

MANUALIA

Davor Pećnjak

**UVOD U
FILOZOFIJU
UMA**



SVEUČILIŠTE U ZAGREBU
HRVATSKI STUDIJI

Davor Pećnjak

UVOD U FILOZOFIJU UMA

MANUALIA (mrežno izdanje)

Sv. 11

© Davor Pećnjak, 2013.

Nakladnik:

Hrvatski studiji Sveučilišta u Zagrebu
Borongajska cesta 83d, Zagreb

Za nakladnika:

Prof. dr. sc. Josip Talanga

Izvršni urednik:

Branko Ivanda

Recenzenti:

Prof. dr. sc. Boran Berčić

Dr. sc. Damir Mladić

ISBN 978-953-7823-32-0

Temeljem odluke Povjerenstva za izdavačku djelatnost Hrvatskih studija objavljivanje skripte odobrilo je Znanstveno-nastavno vijeće Hrvatskih studija 9. travnja 2013. U skladu s člankom 22. Pravilnika o izdavačkoj djelatnosti Hrvatskih studija pribavljene su dvije pozitivne recenzije.

Davor Pećnjak

UVOD U
FILOZOFIJU UMA

Skripta

HRVATSKI STUDIJI SVEUČILIŠTA U ZAGREBU
Zagreb, 2013.

SADRŽAJ

UKRATKO: ŠTO JE UM I ŠTO JE FILOZOFIJA UMA	7
DUALIZMI	10
FUNKCIONALIZMI	19
REPREZENTACIJSKA TEORIJA UMA	35
FIZIKALIZMI	41
SUPERVENIJENCIJA	50
INTENCIONALNOST	52
BIBLIOGRAFIJA	56

UKRATKO: ŠTO JE UM I ŠTO JE FILOZOFIJA UMA

Filozofija uma danas je jedna od najpropulzivnijih grana filozofije. Glavni problem filozofije uma je takozvani problem um-tijelo. Ovaj problem se sastoji u pitanju kako je realiziran naš um odnosno naša mentalnost. Također, značajno je pitanje kako možemo opisati i objasniti rad samog ljudskog uma (ovo pitanje se može razmatrati i bez obzira na ontološku realizaciju uma).

Danas se pod pojmom “um” razumijeva sveukupnost mentalnog odnosno psihičkog života. Također ga razumijevamo i kao uzročnika našeg ponašanja. U vrlo suvremenom žargonu, um možemo promatrati i kao procesor obrade informacija.

Sveukupnost mentalnog života obuhvaća osjete, percepciju, emocije, vjerovanja, želje, namjere, pamćenje, razmišljanje, zaključivanje. Osjeti i percepcija su tzv. kvalitativna ili fenomenalna stanja. Djelomično su to i emocije. Vjerovanja, želje, namjere su tzv. propozicionalni stavovi. Emocije su također djelomično i propozicionalni stavovi. Većina filozofa danas misli da propozicionalni stavovi nemaju kvalitativni odnosno fenomenalni aspekt. No o tome se može diskutirati. Ovaj kvalitativni, odnosno fenomenalni aspekt, može se karakterizirati da “nekako izgleda biti” u takvom stanju. Postoji nekakav subjektivni doživljaj kada se proživljava neki osjet ili kada se doživljava neka percepcija. Unutrašnja

kvaliteta doživljavanja (npr. doživljaj “crvenosti” crvene boje) u filozofskom žargonu naziva se kvalija. Većina filozofa danas smatra da taj aspekt ne postoji kod propozicionalnih stavova – npr. čak i kad nam je neko vjerovanje svjesno prisutno. Propozicionalni stavovi imaju sadržaj koji se izražava nekom propozicijom (rečenicom). Primjer: “Vjerujem da je Zagreb lijep grad”. Ovdje “Vjerujem da” označava stav dok “Zagreb je lijep grad” jest sadržaj tog stava. Dakle, opća shema propozicionalnih stavova je “‘*Vrsta stava*’ _____” (na mjesto crte se upisuje sadržaj). Različiti propozicionalni stavovi mogu imati isti sadržaj: npr. vjerovanje i želja - “Vjerujem da je prozor otvoren” i “Želim da je prozor otvoren”. Vjerovanje možemo interpretirati da uzimamo kao da je sadržaj istinit tj. kao da je takvo-i-takvo stanje stvari. Želja izražava da subjekt hoće da se ostvari neko stanje stvari ako trenutno i nije tako. Ukoliko ništa drugo ne stoji na putu, subjekt će pokušati ostvariti svoju želju tj. promijeniti neko stanje stvari.

Rješenje ontološkog problema ostvarenja uma može biti monističko ili dualističko. Monizmi smatraju da je sve što postoji, pa tako i um, ostavreno od samo jedne vrste “tvari”. Monizmi su idealizam i materijalizam, odnosno fizikalizam. Idealizam tvrdi da postoji samo nematerijana tvar i da je sve što postoji mentalno. Suprotno tome, materijalizam (fizikalizam) tvrdi da sve što postoji, pa tako i um, jest realizirano materijalno, fizički. Danas se gotovo uvijek pojmovi “materijalizam” i “fizikalizam” rabe sinonimno. No, strože gledano, to nisu sinonimni pojmovi. Pojam “materijalizam” označava metafizičku doktrinu koja tvrdi da u univerzumu postoji samo jedna “tvar” – materija –

od koje je sve građeno. Fizikalizam pak tvrdi da u univerzumu postoje samo oni entiteti koje može opisati (kompletna) fizika. Naravno da su to materijalni entiteti, ali suvremena fizika danas uključuje u svoj (egzaktni) opis i entitete kao što su npr. polja sila, kao što je elektromagnetsko polje ili gluonsko polje, koji nisu materijalni entiteti. Oni sami nisu sazdana od materije, (nemaju masu, ali jesu energetska tj. imaju određenu energiju; a u fizici zapravo vrijedi ekvivalencija mase i energije – to su dva različita načina pojavljivanja istog) ali utječu na entitete koje smatramo materijalnim (dakle, na one koje imaju masu ali mogu imati i druge karakteristike pokraj nje, kao što je npr. električki naboj itd.) i dadu se eksperimentalno i empirijski proučavati i egzaktno opisati fizikalnom teorijom. Fizikalizam, dakle, smatra da ono što postoji jesu samo fizikalni objekti, stanja i procesi opisivi i objašnjivi fizikom. (Odnosno, u suvremenom žargonu tako se rabi i pojam “materijalizam”. Za nešto drugačiju razradu ovih razlikovanja vidi Robinson (1993, str. 1-3).) Naravno, fizikalni objekti, stanja i procesi mogu biti u različitim stupnjevima složenosti. Fizikalizam je redukcionistička teorija jer isto tako smatra da se kemija, biologija pa onda i psihologija, u krajnjoj konzekvenci, mogu na ovaj ili onaj način reducirati na fiziku.

Dualizmi tvrde da postoje dvije “tvari”: i materijalna i nematerijalna. Naše tijelo i objekti izvan nas bili bi materijalni, dok bi um bio realiziran na neki nematerijalan način. Kada dualizmi tvrde da su psihološka stanja i procesi nematerijalni onda misle da su i ne-fizikalni tj. da ih nikakva, pa ni potpuna fizika, ne može opisati. Dualizam može dopustiti postojanje ne-materijalnih fizikalnih entiteta kao što su

energetska polja sila, ali ona nisu, za dualiste, kao fizikalni entiteti, gradivna “tvar” umskih, psiholoških stanja i procesa.

Također postoji i inicijalno ontološki neutralna teorija a to je funkcionalizam. S obzirom na sadržaj teorije, funkcionalizam može biti kompatibilan i s monizmima i s dualizmima.

DUALIZMI

Prvo ćemo reći nešto o dualizmima.

Postoji nekoliko različitih dualističkih teorija: interakcionizam, epifenomenalizam, paralelizam, okazionalizam. Svako od njih, naravno, zajedničko je to da smatra da postoje dvije osnovne “tvari” – materijalna i nematerijalna. Um, svijest i mentalnost su, prema njima, ostvareni nematerijalno, makar jesu ili mogu biti povezani na neki određeni način s materijalnim, odnosno s fizikalno-kemijskim procesima u mozgu. Paralelizam i okazionalizam zagovaraju potpunu odvojenost uma i mentalnosti od tijela i materijalnog, fizičkog.

Jedan od najpoznatijih argumenata za dualizam jest Descartesov (1911) argument. On počinje od toga da mi možemo zamisliti da zapravo one stvari koje nazivamo materijalnim i koje se nalaze oko nas, zapravo ne postoje. Prisjetimo se da možemo sanjati ili biti u stanju tako jake halucinacije da nam izgleda kao da su neke stvari prisutne oko nas dok to ne odgovara stvarnosti – kada sanjamo mi smo u stvari u stanju spavanja, recimo u našoj vlastitoj sobi u krevetu, a ne kako nam se čini u snu da

upravo prvi prolazimo ciljem na utrci formule 1 u kabini McLarena M 22P4. Tako sistematski možemo zamisliti da ništa za što nam se čini da je oko nas, ne postoji. Budući da je naše vlastito tijelo također materijalno, mi možemo zamisliti da čak niti tijelo nemamo. No, kada mislimo ne možemo zamisliti da upravo tada ne mislimo. To bi bila kontradikcija. Kada mislim ne može biti da ne mislim. S druge strane, materijalni objekti imaju svojstvo protežnosti: oni se nalaze u prostoru, zauzimaju određeni prostor, protežu se kroz prostor. No, ne čini se tako da misao ima neku protežnost – misao nije nešto što zauzima prostor. Ona nema, prema Descartesu, svojstvo protežnosti. Sada imamo po dva potpuno suprotna svojstva koja pridajemo tijelima i mislima – tijela su protežna te uvijek možemo zamisliti da u stvari ne postoje dok je misao neprotežna te ne možemo zamisliti da ne postoji kada mislimo. Tvar od koje su tijela građena jest materija. Kako su tijela građena od materije, onda ono što ne može imati glavna svojstva koja ima materija, nego ima potpuno suprotna svojstva, mora biti nematerijalno. Misao pak pripada umu te je um, dakle, nematerijalan. Dakle, prema Descartesu, postoje dvije vrste “tvari” – materija – od koje su građeni materijalni predmeti u prostoru oko nas (uključujući i naša, ljudska tijela) i nematerijalna “tvar” - od koje su “građeni” nematerijalni umovi.

No, umovi i tijela (ljudska) nisu potpuno odvojeni. Oni su uzročno povezani. Descartes je smatrao, a tako smatraju i neki suvremeni dualisti, da postoji uzročna veza između (nematerijalnog) uma i (materijalnog) tijela. Naime, i zdrvorazumski se čini da sadržaji naših misli određuju i uzrokuju naše ponašanje, i obrnuto, predmeti iz naše okoline utječu na

oblikovanje percepcije i sadržaja misli. Descartes je smatrao da se uzročno međudjelovanje između uma i tijela događa u mozgu u malenom dijelu koji se zove epifiza. Od tuda i dolazi ime za ovu vrstu dualizma a to je “interakcionistički dualizam”.

Međutim, ovdje nastaje problem. Kako je uopće moguće da dvije potpuno različite “tvari” – materijalno i nematerijalno – budu u uzročnom odnosu? Problem se još više pojačao kada su kasnije u fizici otkriveni zakoni očuvanja. Postoji nekoliko zakona očuvanja – zakon o očuvanju energije, zakon o očuvanju električnog naboja, zakon o očuvanju momenta i još neki zakoni. Zakon o očuvanju energije kaže da je ukupna energija unutar sustava uvijek konstantna. Uzročni tijek prijenosa i transformacije energije opažamo uvijek između fizičkih tijela. No, unutar sustava ukupna energija ostaje uvijek ista. Ali, ako nematerijalan um uzročno djeluje (u, recimo, uzrokovanju ponašanja) na tijelo da ono nešto napravi, npr. da se pomakne u prostoru, onda bi bio slučaj da ni iz čega fizičkog imamo utjecaj na fizičko i tako bi se ukupna energija sustava povećala. Time bi bio narušen ovaj zakon očuvanja.

Paralelizam je pokušao odgovoriti na probleme u vezi uzrokovanja na radikalni način – tvrdnjom da zapravo nematerijalan um i materija, pa tako i samo ljudsko tijelo, uopće nisu ni u kakvoj uzročnoj vezi. Samo se čini da je tako. Naime, paralelizam drži da je ono što se mentalno događa u umu potpuno korelirano s onim što se događa u svijetu, ali između fizičkih događaja u svijetu, pa i u samom ljudskom tijelu, i mentalnih događaja nema međusobne uzročne povezanosti: mentalna i fizička domena su nezavisne. Kako objasniti tako nevjerovatnu činjenicu da su dvije tako

različite domene potpuno korelirane? Neki paraleleisti su to pokušali objasniti na taj način da su dalje tvrdili da je to naprosto elementarna činjenica o svijetu koja nema daljnje i dublje objašnjenje; jer kakvu god ontologiju imali i kakav god svijet bio, neka objašnjenja su fudamentalna i zašto je nešto tako kao što to tvrdi neko fundamentalno objašnjenje nema više neko daljnje objašnjenje. No, još uvijek se čini da je ovakav odgovor previše neobjašnjavajući za stalne i pravilne koreliranosti koje primjećujemo između mentalnog i fizičkog. Drugi paralelisti smatraju da se ova koreliranost može objasniti time da je Bog unaprijed odredio potpunu koreliranost mentalnog i fizičkog. Ali zašto Bog ne bi stvorio jednostavniji svijet u kojem nemamo zapravo dva neovisna svijeta (mentalni i fizički) nego jedan u kojem postoji veza između mentalnog i fizičkog? S druge strane, ako nema veze između mentalnog i fizičkog, onda je dovoljno da Bog stvori samo mentalni svijet i tijek mentalnih stanja jer sve za um ostaje isto i svejedno je postoji li ili ne još i materijalan svijet. Materijalan svijet naprosto nije potreban. Tako zapravo paralelizam, ukoliko bi se dosljedno proveo, završava u idealizmu koji je na tim osnovama – neovisnosti mentalnog i fizičkog - prihvatljivija teorija.

Okazionalizam je također pokušaj odgovora na problem međusobnog uzročnog odnosa fizičkog i ne-fizičkog. Kako bi objasnio kako to da se *čini tako jasno* da fizičko uzročno utječe na ne-fizičko i obrnuto (svijet na um i obrnuto), a da u stvari takva uzročna veza uopće nije moguća, okazionalizam smatra da npr. u trenutku kada zagrizem čokoladu, Bog želi pojavljivanje osjeta slatkog kod mene. To proizlazi iz

toga što ovaj pravac, čiji je najpoznatiji predstavnik bio Malebranche (1674, 1688), smatra da je Bog “aktivno odgovoran” (vidi i Heil 1998, str. 29-32) za bilo koje slijedove događaja koji se događaju. Dakle, Božja volicija na ovaj način slična uzrokovanje, iako ne djeluje na onaj način kao što mi obično smatramo da djeluje uzročnost. Naime, kako je Bog “aktivno odgovoran” za sve događaje koji se događaju, to znači da je On taj koji je “odgovoran” za univerzum kao cjelinu. On u svakom trenutku stvara cijeli univerzum ni iz čega. Svaki univerzum je samo malo drugačiji od prethodnog pa se nama čini kao da postoji kontinuirani tijek, dok toga zapravo nema. Postoji mnoštvo diskontinuiranih univerzuma koje jednog za drugim, kao potpune cjeline, ni iz čega stvara Bog. Kako u samom univerzumu ništa ne utječe ni na što drugo, dakle, naš um i cjelokupna mentalnost je takva da i nju u svakom trenutku ni iz čega stvara Bog.

Epifenomenalizam tvrdi da uzročna veza postoji, ali samo u jednom smjeru: fizikalno-kemijski događaji u mozgu mogu uzrokovati nematerijalna mentalna stanja, ali ova nematerijalna mentalna stanja su takva da nemaju nikakve kauzalne moći. Naravno, fizikalno-kemijski događaji u mozgu, osim što mogu uzrokovati nematerijalna mentalna stanja, mogu i uzrokovati i daljnje fizikalno-kemijske događaje u mozgu. Epifenomenalizam je motiviran time da spasi našu jaku intuiciju u prilog tome da su mentalna stanja nesvodiva potpuno na fizičko, ali i time da se to pomiri sa znanstvenom fizikalističkom slikom svijeta. Naime, budući da fizikalno-kemijski događaji u mozgu uzrokuju i daljnje fizikalno-kemijske događaje u mozgu, oni u neprekinutom čisto fizičkom lancu

uzrokuju i naše ponašanje te se i ono može objasniti samo fizikalnim, iako imamo i nefizikalni mentalni dio.

Epifenomenalizam je neuvjerljiv s prihvaćanjem uzročne neučinkovitosti mentalnog. Naša je jaka intuicija da su baš sadržaji mentalnog, sadržaji mentalnih stanja i procesa, glavni uzročnici našeg ponašanja, a ne samo inertni epifenomeni moždanih događaja i procesa.

Problem uzrokovanja sam po sebi je problem na koji ima mnogo različitih odgovora a niti jedan nije širje prihvaćen. Tako i dualist interakcionista može na njega uvjerljivo odgovoriti. Filozofska razmatranja uzročnosti omeđena su i uvjetovana onim načinima na koje uzročnost operira u fizičkom svijetu. Kada jedan fizički događaj uzrokuje drugi fizički događaj, ta dva događaja su prostorno-vremenski u kontinuitetu odnosno oni su koincidentni. Na taj se način čini da uzročnost djeluje i jedino može djelovati kroz prostornu povezanost. Kada se to prihvati, onda je jasno da se više ne bi moglo zamisliti kako bi nefizički događaji koji nemaju prostornu protežnost, odnosno prostornu komponentu, mogli biti u uzročno-posljedičnim vezama s fizičkim događajima.

Ipak, na ovo se može odgovoriti s dva razmatranja psihofizičkog uzrokovanja kako ih je dao John Foster (1991, str. 159-163). Moguće je zamisliti kako uzročnost može postojati između fizičkog i ne-fizičkog, čak iako u većini slučajeva uzročnost doista djeluje kroz prostorno-vremensku kontinuiranu povezanost. Na primjer, možemo zamisliti jedan fizički događaj koji uzrokuje drugi fizički događaj a oni su međusobno udaljeni pet kilometara. Niti jedan objekt, fizički, koji je sastavni dio prvog od tih fizičkih događaja (uzrok) nije u prostornom kontinuiranom kontaktu s bilo kojim objektom, fizičkim, koji je sastavni dio drugog od tih dvaju događaja (posljedica). Svjedočanstvo za ovakvo

odvijanje događaja također možemo zamisliti, imamo stalnu empirijsku korelaciju tih dvaju događaja, nemogućnost detekcije bilo kakvog mehanizma koji posreduje između ta dva događaja, nepostojanje bilo kakvog boljeg objašnjenja tih dvaju događaja osim toga da prvi od njih uzrokuje onaj drugi. Dakle, nema konceptualne prepreke da ne bismo mogli zamisliti uzrokovanje koje se ne odvija u strogo prostornom kontinuitetu, pa ako tako prihvatimo mogućnost uzrokovanja na daljinu, nema razloga zašto ne bismo mogli prihvatiti konceptualnu mogućnost uzrokovanja koje također nema striktnu prostornu komponentu – dakle, uzrokovanje fizičkog od strane ne-fizičkog i obrnuto.

Neprijatelji interakcionističkog dualizma katkada problem psihofizičkog uzrokovanja postavljaju tako kao da dualisti ne mogu uopće objasniti kako takvo uzrokovanje može biti operativno, za razliku od toga da u fizikalnim znanostima imamo potpuno razumijevanje kako i zašto radi uzročnost. Navest ću ponovo Fosterov (1991) odgovor, koji pokazuje da niti u fizikalnim znanostima zapravo nemamo takvo potpuno objašnjenje i razumijevanje uzročnosti kakvo se zahtijeva od dualista.

Dualist može reći da fizičko-nefizičko uzrokovanje možemo prihvatiti kao jednostavnu činjenicu koja nema daljnje i dublje objašnjenje, naprosto, određeni neurofiziološki događaji uzrokuju mentalne (svjesne) događaje (nefizičke, nematerijalne) i obrnuto. Dakle, uzrokovanje je izravno i nema nikakvog daljnjeg intervenirajućeg mehanizma. No, pojam izravnog uzrokovanja zapravo uopće nije nikakv problem, kada god imamo neki uzročno posljedični događaj, uzrokovanje jest izravno – jedan događaj uzrokuje drugi. Naravno, neki uzročno-posljedični lanac može biti dug, od prvog uzroka do krajnje posljedice, ali svi međudogađaji koji između njih slijede, a koji su opet u odnosima uzroka i

posljedice, ponovo su međusobno izravno povezani u uzročnom smislu. Što se tiče psihofizičkog uzrokovanja fizičko-nefizičko, mi možemo naprosto dovesti u vezu određene neurofiziološke događaje s mentalnim događajima i tako imati uzročno-posljedične “parove” i prema tome formulirati točne psihofizičke uzročne zakone.

U fizičkom svijetu, koji objašnjavamo fizikom i drugim fizikalnim znanostima, uzročna objašnjenja također negdje prestaju upravo postuliranjem ili izvođenjem kauzalnih zakona i nema daljnjih objašnjenja zašto je baš tako. Na primjer, elektroni ili bilo kakve druge elementarne čestice imaju ta-i-ta svojstva i kada dođu u međudjelovanje, one međudjeluju na taj-i-taj način. Zašto one imaju ta-i-ta bazična svojstva da bi međudjelovale tako-i-tako i zašto je na djelu baš takva kauzalnost između tih fizičkih objekata nema daljnje objašnjenje. Naravno, ovo vrijedi ne samo za mikrofiziku nego i za objekte makrofizike. Promatrajući (među)djelovanja objekata, mi formuliramo što je moguće općenitije zakonitosti koje govore o odnosima kako se u datim uvjetima ti objekti ponašaju, tj. mi formuliramo općenite kauzalne zakonitosti; dakle, u kauzalnim zakonima daje se slijed kako iz određenih stanja/procesa slijede daljnji procesi i/ili stanja, ali zašto je uzročnost upravo baš takva kakva je na djelu, u ovome svijetu, između objekata koji ga ispunjavaju, nema daljnjeg objašnjenja u smislu još fundamentalnijih zakona ili fundamentalnijih objašnjenja. Općeniti kauzalni zakoni su krajnja objašnjenja koja imamo u znanosti.

Naravno, želim samo pripomenuti da nisu svi zakoni kauzalni, već u kvantnoj fizici (Wichmann 1988, Strnad 1985) imamo i statističke zakone te se određene pojave opisuju pomoću vjerojatnosti, ali to ništa ne mijenja na stvari

tamo gdje procesi i događaji jesu kauzalni te se mogu pretpostaviti i formulirati kauzalni zakoni.

Danas je popularan i tzv. "Argument iz znanja" za dualizam (Broad 1925, Jackson 1982). On se temelji na slijedećem misaonom eksperimentu. Zamislimo briljantnu, fenomenalnu, neuroznanstvenicu Mariju. Zamislimo da ona znade cijelu neurofiziologiju pa tako i cjelokupnu neurofiziologiju mozga. Ona zna sve što se fizikalno-kemijski događa u mozgu. Međutim, zamislimo dalje da je ona cijeli svoj dosadašnji život provela u crno-bijeloj sobi, učila je iz knjiga, časopisa i preslika koji su, naravno, takvi da sadrže samo crna slova na bijeloj podlozi, sučelja elektronskih aparata su također bila crno-bijela i sve s čim je bila u kontaktu bilo je crno ili bijelo. Nikada nije imala, dok je živjela u toj svojoj sobi učeći i radeći neurofiziologiju, nikakav kontakt s nekom drugom bojom. No, s druge strane, ona potpuno znade što se događa u mozgu kada ljudi doživljavaju bilo koju boju: plavu, crvenu, narandžastu itd. Dakle, njezino znanje što se fizički događa u mozgu prilikom doživljavanja npr. plave boje je potpuno. No nikad nije sama imala iskustvo viđenja plave boje. Kako joj je neurofiziologija i znanost već pomalo dosadila, ona konačno izađe iz svoje sobe u stvarni svijet i prvo što ugleda je vedro nebo. Oduševi se svojim novim doživljajem i kaže sama sebi: aha, znači tako izgleda plava boja. Sada ona zna kako izgleda biti u stanju doživljavanja plave boje. No, ona je, prema hipotezi, znala sve fizičko što se događa kada ljudi doživljavaju boje. Kako ona sada zna kako izgleda biti u stanju doživljavanja plave boje, odnosno sada zna kako izgleda unutrašnja kvaliteta (kvalija) doživljavanja plavog, onda mora biti da je to ne-fizičke prirode. Kada bi

ovaj mentalni događaj bio fizički, a ona zna sve fizičko, morala bi znati i kako izgleda biti u stanju doživljavanja plave boje makar ju prije nikad nije vidjela. No, ona to ne zna, odnosno nije znala do trenutka kada je izašla iz svoje crno-bijele sobe i ugledala nebo. Sada, kada ona zna kako izgleda biti u stanju viđenja plave boje, a prije nije znala, ali je imala potpuno fizikalno znanje, onda to novo znanje mora biti o ne-fizičkom. Prema tome, mentalni doživljaji su ne-fizičke prirode.

Ovaj argument je izazvao puno rasprava i puno odgovora. Neki od zanimljivih odgovora mogu se naći u slijedećoj literaturi: Lewis (1990), Papineau (2002, poglavlje 2.), Churchland (1985).

FUNKCIONALIZMI

Funkcionalizam je teorija koja, kao što joj samo ime kaže, nastoji objasniti svijest, mentalne procese, mentalne događaje i mentalna stanja funkcionalno, tj. prema funkcijama koje ta svjesna stanja i procesi te ostala mentalna stanja i procesi imaju. Mali broj funkcionalista je eliminativistički raspoloženo, poput Reya (1997) ili Sticha (1983) što znači da smatraju da uspješna funkcionalistička ili kompjutacijska teorija može zapravo eliminirati “mentalna stanja” iz ontologije (za kritiku takvog stava vidi npr. Pećnjak (2001)); međutim, velika većina ili gotovo svi ostali funkcionalisti nisu eliminativistički raspoloženi. Funkcionalisti smatraju da svjesna i mentalna stanja i procesi postoje i da se njihova bit sastoji i dade karakterizirati specifičnim funkcijama koje ona obavljaju.

Te su funkcije karakterizirane kauzalnim vezama koje postoje između raznih mentalnih stanja i mjestima koja pojedina stanja ili procesi zauzimaju unutar ovakve razgranate uzročno-posljedične mreže. Dakle, funkcionalisti smatraju da postoji problem što je to svijest i što je to specifično mentalno te da se svijest i mentalno može objasniti specifičnim vrstama kauzalnih veza. Neko mentalno stanje možemo karakterizirati uzročnom ulogom koju ima. Mentalna stanja mogu biti uzročno-posljedično povezana s nekim osjetilnim ili perceptivnim ulaznim podacima, s drugim mentalnim stanjima i s izlaznim podacima koji se, u krajnjoj konzekvenciji, manifestiraju kao ponašanje. Dakle, osnovna apstraktna shema funkcionalizma je jednostavna, što mu odmah daje i veću prihvatljivost: ulazni podatak – obrada – izlazni podatak.

Neki funkcionalisti prave razliku između onog što konkretno ostvaruje navedenu kauzalnu ulogu od samog ispunjenja te kauzalne uloge. Za neke funkcionaliste mentalno stanje je karakterizirano time što je neka specificirana kauzalna uloga ispunjena bez obzira što je to što je konkretno ostvaruje, dok drugi funkcionalisti smatraju da je mentalno stanje karakterizirano upravo onim konkretnim što tu ulogu ostvaruje. Postoji, dakle, nekoliko vrsta funkcionalizama: zdravorazumski ili analitički funkcionalizam, referencijalni funkcionalizam i strojni funkcionalizam. Ukratko ćemo opisati svaki od ovih funkcionalizama a vidjet ćemo da ni jedan od njih ne daje potpuno zadovoljavajuća objašnjenja, jer iste vrste prigovora obuhvaćaju sve ove vrste.

Zdravorazumski funkcionalizam ukazuje na to kako mi mnoge stvari određujemo, prepoznajemo i razlikujemo od drugih upravo

specificirajući ih prema onome što obavljaju odnosno prema onome kako i za što su dizajnirani da obavljaju. Ista takva strategija, smatraju pristaše zdravorazumskog funkcionalizma, treba se primijeniti i na objašnjavanje rada uma, dakle na cjelokupnu mentalnost i na svijest. Pogledajmo jedan jednostavan primjer kako funkcionira neki artefakt i kako se onda primjenjuje ova analogija. Ovo su manje-više standardni primjeri (vidi npr. Braddon-Mitchell and Jackson 1996, poglavlje 3.)

Termostat je naprava koja služi za kontroliranje temperature u nekom prostoru. Ova sprava to postiže time što uključuje ili isključuje neki drugi stroj – za hlađenje ili grijanje – na koji je vezan. On prima podatak o tome kolika je temperatura u okolini; ovo uzrokuje neko interno stanje termostata koje pak, prema tome kako je termostat baždaren, ostavlja stroj da radi ukoliko radi, prekida njegov rad, uključuje ga ili ostavlja stroj da i dalje ne radi. Naravno, termostat može obavljati i razne druge funkcije, npr. on služi kao primjer za raspravu o funkcionalizmu, može poslužiti kao umjetnički objekt koji izaziva umjetnički doživljaj na postmodernističkoj konceptualnoj izložbi ili objesnom dječaku može služiti kao sredstvo napada ili prijetnje starijim članovima susjedstva. No, za što god on bio upotrijebljen, prava funkcija termostata, te ono što ga razlikuje od ostalih stvari, jest i dalje samo ona prva opisana djelatnost – kontrola temperature u nekom prostoru uključivanjem ili isključivanjem nekog daljnjeg aparata. Opis takve djelatnosti daje smisao spravi i pojmu termostata.

Na taj način i mi prepoznamo što je bitno da netko ili nešto obavlja da bismo tom sustavu pripisali um i svijest. Ukazujemo na glavne

kauzalne uloge, oslanjajući se na to zajedničko zdravorazumsko znanje o tome što je to bol, percipiranje, vjerovanje, želja, namjera itd. Pomoću tog specificiranja uloga za pojedina vjerovanja, želje itd. mi zapravo dajemo smisao tim pojmovima. Evo kako to izgleda na jednom konkretnom primjeru. Kada traje dugo toplo ljeto, čovjek ožedni. Klauzulu ulaznog podatka možemo ovako specificirati: “vrućina i suhoća u okolini uzrokuje žeđ u čovjeka”. Klauzule unutrašnjih stanja specificiramo pak ovako: “žeđ uzrokuje želju za njezinom eliminacijom” – ova klauzula povezana je s vjerovanjem “žeđ se eliminira uzimanjem osvježavajućeg pića”. Općenita želja za eliminacijom žeđi i vjerovanje da osvježavajuće piće eliminira žeđ, uzrokuju dalje konkretnu želju za potragom i uzimanjem nekog osvježavajućeg pića. Dalje: “percipiranje bočice osvježavajućeg pića uzrokuje vjerovanje da je bočica osvježavajućeg pića ispred subjekta”. Izlazna klauzula onda izgleda ovako: “želja za osvježavajućim pićem (u tipičnom slučaju) uzrokuje ponašanje koje vodi uzimanju i konzumiranju osvježavajućeg pića”. Ako ovakav način funkcionalističkog razmišljanja želimo poopćiti onda možemo za slijedove propozicionalnih stavova dati i slijedeće specifikacije klauzula: “subjektovo vjerovanje da vrijedi ‘ako p onda q’, tipično uzrokuje vjerovanje da q ukoliko je subjekt upoznat da je p ostvareno” (vidi npr. Johnson-Laird 1993, poglavlje 12.). Na ovaj formalniji način mogu se tipizirati ne samo deduktivno i induktivno mišljenje već i raznorazne veze između ulaza, unutrašnjih stanja i izlaza. Ali, ovdje iskače prvi problem za funkcionaliste. Vidimo da se cijela mreža definira međusobnim odnosima, npr. apstraktno gledano: X-S-P-Q-R-Y. X je ulazni podatak, Y je izlazni podatak, a ostalo su unutrašnja

mentalna stanja. Crtica gledana s lijeva na desno može se čitati i tumačiti kao: *(tipično) uzrokuje*, a gledana s desna na lijevo, može se čitati i tumačiti kao: *(tipična) posljedica od*. Stanje Q npr. definirano je kao posljedica od P i kao uzrok stanju R. Time se daje značenje pojmu i stanju Q. Međutim, ako želimo nešto reći o stanju P, onda kažemo da je ono posljedica od S i uzrok od Q. Primjećujemo odmah da su ove definicije cirkularne! Jedno stanje se definira pomoću drugog: Q pomoću P i obratno: P pomoću Q! Da li je to razarajuća kritika zdravorazumskog analitičkog funkcionalizma? Jest i nije. U onim prvobitnim varijantama kako je funkcionalizam ove vrste definiran to jest ozbiljna primjedba, barem što se tiče formalne konzistentnosti teorije, iako se i uz tu cirkularnost mnogo toga može zapravo razumjeti i objasniti. Naravno, teorije koje sadrže cirkularnosti ipak ne mogu biti temelj potpunog objašnjenja onog što želimo objasniti, u ovom slučaju rad uma i svijest. No, cirkularnosti koje su bile obilježje prvobitnih formulacija funkcionalizma mogu se pažljivijim formulacijama otkloniti. Najbolji formalni način da se to učini jest preformulacija definicija i međusobnih veza ulaznih podataka, zatim naročito unutrašnjih mentalnih stanja, a također i izlaznih podataka, pomoću takozvanih Ramsey-rečenica (Braddon-Mitchell and Jackson 1996, str.51-52., Sesardić 1984, str.52.). Kako to učiniti? Jednostavno – sve implicitne definicije unutar neke funkcionalističke teorije treba preformulirati u eksplicitne definicije. Tako ćemo jednostavno izbjeći da se definicija jedne vrste pojmova ili objašnjenje jedne vrste mentalnih stanja ili procesa, poziva na drugu vrstu i obratno. Osim toga, neki mentalni pojam koji uvodimo, moramo

definirati bez toga da ga u definiciji upotrijebimo. To bi izgledalo ovako. Kada želimo reći da je neki subjekt u nekom mentalnom stanju moramo upotrijebiti rečenicu slijedećeg oblika:

“S je u stanju M_y ako i samo ako _ _ _ _ .”

u kojoj se prazna mjesta na kraju ispunjavaju sa ne-mentalnim pojmovima, a M_y označava neki mentalni pojam koji definiramo. Vidimo da M_y sada ima oblik imenice tj. više ne kažemo da S vjeruje ili S želi itd., nego kažemo S ima vjerovanje, odnosno S ima želju itd. Recimo da \mathbf{M} označava konjunkciju svih klauzula zdravorazumskog funkcionalizma koje se mogu primijeniti na našeg S-a. Sada imenice tj. mentalne pojmove koje smo tako uveli i definirali zamijenimo varijablama:

$\mathbf{M}(x_1, x_2, \dots, x_n)$

Sada možemo napisati Ramsey-rečenicu koja ima oblik predikata prvog reda:

$\exists(x_1)\exists(x_2)\dots\mathbf{M}(x_1, x_2, \dots)$

Ovaj nam oblik rečenice daje sadržaj zdravorazumskog funkcionalizma kako se može primijeniti na neki subjekt koji smo ovdje nazvali S. Ovo nam kazuje da postoje S-ova stanja koja stoje u specificiranim relacijama.

Sada možemo potpuno formulirati opću formu uvjeta pod kojim S ima neko stanje M_y . Ovaj oblik glasi ovako:

S ima stanje M_y ako i samo ako $\exists(x_1)\exists(x_2)\dots(S \text{ ima } x_1 \& \dots \mathbf{M}(x_1, x_2, \dots))$.

Svi pojmovi za mentalna stanja su zamijenjena varijablama tako da desna strana definicije nema nikakve mentalne pojmove. Prema tome, ono što definiramo nije više upotrijebljeno u samoj definiciji i tako izbjegavamo cirkularnost. Više nema pozivanja jednih mentalnih pojmova na druge i obrnuto. Na taj način mogao bi se formulirati funkcionalizam bez cirkularnosti.

Ipak, ni tu ne prestaju problemi za zdravorazumskog funkcionalista. Prvo, oni su htjeli imati jednostavnu teoriju i objašnjenje mentalnosti i svijesti. Dosljedno prevođenje iskaza funkcionalizma u iskaze tipa Ramsey-rečenica znatno komplicira samu stvar. Drugo, više nije dostatan zdravorazumski jezik i jezik psihologije nego se uvodi, u značajnoj mjeri, i jezik logike. Osim toga, u psihološkoj literaturi to nitko nikada nije dosljedno proveo u praktičnom smislu. To još samo po sebi nisu stvari koje apsolutno diskvalificiraju reformulirani zdravorazumski funkcionalizam, ali on time polako gubi na početnoj snazi i kredibilitetu.

Ono s čime niti reformulirani zdravorazumski funkcionalizam ne može izaći na kraj je slijedeće: ukoliko se još i mogu pomoću Ramsey-rečenica bez većih problema formulirati ulazni podaci i sami odnosi među unutrašnjim mentalnim stanjima – vjerovanjima, željama i ostalim propozicionalnim stavovima – ne vidi se kako bi se određeni tip

intencionalnih djelovanja formuliralo bez pozivanja na mentalni rječnik. Jer, ukoliko opisujemo izlazne podatke koji su određeno ponašanje koje nije slučajno nego racionalno ponašanje usmjereno nekom cilju, onda je ono takvo upravo radi toga jer je određeno nekom namjerom, zatim vjerovanjem da će se čineći to-i-to postići to-i-to i ti ciljevi su već unaprijed projicirani mentalno kod subjekta. U teoriji djelovanja, neko djelovanje je djelovanje upravo i samo u svijetlu intencionalnih predikata. To je ono što razlikuje djelovanje od radnji, pogotovo bazičnih radnji. Isti pokreti - mahanje rukom - može biti samo to, čisto fizičko opuštanje ili čak nesvjesna radnja, zatim to može biti pozdrav voljenoj osobi koja se upravo otiskuje Titanicom preko Atlantskog oceana, a može biti i dozivanje taksija. Pozdrav i dozivanje taksija jest djelovanje te se sastoje u glavnom dijelu i u intenciji zajedno sa mahanjem ruke i ne mogu se opisati bez pozivanja na mentalne pojmove – intencije, percepcije objekta, vjerovanja i želja koje ta percepcija uzrokuje (da nam se približava taksijem, npr., pa zatim da želim ostvariti vožnju taksijem itd.) te tako ne možemo izbjeći pozivanje na mentalne pojmove opisujući i objašnjavajući funkcionalistički izlazne podatke, odnosno ponašanje kao djelovanje.

Ukoliko funkcionalizam još i može pružiti zadovoljavajući opis i objašnjenje mentalnosti dajući značenje i sadržaj propozicionalnim stavovima i kako se izvode operacije nad njima, teškoće nastaju prilikom opisa i objašnjenja iskustava - osjeta i zamjedbi. Tradicionalni problem je problem izvrnutog spektra koji datira još od Locke (1659, 1690), a koji funkcionalizam nije uspio riješiti unatoč nekim zanimljivim pokušajima (Shoemaker 1997, White 1997, Hardin 1987).

Boje se međusobno kvalitativno razlikuju. Crveno je različito od zelenog, zeleno je različito od plavog, crveno je različito od plavog itd. Kao što smo rekli, ovu kvalitetu boje, npr. crvenost crvenog, nazivamo kvalijom. Ovdje nije mjesto i nije potrebno ulaziti u iscrpnu ontološku raspravu o bojama (za to vidi npr. Maund 1995), nego kada govorimo o bojama preventivno govorimo o osjetu boja tj. onako kako ih zamjećujemo u svijesti. Dakle, kada imamo zamjedbu zelenog, ona je, u svijesti, različita od zamjedbe crvenog i tako dalje po spektru vidljivih boja. Sada vrlo lako možemo zamisliti dvije osobe, odnosno dva subjekta koji su potpuno isti u svim psihološkim funkcionalističkim opisima, osim što imaju drugačija kvalitativna iskustva što se tiče boja. Kada jedan od njih ima kvalitativna iskustva crvenog, onaj drugi ima kvalitativna iskustva zelenog i obratno. Naravno, radi se o promatranju istog objekta. Dakle i jedan i drugi vide zalaz Sunca i zastavu Narodne Republike Kine kao istovjetne po boji, međutim, njihove unutrašnje kvalitete, iliti kvalije, tih doživljaja međusobno su potpuno različite. Možemo to i ovako formulirati: za oba subjekta individualno, boja zalazećeg Sunca razlikuje se od boje trave, međutim jedan vidi boju zalazećeg Sunca na onaj kvalitativni način kako onaj drugi vidi boju trave i obratno. Ništa u ponašanju i njihovim drugim mentalnim propozicionalnim stanjima ne treba otkriti tu razliku. Oba subjekta zovu boju zalazećeg Sunca “crveno” i oba subjekta zovu boju trave “zeleno”. Oba subjekta su deprimirana kada vide kinesku zastavu i oba subjekta uživaju u zalazu Sunca... i tako sustavno dalje. Funkcionalistički opisi tih dvaju subjekata bit će potpuno isti, međutim njihove kvalije što se tiče doživljavanja boja sustavno se

razlikuju. Iz funkcionalističkog opisa koji se sastoji iz uzročno-posljedičnih veza i mreže mentalnih stanja i procesa, nigdje se neće moći pojaviti razlika koja će upućivati da su, barem *iskustva* koja tih dvaju subjekata različita. To znači da funkcionalizam nije u stanju adekvatno opisati i objasniti svjesna iskustva - osjete i dijelove percepcije. Psihološka teorija koja to nije u stanju napraviti na zadovoljavajući način sigurno nije adekvatna znanstvena teorija.

Funkcionalisti su pokušali otkloniti neke od ovih teškoća i adekvatnije i preciznije opisati rad uma i dati odgovor na pitanje što je to i kako je realizirana svijest. No, to ih je otklonilo od zdravorazumskog analitičkog formuliranja funkcionalizma i priklonilo apstraktnijem pokušaju rješenja. Rad ljudskog uma uvijek je bio uspoređivan s radom nekog umjetno napravljenog objekta. Naravno, taj objekt je morao biti neki malo sofisticiraniji objekt – ne neka lopta nego neki stroj. Dakle, ponovo se pokušalo sa strojnom analogijom. Strojevi koji su bili u modi prije tridesetak godina i koji su još uvijek u modi su računala, kompjuteri, odnosno elektronski strojevi. Oni rade prema svojim programima. Programi se specificiraju prema tzv. dijagramu tijeka. Ideja je bila, dakle, da se za opis upotrijebe dijagrami tijeka (Dennett 1978). Oni sadrže ulazne podatke, razna unutrašnja stanja i procese i izlazne podatke. Unutrašnja stanja i procesi su u stvari strojne obrade podataka prema nekom određenom skupu računalnih pravila što znači da su dijagrami tijeka čisto sintaktički. To je formalno operiranje nad formaliziranim simbolima. Vidimo da je ovakav funkcionalizam zapravo sofisticiranija verzija zdravorazumskog funkcionalizma. U sintaktički napisanim programima i

dijagramima tijeka također nikako ne možemo dobro objasniti unutrašnje sadržaje fenomenalnih iskustava – kvalitetu bola (Dennett 1978), iskustva boja itd., nego možemo imati samo njihove uzroke i posljedice, što nam naravno, nije sa stajališta psihološkog objašnjenja, dovoljno. No imamo i jedan daljnji problem. Naime, u analitičkom funkcionalizmu smo propozicionalnim stavovima mogli dati, i to svakom posebnom, njihove specifične sadržaje koji ih određuju i koje imaju. Kada propozicionalne stavove prikažemo kao formalne simbole nad kojima se vrše formalne operacije, više uopće nije jasno koji im sadržaji pripadaju, tj. koji sadržaj pripada kojem simbolu. Budući da je dovoljno da se proces odvija samo prema formalnim karakteristikama simbola i formalnim uputama kako transformirati ove formalnosti, sadržaj kao da više nije potreban, odnosno, kao da ga u dijagramu tijeka nema. No, recimo, vjerovanje da pada kiša ima svoj sadržaj, on nije samo formalno i sintaktički određen kako se uklapa u neki proces i kakve su veze s drugim mentalnim objektima. Čini se da se on uklapa u veze s drugim mentalnim objektima upravo i na osnovi svog intrinzičnog sadržaja a ne samo na osnovi svoje forme. Prema tome, propozicionalni stavovi imaju i semantički dio koji strojna analogija ne obuhvaća zadovoljavajuće, odnosno iz same sintakse ne proizlazi i semantika. Dakle, teorija, odnosno objašnjenje, koja to ne obuhvaća, također ne može biti potpuna i zadovoljavajuća.

U pozadini dijagrama tijeka i rada računala, kao i bilo kojeg jednostavnijeg i kompliciranijeg konkretnog stroja stoji još jedan stroj koji zapravo omogućava da ovi strojevi rade. To je Turingov stroj. Turingov stroj je najapstraktnija matematička formulacija rada stroja (Ershov and

Palyutin 1984, Trahtenbrot 1978, Penrose 1989, Putnam 1975, Pećnjak 1995). Ukoliko se pomoću nekog Turingovog stroja može opisati bilo koji rad koji želimo, odmah se funkcionalistu nameće ideja da Turingov stroj upotrijebi za opis rada uma i tako pokuša objasniti što je um i kako on ima i svijest. Svijest bi bila jedna posebna realizacija nekog Turingovog stroja. Kakav je zapravo izgled rada Turingovog stroja? Ovdje ćemo to objasniti u osnovnim crtama. Turingov stroj operira nad beskonačnom trakom koja je podijeljena na beskonačni broj dijelova – obično se ti dijelovi zamišljaju kao neki pravokutnici. Na tim dijelovima su upisani neki podaci – u obliku jedinica (1) i nula (0) recimo. U svakom slučaju broj simbola koji se upisuju na traku mora biti konačan. Stroj može biti u nekom unutrašnjem stanju, može ostati u tom stanju ili prijeći u neko drugo stanje. Stroj ima konačan skup različitih mogućih unutrašnjih stanja. Zatim može pročitati što je upisano na traci, ostaviti to takvim ili izbrisati ili izbrisati i upisati novi simbol. Može ostati na mjestu na kojem je bio, pomaknuti se jedno mjesto u desno ili jedno mjesto u lijevo. U svakom trenutku ili bolje reći koraku, stroj izvršava jednu operaciju. Operiranje svakog stroja određuje se strojnom tablicom. Strojna tablica sastoji se od redaka i stupaca. Redak predstavlja alfabet, a stupac stanja stroja. Popis instrukcija isto tako je konačan. Svaki slijedeći korak točno je određen prethodnim. To izgleda, na primjer, ovako:

s_1DP

Ovdje, interpretacija može biti slijedeća: s_1 znači “upisati simbol ‘1’ u pravokutnik koji se trenutno pregledava”, D znači “prijeći na pravokutnik koji se nalazi odmah desno od trenutno pregledavanog pravokutnika”, a P znači “prijeći u unutrašnje stanje P”. Sada, ovisno o tome što će se pročitati na tom novo pregledavanom pravokutniku i što je u tablici dato pod stanjem D, stroj će izvršiti slijedeću operaciju, i tako dalje, dok ne dođe do instrukcije koja će mu reći da stane. Tada je zadatak završen. Turingov stroj koji može izvršiti bilo koji zadatak koji može izvršiti bilo koji drugi mogući Turingov stroj zove se Univerzalni Turingov stroj. Vidimo da je ova formulacija dovoljno apstraktna da je zapravo svaki konkretan stroj realizacija nekog Turingovog stroja. U tome je ležala i dodatna motivacija pokušaja opisa uma pomoću ove vrste (zapravo matematičkog) stroja. Ne samo da je svaka konkretna realizacija nekog stroja zapravo instancijacija Turingovog stroja nego izvedba Turingovog stroja ne ovisi niti o jednoj konkretnoj instancijaciji. Jedan te isti Turingov stroj može se realizirati na mnoštvo različitih načina. Ne samo to, nego niti supstancija ne igra nikakvu ulogu. Turingovi strojevi mogu biti realizirani materijalno, ali i nematerijalno: apstraktno ili u kakvoj supstanciji koja nije materijalna. Ništa u formulaciji Turingovog stroja ne određuje kako će on konkretno biti realiziran. Ako, dakle, mentalna stanja i procese odredimo apstraktnim operacijama i ako tvrdimo da je to njihova bit – funkcionalno operiranje apstraktno definirano, onda je staro pitanje o tome da li je naš um realiziran materijalno ili nematerijalno zapravo krivo postavljeno i ne treba nas puno zabrinjavati; pravo pitanje je autonomija uma i njegovog objašnjenja. Tako je to formulirao Putnam (1975a).

No, niti to nije zapravo riješilo problem zbog toga što se nameće moguća izomorfnost dvaju ili više sustava. Možemo pronaći dva sustava koji zadovoljavaju istu strojnu tablicu, a koji zapravo rade potpuno različite stvari. Jedan sustav može biti radio-aparat a drugi centralno grijanje deseterokatnice. Isto tako, ako i napravimo potpuni opis rada uma pomoću Turingova stroja, ta komplicirana i dugačka tablica može biti i opis rada cijele nuklearne podmornice. Nitko neće tvrditi da su čovječji um i nuklearna podmornica ista stvar! Još manje će netko tvrditi da nuklearna podmornica kao podmornica jest u stvari neki um! Dakle, ono što se prvotno nudilo kao demokratsko moguće rješenje – mogućnost višestruke realizacije istog – okrenulo se u eksces jer dozvoljava previše. Nema nikakvog argumenta – niti jedan funkcionalist ga nije ponudio – da je um, a pogotovo ne svijest, neki posebni Turingov stroj koji vrijedi samo za opis tih dvaju entiteta i razlikuje ih od ostalih. Za sve što bi se tvrdilo da je bitno funkcionalistički opis svjesnih operacija uma, može se pronaći neka interpretacija za koju se vidi da to naprosto nije svijest niti da u tome ima išta svjesnog.

U ovakvoj situaciji, funkcionalisti su se pokušali grčevito spasiti pokušavajući naći nešto što bi ipak moglo biti jasna realizacija svijesti i samog uma. Logičan izbor bio je sam mozak, barem ljudski mozak. Tvrdnja je bila da je funkcionalni opis rada mozga onaj specifični funkcionalistički opis i objašnjenje koje se traži. Međutim ni to nije spasilo funkcionaliste. Za ovakav opis također vrijedi prigovor da se lako nađe interpretacija koja je istovjetna funkcioniranju ljudskog mozga a koja vrlo jasno ne predstavlja nikakv um niti ne ostvaruje svijest. Takav

protuprimjer je zanimljivo konstruirao Block (1978) (vidi i Braddon-Mitchell and Jackson 1996, poglavlje 7.). Zamislimo da imamo potpuni funkcionalni opis rada neurona u mozgu. Slijedeće, izgradimo robota – androida koji u sebi nema ništa od materijala od kojih su izgrađeni ljudi nego je izgrađen od plutonija, urana, lutecija, titanovog oksida itd. S funkcionalističkog stajališta nije važno da li je centralni procesor koji općenito upravlja androidom u samom androidu ili izvan njega. Prema tome, zamislimo da je izvan njega. Procesor je povezan s robotom duplex radio vezom. Dalje, umjesto nekog organskog ili neorganskog objekta, odnosno silikonskog ili već kakvog god računala, simulaciju mozga i rada neurona izvršava cjelokupna populacija Narodne Republike Kine s Mao Tse Tungom! Svakom pojedincu dodijeljena je funkcija simulacije jednog neurona. Svakom pojedincu je pridodan telefon i neki broj, a telefon pokazuje s kojeg je broja taj pojedinac pozvan. U određenoj kombinaciji brojeva, ovaj nazvani pojedinac treba dalje nazvati neki broj. Svi imaju vrlo precizne upute tako da se točno zna što se treba pri bilo kojoj kombinaciji napraviti. Ovo osigurava da svaki pojedinac točno radi i modelira ono što radi pravi neuron. Tako imamo cjelokupno funkcionalističko detaljno simuliranje rada neurona i mozga. Pretpostavimo da ulazni okolinski podaci dolaze do središnjeg procesora kao i kod ljudi, te da izlazni podaci koji dolaze do tijela androida pokreću njegovo tijelo i udove tako da se on ponaša isto kao što se i mi ponašamo. U svim stvarnim situacijama nema razlike u ponašanju ovog androida – robota i pravih ljudi. Jedino, kod androida središnje procesiranje ne radi mozak nego strogo organizirana populacija Kineza.

U ovom slučaju, teško da bi se reklo da android zapravo ima mentalna stanja kao što su naša. Teško da bismo mogli tvrditi da ima bilo kakva mentalna stanja. Prema tome ne bismo ovom androidu, kojim upravlja kineska nacija, pridali posjedovanje mentalnosti i uma. Još manje bismo rekli da je i svjestan. Naše intuicije su vrlo snažne u prilog poricanja uma, mentalnih stanja i svijesti ovome androidu. Ne vidi se kako bi se zapravo najednom, operiranjem discipliniranih Kineza, pojavila neka individualna svijest koja nije svijest niti jedne jedinice Kineza, nego je ona jedinstvena svijest samog androida kao jedne nove i drugačije jedinice. Ovaj argument u formi misaonog eksperimenta čvrsto govori u prilog tome da niti funkcionalistički opis rada ljudskog mozga nije dovoljan da bi potpuno objasnio što je to ljudski um, i kako se realizira svijest.

Ovdje smo polako došli do referencijalnog funkcionalizma koji kaže da prvo specificiramo kauzalne funkcionalne veze između zdravorazumski ili strojno određenim specifikacijama mentalnih stanja i da onda pogledamo *što konkretno* realizira ovu funkcionalnost. Kod čovjeka je to ljudski mozak. Naravno, ukoliko ostanemo samo na toj tvrdnji da opišemo funkcionalistički rad ljudskog mozga, odmah se primjenjuje gore navedeni argument koji pokazuje da to nije dovoljno. No, tada se teoretičar može pokušati izvući tražeći da opis rada mozga mora biti širi i detaljniji nego što je to samo sa funkcionalističkog aspekta i nastojati tvrditi da su i specifične biokemijske, neurofiziološke i fizičke karakteristike mozga važne da bismo imali cjelokupno objašnjenje mentalnosti i svijesti. Tako se samo, u prvom koraku, funkcionalistički fiksira referencija na što usmjeriti istraživanje, a zatim to istraživanje

treba provesti detaljno, uvodeći i različite vrste objašnjenja koje su različite od samo funkcionalističkih objašnjenja. Tada ova vrsta funkcionalizma zapravo kolapsira u redukcionizam, fizikalizam, odnosno u neku vrstu teorije identiteta um-mozak. To je tema slijedećeg poglavlja.

Mislim da možemo reći da je funkcionalizam, u bilo kojem svojem obliku zapravo teorija višeg reda; ona na jednom apstraktnijem nivou može biti relativno dobra – možemo uočiti određene veze između raznih mentalnih entiteta, vidjeti kauzalne puteve između njih, fiksirati referenciju gdje zapravo istraživati, tako da zasluge funkcionalizma nisu male. Ali, kao teorija koja pretendira potpuno dati objašnjenje i opis djelovanja uma, te opis i objašnjenje svijesti, funkcionalizam ni iz daleka nije zadovoljavajući i zapravo nije ispunio svoju namjeravanu funkciju.

REPREZENTACIJSKA TEORIJA UMA

Reprezentacijska teorija uma najznačajnija je teorija uma koja ima za podlogu funkcionalizam i možemo reći da je ona zapravo najpotpunija interpretacija funkcionalizma. Kao što samo ime ove teorije kaže, ona uvodi i razrađuje pojam *reprezentacije*. Jasno je već samo po sebi da um odnosno njegovi sadržaji nisu identični vanjskom svijetu, nego da ga on zahvaća i *predočuje*. No, kako ga on to predočuje, kako ga reprezentira? Kakva je struktura reprezentacija i kakav je njihov odnos? Kako i kakve operacije se izvode nad njima? Sve ovo pokušava detaljno dati reprezentacijska teorija.

Um, kao što smo vidjeli u prvom poglavlju, ima razne sposobnosti. Vjerovanja i želje su propozicijski stavovi. To znači da se njihov sadržaj izražava propozicijom (dok se propozicija izražava rečenicom u nekom jeziku). Dakle, *stav* prema propoziciji može biti različit. Sukladno tome, i ponašanje će biti različito. No, vjerovanja i želje nikad ne dolaze izolirano, pogotovo što se tiče uzrokovanja a onda i objašnjenja ponašanja. Zato možemo zamisliti da um ima jedno „spremište“ vjerovanja, zatim jedno „spremište“ želja itd., ali su ona povezana međusobno na neke načine kao što su povezana i razna vjerovanja u „spremištu“ vjerovanja (isto vrijedi i za želje i ostale propozicijske stavove unutar njihovih „spremišta“). Ista propozicija može se naći npr. i u „spremištu“ želja i u „spremištu“ vjerovanja. Uzmimo primjer iz uvodnog poglavlja: propozicija je – „prozor je otvoren“. Vjerovanje se izražava: “Vjerujem da – prozor je otvoren” a želja “Želim da - prozor je otvoren”. Vjerovanje možemo interpretirati da uzimamo kao da je sadržaj istinit tj. kao da je takvo-i-takvo stanje stvari u svijetu. Želja izražava da subjekt hoće da se ostvari neko stanje stvari ako trenutno i nije tako. Ukoliko ništa drugo ne stoji na putu, subjekt će pokušati ostvariti svoju želju tj. promijeniti neko stanje stvari. Naravno, ako želim da prozor bude otvoren, onda vjerujem da je prozor zatvoren, te će, uz druga pozadinska vjerovanja zajedno (kao što je npr. vjerovanje da mogu ući u sobu u kojoj je prozor otvoren, da imam snagu zatvoriti ga, da je pod u sobi stabilan itd.) ova želja uzrokovati (odnosno, bit će glavni uzrok) moje ponašanje. Naravno, ponašanje će se sastojati od odlaska u sobu u kojoj je otvoreni prozor i poduzimanja radnji od kojih se sastoji zatvaranje

prozora. No, kakva je struktura samog tog vjerovanja, struktura odnosa ovog vjerovanja s drugim vjerovanjima te struktura odnosa s drugim propozicijskim stavovima, da bi se uopće ovo uzrokovanje ponašanja moglo objasniti? U pozadini mora biti neka uređena struktura. Da bi se objasnila ova struktura reprezentacija i njihovih odnosa, postavljena je tzv. hipoteza „jezika misli“. Prvi ju je iznio Jerry Fodor (1975) u djelu *Language of Thought*, te je razradio u knjizi *LOT 2* (2008).

Kao što samo ime već govori, uporabljena je analogija s prirodnim jezikom. Jedna od najvažnijih komponenata prirodnog jezika tj. njegove gramatike je sintaksa. Ona se sastoji od pravila pomoću kojih se kombinira i sastavlja redosljed jezičnih jedinica. Npr. rečenica nekog prirodnog jezika sastoji se od fraza i riječi, a kako se te riječi mogu ili moraju staviti u poredak te kako se fraze stavljaju u međusobni odnos i poredak određuju sintaktička pravila. Sintaktička pravila su takva da omogućavaju određene poretke i isključuju druge poretke. Znači, nije svaki poredak fraza i riječi moguć (da bismo imali rečenicu koja je smisljena). Sintaktička pravila operiraju nad formalnim, odnosno strukturalnim, obilježjima jedinica bez obzira na njihove sadržaje odnosno značenja. (No, naravno, ovo ipak nije samorazumljivo i o tome se u lingvistici može dosta diskutirati.) Znanje sintaktičkih pravila je za izvornog govornika nekog jezika *implicitna*. Znanje tih pravila koja se rabe u produkciji govora nije svjesno znanje. Analogno ovome, Fodor i ostali zastupnici hipoteze „jezika misli“ objašnjavaju i reprezentacije. Naime, propozicionalni stav je reprezentacija jer nam predočuje nešto na neki način: npr. predočuje nam svijet ili stanje svijeta, odnosno neki

njegov fragment, da je takav-i-takav. Elementi od kojih je sastavljena reprezentacija moraju biti poredani na neki način a taj poredak određuje sintaksa „jezika misli“. Elementi od kojih se sastoji propozicija su pojmovi. Uzmimo primjer. Recimo da osoba XXYY vjeruje da je Zagreb lijepi grad podno brda Medvednice. (Ovdje je propozicija uzražena rečenicom u Hrvatskom jeziku, no ona – jedna te ista propozicija – može se izraziti u bilo kojem drugom jeziku /npr. engl. Zagreb is a beautiful town under the hill Medvednica/.) Propozicija sadrži nekoliko pojmova – „ZAGREB“ „LIJEPO“ „GRAD“ „ISPOD“ „BRDO“ „MEDVEDNICA“ koji stoje u međusobnim odnosima. „GRAD“ se odnosi na „ZAGREB“ ali ne i na „MEDVEDNICA“. Također, „ISPOD“ locira „ZAGREB“ prema „MEDVEDNICA“ a ne obrnuto. Dakle, prema reprezentacionalistima, važan je oblik i poredak elemenata; odnosno, pravila pomoću kojih se operira su takva da ona operiraju uzimajući oblik elementa, i s obzirom na to, stavljaju ga u odgovarajući poredak. To možemo usporediti sa slagalicom. Svaki element u igri slagalice paše na točno određeno mjesto i to prema svojim formalnim odnosno strukturalnim obilježjima. Svaki pojedini element slagalice nema nikakav sadržaj – tek kada se sastavi cijela slagalica, onda tek imamo posla s nekim sadržajem – nekom slikom na primjer. Dakle, mogli bismo napisati pravila za slaganje slagalice koja bi vodila računa samo o obliku i poretku elemenata bez ikakvog pozivanja na bilo kakav sadržaj. No, na kraju bi sadržaj „izronio“. Dakle, dovoljna bi bila samo formalna ili strukturalna obilježja i sintaktička pravila. Isto tako, prema reprezentacionalistima, kako um slaže pojedinu reprezentaciju, tako um operira i nad odnosima među različitim

reprezentacijama. Ponovo, dovoljna su samo strukturalna obilježja i sintaktička pravila. Nije potrebno nikakvo pozivanje na sadržaje i značenja. Lijep primjer daje Heil (2004, str. 109): Uzmimo pravilo *modus ponens*: $p, p \rightarrow q \mid q$. Ono kaže „Tamo gdje su p i q rečenice, iz p i ako p onda q , slijedi q . Dakle, ako imamo napisane simbole p i $p \rightarrow q$, onda nam pravilo dozvoljava da napišemo novi simbol: q . To vrijedi za bilo koje rečenice. Nije nam potrebno poznavanje sadržaja i značenja rečenica, ali slijedeći ovo pravilo dobivamo rezultat koji je sukladan smislenim sadržajima i značenjima. Evo primjera: Ako pobijedim na zadnjoj utrci, onda ću biti svjetski prvak; pobijedio sam na zadnjoj utrci; dakle, ja sam svjetski prvak.

Da polako zaključimo: prema reprezentacionalistima, um proizvodi i operira nad reprezentacijama. To operiranje se izvodi samo prema sintaktičkim pravilima na osnovi formalnih odnosno strukturalnih svojstava. Nije potrebno uvoditi nikakva pozivanja na sadržaje i značenja, odnosno na *semantiku*. No, sintaktičko operiranje vjerno odražava sadržaje i značenja.

Ovoj teoriji, kao i manje-više svemu drugome u filozofiji, izneseni su razni prigovori. Ovdje ćemo prikazati najjači i najviše diskutirani. Taj argument protiv reprezentacionalizma, odnosno, sintaktičke teorije uma, zove se „Argument kineske sobe“. Iznio ga je John Searle (izvorno u 1980; vidjeti i 1984, poglavlje 2.). Zamislimo nekog čovjeka koji se nalazi u jednoj finoj prostranoj sobi. Čovjek je izvorni govornik hrvatskog jezika i ne zna nikakav drugi jezik. Soba ima dva otvora. Kroz jedan otvor u sobu mu stižu pločice s kineskim

simbolima. Na raspolaganju mu je knjiga na hrvatskom koja sadrži pravila kako operirati s kineskim simbolima. Također, na raspolaganju mu je i spremište pločica s kineskim simbolima. Knjiga pravila operiranja s kineskim simbolima kazuje ovako: kada dobiješ znakove tog-i-tog oblika i u tom-i-tom redosljed, onda uzmi te-i-te pločice i sastavi ih po ovakvom redosljed. Kada čovjek to učini, kroz drugi otvor šalje ono što je sastavio van. Ono što on dobiva na ulazu, sastavili su kinezi kojima je kineski materinji jezik. Oni koji znaju kineski jezik i čitati kineske simbole, razumiju da su to pitanja iz slavne kineske povijesti do 1949. godine. Također, oni koji znaju kineski jezik i čitati kineske simbole razumiju da su ono što izlazi iz sobe točni odgovori na ova pitanja. Dakle, izvorni govornik ili govornik koji zna kineski, rekao bi da netko tko sastavlja ove odgovore, razumije sadržaje pitanja koja dobiva i onda na njih smisljeno odgovara. No, kakva je stvarna situacija? Čovjek u sobi, sjetimo se, ne zna kineski. On samo slijedi sintaktička pravila – pravila koja operiraju nad formalnim odnosno strukturalnim svojstvima znakova koje dobiva i koje sastavlja. Čovjek u sobi uopće nema nikakvo razumijevanje niti shvaća *sadržaje* onog što dobiva a niti onog što sam sastavlja. Prema tome, zaključak koji izvlači Searle iz ovog misaonog eksperimenta jest da sintaksa nije dovoljna da bi mogla objasniti i to kako ljudskii umovi imaju i razumijevanje sadržaja i značenja onoga s čim operiraju. Prema tome, reprezentacijska teorija koja se služi u objašnjenju rada uma samo sintaktičkim pravilima ne bi bila potpuna teorija.

FIZIKALIZMI

Uspjesi fizikalnih znanosti, naročito same fizike, ponukao je filozofe da bi se, u krajnjoj konzekvenci, sve što postoji moglo reducirati na fiziku; fizika bi imala status temeljne znanosti i sve što postoji postojalo bi u svijetlu fizike. Neuroznanosti su dale prirodnu ideju da bi se i mentalno, dakle psihičko, također moglo reducirati na neurofiziološko i, budući da se neurofiziološko sastoji od fizikalno-kemijskih objekata i procesa, u krajnjoj konzekvenci, na fizičko. Redukcionisti priznaju postojanje mentalnog, psihičkog i onog što to uključuje kao što su propozicionalni stavovi, emocije, osjeti, percepcija, svijest itd., ali sve to je realizirano na neki specifični neurofiziološki, dakle fizički, način. Specifičnost mentalnog bi se sastojala u tome što bi ono bilo ostvareno na neki specifični fizički način, koji se daje potpuno opisati fizikalno.

Dakle, materijalisti tvrde da ne postoji mentalno i psihičko koje bi bilo nezavisno od materijalnoga i da je ono izgrađeno isključivo iz materijalnih, odnosno, fizičkih sastavnica. Mentalno koje bi bilo realizirano u nekoj drugačijoj supstanciji od materijalne, ne postoji u našem svijetu. Mentalno se sastoji od fizičkog i daje se opisati fizikalno i ne postoji nematerijalna supstancija pomoću koje bi se ono realiziralo. Um i svijest realizirani su moždanim stanjima i procesima. Materijalistička doktrina u suvremenoj varijanti postupno se razvijala i zadobila nekoliko varijanti i imena.

Za naše svrhe ovdje je najzanimljivije razlikovanje dvije vrste fizikalizma odnosno materijalizma tj. dvije vrste teorija identiteta umozak a to su, *Type-type* teorija identiteta (ili fizikalizam) i *Token-token* teorija identiteta (ili fizikalizam). Kako je, naročito izraz *token*, vrlo teško jasno i jednoznačno prevesti na hrvatski jezik, to sam ova dva pojma ostavio u originalnoj engleskoj verziji. Razlaganjem argumenata za ove pozicije bit će razvidno što je što. Pioniri fizikalističkog pogleda, odnosno teorije identiteta, su Place (1956), Smart (1959), Lewis (1966) i Armstrong (1968), dok ih je slijedilo mnoštvo drugih filozofa razvijajući ove ili one detalje (vidi Sesardić 1984 i Hill 1991).

Type-type teorija identiteta kazuje da su tipovi mentalnih stanja/procesa identični s tipovima fizičkih stanja/procesa. Na primjer, svjesna stanja/ procesi realizirani su fizičkim stanjima/procesima koji se odvijaju u mozgu. Kod ove vrste identiteta, uvijek je isti tip mentalnog stanja ostvaren istim tipom fizičkog stanja. Dakle, trebala bi postojati korelacija jedan – jedan između (tipova) mentalnih stanja/procesa i (tipova) fizičkih (moždanih) stanja/procesa. Drugim riječima, ako ja vjerujem da se državni udar može izvesti samo oklopnom divizijom i ako pukovnik Tejero vjeruje da se državni udar može izvesti samo oklopnom divizijom, onda i on i ja imamo ne samo isto vjerovanje, nego je ono ostvareno istim fizičkim stanjima/procesima u njegovom i u mom mozgu. Također, ako ja to vjerujem u jednom trenutku t_1 , i kasnije u nekom trenutku t_2 , u oba ova vremena vjerovanje je ostvareno istim fizičkim procesom u mozgu. Isto vrijedi i za pukovnika Tejera. Da bismo do kraja razjasnili stvari, ako ja imam navedeno vjerovanje u t_1 , pa onda prestanem

u to vjerovati neko vrijeme u periodu t_2 - t_3 , a nakon tog vremena, u t_4 , ponovo počnem u to vjerovati, realizacija tog vjerovanja u t_1 i u t_4 jest ostvarena istim fizičkim događajem u mozgu. Prema tome, kod *type-type* identiteta, svaka pojedinačna instanca nekog mentalnog stanja ili procesa ostvarena je istim stanjima ili procesima u mozgu. Isto stanje ili proces u mozgu, uvijek je realizacija istog mentalnog stanja ili procesa. Također, druga važna tvrdnja jest ta da je psihološki karakter nekog mentalnog događaja, zapravo doslovno, fizički aspekt fizičkog događaja u mozgu. Ono što me boli kad sam u stanju bola ili instanca vjerovanja da je valcer lijepi ples, jest u stvari identično određenom fizičkom (moždanom) događaju.

Ukoliko su svjesna i ostala mentalna stanja fizička (fizičko-kemijska) stanja i procesi u mozgu, onda se ona, kako smo već rekli, dadu opisati pomoću fizikalnih teorija koje uključuju samo objekte (entitete) i zakone fizike. Malo pojednostavljeno rečeno, fizikalne teorije se sastoje od specificiranih veza između objekata te formula i algoritama koji opisuju i objašnjavaju njihovo ponašanje te njihova svojstva (Rothman 1966). Zakoni su dakle formulirani matematički i to na dva načina – jedna vrsta zakona je kauzalna koja opisuje razvoj nekog sustava iz trenutka u trenutak s obzirom na početne uvjete i za svaki trenutak daje točno određeno stanje razvoja procesa ili točno karakterizira stanje (Kittel, Knight, Ruderman, 1982), dok je druga vrsta statističke prirode i daje određene vjerojatnosti za određene događaje (Strnad 1985, Wichmann 1988). Većina događaja u mozgu su takve prirode da se mogu opisivati prvom, kauzalnom, vrstom zakonitosti. No, sigurno je da se određeni

procesu opisuju ili će se morati opisati i statističkim zakonitostima ili nedeterminističkim formulama kvantne mehanike (Nunn 1996, Penrose 1989). U svakom slučaju, ono što želim reći je to da ona fizikalna teorija koja opisuje stanja i procese, u mozgu, formulama i algoritmima koji karakteriziraju uzročno-posljedično odvijanje procesa uključuje samo fizikalne tj. fizičke komponente. Na neki način ona je “sintaktička” – iz jednog događaja koji sadrži određena fizikalna/fizička svojstva razvija se drugi događaj koji također sadrži samo daljnja fizikalna/fizička svojstva. Kako je to onda još i nešto drugo, vrlo različito od toga - npr. vjerovanje sa sadržajem “miš je mali glodavac” ili vizualna percepcija nosača aviona koji isplovjava iz San Diega? No, dalje, ukoliko imamo opis i objašnjenje događaja u mozgu pomoću teorije koja daje samo slijed fizikalnih/fizičkih događaja koji uključuju samo fizikalne objekte i fizikalne zakone, tada doista dobivamo neku vrstu “fizikalističkog funkcionalizma”, odnosno ovakav pokušaj objašnjenja, zbog prirode formula i algoritama koji se rabe u fizici, doista kolapsira u funkcionalistički oblik iz kojeg se uopće ne vidi kako mi imamo mentalna sadržajna stanja i koji su to uopće sadržaji. Naime, funkcionalizam specificira da je bit mentalnog stanja ili procesa to što on funkcionalno obavlja tj. kakva mu je funkcionalna (uzročna) uloga. Tako više nije identitet mentalnog i fizičkog (specifičnog neurofiziološkog) ono što karakterizira objašnjenje mentalnosti nego sama funkcionalna uloga. No, ta uloga može biti ostvarena na bezbroj različitih načina – i fizički, ali nema u principu prepreke da može biti ostvarena i nefizički. Dakle, funkcionalna, sintaktička, uloga nije karakterizirana fizičkim, materijalnim, pa tako više nije zanimljiva *type-*

type teza u smislu identificiranja mentalnog i fizičkog. Osim toga, vrlo je teško samo pomoću slijeda operacija, dakle sintaktički, karakterizirati i sadržaje mentalnih i/ili svjesnih stanja koji su semantičke prirode. No to je problem za sebe i neću o tome ovdje duljiti.

Također, ne vidi se ni razlika između svjesnog i nesvjesnog. Mnogi procesi u mozgu nisu niti mentalni niti svjesni – kako to da neko odvijanje procesa koji uključuju samo fizikalne objekte koji se ponašaju po fizikalnim zakonima jesu svjesni, a drugi koji uključuju iste objekte i zakone u istom organu, to nisu? Na vidiku nije neki principijelni odgovor na to pitanje i to smatram velikom teškoćom za fizikaliste bilo koje vrste. Jasno, može biti da bi se pokazalo da postoje *tipovi* fizikalnih procesa koji su isključivo vezani za svijest dok niti jedan takav proces nije sudionik ne-svjesnih procesa u mozgu i obratno, *tipovi* fizikalnih procesa koji su isključivo vezani uz ne-svjesno i ne pojavljuju se nikada kada imamo svjesne procese. Ali, i tada mi se čini da može postojati samo korelacija tipa svjesno – određeni tip fizičkog događaja; i dalje nije jasno kako bi se iz samih teorija, formula i objekata koje oni uključuju, vidjelo da jedna vrsta tih teorija i formula opisuje svjesni proces a da druga vrsta formula to ne opisuje. Morali bismo korelirati, *bona fide*, izjave ispitanika s pronađenim i registriranim što se tada u mozgu ispitanika fizički događa.

To i jest u stvari ono što se radi u neurofiziološkim istraživanjima i eksperimentima. Recimo da se ispitaniku pokazuje neki jednostavni objekt i da ispitanik mora imenovati taj objekt čim ga postane svjestan. Postoje samo korelacije onog što ispitanik kaže i zamijećenih fizičko-kemijskih procesa u mozgu u tim trenucima. Reakcije ispitanika

obično, u većini slučajeva kasne oko 200 - 500 milisekundi; tj. ispitanik kaže da je svjestan nečega, u prosjeku, u navedenoj odgodi od 200 - 500 milisekundi ili čak i više, nakon što se fizičko-kemijski procesi već odvijaju nakon pokazivanja stimulusa (Nunn 1996, poglavlje 2.). Ne samo to, vrijeme od oko 200 milisekundi je najmanje vrijeme koliko mora trajati stimulus da bi subjekt uopće mogao postati svjestan toga što mu se pokazuje. Ukoliko se stimulus pokazuje kraće vrijeme, koliko god bio intenzivan, subjekt ga uopće ne postaje svjestan. Prema tome, fizički procesi *prethode* svjesnosti ispitanika. Oni se ne pojavljuju identično u vremenu. To sugerira razliku između fizičkih neuro-procesa i samih sadržaja svijesti. Dakle, iz samih neuroznanstvenih istraživanja čini se kao da se sugerira da veza između svijesti i fizičkog u mozgu nije identitet nego da ima nekakav drugačiji oblik.

Fizikalisti su nastojali pokazati da pojmovi i propozicije koje se referiraju na mentalno i fizičko (specifično moždano) imaju isti status kao i ostali pojmovi i propozicije u znanosti. Pomoću fenomenalnog aspekta boli mi određujemo referenciju na ono što ostvaruje bol. To što ostvaruje bol i što je dakle u biti bol, jest skriveno i prepušta se empirijskoj znanosti da to otkrije. Pogledajmo analogije s drugim pojmovima i propozicijama i načinu na koji znanost otkriva što je što. Standardni primjeri su “munja” i “električno pražnjenje” te “voda” i “H₂O” (Smart 1959). Prvo, određujemo referencu ostenzivno pokazujući, na primjer, na ono što ispunjava korita jezera, rijeka, bare itd. i pomoću opisa “prozirna tekućina bez okusa i mirisa”; tome, što god da je, dajemo ime “voda”. Kasnije, kada znanost razvije svoje metode, i kada možemo detaljno analizirati od čega se sastoji

“prozirna tekućina bez okusa i mirisa”, odnosno voda, mi ustanovljujemo da se sastoji od povezanih atoma vodika i kisika i to u omjeru 2:1. Dakle, sada imamo vezu između pojmova “voda” i “ H_2O ”, odnosno možemo reći “voda je H_2O ”. Karakteristika znanstvenih identifikacija smatrana je kontingentnom. To bi značilo da veza između “vode” i “ H_2O ” nije nužna, već je slučajno to tako. U ovom našem svijetu, voda je ostvarena s H_2O . To ne mora biti tako, smatrali su fizikalisti (Smart 1959). Ovaj način određivanja referencije i znanstvenog istraživanja poslije toga možemo primijeniti i na pojmove koji označavaju svjesne procese ili, općenitije, mentalne procese. Uzeli smo bol kao primjer. Fenomenalni aspekt boli tj. kako mi osjećamo neki bol, odnosno kako izgleda biti u stanju bola određuje referenciju. Zatim se potraže fizikalni događaji koji se događaju uvijek kada netko osjeća tu bol. Fizikalni događaji koji se otkriju da se događaju u mozgu su u stvari ostvarenje te boli. Sada imamo vezu između pojma “bol” i pojma “fizički moždani proces $\alpha\beta\gamma\delta$ ”. “ $\alpha\beta\gamma\delta$ ” je skraćenica za neku formulu ili niz formula koje opisuju fizičke događaje u mozgu. Dakle, bol je fizički moždani proces $\alpha\beta\gamma\delta$. Veza između boli i $\alpha\beta\gamma\delta$ bila bi kontingentna. Moguće je da bol bude i nešto drugo.

No, pažljivija analiza (Putnam 1975) pokazala je da je voda *nužno* H_2O . Iako je ovo otkriveno empirijski tj. aposteriorno, ova identifikacije je nužna. Nije moguće da išta što jest voda bude nešto drugo nego H_2O . Nešto što jest “prozirna tekućina bez okusa i mirisa” može biti i nešto drugo nego voda, npr. supstanca s kemijskim sastavom XYZ (Putnam 1975), jer je ovo u stvari neodređen opis i može referirati na više različitih stvari. Međutim, “voda”, kada je referencija jednom određena, a

ona može biti određena, i to je tako i bilo, i kada još uopće nemamo nikakvu ideju o možebitnom mikro, odnosno kemijskom sastavu, ona referira samo na jednu određenu supstancu; Fregeovskim ili Russellovskim rječnikom rečeno, ona je vlastito ime koje točno određuje jednu i samo jednu supstanciju (ovdje rabim pojam “supstancija” u kemijskom smislu), a to je voda; a voda je H_2O i ništa drugo! Voda je H_2O i kada mi to još ne znamo! Supstancija s kemijskim sastavom XYZ, iako ima iste površinske karakteristike kao i voda – nema okus, miris niti boju - svakako nije voda, jer ima kemijski sastav drugačiji od H_2O , a za nešto što ima drugačiji sastav od H_2O ne bismo niti slučajno rekli da je voda.

Isti postupak bi sada morao vrijediti i za neki fenomenalni mentalni pojam, npr. bol. Bol ne bi dakle kontingentno, nego nužno bio $\alpha\beta\gamma\delta$. Ništa drugo osim $\alpha\beta\gamma\delta$ ne bi bilo bol. Veza između pojma “bol” i “ $\alpha\beta\gamma\delta$ ” morala bi biti nužna. “ $\alpha\beta\gamma\delta$ ” bi izričalo bit, esenciju, bola isto kao što “ H_2O ” izriče bit, esenciju, vode.

No, ne vidi se kako bi ikakva nužnost postojala između fenomenalnog aspekta boli i čisto fizikalnih događaja $\alpha\beta\gamma\delta$ te, dakle, pojma “bol” i opisa događaja u mozgu nizom formula “ $\alpha\beta\gamma\delta$ ”. Zar je bit bola ili bilo kojeg drugog svjesnog događaja izrečena bezličnim nizom fizičkih događaja i formulama kojima se ti fizički događaji opisuju? Zar upravo bit bola nije u tome kako izgleda biti u stanju tog bola odnosno kakvu unutrašnju kvalitetu mi osjećamo kada nas nešto boli. Bit je upravo u kvaliji bola.

Mi nemamo introspekciju $\alpha\beta\gamma\delta$ nego kvalitetu stanja odnosno kvalije. To je ono što je za nas važno kod bola ili bilo kojeg drugog

svjesnog događaja. Kod fenomenalne svijesti – osjeta i percepcije važno nam je kako izgleda biti u nekom osjetilnom stanju. Kod propozicionalne svijesti važno nam je svjesno prepoznavanje i intencionalno baratanje sadržajem kojeg sadrže reprezentacije propozicija određenog propozicionalnog stava. Kako je to neki fizički proces εζηλ?

S druge strane, *token-token* teorija zastupa slabiju tvrdnju od *type-type* teorije, ona tvrdi da je svako mentalno stanje ostvareno nekim fizičkim stanjem. Vidimo da *type* teorija implicira *token* teoriju, ali da obrnuto ne vrijedi. Isti tip mentalnog stanja, u *token-token* teoriji, ne mora uvijek biti realiziran istim fizičkim procesom. No, on jest realiziran nekim fizičkim procesom, ovakvim ili onakvim, koji se dakle, opisuje fizikalno. Gornji argument, kojim se pokazuje kako iz fizikalnih formula ništa ne slijedi o svjesnosti i mentalnim sadržajima, bilo fenomenalnim, bilo propozicionalnim, primjenjuje se i na *token-token* teoriju. No ipak, u jednom *token-token* teorija ima veću prihvatljivost nego *type-type* teorija. Sada je moguća tvrdnja da je mentalno i svjesno samo kontingentno vezano za fizičko. Isti mentalni proces ili stanje može biti ostvareno različitim fizičkim procesima. Ali, ne znači li to ipak da je onda bit mentalnog u nečem drugom tj. pomoću nečeg drugog određujemo bit mentalnog ukoliko *različiti* procesi ostvaruju jedno te isto u mentalnom smislu. Kako identificiramo iste mentalne događaje kada nisu uvijek identični istim fizičkim procesima, a *token* teoretičari tvrde da mentalno jest identično s nekim(!) fizičkim procesom? Mislim da je to problem sam za sebe i da nije dat niti jedan zadovoljavajući odgovor. Standardni odgovor je da se mentalno identificira prema funkcijama, odnosno prema

uzrocima i posljedicama, koje ova *token* mentalna stanja imaju (npr. Lewis 1966). Ali, sada imamo *izravni* kolaps *token-token* fizikalizma u funkcionalizam. Isto razmatranje, gore navedeno, o kolapsu *type* teorije može se na ovom mjestu primijeniti i na *token* teoriju.

SUPERVENIJENCIJA

Teorija supervenijencije (Berčić 1988, Kim 1993, Horgan 1981, 1993, Haugeland 1982, Melnyk 1991), jest teoretski aparat koji je izgrađen kako bi se na stroži i bolji način odredile veze između određenih bazičnih entiteta i viših ili većih cjelina koje čine strukturirane kombinacije ovih bazičnih entiteta. Ona ima širju teoretsku primjenu, može se primijeniti na uzrokovanje, u određivanju moralnih vrijednosti i pojmova, u ontologiji itd. Ovdje ću prikazati osnovnu primjenu na problem um-tijelo.

Ponekad se supervenijencija u smislu superveniranja mentalnog nad fizičkim karakterizira slijedećim iskazima (Melnyk 1991):

1. Nužno, ukoliko postoji razlika u mentalnom, postoji i razlika u fizičkom.
2. Nužno, ukoliko nema razlike u fizičkom, nema razlike niti u mentalnom.
3. Nije moguće da postoji razlika u mentalnom, ali da nema razlike u fizičkom.

Sada ću ukratko objasniti što to znači.

Teorija supervenijencije je trebala poslužiti fizikalistima kao oruđe pomoću kojeg bi se točno odredile veze između temeljnih entiteta i svojstava, koji

su fizički entiteti i/ili fizička (fizikalna) svojstva, i svojstava višeg reda, a za koja fizikalisti smatraju da ne postoje kao specijalne supstancije i neovisni entiteti od dijelova i načina kako su ti elementarni dijelovi povezani u cjelinu. Dakle, prema fizikalistima, fizičko determinira sve što postoji u našem svijetu. Preciznije rečeno, sve što postoji izrasta iz temeljnih fizičkih entiteta, a to su mikrofizički objekti i fizikalni zakoni koji vladaju među njima. Teoretičari supervenijencije ovo određuju kao bazu supervenijencije, odnosno, mikrofizički objekti i zakoni koji vrijede za mikrofizičke objekte, događaje i procese su baza supervenijencije, a njihova su svojstva subvenijentna svojstva. Sve što je na intermedijarnom i makro nivou determinirano je bazom supervenijencije. Na primjer, čak i fizikalni zakoni koji vrijede u svijetu makrofizike, posljedica su bazičnih zakona mikrofizike. Sami fizičari kažu da zakoni mikrofizike objašnjavaju i mikrosvijet i određuju kako izgleda svijet makrofizike - on je ovisan o kompleksnim odnosima u mikrosvijetu (Wichmann 1988). Ipak, formuliranje supervenijencije nije jedinstveno i jednoznačno. Ovisi točno koju svrhu želimo postići i tako se mogu formulirati slabije i jače varijante supervenijencije. Tako postoji nekoliko glavnih oblika supervenijencije, slaba, jaka, lokalna, globalna, supervenijencija Nad-svojstava, superdupervenijencija. No, ovdje nije cilj iscrpno razmotriti sve moguće formulacije supervenijencije i njihovih posljedica – samo to bi moglo biti dostatno za jedan cjeloviti rad – nego samo želim ukazati primjenu u filozofiju uma.

Pogledajmo navedene tri propozicije pomoću kojih smo karakterizirali supervenijenciju. Supervenijentnim rječnikom rečeno, prva propozicija bi odredila da u bazi supervenijencije mora postojati fizička razlika ukoliko imamo različite mentalne događaje. Ukoliko imamo dva različita mentalna događaja, moramo

imati dva različita (mikro)fizička događaja kao njihove subvenijente. Ovi (mikro)fizički događaji, svaki za sebe, konstituiraju mentalni događaj; mentalni događaj se sastoji samo u svijetlu ostvarenja nekih (mikro)fizičkih procesa. Drugo, isti (mikro)fizički proces uvijek je baza istog mentalnog procesa. Treće, moguće je da jedna te ista vrsta mentalnih procesa supervenira nad različitim (mikro)fizičkim događajima. Čini se da ovako formulirana supervenijencija, najčešća u filozofiji uma ide više u smjeru *token* fizikalizma, no, naravno, dade se formulirati i oblik koji bi bio za upotrebu *type* fizikalistima. Neki teoretičari supervenijencije se pak, zbog posebnosti teorije, ne žele svrstati niti u jedan tabor smatrajući samu supervenijenciju drugačijom od ovih fizikalizama.

INTENCIONALNOST

Intencionalnost je svojstvo mentalnih stanja da se *odnose na nešto*, mentalna stanja su o *nečemu*. Možemo isto tako reći da su mentalna stanja usmjerena na nešto, usmjerena su prema nečemu. Često kada pogledamo prijatelja i zamjetimo odsutan izraz lica, priupitamo ga: „O čemu sada misliš?“. Tu je jasno implicirano da misao, kao mentalno stanje (odnosno proces), ima neki sadržaj koji se odnosi na neke objekte. Objekti na koje se mentalna stanja odnose najčešće su neki objekti u vanjskom svijetu; no, mentalna stanja mogu se odnositi, tj. biti i o nepostojećim objektima, kao i o drugim mentalnim stanjima, a neko mentalno stanje može se odnositi na samog sebe tj. može biti autoreferencijalan. Ne rabeći pojam intencionalnost, već smo implicitno nešto o tome rekli kada smo govorili o reprezentacijama – reprezentacije nam, na primjer, predočuju da je neki dio svijeta

takav-i-takav. To je *perspektiva* koju mentalno stanje (a tako i um kao cjelina) ima prema svijetu (vidi npr. Crane 2001, poglavlje 1). Daljnje svojstvo perpektive mentalnog stanja, odnosno njegove intencionalnosti, jest tzv. *aspektualni oblik* (vidi Searle 1992, poglavlje 7). Objekti, bilo perceptivno, bilo propozicionalno, uglavnom su nam predloženi zapravo samo djelomično. Kada gledamo avion lovac Suhoj Su-27 vidimo, u kojem god položaju u odnosu na njega bili, samo jedan dio. Vidimo ga sprijeda ili sa strane, odozgora ili odozdola, pod nekim kutom itd. Nikada ne možemo vidjeti istovremeno sve dijelove i istovremeno cijeli avion. To vrijedi za sve trodimenzionalne objekte. Isto tako, svako pojedinačno vjerovanje čiji je sadržaj izražen nekom propozicijom, govori uglavnom samo određene stvari o objektu. Dakle, jedan te isti objekt može biti predloženi na razne načine.

Možemo razlikovati izravnu i neizravnu, odnosno izvedenu intencionalnost. Mentalna stanja imaju izravnu intencionalnost. Svojstvo intencionalnosti kao usmjerenosti na objekt, odnosno to da je mentalno stanje o nečem, intrinzično je svojstvo tog mentalnog stanja. Samom imatelju tog mentalnog stanja ono nešto znači. Sadržaj tog stanja odnosno ono kako je objekt prezentiran pomoću intencionalnosti stanja, izravno je dostupno subjektu tog stanja. Neizravna, izvedena (derivirana) intencionalnost je intencionalnost objekata koji se moraju interpretirati od strane nekog interpretatora. Na primjer, prometni znakovi nešto znače, ali samo za onog tko ih interpretira. Stanja samog prometnog znaka njemu samom ništa ne znače.

Može se također postaviti pitanje da li sva mentalna stanja imaju svojstvo intencionalnosti? Što je s elementarnim mentalnim stanjima kao što su osjeti – npr. škakljanje ili bol – da li su oni usmjereni na nešto ili o čemu bi oni

uopće bili? Neki autori će reći da bol nema svojstvo intencionalnosti. Bol naprosto nema nikakv sadržaj, nije ni o čemu. Bol ima fenomenalno svojstvo tj. izgleda nekako biti u stanju bola; bol, za onog tka ga doživljava, ima nekakvu unutrašnju kvalitetu koja se doživljava, ali on nema neki sadržaj koji bi bio o nečemu. No, oni autori koji zastupaju tzv. *intencionalizam* tvrdit će da i tako elementarno mentalno stanje kao što je bol ima intencionalnost. Intencionalisti su oni autori koji smatraju da sva mentalna stanja, bez iznimke, imaju intencionalnost. Prvi koji je to tvrdio bio je Franz Brentano, pa se teza intencionalizma naziva još i Brentanova teza. Suvremeni branitelji takve teza su npr. Tye (1997) i Crane (1998, 2001). Kako oni argumentiraju da i bol ima svojstvo intencionalnosti? Kada nas nešto boli, onda nas manje-više boli neki određeni dio tijela. Uzmimo primjer kada nas boli koljeno. Intencionalist će onda reći da je u stvari nama koljeno (kao dio tijela) prezentirano na jedan specifičan način; osim toga bol je osjet koji nas upozorava da na onom mjestu gdje ga osjećamo, nešto nije u redu. To bi, dakle, prema intencionalistima, bio intencionalni sadržaj boli.

Napomena: Dva poglavlja ove skripte, *Fizikalizmi* i *Supervenijencija*, objavljeni su prvi put kao prvi dio mog članka “Ljudski um, ljudski mozak i praktičnost” koji je izašao u knjizi Žebec, M. S., G. Sabol, M. Šakić i M. Kotrla Topić (urednici), *Mozak i um – trajni izazov čovjeku*, Zagreb, Institut društvenih znanosti “Ivo Pilar”, 2006. Zahvaljujem se urednicima za dopuštenje da dio tog teksta upotrijebim i ovdje. Poglavlje *Funkcionalizmi* sastoji se od većeg dijela teksta koji je pod naslovom „Nefunkcionalnost funkcionalizma“ izrađen u koautorstvu s Tomislavom Janovićem i objavljen u zborniku *Aspekti uma* koji su uredili Maja Hudoletnjak Grgić, Filip Grgić i Davor Pećnjak (Zagreb, Institut za filozofiju, 2011). Zahvaljujem se urednicima i koautoru za dopuštenje da dio tog teksta upotrijebim i ovdje.

Komentari: Ukoliko želite, komentare i sugestije možete poslati na adresu:

davor@ifzg.hr

BIBLIOGRAFIJA

Odlični udžbenici filozofije uma su slijedeći:

- Braddon-Mitchell, D. and F. Jackson (1996), *Philosophy of Mind and Cognition*, Oxford, Blackwell
- Crane, T.(2001), *Elements of Mind*, Oxford, Oxford University Press
- Heil, J.(1998), *Philosophy of Mind*, London, Routledge
- Heil, J.(2004, 2nd ed.), *Philosophy of Mind*, London, Routledge
- Lowe, E.J.(2000), *An Introduction to the Philosophy of Mind*, Cambridge, Cambridge University Press

Vrlo korisne zbirke članaka su:

- Block, N., O. Flanagan, G. Guzeldere(eds.)(1997), *The Nature of Consciousness*, Cambridge, Mass., MIT Press
- Heil, J.(ed.)(2003), *Philosophy of Mind: A Guide and Anthology*, Oxford, Oxford University Press
- O'Connor, T., D.Robb(eds.)(2003), *Philosophy of Mind: Contemporary Readings*, London and New York, Routledge

Slijede navedene bibliografske jedinice:

- Armstrong, D.(1968), *A Materialist Theory of Mind*, London, Routledge and Kegan Paul
- Berčić, B.(1988), Supervenijencija, *Filozofska istraživanja*, 25 (2): 507-519.
- Block, N.(1978), Troubles with Functionalism, u Savage, C. W. (ed.): *Perception and Cognition*, University of Minnesota Press

- Broad, C. D.(1925), *Mind and Its Place in Nature*, London, Routledge and Kegan Paul
- Churchland, P.M.(1985), Reduction, Qualia, and the Direct Introspection of Brain States, *Journal of Philosophy*, 82 (1): 8-28.
- Crane, T. (1998), Intentionality as the Mark of the Mental, u O'Hear (ed.)(1998)
- Dennett, D.(1978), *Brainstorms*, Hassocks, Harvester Press
- Descartes, R.(1911), *The Philosophical Works of Descartes*, prijevod na engleski, Haldane, E.S. i Ross, G.R.T., Cambridge, Cambridge University Press
- Ershov, Yu. L., E.A. Palyutin(1984), *Mathematical Logic*, Moscow, Mir Publishers
- Fodor, J.(1975), *The Language of Thought*, Hassocks, Harvester Press
- Fodor, J.(2008), *LOT 2*, Oxford, Clarendon Press
- Foster, J.(1991), *The Immaterial Self*, London and New York, Routledge
- Hardin, C.L.(1987), Qualia and Materialism: Closing the Explanatory Gap”, *Philosophy and Phenomenological Research*, 48 (2): 281-298.
- Haugeland, J.(1982), Weak Supervenience, *American Philosophical Quarterly*, 19 (1): 93-104.
- Hill, C.(1991), *Sensations*, Cambridge, Cambridge University Press
- Horgan, T.(1981), Token Physicalism, Supervenience, and the Generality of Physics, *Synthese*, 49 (3): 395-413.
- Horgan, T.(1993), From Supervenience to Superdupervenience, Meeting the Demands of a Material World, *Mind*, 102 (408): 555-586.
- Jackson, F.(1982), Epiphenomenal Qualia, *Philosophical Quarterly*, 32 (2): 127-136.

- Johnson-Laird, P.(1993, 2nd ed.), *The Computer and the Mind*, London, Fontana Press
- Kim, J.(1993), *Supervenience and Mind*, Cambridge, Cambridge University Press
- Kittel, C., Knight, W., Ruderman, M.(1982), *Mehanika*, prijevod s engleskog: Nikola Zovko, Zagreb, Tehnička knjiga
- Lewis, D.(1966), An Argument for the Identity Theory, *Journal of Philosophy*, 63 (1): 17-25.
- Lewis, D.(1997), What Experience Teaches, u Block, N., O. Flanagan, G. Guzeldere(eds.)(1997)
- Locke, J.(1959, 1690), *An Essay Concerning Human Understanding*, edited by Fraser, A.C., London, Constable
- Malebranche, N. (1997, 1688), *Dialogues on Metaphysics and on Religion*, edited by Jolley, N. and translated by Scott, D., Cambridge, Cambridge University Press
- Maund, B.(1995), *Colours*, Cambridge, Cambridge University Press
- Melnyk, A.(1991), Physicalism: From Supervenience to Elimination, *Philosophy and Phenomenological Research*, 51 (3): 573-587.
- Miller, A.(1998), *Philosophy of Language*, London, University College London Press
- Nunn, C.(1996), *Awareness*, London, Routledge
- O'Hear, A. (ed.)(1998), *Current Issues in the Philosophy of Mind*, Cambridge, Cambridge University Press
- Papineau, D.(2002), *Thinking About Consciousness*, Oxford, Oxford University Press
- Penrose, R.(1989), *The Emperor's New Mind*, Oxford, Oxford

University Press

- Pećnjak, D.(1995), Turingovi strojevi, Gödelov teorem i Searleova soba, *Treći program Hrvatskog radija*, 48: 11-15.
- Pećnjak, D.(2002), Eliminacija eliminativizama, *Prolegomena*, 1 (1): 19-33.
- Place, U.T.(1956), Is Consciousness a Brain Process?, *British Journal of Psychology*, 47: 44-50.
- Putnam, H.(1975), *Mind, Language and Reality, Philosophical Papers, Volume 2*, Cambridge, Cambridge University Press
- Rey, G. (1997), A Question About Consciousness, u Block, N., O. Flanagan, G. Guzeldere(eds.)(1997)
- Robinson, H.(1993)(ed.), *Objections to Physicalism*, Oxford, Clarendon Press
- Rothman, M.(1966), *The Laws of Physics*, Harmondsworth, Penguin
- Searle, J.(1980), Minds, Brains and Programs, *Behavioral and Brain Sciences*, 3: 417-424
- Searle, J.(1984), *Minds, Brains and Science*, London, BBC
- Sesardić, N.(1984), *Fizikalizam*, SSOS
- Shoemaker, S. (1997), The Inverted Spectrum, u Block, N., O. Flanagan, G. Guzeldere(eds.)(1997)
- Smart, J. J. C.(1959), Sensations and Brain Processes, *Philosophical Review*, 68 (2): 141-156.
- Stich, S. (1983), *From Folk Psychology to Cognitive Science*, Cambridge, Mass., MIT Press
- Strnad, J.(1985), *Mala kvantna fizika*, prijevod sa slovenskog: Krešimir Adamić, Zagreb, Školska knjiga

- Trahtenbrot, B. A.(1978), *Što su algoritmi*, Zagreb, Školska knjiga
- Tye, M. (1997), A Representational Theory of Pains and Their Phenomenal Character, u Block, N., O. Flanagan, G. Guzeldere(eds.)(1997)
- White, S.(1997), The Curse of the Qualia, u Block, N., O. Flanagan, G. Guzeldere(eds.)(1997)
- Wichmann, E.(1988), *Kvantna fizika*, prijevod s engleskog: Nikola Zovko, Zagreb, Tehnička knjiga

ISBN 978-953-7823-32-0

