je to zato, ker je bolj zaželeno biti prepričan na občudovanja vreden

način. Toda potem ne vem, kako bi lahko to zagovarjali brez splošnejše teorija *eudaimonic* ali srečnega življenja. To pomeni, da so razprave v sodobni vrlinski etiki o razmerju med vrlo dejavnostjo in dobrim življenjem pomembne za razumevanje intelektualno dobrega življenja, kakor tudi za razumevanje dobrega življenja *simpliciter*.

4. Zaključek

Vprašanje "Kaj je védenje?" ni nepovezano s vprašanjem "Zakaj vrednotimo védenje?". Tisti, ki jim je prvo vprašanje važnejše, primerjajte še "Kaj je védenje?" in "Ali ga lahko pridobimo?". Za antiskeptično naturalistično spoznavno teorijo je značilno, da moramo védenje, ne glede na to, kaj le to je, opredeliti kot nekaj, kar imamo. Neobstoječ pojav oz. predmet nas ne zanima. Trdim, da moramo védenje opredeliti kot nekaj, kar vrednotimo. Ne zanima nas pojav, ki vrednosti nima ali pa je ima zelo malo. Možno je, da noben pojav, ki se bolj ali manj prekriva s tradicionalnim pojmov védenja, nima vrednosti, ki sem jo iskala v tem članku. Če je temu tako, potem bi morala preiti v teorijo zmote, kakor jo iz etike poznamo pri J. L. Mackieju. Toda zaenkrat ne vidim, da bi bilo to neizogibno, ker je možno podati opredelitev védenja, ki zadosti tako tradicionalnim zahtevam in ki ta pojav poveže z vrednostjo. Prav tako menim, da bi morali sprejeti, da če je védenje stanje vredno nenehne pozornosti, ki je bila izkazano skozi vso zgodovino filozofije, je temu tako, ker njegova vrednost močno presega spoznavno vrednost resničnosti, ki vodi do resničnega prepričanja. Védenje je pomembno zaradi njegove tesne povezanosti z moralno vrednostjo in širšimi vrednotami dobrega življenja. Malo verjetno je, da je spoznavna vrednost v kakršnem koli zanimivem smislu povsem avtonomna.

Zahvala

Za zanimive in koristne pripombe se zahvaljujem Philipu Percivalu – komentatorju na konferenci na univerzi v Stirlingu. Začetne različice članka so bile predstavljene na univerzah California, Riverside, Tulane, Oklahoma in na Eastern Division Meetings of the American Philosophical Association (december 2001). Zahvaljujem se občinstvom in tudi komentatorju na APA srečanju, Michaelu DePaulu, za njegovo pomoč pri izboljšanju članka.

Literatura

Benson, Hugh (2000). *Socratic Wisdom.* Oxford: Oxford University Press.

Bonjour, Laurence (1985). *The Structure of Empirical Knowledge.* Cambridge, Mas.: Harvard University Press.

DePaul, Michael (1993). *Balance and Refinement.* New York: Routledge.

DePaul, Michael (2001). "Value Monism in Epistemology". V M. Steup (ur.),

Knowledge, Truth, and Duty, Oxford: Oxford University Press.

Foley, Richard (2001). "The Foundational Role of Epistemology in a General Theory of Rationality". V A. Fairweather in L. Zagzebski (ur.), *Intellectual Virtue: Essays on Epistemic Virtue and Responsaibility,* Oxford: Oxford University Press.

Goldman, Alvin (2001). "The Unity of the Epistemic Virtues". V A. Fairweather in L. Zagzebski (ur.), *Intellectual Virtue: Essays on Epistemic Virtue and Responsaibility,* Oxford: Oxford University Press.

Greco, John (v pripravi). "Knowledge as Credit for True Belief". V L. Zagzebski in M. DePaul (ur.), *Intellectual Virtue: Perspectives from Ethics and Epistemology,* Oxford: Oxford University Press.

Greco, John (1999). "Agent Reliabilism". *Philosophical Perspectives* 13, str. 273-296.

Montmarquet, James (1993). *Epistemic Virtue and Doxastic Responsibility.*

Lanham, Md.: Rowman and Littlefield.

Riggs, Wayne (v pripravi) "Understanding Virtue and the Virtue of Understanding". V L. Zagzebski in M. DePaul (ur.), *Intellectual Virtue: Perspectives from Ethics and Epistemology,* Oxford: Oxford University Press.

Riggs, Wayne (1998). "What Are the 'Chances' of Being Justified?". *The Monist* 81 (3), str. 452-472.

Riggs, Wayne (2002). "Reliability and the Value of Knowledge". *Philosophy and Phenomenological Research* 64(1), str. 79-96.

Sartwell, Crispin (1992). "Knowledge as Merely True Belief". *American Philosophical Quarterly* 28, str. 157-165.

Sosa, Ernest (v pripravi) "The Place of Truth in Epistemology". V L. Zagzebski in M. DePaul (ur.), *Intellectual Virtue: Perspectives from Ethics and Epistemology,* Oxford: Oxford University Press.

Sosa, Ernest (2001). "For the Love of Truth?". V A. Fairweather in L. Zagzebski (ur.), *Intellectual Virtue: Essays on Epistemic Virtue and Responsaibility,* Oxford: Oxford University Press.

Zagzebski, Linda (v pripravi-a). "Intellectual Motivation and the Good of Truth". V A. Fairweather in L. Zagzebski (ur.), *Intellectual Virtue: Essays on Epistemic Virtue and Responsaibility,* Oxford: Oxford University Press.

Zagzebski, Linda (v pripravi-b). "Epistemic Value Monism". V J. Greco (ur.), *Sosa and His Critics,* Oxford: Basil Blackwell.

Zagzebski, Linda (1996). *Virtues of the Mind: An Inquiry into the Nature of Virtue and the Ethical Foundations of Knowledge.* Cambridge: Cambridge

University Press.

Zagzebski, Linda (2000). “From Reliabilism to Virtue Epistemology”. V G. Axtell (ur.), *Knowledge, Belief, and Character*, Lanham, Md.: Rowman and Littlefield.

Zagzebski, Linda (2001a). “Must Knowers Be Agents?”. V A. Fairweather in L. Zagzebski (ur.), *Intellectual Virtue: Essays on Epistemic Virtue and Responsaibility*, Oxford: Oxford University Press.

Zagzebski, Linda (2001b). “Recovering Understanding”. V M. Steup (ur), *Knowledge, Truth, and Obligation*, Oxford: Oxford University Press.

Vojko Strahovnik
Filozofska fakulteta Univerze v Ljubljani

UVOD V VRLINSKO EPISTEMOLOGIJO

Članek je uvod v vrlinsko spoznavno teorijo. Na kratko predstavi njeno zgodovino in njen razvoj. Obravnava naravo spoznavnih vrlin in njihovo vlogo pri odgovorih na osrednja vprašanja spoznavne teorije. Ukvarja se tudi z razlikovanjem med reliabilističnimi in responsibilističnimi različicami vrlinske epistemologije ter med njenimi šibkimi in močnimi različicami, z odnosom med vrlinsko etiko in vrlinsko epistemologijo ter si zastavi vprašanje, ali se lahko s pomočjo vrlinske epistemologije izognemo gettierovskim problemom.

Ključne besede: vrlinska spoznavna teorija, spoznavne vrline, moralne vrline, reliabilizem, responsibilizem, gettierovski problemi.

1. Uvod

Sodobna *vrlinska epistemologija* je dokaj mlada filozofska teorija, katere začetki segajo v osemdeseta leta dvajsetega stoletja. Gre za spoznavnoteoretski pristop, ki *spoznavnim ali intelektualnim vrlinam* ali krepostim namenja eno od osrednjih vlog pri odgovorih na vprašanja in probleme v spoznavni teoriji. O intelektualnih vrlinah je že v antiki razpravljal Aristotel. Vrlinska epistemologija je nov preporod doživela v času največjih nepremostljivih ovir med spoznavnim internalizmom in eksternalizmom in ponudila most, ki bi lahko povezal oba bregova. Sčasoma pa je prerasla to vlogo in se razvila v eno izmed avtonomnih oblik spoznavne teorije in pričela odpirati spoznavne probleme in vprašanja, ki se jih je "tradicionalna" spoznavna teorija potiskala ob rob ali pa se jih sploh ni lotevala. V njenem razvoju smo priča točki, ko jo avtorji poskušajo razviti predvsem v tej slednji smeri in jo na nek način zopet povezati z njenimi antičnimi izvori.

Članek bo v obrisih predstavil jedro vrlinske epistemologije in motivacijo zanjo. Kratko bo predstavil njeno zgodovino in si zastavil vprašanje, kaj in katere so spoznavne vrline. Poleg tega se ukvarja še s tremi vprašanji oz. tematskimi sklopi, to so: (i) razlikovanje med *reliabilizmom* in *responsibilizmom* v vrlinski epistemologiji, (ii) povezava med moralnimi in spoznavnimi vrlinami oz. med *vrlinsko etiko* in *vrlinsko epistemologijo* in (iii) vrlinska epistemologija kot *odgovor na izziv gettierovskih problemov.*

2. Kratka zgodovina vrlinske epistemologije

Med bolj znane sodobne vrlinske spoznavne teoretike lahko uvrstimo naslednje avtorje: Ernesta Soso, Johna Greca, Alvina Goldmana, Guya Axtella, Jamesa Montmarqueta, Lindo Zagzebski, Julio Driver, Jonathana Kvanviga in Christopherja Hookwaya. V zgodovini filozofije pa so z vrlinsko epistemologijo povezana imena kot so Aristotel, Tomaž Akvinski, Thomas Reid, John Dewey in C. S. Peirce.

Aristotel je pojmu *vrline [arete] v* svoji filozofiji namenil pomembno mesto. Glede na delitev razumskega dela duše v širšem smislu je razlikoval med *moralnimi ali nravstvenimi* vrlinami in *intelektualnimi ali razumskimi* vrlinami, (glej npr. *NE* I, 13, za opredelitev vrline pa *NE* II, 6)

> "Na isti način, se delijo tudi vrline: ene imenujemo razumske, druge nravstvene. Med razumske vrline spadajo modrost, bistrost in pametnost, med nravstvene pa plemenitost in umerjenost. Zakaj če govorimo o nraveh, ne pravimo, da so modre ali bistre, ampak da so blage ali umerjene. Modrega človeka hvalimo zaradi njegove lastnosti, hvalevredne lastnosti pa imenujemo vrline." (Aristotel, 2002: 73)

Intelektualne vrline so nam po Aristotelu deloma dane ob rojstvu, deloma pa se jih kasneje lahko naučimo, medtem ko so moralne vrline pridobljene z navado[1]. Oboje so oblikovane v skladu z našo naravo. *(NE* II, 1) Aristotel je ločeval med petimi osnovnimi intelektualnimi vrlinami. Te so:

- umetnost oz. umetnostno in tehnično znanje *[techne]*
- znanje oz. znanstvena modrost *[episteme]*
- pametnost oz. praktična modrost *[phronesis]*
- modrost oz. filozofska modrost *[sophia]*
- razumnost oz. intuitivni razum *[nous]*[2] (*NE* VI, 3)

To je pet osnovnih zmožnosti, po katerih lahko duša doseže spoznanje resnice. Poleg teh Aristotel na nekaterih mestih omenja še druge intelektualne vrline, npr. izkušenost, bistrost, pozornost in še nekatere druge *(NE* VI). Za Aristotela so vrline, tako moralne kot intelektualne, *zadržanja oz. osebnostne lastnosti.* So nekaj, za kar nas lahko označujejo kot dobre ali slabe in za kar lahko prejmemo hvalo ali grajo.[3]

---

[1] Navada tukaj pomeni širši pojem, ki zajema učenje preko navad, običajev, sprejemljivih oblik obnašanja v družbi in prakso, ne pa npr. učenja z razlago, učenja iz knjig ipd.

[2] Umetnostno in tehnično znanje je znanje o tem, kako izdelati stvari oz. zunanje objekte in kako razviti obrt. Znanstvena modrost temelji na znanju tega, kar je nujno in obče. Intuitivni razum vzpostavi prva načela znanja, id so izhodišče vsega, a jih ni mogoče logično dokazati, ampak te intuitivno umeti. Praktična modrost je sposobnost delovati v skladu z vrlino in splošno srečnostjo. Filozofska modrost je kombinacija intuitivnega razuma in znanstvene modrosti, katere predmet so prapočela in nujne ter nespremenljive bitnosti. (glej *NE* VI, 3-13)

[3] "Vse duševne pojave lahko razvrstimo v strasti, zmožnosti in zadržanja; vrlina spada nedvoumno v eno od teh

V sodobne razprave spoznavne teorije je vrlinsko epistemologijo uvrstil Ernest Sosa s svojim člankom "The Raft and the Pyramid" iz leta 1980, kjer je vrlinski epistemologiji namenjena vloga razrešitve tradicionalnih nasprotij med *fundacionalizmom* in *koherentizmom. V* članku z znanima metaforama o piramidi in splavu prikaže strukturi upravičenja, ki nam jih ponujata fundacionalizem in koherentizem. (o tem glej tudi Miščevič, 1999) Pri fundacionalistični strukturi upravičenja, ki si jo lahko predstavimo z modelom na glavo obrnjene piramide, se pojavi problem upravičenja temeljih prepričanj in načel kot temeljnih virov upravičenja, ki ta povezujejo z ostalimi prepričanji. Pri koherentizmu pa se, poleg drugih, pojavi problem upravičenja najbolj primarnih oblik védenja, ki ga pridobimo npr. z zaznavnimi procesi oz. vsakega drugega védenja, ki se pojavi na periferiji "splava" in ki ga z ostalimi prepričanji ne veže mnogo relacij. Na spoznavnih vrlinah temelječi pristop naj bi po Sosovem predlogu pomagal pri preseganju razlik med teorijama in tako odpravil tudi njune pomanjkljivosti. Najprej tak pristop omogoči, da zaznavne procese lahko vidimo kot spoznavne vrline, ki so zanesljive in jim tako podeli dober spoznavni status, za koherenco pa lahko rečemo, da je sama spoznavna vrlina, ker povečuje zanesljivost sistema prepričanj in omogoča prehod k reflektivnemu védenju, (glej Greco, 2002)

Zelo kmalu po objavi Sosovega članka, na začetku osemdesetih let, so se pojavili tudi odzivi na ta njegov predlog, pri čemer sta se v grobem že izoblikovali dve skupini teoretikov. V prvo lahko uvrstimo tiste, ki so v tem predlogu videli predvsem *nadgradnjo standardnega reliabilizma* in dobrodošlo priložnost za združitev prednosti tako eksternalizma kot internalizma v spoznavni teoriji. Tudi Sosa sam je svojo spoznavno teorijo razvijal v tej prvi smeri. Drugo skupino[4] sestavljajo ti, ki so Soso podprli pri ideji spoznavnih vrlin, a mu očitali, da pri svojem poskusu ni šel dovolj da leč, da bi zares izgradil novo vrsto spoznavne teorije temelječo na vrle nah. Slednji so se zgledovali predvsem po vrlinski etiki, ki je v tistem času že bila v vzponu svojega preporoda, segli pa so tudi nazaj k zgledom antične vrlinske teorije.[5]

3. Kaj in katere so spoznavne vrline?

Najenostavneje bi lahko spoznavne vrline opredelili na naslednji način:

Spoznavne (ali intelektualne) vrline so lastnosti (sposobnosti, dispozicije, naučene navade, osebnostne značilnosti ...) posameznikov, ki jim pomagajo pri uresničevanju njihovih spoznavnih ciljev (resničnih prepričanj o svetu, védenja,

zvrsti." (Aristotel, 2002: 82) Ker niti strasti niti zmožnosti niso tiste, zaradi katerih bi nam kot akterjem lahko pripisovali evaluativne lastnosti, vrline ne morejo biti te vrste entitet. "Če torej vrline niso ne strasti ne zmožnosti, je nujno, da so zadržanja." (Aristotel, 2002: 83)

[4] Ker je eden izmed osrednjih pojmov znotraj te teorija spoznavna odgovornost, je pogosto poimenovana tudi responsibilistična vrlinska epistemologija ali kar responsibilizem.

[5] V luči tega dejstva, se slednja skupina teorij pogosto imenuje tudi neoaristotelska vrlinska epistemologija. Imena aristotelska ji ne moremo nadeti, ker *je* v njenem pojmovanju spoznavnih vrlin in njihovega odnosa do moralnih vrlin kar nekaj pomembnih razlik glede na Aristotela, se pa njegovi teoriji približa predvsem z obuditvijo nekaterih pojmov, kot so npr. *eudaimonija* in spoznavno dobro.

znanja, razumevanja, modrosti ...).[6]

Kot smo videli že zgoraj, je že Aristotel razlikoval med moralnimi vrlinami in intelektualnimi vrlinami. Pri prvih naj bi šlo za osebnostne značilnosti, kot sta npr. darežljivost in zmernost, pri drugih pa za npr. filozofsko modrost, razumevanje in praktično modrost. Aristotel je v svoje pojmovanje vrlin že vgradil določeno napetost. Na nekaterih mestih o njih govori kot o čistih *osebnostnih lastnostih in odlikah*, na drugih pa izpostavi tudi njihovo "reliabilistično" plat in jih približa *uspešnim zmožnostim.* (glej tudi 3. opombo in 4. razdelek) Kot bomo videli v sodobni vrlinski epistemologiji, je ta dvojnost dobila še jasnejši izraz. Sedaj pa poskusimo zarisati natančnejšo mejo te dvojnosti.

Kot prvo vrsto pojmovanja spoznavne vrline (pa tudi vrline nasploh) bi lahko oddvojili pojmovanje, ki vrline vidi kot človekove zmožnosti, sposobnosti ali spretnosti, ki so lahko prirojene ali naučene. Teoretiki, ki prevzemajo takšno pojmovanje narave spoznavnih vrlin, kot te omenjajo vrline kot: dobro razvit vid, dober spomin, dobra zmožnost argumentacije, razpoznavanje subtilnih zaznavnih dražljajev ipd. To so nedvomno spoznavne odlike[7], ki pripomorejo k zanesljivosti in uspešnosti naših spoznavnih mehanizmov.

Drugo pojmovanje pa sledi vrlinam kot osebnostnim oz. karakternim značilnostim spoznavnih akterjev v celoti. V tem smislu so spoznavne vrline podobne moralnim (kot so npr. poštenost, pogum, darežljivost). V to skupino spoznavnih vrlin bi lahko uvrstili vrline kot npr. domiselnost, spoznavni pogum, spoznavna odgovornost, ustvarjalnost ipd. Kot vidimo tukaj ne gre za preproste lastnosti "dobro delujočih" spoznavnih mehanizmov spoznavajočega, temveč za bolj kompleksne vrline, pri katerih je pogosto odločujoča tudi *motivacija* zanje oz. za *spoznavna dejanja.* Te vrline pripisujemo spoznavnemu akterju v celoti in njih delovanje si težko zamislimo izven posamičnih spoznavnih situacij.[8] Pri tej vrsti spoznavnih vrlin lahko govorimo zanesljivosti in uspešnosti spoznavnih akterjev. V nadaljevanju bomo videli, da je to razlikovanje v pojmovanju spoznavnih vrlin zelo pomembno za razlikovanje med različnimi vrstami vrlinske

[6] Seveda se opredelitev spoznavne vrline ali spoznavnih vrlin od avtorja do avtorja razlikuje, kakor se razlikuje tudi spoznavna vloga teh vrlin. Tukaj sem poskušal podati kar najbolj široko in univerzalno opredelitev.

[7] "Npr.: lahko gre za sposobnost dobrega vida nekoga pri dnevni svetlobi, ki proizvede prepričanje o beli barvi in okroglosti snežene kepe, ki se mu bliža. Ali je posedovanje takšne sposobnosti 'vrlina'? Seveda ne v ozkem Aristotelovem smislu, ker ne gre za dispozicijo za premišljene izbire. Toda tu je tudi širši, a še vedno tudi grški, smisel 'vrline', kot nečesa, kar lahko pripišemo vsaki stvari z določeno funkcijo, naravno ali umetno. Tudi oko ima potemtakem svoje vrline, kot jih ima nož. In če vključimo dojetje resnice o posameznikovem okolju med prave cilje človekovega obstoja, potem lahko sposobnost vida dojemamo kot človekovo vrlino v širšem smislu. In če je dojetje resnice stvar spoznanja, potem je ta vrlina v enostavnem smislu tudi spoznavna vrlina." (Sosa, 1991: 271; navedeno po Brady in Prichard, 2003: 3)

[8] Kakor pri podobnih moralnih vrlinah (npr. pogumu ali darežljivosti) je za pripis le teh pomembna določena konstantnost v našem delovanju in pravilna motivacija v ozadju naših moralnih ali spoznavnih dejanj. Okoliščine teh dejanj so za pripis prav tako pomembne (npr. poguma ne moremo izkazovati pogosto, darežljivost je odvisna od tega, kolikor dobrin imamo sami ipd.). Spoznavna odgovornost terja pravilne odločitve v določeni situaciji, ki je za nas npr. moralno in spoznavno pomembna. Tudi glede na zdravo-razumski pogled na moralnost je pojem dobre osebe mnogo širši od zgolj moralne pravilnosti njenih dejanj in mnogokrat vključuje določena občutja, čustva in naravnanosti, ki jih tradicionalne moralne teorije usmerjane zgolj na značilnosti dejanj težko pojasnijo.

epistemologije.[9]

Sodobne vrlinske teorije imajo med drugim tudi ta cilj, da bi *poenotile* seznam spoznavnih vrlin. Aristotelov seznam spoznavnih vrlin smo si že predstavili, tukaj pa si bomo pogledali dva izmed teh poskusov, in sicer seznam spoznavnih vrlin, ki ga je oblikoval James A. Montmarquet, in seznam, ki ga je svojo pluralistično teorijo spoznavnih vrlin vključila Juli Eflin.

Montmarquet govori o štirih osrednjih spoznavnih vrlinah:

– spoznavna želja po resnici in izogibu zmoti *[epistemic conscientiousness]*
– nepristranskost *[impartiality]*
– intelektualna treznost *[intellectual sobriety]*
– intelektualni pogum *[intellectual courage]* (Montmarquet, 2003)

Prva vrlina je torej želja po resnici in izogibu zmoti; druga, nepristranskost, vključuje željo po izmenjavi idej in učenju od drugih, brez da bi bili enim bolj naklonjeni kot drugim ali bili zaverovani s svojo nezmotljivost. Pri tretji, intelektualni treznosti, gre za pazljivost in previdnost ter izogibanje kapricam ali nerazumnim strastem do novih idej in spoznanj, ki bi lahko zasenčile željo po resnici. Intelektualni pogum vključuje pripravljenost za dojemanje stvari tudi na drugačne načine, raziskovanje alternativ in določeno nepopustljivost, ko gre za zagovarjanje svojih stališč. Motivacija, ki je v ozadju vseh teh vrlin, je za Montmarqueta želja po razumevanju, (glej Montmarquet, 2003)

Eflinova kot osrednje spoznavne vrline izpostavi naslednje:

– sinoptični dojem oz. sinoptično razumevanje *[synoptic grasp]*
– smisel za alternativen pogled (tako v zaznavnem kot v teoretičnem smislu) *[sense for alternative points of view (perceptual and theoretical)]*
– spoznavna želja kot motivacijsko jedro *[epistemic conscientiousness based motivation]*
– smisel za usmerjanje na bistveno *[sailence recognition and focus]*
– praktična modrost *[ability to exercise the right epistemic virtue at the right time, practical wisdom]* (Eflin, 2003: 63)

Eflinova dodaja, da lahko posamezne spoznavne vrline izdvojimo, a da le te nikoli ne delujejo izolirane od drugih, ampak morajo nastopati medsebojno uravnoteženo in spoznavni situaciji primerno. Prva spoznavna vrlina tako sestoji iz treh drugih vrlin: ozadnjem znanju o določenem področju, usmerjanjem na

[9] Poleg teh dveh vrst spoznavnih vrlin bi lahko o slednjih govorili vsaj še v dveh drugih smislih. Prvi se veže na spoznavne *vrtine družbenih skupin oz. skupnosti*, ki jih ne moremo nujno reducirati na vrline, ki jih srečujemo pri posameznikih. Skupnostim bi lahko tako pripisali vrline, kot so: odprtost, spodbujanje razprave, podpiranje raziskovanja, ohranjanje znanja ipd. V tej smeri je vrlinsko epistemologijo razvijal Jonathan Kvanvig. (o tem glej Greco, 2002) Drugi pa je smisel, ki spoznavne vrline povezuje z *značilnostmi znanstvenih teorij*, ki jim prav tako lahko pripisujemo določene odlike.

prave probleme in zmožnostjo asociiranja, vključuje pa tudi sposobnost abstrakcije. Druga spoznavna vrlina na njenem seznamu je tista, ki bi jo lahko poimenovali tudi kreativnost oz. ustvarjalnost, vključuje pa sposobnost alternativnih pogledov na probleme in hipotetičnega razmišljanja. Strukturo tretje sestavljajo težnja po védenju, intelektualni pogum in intelektualna treznost. Pri slednji vrlini pa gre za to, da v situaciji znamo izbrati primerno in ustrezno spoznavno vrlino ali uravnoteženo kombinacijo le teh.[10] (Eflin, 2003: 62-64)

Če sklenemo, lahko kot najpogostejše spoznavne vrline, ki se pojavljajo v različnih teorijah, tako izpostavimo naslednje: oster vid, dober spomin, dobro razvita čutila, dobro razvite sposobnosti logičnega mišljenja oz. argumentiranja, bistroumnost, sposobnost abstrakcije, zaznavna občutljivost, sinoptični dojem oz. sinoptično razumevanje, domiselnost, ustvarjalnost, smisel za pomembnost in za bistveno, smisel za alternativen pogled (tako v zaznavnem kot v teoretičnem smislu), usmerjenost k pravim oz. pomembnim ciljem, intelektualni pogum, vztrajnost, pazljivost, praktična modrost, intelektualna integriteta, potrpežljivost, intelektualna temeljitost, intelektualna treznost, intelektualna poštenost, nepristranskost, spoznavna želja po resnici in izogibu zmoti, spoznavna odgovornost.[11]

4. Reliabilizem in responsibilizem

V sodobni literaturi oz. pregledih vrlinske epistemologije (npr. Greco, 2002) najpogosteje zasledimo delitev vrlinske epistemologije v dve osnovni obliki, in sicer (a) *reliabilizem* in (b) *responsibilizem.* Pri prvi obliki seveda ne gre za spoznavni reliabilizem na splošno, ampak samo za tiste njegove različice, ki védenje opredeljujejo tudi v odnosu do spoznavnih vrlin. Kakor nam razkrivata že imeni teh teorij, gre pri reliabilizmu za spoznavno teorijo, ki upravičenje naših prepričanj vidi v *zanesljivosti spoznavnih mehanizmov*, s katerimi smo tvorili določeno prepričanje. Ti spoznavni mehanizmi so stabilne in uspešne dispozicije za zanesljivo doseganje nekega dobrega, v tem primeru spoznavnih dober oz. spoznavnih ciljev, kot so npr. resnično prepričanje, védenje, razumevanje ali modrost, in so v tem smislu spoznavne vrline. Ta oblika teorij je uspešna pri pojasnjevanju upravičenja, npr. temeljnih zaznavnih prepričanj, in ji ni nujno potrebno posegati k pojmovanju prepričanja kot dejanja. Standardnemu

[10] Eflinova ima z *pluralistično vrlinsko epistemologijo v* mislih teorijo spoznavnih vrlin, ki ni nujno niti *hierarhična* niti *sklenjena.* Pri hierarhičnosti gre za to, da lahko osnovne spoznavne pojme uredimo (z spoznavnimi načeli, pravili in opredelitvami) od najbolj temeljnih in osrednjih do tistih, ki so vzpostavljeni v odnosu do prvih. Sklenjenost neke teorije za Eflinovo pomeni, da lahko le-ta pojasni vse možne primere spoznavnih dejanj in jim določi njihovo vrednost. Eflinova pravi, da so prav gotovo vse tradicionalne spoznavne teorije usmerjene k zadovoljitvi obeh kriterijev in da podobno težnjo kažejo tudi nekateri vrlinski teoretiki, tudi Linda Zagzebski. Če pogledamo na polje moralne filozofije, potem lahko rečemo, da poskuša Eflinova vzpostaviti nekakšen korelat pluralistični etiki, kot jo je zagovarjal npr. W. D. Ross, kjer gre prav tako za zanikanja hierarhičnosti in v manjši meri tudi sklenjenosti moralnih dolžnosti oziroma moralnih načel. Pomembna razlika pa je v tem, da je Rossova pluralistična moralna teorija usmerjena k lastnostim dejanj in le v manjši meri k moralnim akterjem in njihovim vrlinam. (o tem glej Eflin, 2003: 52-63, in Ross, 1930).

[11] Vrline sem tukaj poskušal razvrstiti od tistih, ki se najpogosteje pojavljajo v reliabilistični vrlinski epistemologiji, do tistih, ki sodijo bolj v responsibilizem.

reliabilističnemu pogoju zanesljivosti se v vrlinski različici reliabilizma priključi še pogoj, da mora biti resničnost prepričanja oz. zanesljivost mehanizmov, ki so do nje privedli, *akterjev* kognitivni dosežek.

> "Temeljna ideja v ozadju teh vrst reliabilizma akterja, kot je Grecova ali podobne teorije Sose, poznega Goldmana in drugih, je v omejitvi obsega zanesljivih procesov, ki lahko podpirajo védenje, na način, da se usmerijo na stabilne in uspešne dispozicije akterjev, ki so pomembne za védenje in upravičenje – na njihove intelektualne zmožnosti." (Prichard, 2003: 110)

Pri responsibilizmu pa so v ospredju predvsem pojem spoznavne odgovornosti spoznavajočega subjekta in drugi pojmi, ki so bolj tesno kot zanesljivost procesov povezani s spoznavnim akterjem. Spoznavne vrline so pojmovane kot *dobre* (v smislu spoznavnega dobra) *osebnostne značiln* prepričanje pa se približuje pojmu dejanja oz. delovanja.[12]

> *Vrlinski reliabilizem: védenje* je resnično prepričanje, ki je rezultat spoznavne vrline; *spoznavne vrline* so stabilni kognitivni procesi in mehanizmi, ki spoznavajočemu zagotavljajo visok delež resničnih prepričanj (npr. zaznavni aparat, spomin, introspekcija ipd.). (Sosa, Greco)
>
> *Vrlinski responsibilizem: védenje* je prepričanje, ki izvira iz spoznavno vrlih dejanj; *spoznavne vrline* so dobre intelektualne značilnosti spoznavajočega (npr. spoznavna odgovornost, pazljivost, premišljenost, pogum ipd.). (Zagzebski, Hookway)

Zgornje razlikovanje pa bi lahko opisali tudi na drugačen način, ki še z druge plati osvetli vlogo spoznavnih vrlin pri opredelitvi védenja. Simon Blackburn predlaga, da moramo razlikovati med (i) *šibkimi* in (ii) *močnimi* različicami vrlinskih spoznavnih teorij. (glej npr. njegov članek "Reason, Virtue, and Knowledge", 2001) Pri šibki vrlinski epistemologiji gre, v smislu pojmovnega primata, za zgolj sekundarno vlogo spoznavnih vrlin, ki v teorijo vstopijo naknadno, ko smo poprej že opredelili pojme, kot so resnica, védenje in upravičenje. Spoznave vrline so nato pogosto opredeljene kot te, ki z določeno zanesljivostjo postavijo naša prepričanja v odnos do resnice.[13] Za razliko od takšnih teorij močne vrlinske teorije vidijo spoznavne vrline kot osrednji pojem in glede na njih opredelijo pojem pojme resnice, védenja in upravičenja.[14]

---

[12] Kot eno izmed prednosti responsibilizma pred reliabilizmom Zagzebska izpostavi odgovor na problem vrednosti. Gre za problem, kaj naredi védenje vrednejše od zgolj resničnega prepričanja. "Glede na standardne reliabilistične modele je védenje resnično upravičeno prepričanje, ki je rezultat zanesljivih kognitivnih *[belief-forming]* procesov in sposobnosti. Toda zanesljivost izvora prepričanja ne more pojasniti razlike v vrednosti med védenjem in resničnim prepričanjem. Eden izmed razlogov je v tem, da zanesljivost per se nima vrednosti ali nevrednosti *[disvalue]*." 2003: 13; slovenski prevod: 85) Zagzebska predlaga, da lahko problem rešimo na način, da na spoznavna dejanja pričnemo gledati podobno kot na moralna dejanja. Kakor lahko moralno pomembna dejanja dobijo evaluativne lastnosti od akterja (npr. njegovih motivov), lahko tudi prepričanje dobi evaluativne lastnosti, in s tem dodatno vrednost, ki jo iščemo, od značilnosti spoznavnega akterja. (o tem glej Zagzebski, 2003)

[13] Blackburn pravi, da so takšne teorije preprosto prešibke in so zgolj "figov list", ki prekrije navaden reliabilizem. (glej Blackburn, 2001: 15-18)

[14] V literaturi tako srečamo tudi pojma vrlinsko osredinjena *[virtue-centered]* teorija in k vrlinam orientirana *[virtue-oriented]* teorija, kar se v grobem sklada z zgornjo delitvijo med močnimi in šibkimi vrlinskimi

Zagzebska tako teorijam te vrste pravi tudi "čiste" vrlinske teorije in jih opredeli na naslednji način:

> "S čisto vrlinsko teorijo *[pure virtue theory]* imam v mislih teorijo, pri kateri je pojem moralno pravilnega dejanja izpeljan iz pojma vrline ali nekega drugega notranjega stanja osebe, ki sestavlja vrlino. To je mišljeno tako v smislu pojmovne primarnosti kot moralne ontologije. Čista vrlinska teorija opredeli pojem moralno pravilnega dejanja s pomočjo pojma vrline ali gradnika vrline, kot je npr. motivacija. Nadalje, lastnost moralne pravilnosti *[rightness]* je nekaj, kar vznikne iz notranjih značilnosti oseb." (Zagzebski, 1996: 79)

Čeprav tukaj Zagzebska govori o moralni teoriji in moralnih vrlinah, velja to tudi za področje spoznanja, saj je za Zagzebsko védenje isto kot delovanje. Blackburn v luči zgornjega razlikovanja primerja vrlinsko etiko z vrlinsko epistemologijo in poda naslednjih pet opredelitev, ki jih lahko najdemo znotraj njiju:

(1) Dejanje proizvede (ali teži k temu, da proizvede) največji preostanek koristi oz. dobrih posledic nad slabimi kot katera koli druga alternativa, če in samo če je to dejanje, ki bi ga izvršil vrl akter.

(2) Dejanje je moralno pravilno v določenih okoliščinah, če in samo če je to dejanje, ki bi ga v teh okoliščinah izvršil vrl akter.

(3) Propozicija je upravičena v okoliščinah *C*, če in samo če bi jo v teh okoliščinah kot upravičeno oz. verjetno smatral spoznavno vrl akter.

(4) *S* ima védenje o resnični propoziciji v okoliščinah *C*, če in samo če *S* v okoliščinah *C* pri njenem sprejemanju izkazuje spoznavne vrline.

(5) Propozicija je resnična, če in samo če bi jo sprejemal spoznavno vrl akter v primeru, da je pravilno uporabil spoznavne vrline. (Blackburn, 2001: 16-17)

Blackburn nadaljuje, da šibke vrlinske teorije, tako moralne kot spoznavne, berejo zgornje trditve *od leve proti desni*, da pa je edino zanimivo branje, če se hoče neka teorija zares označiti kot vrlinska, tisto, ki ga ponujajo močne vrlinske teorije, in sicer *od desne proti levi*. Tukaj gre za vprašanje *pojmovnega primata*. Šibke vrlinske teorije morda resda na prvi pogled ne ponujajo zanimivih in neodvisnih rešitev spoznavnih problemov, vprašanje pa je ali so močne vrlinske teorije sploh dobra alternativa. Od zgornjih trditev se po "močnem branju", tj. od desne proti levi strani, zdita kolikor toliko sprejemljivi le trditvi (3) in (2). Pri vseh ostalih trditvah pojmu spoznavne ali moralne vrline naložimo preveliko breme, saj ima tako celo primat pred pojmi, kot so resničnost, védenje in korist. To pa je stališče, ki ga je izrazito težko zagovarjati in ubraniti. Na težave namreč naletimo pri iskanju zadostne in neodvisne utemeljitve posameznih spoznavnih, in tudi moralnih, vrlin, ki bi zaobšla te pojme.

---

teorijami.

5. Vzporednice med vrlinsko etiko in vrlinsko epistemologijo

Med *vrlinsko epistemologijo* in *vrlinsko etiko* lahko najdemo kar nekaj stičnih točk. Obe seveda gradita na pojmu vrline in ga skušata poenotiti, kar bomo obravnavali na koncu tega razdelka. Pri obeh se težišče raziskovanja od lastnosti *prepričanj in dejanj* prenese k značilnostim spoznavnih in moralnih *akterjev*. Tu pa so še druge podobnosti.

Najprej je tu odnos do tradicionalnih oz. standardnih teorij, tako v epistemologiji kot v etiki. Vrlinska etika je svoj ponovni vzpon začela nekoliko prej kot vrlinska epistemologija, v zadnji četrtini dvajsetega stoletja. Stališčema, kot sta npr. kantovska etika ali utilitarizem, so vrlinski teoretiki očitali, da se primarno osredotočata na težke *moralne dileme* in poskušata podati na *načelih in pravilih* temelječe teorije normativne presoje < ovrednotenja dejanj. Podobna ugotovitev drži tudi na področju epistemologije, kjer mnogi vrlinski teoretiki očitajo standardnim spoznavnim teorijam (npr. fundacionalizmu in koherentizmu), osredotočanje na idealnega spoznavalca, skeptične izzive in opredelitev *nujnih in zadostnih pogojev za znanje.* Pri tem so v ozadju ostajale teme, kot npr. razumevanje, končni spoznavni smoter, spoznavne vrline, ki bi lahko vodile tudi do teorije uporabne za vsakdanjega spoznavalca, ki ni le zmotljiv, pač pa tudi operira znotraj posamičnih in omejenih ciljev, spretnosti, želj in okoliščin. (glej Eflin, 2003: 49-52) Kakor je vrlinska etika prestavila težišče zanimanja od značilnosti dejanj k značilnostim moralnih akterjev, tako je tudi vrlinska epistemologija prvenstveno osredotočena na spoznavne akterje in njihove značilnosti, šele v drugi vrsti pa na spoznavna dejanja oz. prepričanja.

Pomembno je tudi vprašanje povezanosti moralnih in spoznavnih vrlin. Mnogo vrlinskih teoretikov trdi, da sta kar najtesneje povezani *spoznavna in moralna odgovornost.* Tako npr. lahko slednji temelji na prva, če gre za primer praktičnega oz. moralnega dejanja, ki je v temelju odvisen od nekaterih resničnih prepričanj in torej terja spoznavno odgovornost. To povezanost je v *Nikomahovi etiki* izpostavil že Aristotel.

> "Nravstvena vrlina je zadržanje, ki je usmerjeno k neki odločitvi, odločitev pa je premišljeno hotenje. Iz tega sledi: če naj bo odločitev dobra, mora biti misel resnična in hotenje pravilno ter hotenje mora biti usmerjeno k temu, kar misel zatrjuje."[15] (Aristotel, 2002: 187)

Linda Zagzebski je povezanost moralne in spoznavne odgovornosti opredelila preko pojma *nedopustnosti:*

> "Ko je v igri nekaj, kar je moralno pomembno pri izvršitvi dejanja S; in če je S primer delovanja glede na prepričanje *B*, potem je moralno pomembno, da je *B* resnično. Tako je nedopustno za akterja biti prepričan

[15] V nadaljevanju Aristotel z moralnim delovanjem poveže predvsem pametnost oz. praktično modrost "Iz naših izvajanj je jasno, da sploh ni mogoče biti dober brez pametnosti in da tudi ni mogoče biti pameten brez nravstvene vrline." (Aristotel, 2002: 205)

> na način, ki mu ne uspe spoštovati pomembnosti resničnosti *B.*" (Zagzebski, 2003: 18-19; slovenski prevod: 91)

Pri vprašanju povezanosti spoznavnih in moralnih vrlinskih teorij naletimo tudi na "problem sovpadanja" *[conflation problem]* moralnih in spoznavnih vrlin.

> "Mnogo vrlinskih teorij stremi k neki vrsti poenotenja. Tj., obstaja ena ali več skupnih značilnosti sposobnosti, ki jih narede za 'vrline' oz. osebnostne 'odlike'. Problem sovpadanja se pojavi, ko takšna vrsta teorije ne more več ohraniti intuitivnih razlik med tipi vrlin. Nekateri avtorji tega ne dojemajo kot problem, a mnogo vrlinskih teoretikov razlikuje med moralnimi in intelektualnimi ali spoznavnimi vrlinami. Ideja je v tem, da lahko ljudje izkazujejo več različnih vrst odlik, od katerih so nekatere po svoji naravi intelektualne, druge pa bolj socialne. Primeri intelektualnih ali spoznavnih vrlin so takšne značilnosti kot inteligentnost in odprtost, moralnih pa npr. radodarnost, dobrodelnost in poštenost. Ne glede na to, da lahko v nekaterih primerih pride do interakcije ali celo delne odvisnosti med teme dvema skupinama, sta ne glede na vse docela različni." (Driver, 2003: 367)

Takoj ko se oddaljimo od strogega reliabilističnega pojmovanja spoznavnih vrlin kot zanesljivih in uspešnih mehanizmov se znajdemo sredi ti. problema sovpadanja (glej npr. Driver, 2003) (spoznavnih in moralnih vrlin). Če so vrline osebnostne značilnosti, ki temeljijo na značilni motivaciji, namreč izgubimo *temelj za razlikovanje* med obema vrstama vrlin, za katerega se intuitivno zdi, da vendarle obstaja. Poglejmo nekaj načinov, na katere so vrlinski teoretiki poskušali razlikovati med obema vrstama vrlin.

kriterij razlikovanja	moralne vrline	spoznavne vrline
volja	prostovoljne	naravne
izvor vrlin	navada	učenje
deli duše	nerazumske/emocionalne	razumske
orientacija	"k drugim"	"k sebi"
motivacija	skrb za druge	skrb zase, lastni interes
smotri, cilji	moralni	spoznavni
možne posledice	le dobre	dobre ali slabe

Poglejmo si nekaj težav, s katerimi se srečajo zgornji poskusi razlikovanja. Moralne vrline naj bi bile pomembnejše od spoznavnih, ker naj bi zavisele od *volje akterja*, *v* smislu, da so prostovoljne in temeljijo na svobodni odločitvi. Na drugi strani naj bi bile spoznave vrline naravne. Vendar se zdi, da so lahko tudi nekatere moralne vrline v človekovi naravi (npr. darežljivost) in da lahko gre pri nekaterih spoznavnih vrlinah za zavestno odločitev in rezultat trdega dela. Aristotel je kot točko razlikovanja izpostavil *izvor vrlin*, pri čemer naj bi bile

spoznavne prirojene ali pridobljene z učenjem, moralne pa pridobljene z navado in vršenjem vrlih dejanj. Toda tudi pri tem kriteriju lahko zlahka najdemo primere vrlin, za katere to ne velja nujno (npr. analitično in argumentativno razmišljanje kot spoznavna vrlina, ki jo lahko pridobimo z navado oz. ponavljanjem in analiziranjem argumentov). Aristotel je poleg tega postavil še drug kriterij, to je kriterij glede na *del duše,* iz katerega vrline izvirajo. Moralne vrline izvirajo iz nerazumskega dela duše, spoznavne pa iz razumskega (NE VI, 2) Toda nekatere spoznavne vrline lahko gotovo vsebujejo določene emocionalne komponente (npr. radovednost, ustvarjalnost). Morda boljši kriterij za razlikovanje *orientiranost vrlin.* Moralne vrline naj bi bi usmerjene k drugih moralnim subjektom, medtem ko naj bi bile spozna; ne vrline usmerjene k nam samim kot spoznavajočim. Temu kriteriju lahko pridružimo tudi kriterij *motivacije,* saj naj bi bila pri vrlinah prve vrste značilna motivacija, skrb za druge oz. benevolenca, pri slednjih pa lastni interes ali pa želja po resnici.[16] Vendar pa se tudi ta kombinirani kriterij kaj hitro izkaže za pomanjkljivega. Če priznavamo, da spoznavna dejavnost prispeva k občečloveškemu interesu in razvoju, potem moramo tudi spoznavnim vrlinam priznati orientiranost, ki je širša od lastnega interesa.[17] Naslednja točka razlikovanja bi lahko bili *cilji in smotri* vrlin, pod pogojem seveda, če bi nam uspelo najti značilno moralne in značilno spoznavne cilje. Julia Driver trdi, da je to najboljši kriterij razlikovanja in s tem najboljša rešitev problema sovpadanja.

> "Spoznavno in moralno dobro, ali cilja dveh vrst vrlin, se zdita različna. Tukaj moramo razlikovati med ciljem oz. smotrom vrline in tem, kar ima pri njenem udejanjanju v mislih akter, saj ne gre za eno in isto. Ko govorim o cilju vrline, mislim na tisti dejavnik, ki je odgovoren, da ima določena osebnostna lastnost sploh status vrline. Trdim, da moralne in spoznavne vrline črpajo svojo vrednost iz različnih ozirov oz. nagibov – upravičenega prepričanja ali resnice v primeru prvih in srečnosti ali dobrobiti drugih ljudi v primeru slednjih." (Driver, 2003: 377)

Še en poskus razlikovanja, ki ga je zagovarjal tudi Kant (Kant, 2001, uvoden razdelek 1. dela), pa se opira na *posledice dejanj,* ki izvirajo iz obeh vrlin. Dejanja, ki imajo v podlagi moralno vrlino ali moralne vrline, naj bi vodile le do dobrih dejanj, medtem ko bil lahko akter spoznavne vrline izkoristil tudi v zle namene in za slaba dejanja.[18]

6. Ali se vrlinska epistemologije izogne gettierovskim problemom?

[16] Pomembno je, da ne glede na to, da želje po resnici ne moremo enačiti z lastnim interesom, pri tej želji ne gre za motivacijo podobno tisti pri skrbi za druge.

[17] Linda Zagzebski nasprotuje drugi komponenti te strategije, tj. razliki v motivaciji. Trdi namreč, da obe vrsti vrlin v temelju izhajata iz iste vrste motivacije, tj. iz *želje po srečnem življenju in ljubezni do sveta.* (glej Zagzebski, 1996 in 2003).

[18] Kot protiprimer bi lahko tu navedli znan primer, ali je govoriti resnico (v tem primeru moramo reči "biti odkritosrčen in pošten") vedno nujno dobro dejanje, tudi ko so posledice nedvoumno slabe, npr. v primeru, ko "nacističnim oficirjem" povemo resnico o skrivališču judovskih družin.

Kot eden izmed motivov za sprejemanje vrlinske epistemologije se večkrat omenja tudi ta, da se tako *izognemo gettierovskim problemom.* V tem razdelku si bomo najprej ogledali splošno strukturo gettierovskih problemov, potem pa težave, ki jih srečamo pri ostalih teorijah upravičenja prepričanj, ki ne temeljijo na spoznavnih vrlinah. Na koncu si bomo zastavili vprašanje, ali se vrlinska spoznavna teorija res izogne omenjenim problemom.

Najprej si oglejmo standardno opredelitev védenja in način, na katerega se nanjo vežejo gettierovski problemi. Védenje lahko opredelimo kot resnično upravičeno prepričanje, kar pomeni, da subjekt *S ve,* da P, če in samo če:

– je *P* resnično;

– ima *S* ustrezen, kognitiven oziroma prepričanjski stik s *P*;

– je *S* upravičeno prepričan, da *P*, oziroma, če ima *S* dobre spoznavne razloge, da *P*.

Edmund Gettier je leta 1963 objavil članek z naslovom "Ali je upravičeno resnično prepričanje védenje?", v katerem je izpostavil nekaj tipov primerov, kjer so izpolnjeni vsi zgornji pogoji za védenje, ampak pri katerih *S*-u intuitivno ne bi pripisali pravega védenja. Ti primeri tako razkrijejo, da zgoraj omenjeni pogoji (resničnost, prepričanost in upravičenost) tako niso *zadostni pogoji za védenje.*[19] Do gettierovskih problemov pride, ko se pojavi zgolj *naključna povezava* med resničnostjo in upravičenjem določenega prepričanja, ne glede na to, na kakšno teorijo upravičenja pristajamo.

Če pogledamo najprej *eksternalistično* teorijo upravičenja, kot je npr. *reliabilizem*, lahko vidimo, da se le ta gettierovskim problemom ne izogne. Reliabilizem trdi, da je prepričanje upravičeno, če je rezultat zanesljivih kognitivnih procesov, ki nas v večini primerov privedejo do resničnih prepričanj o svetu. Ker pa ni nobenih zagotovil, da ne more priti do brezna med temi zanesljivimi procesi in resničnostjo, se reliabilizem ne more izogniti gettierovskim problemom. Za primer lahko vzamemo pokrajino z navideznimi skednji, ki so jih prebivalci postavili ob cesti, da bi zavedli popotnike. Fasade skednjev so gledane s ceste povsem nerazpoznavne od dejanskih skednjev, ki pa jih je izredno malo. Če torej popotnik Peter (katerega vid je ob normalnih pogojih oster in zanesljiv), ki po naključju s ceste gleda dejanski skedenj, tvori prepričanje "Tamle stoji zanimiv skedenj", v primeru tega prepričanja ne gre za védenje, saj bi isto mislil tudi, če bi gledal zgolj fasado skednja, in le golo naključje in dejanska sreča je, da je v tem primeru njegovo prepričanje tudi resnično.

[19] Najenostavnejši primer gettierovskega problema je verjetno primer *ustavljene ure.* Osebi A se, sicer zanesljiva in kvalitetna, ura zaustavi natanko ob petih zjutraj. Oseba A popoldan ob petih pogleda na uro, in zatrdi: "Vem, da je ura pet." Njeno prepričanje je resnično, saj je ura res pet popoldan. Oseba ima ustrezno prepričanje in je vsled zanesljivosti ure tudi upravičeno prepričana, da je ura pet. Pa vendar se nam zdi, da oseba A ne ve zares, da je ura pet, saj je zgolj po naključju nanjo pogledala ob petih popoldne, če bi pogledala deset minut prej ali kasneje, njeno prepričanje ne bi bilo resnično. Za primere, ki jih navaja Gettier, glej Gettier (1963). Glej tudi npr. Ule (2001: 122-137).

Tradicionalne *internalistične* teorije upravičenja se ne odrežejo nič bolje. Tudi če so sami razlogi dostopni prepričancu, namreč ni izključeno, da ne obstaja kakšen vidik spoznavne situacije, ki mu ni dostopen, in tukaj je mesto, kjer lahko v internalizem vstopijo gettierovski problemi. To velja tako za *fundacionalizem* kot za *koherentizem.* Za primer lahko vzamemo kar izviren Gettierov primer gospodov Smitha in Jonesa. Odkar Smith pomni, je Jones imel vozilo znamke Ford. Smith ga je venomer videval vozečega se s tem vozilom, Jones je govoril o njem ipd. Zamislimo si tudi, da ima Smith prijatelja Browna, ki ga že dolgo ni videl in nima niti najmanjše ideje, kje bi Brown lahko bil. Na osnovi tega Jones privzame naslednja tri prepričanja:

Jones ima Forda ali pa je Brown v Bostonu.

Jones ima Forda ali pa je Brown v Barceloni.

Jones ima Forda ali pa je Brown v Brest-Litovsku.

Vsako od teh prepričanj sledi iz upravičenega prepričanja, da ima Jones Forda in Smith je tako upravičeno prepričan tudi v (a), (b) in (c). Toda po naključju, Jones sploh nima svojega avtomobila, ampak si ga sposoja od tete. In nadalje, izkaže se, da je Brown zares v Barceloni. Od zgornjih treh prepričanj se tako (b) izkaže za resnično. Vseeno pa se nam intuitivno zdi, da Smith ne ve, da (b), čeprav so izpolnjeni vsi pogoji za védenje: (b) je resnično, Smith je prepričan, da (b), in ima za to upravičena (utemeljena, skladna in njemu notranje dostopna) zagotovila.

Linda Zagzebski v svojem članku o *neizogibnosti gettierovskih problemov* (1994) pravi, da lahko na podlagi teh in podobnih premislekov oblikujemo splošnejšo formulo gettierovskih problemov, katerim podležejo vse vrste spoznavnih teorij, ki obravnavajo védenje kot resnično prepričanje in mu dodajo še dodaten pogoj. védenje = resnično prepričanje + *X*

Kot smo videli v zgornjih primerih, reliabilizem, fundacionalizem in koherentizem kot tretji pogoj dodajo upravičenje, ki pa ga različno pojmujejo (prvi kot *zanesljivost*, drugi kot *utemeljenost*, tretji pa kot *usklajenost).* Recept za gettierovske probleme je sedaj naslednji: v vsakem primeru, ko bo obstajala še tako majhna stopnja neodvisnosti med resničnostjo nekega prepričanja in X-om, tj. tretjim pogojem za védenje, si (lahko) zamislimo napačno, a upravičeno prepričanje. Tu seveda še ne gre za gettierovski problem. Toda sedaj sledi ključen korak. Neresničnost prepričanja ne sme biti sistematično povezana s katero od značilnosti spoznavne situacije (ker bi v nasprotnem, to lahko upoštevali v naši splošni opredelitvi védenja), ampak zgolj po srečnem oz. nesrečnem naključju (na primer, da je naša ura pokvarjena ali da smo v pokrajini s fasadami skednjev). Sedaj dodamo še eno, tokrat bolj srečno (za spoznave teoretike pa verjetno nesrečno) okoliščino[20], ki zopet neodvisno od vsega naredi prvotno prepričanje resnično (na primer, da pogledamo na uro natanko dvanajst

[20] V ozadju gettierovskih problemom je torej ti. struktura *dvojne sreče.*

ur kasneje, kot se je ustavila; ali da smo se zagledali v enega od redkih pravih skednjev; ali da je Brown zares v Barceloni). Tako dobimo gettierovske probleme. (glej Zagzebski, 1994: 69)[21]

Kaj pa lahko sedaj rečemo za vrlinsko epistemologijo in ali tudi ta zapade zgoraj opisanim problemom? Oglejmo si nekaj opredelitev védenja, ki izvirajo iz vrlinskega spoznavnoteoretskega stališča:

- védenje = (resnično) prepričanje, ki je bilo tvorjeno/sprejeto na podlagi dejanj spoznavne vrline oz. spoznavnih vrlin;
- védenje = stanje spoznavnega stika z realnostjo, ki izvira iz dejanj spoznavne vrline;
- *S* ima upravičeno prepričanje, da *p*, če in samo če:

– *S* je motiviran s strani spoznavnih vrlin,

– *S* ve vse relevanten informacije, ki bi jih vedela spoznavno vrla oseba

in

– če bi spoznavno vrla oseba bila prepričana, da *p* (v teh okoliščinah).

Ali lahko tudi tukaj najdemo mesto za gettierovske probleme in spoznavno srečo? Najprej poglejmo k eksternalistični različici vrlinske epistemologije, k reliabilizmu akterja.[22] Tukaj gre za razliko od navadnega reliabilizma, kjer so v igri za upravičenje in védenje zanesljivi procesi (kot izvor prepričanj), za usmeritev bolj k značilnostim spoznavajočega subjekta, k njegovim zanesljivim in stabilnim intelektualnim sposobnostim in zmožnostim. Te naj bi bile takšne, da so zanesljive tudi v vseh bližnjih možnih svetovih, in tako se izognemo določeni množici gettierovskih primerov, npr. primeru sklepanja po *hazarderjevi zmoti* v svetu benevolentnega demona[23], saj ne gre za zanesljivost procesov, temveč za zanesljivost spoznavnih akterjev.[24] Čuden in bežen kognitivni proces, ki je po nekem srečnem naključju zanesljiv v subjektovem aktualnem svetu, odpove pa v vseh bližnjih svetovih, ne more veljati za spoznavno vrlino. Podoben je odgovor internalističnega responsibilizma, kjer je za doseganje védenja potrebno dejanje spoznavne vrline in ne zgolj naključje. Moje prepričanje o tem, da je ura pet (ob gledanju na pokvarjeno uro natanko ob petih) v tem primeru ne zadosti vsem

[21] Zagzebska predlaga dve strategiji za ogib gettierovskih problemov. Prva je, da upravičenje povsem vežemo na resničnost in njegove pogoje postavimo tako, da jih ne more zadovoljiti nobeno neresnično prepričanje. Druga strategija pa gre v nasprotni smeri večanja neodvisnosti med resničnostjo in upravičenjem. To bi imelo za posledico, da bi bili gettierovski primeri povečini smatrani kot védenje, saj bi spoznavna sreča in spoznavno naključje igrala vlogo pri večini včdenja.

[22] Med bolj znanimi zagovorniki te vrste vrlinske epistemologije sta John Greco in Ernest Sosa, čeprav bi pri obeh hitro lahko hitro našli razloge proti tej kategorizaciji. Greco namreč priznava določen pomen, predvsem pri subjektivnemu vidiku upravičenja, tudi spoznavni odgovornosti. (glej Greco, 2002 in 2003)

[23] Pri tem gre za primer, ko npr. oseba *A* sklepa po hazarderjevi zmoti, da se bo številka, ki se na kocki že dolgo ni pojavila, bolj verjetno pojavila ob naslednjem metu kocke. Benevoletni demon, ki stoji v ozadju, pa vsako tako zmotno sklepanje podpre in "uresniči" A-jevo predvidevanje. V njegovem svetu je ta proces popolnoma zanesljiv, v bližnjih svetovih, kjer benevolentnega demona ni, pa seveda ne.

[24] Za reliablistični odgovor na ta argument glej Axtell (2003).

pogojem za védenje, ker resnično prepričanje v tem primeru ni posledica kakršnih koli spoznavno vrlih dejanj, pač pa spoznavne sreče. Tako lahko ugotovimo, da se vrlinski pristop v spoznavni teoriji lahko izogne določenim vrstam gettierovskih problemov na način, da le-ti ne zadostijo vsem pogojem za védenje.

Vendar se izkaže, da se vsaj reliabilizem z vključitvijo ideje spoznavne vrline ne more povsem izogniti gettierovskim problemom.[25] Pretehtajmo naše intuicije ob naslednjem primeru. Mojca dobro vidi, vsaj dovolj dobro, da lahko prepozna svojega moža, ki sedi v delno zatemnjeni dnevni sobi na razdalji npr. štirih metrov. Njena sposobnost vida gotovo ni nezmotljiva ali popolna, ampak dovolj dobra, da upravičeno sprejme prepričanje "Moj mož sedi v dnevni sobi". In sedaj predpostavimo, da Mojca ni pravilno prepoznala moškega, sedečega v naslonjaču, in da tam sedi možev brat, ki mu je zelo podoben. In sedaj smo le še korak od gettierovskega primera, le Mojčinega moža posadimo v nek drug kot dnevne sobe, kjer ga Mojca ne vidi. Kaj reči o tem problemu? Ali Mojčin vid potemtakem ni zanesljiv kognitivni mehanizem in torej tudi ne vrlina? Ali pa to je, in torej Mojca ve, da je njen mož v dnevni sobi? Vsekakor tukaj tudi z vrlinskim pristopom ne moremo najti enostavnega odgovora. Tudi internalistične vrlinske opredelitve védenja ne morejo izključiti določene vrste spoznavne sreče, ki je lahko podlaga omenjenim problemom. Duncan Prichard pravi tej vrsti spoznavne sreče reflektivna spoznavna sreča.[26] Pri njej gre za to, da se ji ne more izogniti noben spoznavni akter, ki nima notranje dostopne refleksije o delovanju svojih spoznavnih procesov. Prichard trdi, da je te vrste sreča problem vsake spoznavne teorije, saj jo na nek način srečamo tudi v radikalnih skeptičnih scenarijih. Zagzebska se izogne gettierovskim problemom, kot je npr. primer zaustavljene ure, s tem, ko glede na njeno opredelitev védenja, le-to izvira iz dejanj spoznavne vrline, pri tem primeru pa gre zgolj za spoznavno srečo, ne pa primerom reflektivne spoznavne sreče. Pri tem se Prichard sklicuje kar na primer, poznan iz razprav o skepticizmu, tj. primer možgan v kadi kot različice reflektivne spoznavne sreče. V tem primeru so nam notranje dostopni razlogi za naša prepričanja povsem enaki, pa najsi gre za dejanski svet brez zlobnega

[25] Pri teoriji Ernesta Sose tako večino dela opravi pogoj varnosti [safety] prepričanja, ne pa pojmovanje zanesljivih mehanizmov kot spoznavnih vrlin. Prepričanje je *varno*, če in samo če ga subjekt ne bi (z določeno lahkoto) sprejel, če ne bi bilo resnično. (glej Sosa, 2003: 39-48)

[26] Prichard (2003) ločuje med veridično in reflektivno spoznavno srečo. Prvo lahko opredelimo na naslednji način: S-ovo prepričanje *Pl* je *veridično srečno*, če je resnično v aktualnem svetu, a neresnično v vseh bližnjih možnih svetovih, kjer je tvorjeno z enakim spoznavnim mehanizmom. Druga vrsta spoznavne sreče je reflektivna spoznana sreča. S-ovo prepričanje *P1* je *reflektivno srečno*, če je resnično v aktualnem svetu, a v skoraj vseh možnih svetovih konsistentnih s tem, kar je S-u dostopno z refleksijo, bi bilo tvorjeno napačno prepričanje. Razlikovanje se kot pomembno pokaže v luči razlik med eksternalizmom in internalizmom, ilustriramo pa ga lahko z naslednjim primerom. Zamislimo si *razvrščevalca piščancev A*, ki je (ne vedoč kako) sposoben z visoko mero zanesljivosti prepoznavati med ženskimi in moškimi primerki majhnih piščancev. Eksternalizem je A-ju pripravljen pripisati védenje, internalizem pa ne. Sedaj si zamislimo še zanesljivega *razvrščevalca piščancev B*, ki pa ima njemu dostopne razloge za prepoznavanje in svoje odločitve in ima védenje o tem, kako in kje deluje proces prepoznavanja. V njegovi spoznavni situaciji je gotovo nekaj, kar pri *A*- umanjka. Prepričanja nobenega od akterjev niso veridično srečna, saj sta njuna mehanizma zanesljiva tudi v bližnjih svetovih. So pa *A*-jeva prepričanja reflektivno srečna, kajti ne poseduje reflektivnega védenja, ki bi lahko izključilo raznorazne možnosti zmot, npr. v svetovih, kjer omenjene zmožnosti sploh ne poseduje.

demona ali pa skeptični scenarij. Dodati pa je potrebno, da se tej vrsti spoznavne sreče težko izogne kakršna koli spoznavna teorija, saj se vse srečajo s težavami pri oblikovanju dokončnega odgovora skeptiku in njegovim izzivom. Sklenemo lahko, da se vrlinska epistemologija izogne nekaterim tipom gettierovskih problemov, ni pa zanje povsem neobčutljiva.

7. Zaključek

Vrlinska epistemologija je poleg kontekstualizma in disjunktivizma[27] ena izmed sodobnih smeri v epistemologiji, na katere lahko gledamo predvsem v luči nadgradnje tradicionalnih stališč (fundacionalizma, koherentizma in reliabilizma). Vsaka izmed teh se je razvila kot določen odziv na težave in probleme, v katere so se zapletle njihove predhodnice. Kot takšna nadgradnja so te nove vrste spoznavnih teorij združljive s starimi. Kot smo videli, gre pri vrlinskem reliabilizmu za spoj standardnega reliabilizma z vrlinsko teorijo, pri responsibilizmu pa za z vrlinsko teorijo nadgrajeni spoznavni internalizem. V tej luči jih gre tudi presojati, pri čemer lahko vrlinski spoznavni teoriji pripišemo pomembno vlogo pri obujanju idej spoznavnega delovanja, spoznavne motivacije, spoznavnih ciljev, spoznavne odgovornosti in pogled preko ozkih mej zarisovanja nujnih in zadostnih pogojev v analizi védenja, k drugim spoznavnim ciljem, kot je npr. razumevanje. Čeprav je paleta stališč, ki bi jih lahko označili ali pa se sama označujejo kot vrlinska, zelo široka, jih kot enotne gre razumeti posebej v luči normativnega ovrednotenja prepričanj glede na normativno ovrednotenje značilnosti spoznavnih akterjev. In čeprav se zdi globel med vrlinskim reliabilizmom in responsibilizmom precej globoka, vidijo mnogi avtorji prihodnost vrlinske spoznavne teorije prav v uspešni združitvi osnovnih idej obeh in v uspešni novi opredelitvi ali zastavitvi temeljnih spoznavnih vprašanj. (o tem glej npr. Hookway, 2003)

INTRODUCTION TO VIRTUE EPISTEMOLOGY

The paper is an elementary introduction to virtue epistemology. It presents in brief its history and development. It deals with the nature of epistemic virtues and their role in epistemological inquiry. The poper also considers the difference between reliabilistic and responsibilistic versions of virtue epistemology, the difference

[27] Kontekstualizem je stališče, po katerem so standardi za pripis védenja kontekstualno spremenljivi (Lewis, Cohen). Osrednja ideja disjunktivizma (McDowell, Williamson) je, da sta védenje ni prepričanje, ki bi zadostilo še nekaterim drugim pogojem, temveč osnovni doživljaj oz. osnovno duševno stanje. Glede na to lahko rečemo, da ali *S* ve, da *P*, ali pa je le prepričan, da *P*. Védenje in zmota tako nimata skupnega gradnika, tj. prepričanja, ki je v prvem primeru resnično, v drugem pa zmotno.

between its weak and strong varieties, the relationship between virtue ethics and virtue epistemology, and poses a question whether a virtue approach in epistemology could provide a means of avoiding Gettier problems.

Keywords: virtue epistemology, epistemic virtues, moral virtues, reliabilism, responsibilism, Gettier problems.

Literatura

Aristotel (2002). *Nikomahova etika.* Ljubljana: Slovenska matica.

Axtell, G. (2003). "*Felix Culpa*: Luck in Ethics and Epistemology". *Metaphilosophy,* 34, str. 331-352.

Blackburn, S. (2001). "Reason, Virtue, and Knowledge". V A. Fairweather in L. Zagzebski L. (ur.), *Virtue Epistemology – Essays on Epistemic Virtue and Responsibility,* New York: Oxford University Press.

Brady, M. S. in Prichard, D. (2003). "Moral and Epistemic Virtues". *Metaphilosophy,* 34, str. 1-11.

Driver, J. (2003). "The Conflation of Moral and Epistemic Virtue". *Metaphilosophy,* 34, str. 367-383.

Eflin, J. (2003). "Epistemic presuppositions and their Consequences". *Metaphilosophy,* 34, str. 48-68.

Gettier, E. L. (1963). "Is Justified True Belief Knowledge?" *Analysis,* 23, str. 121-123.

Greco, J. (2002). "Virtue Epistemology". *Stanford Encyclopedia of Philosophy* (Edward N. Zalta (ur.), URL = http://plato.stanford.edu/archives/fall2002/entries/epistemology-virtue/).

Greco, J. (2003). "Virtue and Luck, Epistemic and Otherwise". *Metaphilosophy,* 34, str. 353-366.

Hookway, C. (2003). "Affective States and Epistemic Immediacy". *Metaphilosophy,* 34, str. 79-96.

Kant, I. (2001). *Groundwork of the Metaphysics of Morals.* Cambridge: Cambridge University Press.

Miščevič, N. (1999). "Uvod v spoznavno teorijo". *Analiza,* 3, str. 3-17.

Prichard, D. (2003). "Virtue Epistemology and Epistemic Luck". *Metaphilosophy,* 34, str.106-130.

Ross, W. D. (1930). *The Right and the Good.* Oxford: Clarendon Press.

Sosa, E. (1991) *Knowledge in Perspective: Selected Essays in Epistemology* (Cambridge: Cambridge University Press).

Sosa, E. (2003). "Relevant Alternatives, Contextualism Included". *Philosophical Studies*, 12, str. 35-65.

Ule, A. (2001). *Logos spoznanja.* Ljubljana: Znanstveno in publicistično središče.

Zagzebski, L. (1994). "The Inescapability of Gettier Problems". *The Philosophical Quarterly*, 44, str. 65-73.

Zagzebski, L. (1996). *Virtues of the Mind.* Cambridge: Cambridge University Press.

Zagzebski, L. (2003). "The Search for the Source of Epistemic Good". *Metaphilosophy*, 34, str. 12-28 (slovenski prevod v pričujoči številki *Analize*, str. 85-100).

Recenzija

Friderik Klampfer

recenziral Vojko Strahovnik

ETIŠKI POJMOVNIK

Aristej, Maribor 2003

Pri založbi Aristej je v zbirki Pojmovniki izšla že druga knjiga s področja filozofije. Pred nami je sedaj poleg *Logiškega pojmovnika za mlade* avtorice Olge Markič tudi knjiga Friderika Klampferja *Etiški pojmovnik za mlade*. Avtor se v njej loteva moralne filozofije, ki jo želi mladim, pa tudi vsem ostalim bralcem, približati predvsem preko obravnave mreže osnovnih moralnih (ali z moralnostjo povezanih) pojmov in odnosov med njimi. Obravnava tem moralne filozofije sledi konkretnim primerom situacij iz vsakdanjega življenja, pri katerih včasih upravičeno, včasih pa tudi neupravičeno sodimo, da so moralno pomembni, da nosijo moralno težo oziroma da so moralno sporni. Skozi celotno knjigo avtor ostaja zavezan spodbujanju bralčeve kritične moralne presoje in hkrati večanju moralne občutljivosti.

Knjiga je razdeljena v posrečeno izbrane tematske sklope. Avtor nas popelje preko osnovnih zarisov etike in morale do vprašanj, ki se vežejo na moralne dileme in moralno presojo, na pojme dobrega in zla, pravic in dolžnosti, pravičnosti in krivičnosti, avtonomnosti in svobode ter vrlin in hib. Knjigo zaključujejo poglavja, ki obravnavajo moralno odgovorne razmerje morale do religije in razuma ter seznam nekaj konkretnih ob moralnih načel. Avtor pri svojem bogatem prikazu polja morale filozof bralca ne obremenjuje z razvitimi in kompleksnimi oblikami moralnih tev rij in argumentov zanje, in mu njihove temeljne značilnosti in usmeritve prikaže le, ko gre za primer določenega celostnega odgovora na določen moralni problem.

Bralec je v posamične problemske sklope uveden z konkretnimi primeri in zgodbami, ki jih srečujemo v naši vsakdanjosti ali jih poznamo iz medijev in ki jih lahko *prima facie* zapopademo kot moralne dileme ali pa tudi ne. Tako uvodno poglavje razpira problem moralne dopustnosti splava, hkrati pa ob tem obravnava vlogo, ki jo moralna filozofija oziroma širše, sploh vsak kritičen premislek, (lahko) odigra pri oblikovanju odgovorov na to in podobna vprašanja. Bralcu je tako predstavljena razlika med posameznimi ravnmi moralne filozofije oziroma etike, od posamičnih moralnih sodb, ki spremljajo naše konkretne odločitve za ali proti dejanjem, do splošnih moralnih sodb, ki oblikujejo naše moralno obzorje in zavezanost ter na koncu še raven teoretične etike ali metaetike, ki si zastavlja vprašanja o naravi naše moralnosti, naših moralnih sodb, moralne resničnosti, moralnega upravičenja ipd., pri čemer pa slednja nikakor ni nepomembna ali zgolj obrobni dodatek praktični etiki. Drugo poglavje obravnava moralne dileme v njihovem odnosu do naših moralnih dolžnosti, sodb in moralnih zakonov. Predstavljena je razlika med pravimi ter zgolj navideznimi in nepristnimi dilemami. Moralne dileme so za bralca, posebej še za mlajšega, ki

mu je knjiga prvenstveno namenjena, zelo dober poligon, kjer se soočijo moralne intuicije in kjer se lahko razkrije tudi njihova nezdružljivost ali nasprotnost ter nujnost razrešitve napetosti med njimi, če želimo priti do pravilne odločitve in jo upravičiti. Na moralne dileme je vezan tudi premislek o racionalnosti moralnega obžalovanja. V naslednjem poglavju avtor bralcu celostno predstavi vidike naše moralne presoje: upravičenje, razlike med vrstami moralnih sodb, vlogo naših moralnih intuicij, moralnih načel in teorij ter moralnih avtoritet. Sledi poglavje o dobrem in zlem. Kaj je dobro, srečno in izpolnjujoče življenje, in kaj nevredno življenje? Gre v ozadju vrednostih sodb za ugodje, blaginjo, morda srečo? Kaj je zlo in od kod izvira? To je le nekaj od vprašanj, ki se jih na tem mestu loteva avtor ter nanje poskuša podati sklenjen odgovor ali pa vsaj zarisati nekaj namigov za iskanje letega. Sledita poglavji o pravicah in dolžnostih ter o pravičnosti in krivičnosti. V vseh svojih dimenzijah sta analizirana pojma pravice in dolžnosti ter tako bralcu predstavljena z vsemi njunimi posledicami. Pojem pravičnosti se tako izkleše v odnosu do enakosti, pravic, zaslužnosti in poštenosti.

V nadaljevanju knjige najdemo sklop poglavij o avtonomnosti in svobodi, vrlinah in hibah ter o moralni odgovornosti. Za idejo avtonomnosti oz. avtonomnega odločanja avtor pravi, da združuje tri od seboj sorodne, a ločene ideje: osebno svobodo, neodvisnost in nadzor. Svoboda in avtonomnost sta sicer izjemno kompleksni temi znotraj moralne filozofije, ki sta na eni strani doživeli tako določene "temeljno življenjske" filozofske interpretacije, hkrati pa tudi mnoštvo izrazito "filozofsko-tehničnih" obravnav (npr. pojem svobodne volje). Avtorju uspe vse te vidike povezati v zelo berljivo poglavje. Kreposti ali vrline se po antičnem razcvetu vse bolj vračajo v moralno filozofijo, zato njihova obravnava vsekakor sodi v knjigo. V nadaljevanju nam avtor predstavi tudi različne dimenzije pojma odgovornosti (npr. za nenamerne posledice naših dejanj, za komulativne posledice, kolektivne vidike odgovornosti, odgovornost za opustitve dejanj ipd.). Knjigo skleneta poglavji o odnosu morale do religije in razuma. Tu je prikazan odnos med moralnimi in verskimi razlogi, pri čemer so prvi vezani na naše moralne dolžnosti, slednji pa na dolžnosti do Boga oziroma verske resnice. Zaključno poglavje raziskuje univerzalno naravo razlogov, ki nam jo narekuje naš razum.

Našo oceno knjige lahko sklenemo z ugotovitvijo, da predstavlja knjiga *Etiški pojmovnik za mlade* izjemno kvalitetno, zanimivo in dobrodošlo branje na področju poljudnih filozofskih del. Gre za celovit in jedrnat, a hkrati natančen uvod v moralno filozofijo, ki ga je slovenski prostor že dolgo potreboval. Sedaj je pred nami ter nas tudi po zaslugi zelo duhovitih, barvitih in premišljenih karikatur Izaija Lunačka kar vabi k branju.

Navodilo avtorjem

Prispevke pošljite v izpisu z dvojnim razmakom na naslov uredništva. Priložiti je treba povzetek (v slovenščini in v enem izmed svetovnih jezikov – angleščini, nemščini, francoščini), ki povzema glavne poudarke dela. Povzetku v tujem jeziku je treba dati naslov v tem jeziku. Povzetek ne sme presegati 150 besed. Na koncu povzetka navedite do 5 ključnih besed (deskriptorjev). Na primer – *Ključne besede*: filozofija jezika, metafora, analogija ... oziroma *Keywords*: philosophy of language, metaphor, analogy ...

Prispevki naj ne presegajo obsega ene in pol avtorske pole (45.000 znakov) vključno z vsemi opombami. Prispevki naj bodo notranje razčlenjeni, torej razdeljeni na razdelke, in opremljeni – če je mogoče – z mednaslovi. V besedilu dosledno uporabljajte dvojne narekovaje (na primer: pri navajanju naslovov člankov, citiranju besed ali stavkov in kadar navajate strokovne in posebne izraze). Izjema so citati znotraj citatov. Pri teh uporabite enojne oklepaje. Naslove knjig, periodike in tuje besede je treba pisati *ležeče*. Primeri: *Kritika praktičnega uma*, *Anthropos*, *a priori*.

Reference se tiskajo v besedilu (“author-date” sistem). Oblikovane morajo biti takole: (avtorjev priimek, letnica: str. ali pogl.); na primer (Davidson, 1967: 312) ali (Blackburn, 1988: III). Na koncu poslanega prispevka morate priložiti popoln, po abecednem redu urejen bibliografski opis citiranih virov. Primeri takih opisov:

Davidson, D. (1967). “Truth and Meaning”. *Synthese*, 17, str. 304-323.

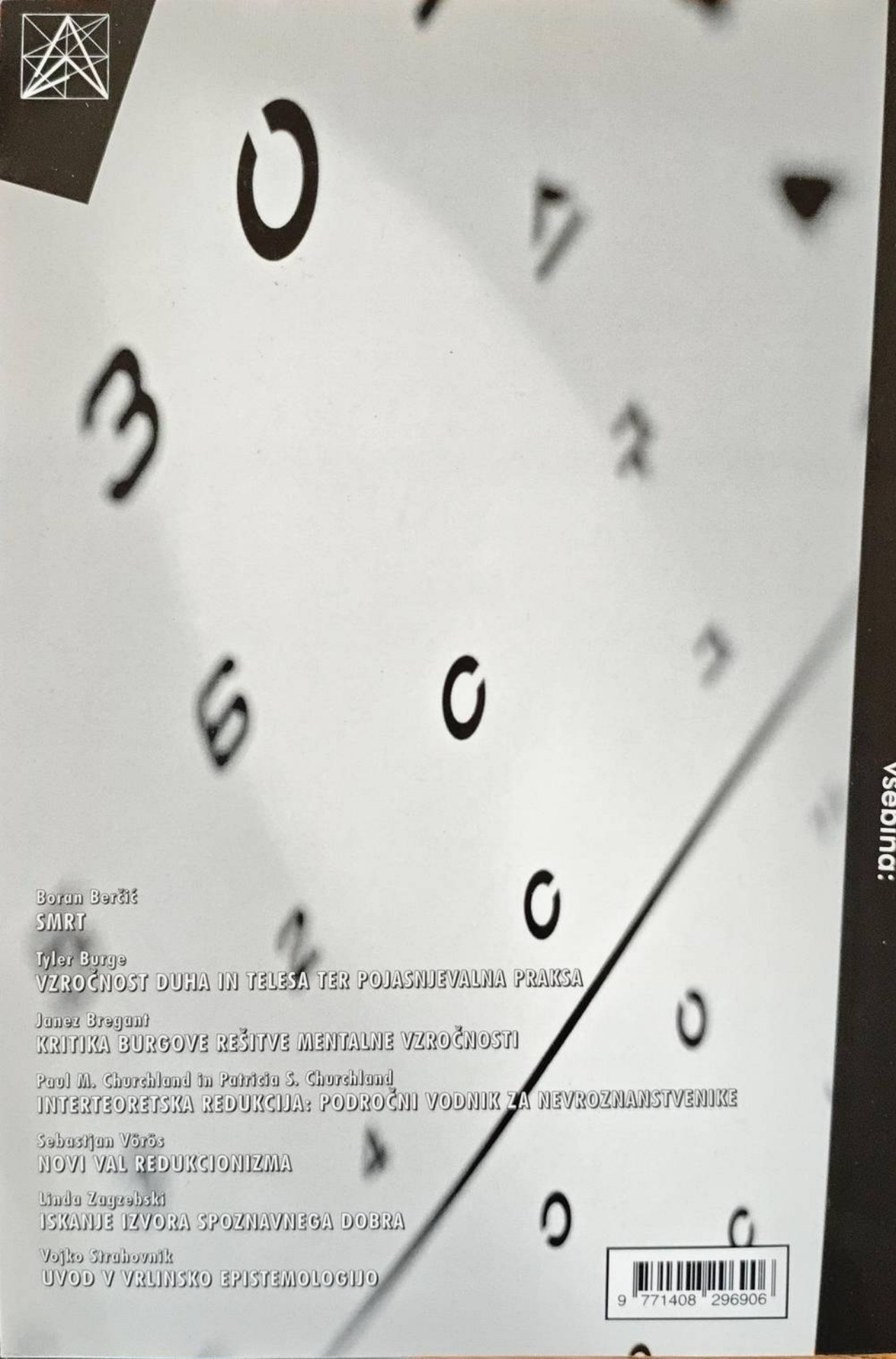
Davidson, D. (1984). *Inquiries into Truth and Interpretation*. Oxford: Oxford University Press.

Davidson, D. (1984). “Truth to the Facts”. V Davidson, D., *Inquiries into Truth and Interpretation*, Oxford: Oxford University Press, str. 37-54.

Davidson, D. (1997). “Radical Interpretation”. V Baillie, J. (ur.), *Contemporary Analytic Philosophy*, New Jersey: Prentice Hall, str. 383-393.

Pod črto navajajte *samo* opombe. V besedilu je treba opombe označiti z dvignjenimi indeksi[1]. Ne uporabljajte opomb za navedbo reference.

Poleg izpisa teksta priložite tudi tekst na 3.5” disketi. Tekst naj bo napisan na IBM združljivem računalniku v urejevalniku besedil Word ali WordPerfect, po možnosti v okolju Windows 95. Poleg datoteke, ki je bila shranjena v Word ali WordPerfect formatu, shranite isto datoteko še v ASCII formatu. Poskrbite, da se bosta besedili na disketi in izpisu natančno ujemali. Nenaročenih rokopisov ne vračamo.



vsebina:

Boran Berčič
SMRT

Tyler Burge
VZROČNOST DUHA IN TELESA TER POJASNJEVALNA PRAKSA

Janez Bregant
KRITIKA BURGOVE REŠITVE MENTALNE VZROČNOSTI

Paul M. Churchland in Patricia S. Churchland
INTERTEORETSKA REDUKCIJA: PODROČNI VODNIK ZA NEVROZNANSTVENIKE

Sebastjan Vörös
NOVI VAL REDUKCIONIZMA

Linda Zagzebski
ISKANJE IZVORA SPOZNAVNEGA DOBRA

Vojko Strahovnik
UVOD V VRLINSKO EPISTEMOLOGIJO

9 771408 296906