

Eliminacija eliminativizama

DAVOR PEĆNJAK

Filozofski fakultet u Zadru – Sveučilište u Splitu
Obala kralja Petra Krešimira IV, 3, HR-23000 Zadar

Izvorni znanstveni članak
Primljeno: 03-04-02 Prihvaćeno: 06-05-02

SAŽETAK: U ovom radu razmatraju se dvije vrste eliminativizama u filozofiji uma – eliminativistički materijalizam i funkcionalni eliminativizam. Pokazano je da, ustvari, zrela neuroznanost mora *objasniti* pojave označene pojmovima kao što su npr. »percepcija«, »um« ili »svijest« i da ti pojmovi nisu uvedeni već kao objašnjenja nečega. Svijest, na primjer, jest činjenica koju se treba objasniti i kao takva se ne može eliminirati. Za strojno funkcionalnu karakterizaciju funkcioniranja uma pokazano je kako je također neadekvatna da bi mogla iz znanosti i ontologije izbaciti svijest i svjesne mentalne procese.

KLJUČNE RIJEČI: Eliminativistički materijalizam, funkcionalni eliminativizam, um, svijest, redukcionizam, pučka psihologija, zrela neuroznanost, analogija um-stroj.

Najpoznatija pozicija u filozofiji uma koja želi eliminirati mentalne entitete i mentalne pojmove kojima se danas služimo jest tzv. eliminativni materijalizam, odnosno, kako ga je zgodno nazvao Foster (1991), mentalni nihilizam. Prema tome, kad bi se ispunio program eliminativnog materijalizma, niti u običnom niti u znanstvenom rječniku ne bi bilo mjesta za pojmove poput »vjerovanje«, »želja«, »osjećaj«, »percepcija« itd., a niti za pojam »svijest«; ovi pojmovi ne bi ni na što referirali, što znači da ne bi postojali, i da ustvari ne postoje, fenomeni za koje se sada čini da postoje. Ukratko, sav inventar uma, a također i sama svijest, prema ovoj teoriji ne postoji kao činjenica odnosno stanje/proces u svijetu, odnosno u našim glavama. To znači da, zapravo, ne postoje niti vjerovanja, želje, percepcije, nego samo neurofiziološki procesi ili stanja, koji se sastoje ili se reduciraju na fizičke procese i stanja. No, ta će redukcija biti takva da neće ostaviti mjesta za svijest, vjerovanja, želje itd., kao *posebne fizičke* entitete, nego će ona biti takve vrste da će se ovi pojmovi, kao što se sada upotrebljavaju, pokazati toliko nekoherentnima da će reduciranje na kraju zapravo eliminirati potrebu za ovim pojmovima te će se na taj način vidjeti da ovi entiteti ni ne postoje.

Osnovni argument temelji se prvenstveno na razlikovanju između tzv. pučke psihologije i znanstvene psihologije (vidi npr. Horgan i Woodward, 1990, Mišćević, 1990, Stich, 1983). Na strani znanstvene psihologije su i oni znanstveni pristupi koji se mogu nazvati kognitivne znanosti, među kojima je, pogotovo za eliminativiste, najzanimljivija i najistaknutija neuroznanost. U zreloj i cjelovitoj kognitivnoj znanosti, a pogotovo u neuroznanosti, više neće biti niti jednog pojma iz pučke psihologije, a to implicira da neće biti mjesta niti za ono na što bi se ti pojmovi referirali; dakle nema niti entiteta (vidi Rorty, 1965, 1970, Churchland, P. S. 1986, 1997, Churchland, P. M. 1979, 1981, 1984, 1985, 1989, Stich, 1983, Wilkes, 1984, 1988). No, što je to pučka psihologija i kako je uzimaju navedeni autori?

Pučka psihologija je zdravorazumski pogled na ponašanje ljudi i zdravorazumski pokušaj njegova sistematiziranja. Da bismo objasnili i predvidjeli ponašanje drugih ljudi moramo imati nekakvu koncepciju o tome. U takvoj koncepciji zdravog razuma običnih ljudi javljaju se pojmovi kao što su vjerovanja, želje, strahovi, djelovanje, svjesno, nesvjesno, osjeti itd. Ti pojmovi su također međusobno povezani u zakonoliku mrežu odnosa, jer za svako objašnjenje i predviđanje potrebne su neke korelacije odnosno zakonitosti. Mi, kao obični ljudi, možemo s velikom i katkada zapanjujućom točnošću predviđati reakcije, postupke, raspoloženja, djelovanja itd. drugih ljudi (Pinker, 1997). To je nesumnjivo točno i to nam je potrebno radi čisto praktičnih ciljeva i u nekom dubljem smislu (ili plićem, rekli bi neki) radi vlastitog preživljavanja (vidjeti Levin, 1984, radi zanimljivog pogleda na problem tuđih svijesti odnosno tuđih umova). Jezgru ovakve koncepcije čine propozicionalni stavovi – vjerovanja, želje, strahovi, namjere, itd. Također, u takvom predviđanju i razmišljanju zdravog razuma, pojavljuju se i pojmovi zamjećivanje, osjećanje, percipiranje, svjesno, nesvjesno, refleksno itd. No, da se pokazati da ova pučka upotreba i raščlamba nije uvijek dosljedna i da se ponekad u jednu kategoriju stavljaju entiteti, tj. pojmovi, koji spadaju u neku drugu (Wilkes, 1984). Osim toga, pučka psihologija je i ograničena – mnoge pojave ne može objasniti. Dublje i potpunije objašnjenje npr. vizualnih iluzija, razlike između sna i jave, porijekla sadržaja nekih propozicionalnih stavova kao što su neka specifična vjerovanja, pučka psihologija jednostavno ne može objasniti svojim sredstvima. Prema tome, smatraju materijalistički eliminativisti, ovo je degenerativni program koji treba potpuno zamijeniti nečim drugim, tj. kognitivnom znanostima i na kraju neurofiziologijom. Ona će upotrebljavati sasvim drugu pojamovnu aparaturu i tako će svi ovi entiteti koji vuku porijeklo iz degenerirane pučke psihologije kao degenerirane teorije uma, biti izbačeni iz, prvo, zrele znanosti, a onda i iz ontologije, tj. postojanja. Sudbina pojmova kao što su »svijest«, »propozicionalni stav«, »vjerovanje«, »osjećanje« i tako redom, bit će ista kao i sudbina pojmova »flogiston«, »elán vital« i sličnih, u fizici, kemiji ili biologiji. Ovi pojmovi su izbačeni iz tih znanosti i ne postoje entiteti, procesi ili pojave za koje se neko vrijeme mislilo da ovi pojmovi označavaju.

No, analogija nije dobra. Pomoću flogistona pokušalo se nešto objasniti, a to je bila pojava sagorijevanja. Flogistonska teorija to nije uspjela te je »flogiston« kao pojam izbačen iz znanosti, a kao entitet izbačen je iz ontologije. No, pojava koju se trebalo objasniti i dalje je ostala – sagorijevanje nije bilo izbačeno iz fizike ili kemije ili iz inventara pojava koju i »pučka fizika« poznaje da se događa. Flogiston je zamijenjen drugim objašnjenjem – kisikom – i tako je sagorijevanje zadovoljavajuće objašnjeno. Ni tada sagorijevanje nije prestalo postojati (vidi i npr. Flanagan, 1992).

Još nešto što onda zbunjuje eliminativističke materijaliste jest da navedenim pojmovima pridaju teoretski status, jer im se čini da su oni ovdje da bi nešto objasnili – kao i flogiston. No ovi naši mentalni pojmovi nisu teoretski pojmovi niti teoretski entiteti (o teoretskim pojmovima i entitetima vidi npr. Hacking, 1983, Sesardić, 1986, O'Hear, 1989, Berčić, 1995).

Svijest jest neopažljivi entitet, ali nije teoretski entitet niti je »svijest« teoretski pojam. Teoretski entiteti uvode se u znanost da bi se objasnile razne pojave, ali svijest nije pojam pomoću kojeg se nešto objašnjava; naprotiv, svijest jest nešto što treba objasniti. Zato eliminativisti misle da se svijest treba izbaciti iz rječnika zrele znanosti, tj. neurofiziologije, ili čak, ako se pokaže da se neurofiziologija može reducirati na fiziku, onda naravno iz fizike. No, jasno je da taj pojam neće, kao niti propozicionalni stavovi ili osjeti, biti upotrijebljen u završnoj teoriji, jer završna će teorija biti *explanans*, dok će pojmovi kao što su svijest, osjeti, iskustvo, percepcija, propozicionalni stavovi itd. uvijek ostati ono što jesu – *explanandumi*, dakle entiteti koje treba objasniti. Loše bi bilo da *explanans* sadrži iste one pojmove koje je htio ili trebao objasniti. Kada bi bilo tako, onda ustvari objašnjenja ne bi niti bilo jer bi ono bilo cirkularno.

Svijest je entitet posebne vrste: subjekt ima pristup samo svojoj svijesti. Mi ne opažamo niti introspektiramo svijesti drugih svjesnih bića. Naravno, ova posebnost jest i korijen filozofskog problema tuđih svijesti. Ali, zanemarujući taj problem, mi možemo reći da ne opažamo direktno niti svoju svijest, već kada opažamo svijest, mi smo angažirani u samo-svjesnoj aktivnosti koja nije istovjetna samoj svijesti. Predmet samo-svijesti jest svijest, ali predmet svijesti nije svijest.

Prava analogija bila bi slijedeća: zamislimo predznanstvenu zajednicu koja igra biljar. Oni su izvrsni igrači, ali nemaju pojma o mehaničkim fizikalnim zakonitostima niti o bilo čemu važnom iz fizike. I većina današnjih igrača biljara, a i igrača u drugim sportovima, u stvari je u istoj takvoj situaciji kao što je i ova naša predznanstvena zajednica. Biljarske kugle odbijaju se nakon udaraca međusobno i od zidova onako kako se već odbijaju itd. Iako su vrsni igrači, ovim predznanstvenjacima je to ipak jako čudnovato kako se sve to događa baš kako se događa. Ipak, naslućuju da tu moraju postojati neka pravila odnosno zakonitosti kako i zašto se na biljarskom stolu sve događa kako se događa. Oni bi mogli ovako eliminativistički razmišljati: iza kretanja kugli i njihovih međusobnih sudara, odbijanja i kretanja nakon sudara mora

postojati nešto što to objašnjava – to su neke zakonitosti (fizikalno-mehaničke) koje na njih prirodno djeluju. Jednog dana netko će uspjeti formulirati te zakonitosti; to će biti zrela znanost koja će u potpunosti objasniti kretanja kugli i njihove međusobne sudare, odbijanja i kretanja nakon sudara. U rječniku te znanosti neće, naravno, biti mjesta za pojmove biljarska kugla, biljarski stol, sudar, odbijanje itd. Takva znanost imat će svoje precizne pojmove kao što su masa, brzina, impuls, količina gibanja, kut, sinus, kosinus itd., te matematički formulirane odnose, pomoću kojih će se potpuno objasniti sve ono što se događa na biljarskom stolu. Tako znači, biljarski stolovi, biljarske kugle, te njihovi sudari, odbijanja itd., zapravo ne postoje i te entitete treba eliminirati.

Naravno, nitko živ ili barem rijetko tko bi tako razmišljao. Stvar je u tome da se pojave na biljarskom stolu objasne, a ne da se eliminiraju; pojave na biljarskom stolu su ono što imamo kao »grube« činjenice za koje tražimo objašnjenje, po mogućnosti jedinstveno, koje će objasniti sva kretanja na biljarskom stolu, a možda i šire – sve vrste kretanja ili barem neke klase kretanja i izvan biljarskih stolova; tako je i npr. svijest te njen fenomenalni karakter »gruba« činjenica, kao uostalom i percepcija i propozicionalni stavovi, koja traži svoje objašnjenje i razlikovanje u odnosu na ne-svjesno. Objašnjenjem kretanja na biljarskom stolu ne eliminiramo objekte od kojih se sastoji to gibanje pa ni objašnjenjem svijesti i mentalnosti ne eliminiramo svijest, propozicionalne stavove, emocije, osjete, percepciju.

Ono što zavodi eliminativističke materijaliste jest to što mi imamo introspektivni pristup samo vlastitoj svijesti. Na ovom mjestu neću istraživati logičke mogućnosti je li moguće imati introspektivni pristup i tuđoj svijesti – možda bi to opet bio samo *vlastiti* doživljaj *tuđeg* doživljaja, a ne direktno doživljavanje tuđeg doživljaja (Levin, 1984) – nego ćemo pretpostaviti da je to tako. Tu se na mala vrata ušuljao problem tuđih svijesti, odnosno tuđih umova, koji je težak i neriješen filozofski problem: kako mi doživljavamo samo vlastitu svijest i imamo introspektivni uvid samo u svoj vlastiti um, pa tako ne znamo da li i drugi subjekti imaju umove i da li su na isti način *svjesna* bića, dovoljno je i nije zabrinjavajuće ako imamo čisto fizikalno-neuroznanstvene opise ljudskog »mišljenja«, »osjećanja«, »percipiranja« i »djelovanja«, koji u sebi nema ništa različito od bilo kakvih drugih znanstvenih opisa događaja u prirodi i svemiru koji uopće ne uključuju bilo što posebno psihološki i čisto su »mehanički« ili statistički. Vjerujem da eliminativistički materijalisti imaju introspektivni pristup svojim vlastitim svijestima, tj. vjerujem da i oni imaju svijest, ali oni su spremni žrtvovati vlastite svijesti koja je za svakog od njih, kao uostalom i za svakoga tko ima svijest, samo jedna svijest, samo jedan takav entitet, a budući da većinu ostalih svijesti, ako uopće postoje, ionako ne introspektiramo, onda je dovoljan eliminativistički pristup svijesti, tj. ona nam nije potrebna. Jedna svijest se lako žrtvuje, pogotovo kada znanost tako lijepo napreduje.

No, ako postoji introspektivni pristup svijesti, onda je dovoljno da postoji barem jedno biće i barem jedan introspektivno-svjesni događaj kod tog bića

da bi barem ono zahtijevalo objašnjenje tog događaja, odnosno da je taj događaj drugačiji od ostalih događaja i zahtijeva drugačije objašnjenje (Foster, 1991). Barem jedno takvo biće postoji – to je pisac ovih redaka – te tako fizikalno-kemijsko-neurološko objašnjenje koje eliminira svijest, svjesne propozicionalne stavove i svjesno razmišljanje na osnovi njih, te ostalo specifično mentalno, nije zapravo zadovoljavajuće jer nije potpuno. Ovdje imamo situaciju kada smo suočeni s jednom izvjesnom pojavom koja zahtijeva objašnjenje, a ne eliminaciju. Svjesnost jest »gruba« činjenica, a ne objašnjenje nečeg.

Neurofiziologija i biofizika mozga danas su vrlo napredovale te eliminativistički materijalisti imaju određeni kredit kada toliko povjerenja polažu u te fizikalne znanosti; međutim, iz samih takvih objašnjenja koja danas imamo, još se ne vidi kako svjesnost izranja iz ne-svjesnog. Ovdje pod ne-svjesnim mislim na čisto neurofiziološki ili fizikalno-kemijski opis raznih događaja. Iz pojmova šiljati val, akcijski potencijal, razlika potencijala, norepinefrin, 5-hidroksitriptamin, izmjena molekula natrija s molekulama kalija na taj-i-taj način, u tom-i-tom vremenskom intervalu i u toj-i-toj količini (Guyton, 1963, Greenfield, 1995), itd., zaista se ne vidi kako slijedi zamišljanje ili doživljavanje zalaza sunca sa zadarske rive ili osjećaj jake pospanosti ili razmišljanje o crnim rupama u svemiru. Iz toga eliminativisti zaključuju da će to biti tako i u zreloj ili cjelovitoj neurofizikalnoj znanosti što, dakle, automatski eliminira svijest i sve ono što spada u mentalnost – vjerovanja, emocije, osjete, percepciju itd., iz ontologije. Međutim, čini mi se da možemo zaključivati i ovako: danas je neurofiziologija vrlo napredovala, ali još nedovoljno da bi cjelovito objasnila mentalnost (što je točno – i sami neurofiziolozi sasvim lako priznaju takvo stanje stvari). No, cjelovita neurofiziologija će na kraju imati objašnjenje, tj. taj dodatni argument, ne samo da mentalnost jest realizirana fizikalno, nego i *kako* svijest »izranja« odnosno *kako* se realizira iz ne-svjesnog, tj. čisto fizikalnog; zrela znanost će pružiti specifični argument kako i zašto upravo određeni neuro-fizikalno-kemijski procesi i stanja jesu mentalni i svjesni procesi i stanja, a zašto ostali to nisu. U biologiji, živo se objašnjava pomoću neživog, ali time živo nije izbačeno iz znanosti i ontologije. Prema tome, ako svijest i jest realizirana čisto fizikalno,¹ onda fizikalistički opis mora dati uvjerljivu vezu između fizičkog i psihološkog, tj. mora se precizno vidjeti kako se psihološko odnosno mentalno i svijest realizira pomoću ne-psihološkog, tj. pomoću bazičnijih fizikalnih svojstava. Sadašnje nepotpuno stanje znanosti nije dovoljan razlog za tvrdnju da zrela znanost također neće imati jasnu vezu između pojma svijesti i fizikalnih pojmova koji izražavaju njezinu realizaciju, ukoliko je ta realizacija čisto fizičke prirode.

¹ Ja, doduše, mislim da svijest i um nisu realizirani čisto fizički, jer sam sklon zastupati neku varijantu kartezijanskog dualizma u problemu um-tijelo, no u *ovom tekstu* pokušavam pokazati *samo* neadekvatnost eliminativističkih pristupa i ne razmatram problem um-tijelo u kontekstu da li su ili nisu, i ako nisu, zašto nisu, ostale vrste fizikalističkih i redukcionističkih pristupa dobre kao rješenje ovog problema.

Na ovakav ili sličan način zapravo razmišljaju materijalisti ili fizikalisti koji nisu eliminativisti nego su redukcionisti ove ili one vrste. No, raspravu o neeliminativističkim varijantama materijalizma odnosno fizikalizma ovdje ostavljamo sa strane.

Kada pogledamo što kaže Wilkes (1984, 1988) u vezi samog pojma »svijest«, vidimo ustvari da ona opisuje jasne slučajeve svijesti i jasne slučajeve onog što nije svjesno. No ona također opisuje te inzistira na slučajevima koji su u sredini između tih jasnih polova, koji zaista jesu problematični, kao da su oni glavni i krucijalni. No oni to nisu. Wilkes razmatra npr. subliminalnu percepciju ili hipnotička stanja. Oni jesu zanimljivi za znanost da bi se objasnili, ali o njima ne ovisi velika većina slučajeva koji ili jesu ili nisu slučajevi svijesti vrlo jasno. Prvo možemo pronaći za jasne slučajeve što ih čini svjesnima za razliku od nesvjesnih – jasno nesvjesnih. S vremenom ćemo se pomicati, s napredovanjem relevantnih znanosti, po spektru, pa ćemo i za nejasne slučajeve jednog dana znati jesu li to slučajevi u kojima ćemo subjektu pridati svjesnost ili nesvjesnost. Osim toga, slučajevi u sredini spektra javljaju se puno rjeđe nego slučajevi na krajevima spektra – koji su uobičajeni i jasni i koji nas najviše interesiraju da ih objasnimo kako to i što ih to čini svjesnima. Ako pogledamo još neke primjere koje daje Wilkes, vidimo da su to patološki slučajevi – npr. slučajevi kada su oštećeni vizualni dijelovi mozga ili neka mentalna bolest. Ne možemo u žarište staviti nekoliko nejasnih slučajeva i na osnovi njih tvrditi da je općenito pojam svijesti nejasan ili heterogen kada imamo mnoštvo uobičajenih slučajeva kod kojih točno znamo napraviti i zdravorazumsku, pučku, i znanstvenu klasifikaciju. Patološki i nejasni slučajevi ne mogu, metodološki, biti *kameni temelji* prema kojima razlikujemo, klasificiramo i objašnjavamo veliku većinu uobičajenih slučajeva. Patološki slučajevi, tj. istraživanje takvih slučajeva, samo nam pomaže da što jasnije odredimo za što više slučajeva kako riješiti problem i da bi za normalne slučajeve dali što precizniju klasifikaciju, a ne da bismo, neriješenim patološkim slučajevima, još više zamutili i ono što nam je prilično jasno bez njih. Tehnika lezija (Sekuler i Blake, 1994, str. 17–18) koristi se da bi se ono što još nije adekvatno objašnjeno, razjasnilo *usporedbom* s normalnim i uobičajenim, kada ipak već dobro poznajemo to normalno i uobičajeno funkcioniranje. Naravno da pomoću te tehnike možemo i za uobičajeno funkcioniranje izvući neke zaključke i rafiniranija objašnjenja, ali takvi slučajevi nisu primarni za opća objašnjenja kako uobičajeno funkcionira svijest.

Čak i ukoliko, a to se uvijek događa, ostane nekoliko slučajeva u sredini, mislim da je čak i zrela znanost spremna prihvatiti i nekoliko nerazjašnjenih slučajeva koje možemo klasificirati ili ovako ili onako, ukoliko većinu problema možemo jasno objasniti. Prema tome, nekoliko središnjih slučajeva između svijesti i ne-svijesti ne može biti temelj za to da svijest kao pojam i pojava trebaju biti eliminirani.

Također, Wilkes daje funkcionalističke karakterizacije uloga koje mentalna stanja koja navodi imaju – pa njezini zaključci više pogađaju funkcionalističko stajalište koje bi onda bilo neadekvatno za objašnjenje svijesti, nego što pogađaju to da je sam pojam ili da su sama ta stanja koja karakteriziramo kao svjesna stanja, nejasna. Prema funkcionalističkoj teoriji, navodi Wilkes, neka stanja se ne mogu razlikovati – npr. ona ne može razriješiti problem izvrnutog spektra (pa možda niti »izvrnutih misli« – vidi Bealer, 1984). To je argument protiv funkcionalizma! Ovim se ne eliminira sam problem *zato jer ga teorija ove vrste nije riješila!*

Drugu vrstu eliminativističkog pogleda koju bismo mogli nazvati objašnjivački ili, možda još bolje, funkcionalni eliminativizam, zastupa Georges Rey (1983, 1995, 1996, 1997), a u nekim aspektima mu je blizak i Stich (1983). Pogledajmo kako Rey gradi argument za eliminaciju svijesti, ne samo kao pojma iz znanstvenih teorija i objašnjenja, nego kao nečega što uopće ne postoji kao fenomen u svijetu. Plan Reyovog argumenta je slijedeći: nakon navođenja različitih mentalnih fenomena kao kandidata za ulogu svijesti i razumnih teorija koje ih pokušavaju objasniti, pita se da li ti mentalni fenomeni mogu biti i nesvjesni; iznenađujući odgovor, smatra Rey, glasi *da*; zajedno s ponuđenim teorijama, koje mogu biti dakle potpuno primjenjive i na sustave za koje ne bismo rekli da su svjesni, imamo dobar razlog smatrati da svijest ne postoji.

Mišljenje, razmišljanje, svakako jest mentalni fenomen, odnosno mentalni proces, koji je izvrstan kandidat da ga proglasimo svjesnim, tj. da se svijest (barem jednim svojim dijelom) manifestira u razmišljanju. Konstitutivni elementi mišljenja odnosno razmišljanja su pojedine misli. Misli imaju svoj sadržaj koji nešto predstavlja, a to nešto je u većini slučajeva nešto drugo nego sama ta misao. Naravno, objekt neke misli može biti druga misao ili čak misao može biti autoreferencijalna, ali u praktičnom razmišljanju, pogotovo ove posljednje misli, rijetko se javljaju. Ukratko, ova usmjerenost misli na svoj objekt zove se intencionalnost. Pomoću ovog svojstva, misao nešto, dakle, reprezentira ili uključuje odnose prema reprezentacijama kao što su to npr. rečenice. U svakom slučaju, reprezentiranje je strukturirano na neki način, ono nije proizvoljno; sustav reprezentacija je strukturirani »posrednik« između subjekta i svijeta. Ova hipoteza omogućava izgradnju objašnjivačke formalne teorije razmišljanja. Takva teorija zove se »Jezik misli« (Harman, 1973, Fodor, 1975, Stich, 1983). Zastupnici ovakve teorije smatraju da su reprezentacije i formalne operacije koje se izvode nad njima dovoljne da se objasni cjelokupna misaona aktivnost. Većina njih su i naturalistički orijentirani te smatraju da je ovakav sustav enkodiran u živčanom sustavu odnosno mozgu. No, ako imamo teoriju koja je formalna, tj. samo su formalni odnosi dovoljni da bi se objasnilo ono što se želi objasniti, tada to daje, u principu, mogućnost konstrukcije stroja koji zadovoljava sve zahtjeve takve teorije. Tada se može postaviti pitanje da li taj napravljeni stroj misli

ili ne i da li jest svjestan ili nije? Bilo koji odgovor na ovo pitanje ima posljedice i za objašnjavanje čovjekovog mišljenja i svijesti.

Slijedećih pet substrukture su bazične u izgradnji nekog sustava koji misli, a za koji se vidi da ne mora biti svjestan, a može objasniti i ljudsko razmišljanje, te je izgrađen na bazi neke varijante teorije jezika misli:

1. alfabet, pravila formacije i pravila transformacije za kvantificiranu modalnu logiku s indeksikalima;
2. aksiomi za sustav induktivne logike s abdukcijskim sustavom hipoteza s razumnom funkcijom koja izabire između njih za neki dani ulazni podatak;
3. aksiomi teorije odlučivanja s nekim skupom temeljnih preferencija;
4. razni prijenosnici (transduktori) ulaznih podataka u 2. gore;
5. sredstva koja omogućavaju cjelokupnom sustavu odnosno stroju da realizira izlaz.

Sada je relativno jednostavno opisati sustav koji »radi« prema gore navedenom. Ulazni podaci ili input dostavljen podsustavom 4. proizvest će tzv. opservacijske rečenice koje se provjeravaju s obzirom na deduktivne posljedice hipoteza koje daje podsustav 2; bit će izabrane one hipoteze čije se posljedice najbolje slažu s opservacijskim rečenicama i onda one služe kao ulaz tj. input za 3; sada se na bazi tih izabranih hipoteza, danih preferencija i aksioma i teorema teorije odlučivanja izabire i generira najbolji opis za djelovanje te se prosljeđuje u 5 koji to djelovanje izvršava (Rey, 1997, str. 464). Pitanje u koje nećemo na ovom mjestu detaljnije ulaziti, a to je pitanje o tome da li ovakav sustav, što god bio, može imati intencionalnost, Rey rješava pomoću Stampea (1977). Naime, Stampe predlaže jednu varijantu kauzalne teorije intencionalnosti što znači da se intencionalnost nekog (mentalnog) stanja/procesa izvodi iz kauzalnog kontakta između objekta i sustava (subjekta) čije je to (mentalno) stanje/proces. Tu se Stampe oslanja i na pojam prirodnog značenja (vidi npr. Grice, 1957), a značenje rečenica u prirodnim jezicima jest vrsta prirodnog značenja. Da bismo ukratko objasnili prirodno značenje, navedimo jedan od najčešćih primjera: broj godina u drvetu znači broj godina starosti tog drveta, jer godovi su proizvedeni kao posljedica tolike starosti drveta. Tako će i kod stroja, kada na osnovi svog rada izabere neku rečenicu koja opisuje neko stanje stvari, ta rečenica biti kauzalna posljedica samog tog stanja stvari. Razjasnimo ovu teoriju malo šire. Naime, vanjski objekt proizvodi pomoću nekog kauzalnog lanca određeno (mentalno) stanje/proces u subjektu i to (mentalno) stanje/proces reprezentira taj objekt upravo radi toga jer je taj objekt na kauzalni način sudjelovao u proizvodnji tog (mentalnog) stanja/procesa. Dva su načina koji zadovoljavaju kauzalnu teoriju. Naravno, subjekt može biti u neposrednom kontaktu s objektom, tj. objekt može neposredno preko subjektovih prijenosnika/transduktora izazvati subjektovo reprezentacijsko stanje. S druge

strane, s većinom objekata ne može se biti u neposrednom kauzalnom kontaktu preko prijenosnika/transduktora. Većina ljudi nikada ne dođe u kauzalni kontakt niti s jednim djelićem kontinenta Amerike, bilo Sjeverne, bilo Južne, a još ih je manje koji su neki ili oba ta kontinenta vidjeli u cijelosti – npr. samo astronauti; vrlo mali broj ljudi bio je u kauzalnom kontaktu s psom moga prijatelja Saše pa ipak mnogi koji i nikada neće biti u bilo kakvoj vrsti kauzalnog kontakta s tim psom, mogu nešto saznati o njemu i upravo o tom psu, a ne nekom drugom. Kako je to moguće? Intencionalnost se može prenositi i posrednim putem. Ja mogu nešto ispričati o tom psu nekome, a budući da sam ja bio u neposrednom vizualnom, dakle perceptivnom kauzalnom kontaktu s tim simpatičnim bićem, moji opisi se odnose upravo na tog psa, tj. jezični iskazi koje formiram ovise o kauzalnom kontaktu s tim psom i na taj način se intencionalnost, tj. usmjerenost jezičnih iskaza na tog određenog psa, dalje prenosi na mog slušatelja. On ih u svom umu obrađuje i pretvara u simbole jezika misli, a ti simboli su kauzalno uzrokovani onim što je slušatelj čuo od mene pa se i na njih prenosi intencionalnost. Naravno, moj slušatelj sada može postati govornik i nekoj trećoj osobi, koja također nikad nije vidjela psa mog prijatelja, ispričati ono što je čuo od mene i što je pohranio u memoriji o tom psu. Lanac ovakvog prijenosa može, naravno, biti proizvoljno dug. Važno je da je barem jednom, na početku tog lanca, netko ili nešto bilo u kauzalnom kontaktu s objektom. Sve ostalo je ovisno o tom prvom kontaktu te se intencionalnost kauzalno prenosi (vidi također npr. Devitt i Sterelny, 1987). Intencionalni sadržaj o nekom objektu najčešće se prenosi jezičnim putem, ali može se prenositi i npr. fotografijama, filmom, videom, tonskim zapisima, preko tipkovnice računala itd.

Što se tiče strojeva koji zadovoljavaju navedene uvjete rada od 1 do 5, oni postoje i uopće ih nije teško tehnološki napraviti. Jedan od takvih strojeva je upravo aktivan: to je računalo na kojem pišem ovaj tekst. No, jasno je da nitko za njega ne bi rekao da je to svjestan sustav niti da on stvarno misli. Ali ako i ljudsko mišljenje možemo opisati pomoću uvjeta 1–5, onda možemo zaključiti, tvrdi Rey, da nema ničeg svjesnog ni u toj vrsti mišljenja pa znači da ljudi također ne posjeduju svijest.

Na ovakav zaključak netko će brzo prigovoriti da ljudi mogu formirati i imati vjerovanja o svojim vlastitim vjerovanjima, želje o svojim željama itd., tj. da intencionalni objekt nekog vjerovanja može biti neko drugo vlastito vjerovanje, objekt misli neka druga misao itd., te da ovo nije obuhvaćeno uvjetima opisa 1–5. Ovdje bi se željelo ukazati na razliku između strojnog operiranja i ljudskog mišljenja i svijesti na taj način što ljudi mogu imati i intencionalnost višeg reda, a strojevi to ne mogu (vidi npr. Dennett, 1978a i Frankfurt, 1971). No to zapravo nije točno. U samim pravilima 1–5, naročito u 3, već je sadržano pregledavanje nekih vlastitih stanja i operiranje nad njima. Osim toga, može se i eksplicitno dodati podsustav koji će to ekstenzivno omogućavati stroju. Tako možemo pridodati novi uvjet:

6. rekurzivni sustav.

Ovakav podsustav omogućava stroju da motri na određenim točkama svoja vlastita stanja, da daje izvješće o tim svojim vlastitim stanjima te da ih upotrebljava u daljnjem svom radu i da na osnovi njih modificira neke svoje operacije ili neka daljnja svoja stanja i procese. Naravno, ne samo da je relativno lako napraviti takav program nego oni i postoje – samo da spomenemo najjednostavnije primjere – svako računalo danas već po *defaultu* daje podatke o nekim svojim stanjima – koliko je memorije ostalo slobodno, koji dio programa se izvršava trenutno, kako su diskovi ili kako je disk raspodijeljen itd. Nema nikakvog problema da se i program koji zadovoljava uvjet 6 vrlo lako, mehanički, »vrti« na stroju, tj. računalu. Ali, niti sada nitko ne bi rekao da je stroj postao svjestan! Dakle, intencionalnost višeg reda sama po sebi nije specifičnost prema kojoj se postiže svijest. Osim toga, nije jasno zašto bi samo ona stanja koja se odnose na druga stanja iste vrste bila svjesna. Takva teorija brzo upada u problem objašnjavanja da li su i intencionalna stanja prvog reda svjesna. Čini se da ja mogu imati vjerovanje koje se odnosi na neko stanje stvari u svijetu i čijeg sadržaja mogu biti svjestan, a da istovremeno ne moram biti svjestan da sam svjestan upravo tog vjerovanja. Naime, ja mogu vjerovati da je Jadransko more plavo, ali ja ne moram istovremeno vjerovati da vjerujem da je Jadransko more plavo, odnosno ja mogu biti svjestan vjerovanja da je Sunce na 45 stupnjeva, a ne moram istovremeno biti svjestan da sam svjestan tog vjerovanja.

Da ne bismo previše duljili, recimo da se današnja računala lako daju opremiti i programima koji im omogućavaju komuniciranje na nekom prirodnom jeziku prema sintaktičkim i semantičkim pravilima toga jezika, a taj posao naročito olakšava transformacijska generativna gramatika (vidi npr. Culicover, 1976). Prigovor da strojevi nemaju i ne mogu imati osjete Rey rješava tako da upućuje na uvjet broj 3: strojevima je omogućeno da imaju prijenosnike koji transformiraju npr. zvučne ili svjetlosne signale iz okoline te onda na osnovi njih i prema svojim programima izvršavaju ili ne izvršavaju neke operacije. Uvjeti 1–5 omogućavaju, već prema vrsti potprograma, da se isti okolinski svjetlosni ili zvučni ulazni podaci različito obrađuju i različito rabe. To isto radi i čovjek – oko, pa zatim očni živac, prenosi i transformira svjetlosne signale, oni se dalje obrađuju u mozgu te su spremni za različite upotrebe. Za Reya ponovo nema razlike između strojnog i čovječjeg rada. Obje vrste rada dadu se opisati na isti način. Malo tko, ili nitko, ipak neće tvrditi da su strojevi postali svjesni kada su opremljeni sredstvima i programima koji izvode sve ono što smo do sada naveli, a današnji strojevi jesu takvi. Nitko ne misli da je kompjuterizirana proizvodna linija Mitsubishija zapravo svjesno biće. No opis koji vrijedi za njezin rad, vrijedi isto tako i kao opis za ljudsko mentalno funkcioniranje. Budući da strojevi – računala – nemaju svijest, odnosno krajnje je nevjerojatno da bi oni imali svijest, onda je to dobar razlog, s obzirom da isti opis vrijedi i za ljude, odnosno za (ljudske)

umove, sumnjati, odnosno negirati da ljudska bića posjeduju svijest. Tako Rey eliminira ne samo pojam »svijest« iz kognitivnih znanosti, nego i svijest kao činjenicu za koju se krivo čini da postoji u našim glavama. Zaključak je – svijest zapravo ne postoji.

Da li je takav zaključak, unatoč uvjerljivosti argumentiranja, zaista opravdan? Mislim da nije i to zbog više razloga.

Prvo, uočavamo da su uvjeti 1–6 dani prilično apstraktno u logičkoj formi i da zbog toga njima možemo opisati mnoštvo vrlo različito realiziranih sustava, tj. mnoštvo različitih sustava može zadovoljiti dane uvjete. Dakle, opis je dan na jednom višem nivou. Kada opisujemo sustave na višem, tj. na apstraktnijem nivou, zanemaruju se mnoge posebnosti sustava. Katkada upravo iz tih posebnosti proizlaze najzanimljivije karakteristike fundamentalnog operiranja sustava i uopće upravo te konstrukcijske posebnosti omogućavaju sustavu da radi. Čak i kada se radi o programima – dakle o *softwareu* – koji su sami po sebi, da tako kažemo, poluapstraktni entiteti, imamo to da do istog rezultata obrade istih ulaznih podataka različiti programi dolaze na različite načine jer imaju različite naredbe, tj. različite cijele sustave i hijerarhije izvršavanja naredbi. Gledano s višeg aspekta programi obavljaju istu operaciju, ali gledano na nižem nivou oni je obavljaju različito. Dakle, do iste stvari može se doći na različite načine. Razlika svjesnog i nesvjesnog ne mora biti u moćima što koji sustav može ili ne može, nego u bazičnijem načinu kako to obavlja. Ali to nije onda objašnjeno niti opisano teorijom višeg aspekta, jer smo upravo zainteresirani za razliku kako dva ili više sustava na različite načine dolaze do istog rezultata.

Osim toga, ako dvije pojave ili dvije stvari objasnimo na isti način ili potpuno analogan način, to ne znači nužno da su te dvije stvari ili pojave iste ili iste vrste. Kako je to moguće? Čini se da bi to moralo biti tako da bi dvije stvari, za koje imamo isto objašnjenje, morale biti iste. Ipak nije tako. Poslušimo se jednostavnim primjerom iz fizike. Formula za silu gravitacijskog privlačenja glasi:

$$F = m_1 m_2 / r^2$$

Formula za elektrostatičko privlačenje glasi ovako:

$$F = q_1 q_2 / r^2$$

Vidimo da su ova dva matematička izraza potpuno jednaka. Međutim nikome ne pada na pamet izjednačiti te dvije sile i tvrditi da je to jedna te ista stvar, a još manje pada na pamet da netko eliminira jednu od njih i kaže kako ona ne postoji i da se tu radi o manifestaciji jedne te iste stvari odnosno sile. Izraz za gravitacijsku silu bio je, od strane Newtona, formuliran ranije (objavljeno 1689. u njegovim *Principia*) nego što je formuliran izraz za elektrostatičku silu od strane Coulomba. Netko je mogao reći, nakon što je formuliran Coulombov izraz, da nema potrebe imati novu vrstu sile, nego da je

to ustvari neka vrsta manifestacije gravitacijske sile budući da je apstraktni matematički opis isti. Naravno, postoje daljnje karakteristike koje razlikuju ove dvije sile i one se, u ovom slučaju vrlo brzo i jasno manifestiraju.

Prema Newtonovom zakonu gravitacijskog privlačenja, gravitacijska sila između dva objekta jednaka je umnošku masa tih dvaju objekata podijeljenom s kvadratom udaljenosti između tih objekata; Coulombova sila jednaka je umnošku električnih naboja podijeljenim s kvadratom udaljenosti između dvaju objekata nositelja tih naboja. Mase objekata nisu isto što i električni naboji. Već tu se vidi razlika u fizikalnim entitetima. Međutim, iste količine davat će iste rezultate izračunavanja i oblik zakona je isti:

$$F = x_1 x_2 / r^2$$

Pa ipak, dakle, te dvije stvari nisu iste i nitko ih ne poistovjećuje. Također i u mnogim daljnjim karakteristikama gravitacijska sila se razlikuje od električne i jasno je da su to dva potpuno različita fizikalna entiteta. Ukoliko te dvije sile i pokazuju iste karakteristike i ukoliko su isti izrazi koji ih opisuju, one ih pokazuju samo do neke granice i izrazi su isti samo do neke granice, a od tih granica one se počinju razlikovati, kao i fizikalni izrazi. Zbog tih djelomičnih ekvivalentnosti u opisu i ponašanju, ne možemo eliminirati jednu nauštrb druge ili, još gore, obje. No, onakvo jasno razlikovanje kakvo imamo između električne i gravitacijske sile, ne mora biti tako očigledno i jasno kod drugih stvari. Neke stvari mogu biti u mnogim dijelovima iste te se počinju razlikovati u nekoj točki koja nije tako odmah očigledna i jasna. Zato isti izrazi mogu opisivati i svjesno i nesvjesno – mehaničko – operiranje, a da negdje ipak postoji razlika. Samo, može biti da je to razlika koju još nismo pronašli. Sasvim plauzibilno možemo tvrditi da je svijest činjenica koju treba objasniti i da ne treba nikakvo daljnje opravdanje za postojanje te činjenice nego da je potrebno objašnjenje za takvu činjenicu – što je omogućava i što je čini takvom kakva jest. Ako neka teorija kao što su to uvjeti 1–6 o kojima smo raspravljali, ne daje razliku između onog što smatramo svjesnim i onoga za što nikada ne bismo tvrdili da je svjesno, tada se čini da ta teorija nije dobra i adekvatna pa nju trebamo eliminirati ili nadopuniti nekim karakteristikama koje će ukazati na razliku svjesno-nesvjesno.

Osim svega, kada eliminativist ili mentalni nihilist izriče i tvrdi svoju poziciju, tada se pretpostavlja da su upotrijebljene sposobnosti i kapaciteti koji su doveli do tih zaključaka, a to su *mentalne* sposobnosti i kapaciteti. Također, samo tvrđenje svoje pozicije je izražavanje svog pogleda, dakle, *vjerovanja* da je taj pogled argumentiran i istinit. Eliminativist *vjeruje* u svoju teoriju. Izlaganje eliminativizma široj publici – znanstvenoj ili neznanstvenoj, svejedno, zahtijeva upotrebu jezika. No sama upotreba jezika pretpostavlja da oni koji upotrebljavaju jezik imaju *razumijevanje* jezika. Razumijevanje jezika jest nešto što dalje pretpostavlja i *svjesnost*, jer smo često svjesno angažirani u razumijevanju tuđih diskursa i proizvodnji vlastitog. Tako se

čini da je sâmo formuliranje i izricanje eliminativističke teorije bilo koje vrste zapravo nekonzistentno. Čak ukoliko smatramo da ovo nije fer kritika, jer, ukoliko bismo doista imali kompletnu znanost u eliminativističkom smislu, koja bi onda pokazala naravno, da eliminativist ne vjeruje, ali niti vjeruje, u svoju teoriju, jer bi ovi mentalni pojmovi bili zamijenjeni drugima, neurofiziološkima i fizikalnima, još uvijek se čini da znanost ne ide u tom smjeru i da ovaj argument ima neku uvjerljivost. Ako se nekome čini da ovaj posljednji argument ipak nema uvjerljivost, mislim da su ostali argumenti ionako dovoljni za odbacivanje navedenih eliminativizama.²

BIBLIOGRAFIJA

- Bealer, G. 1984. *Mind and Anti-Mind: Why Thinking Has No Functional Definition*. U: French, P., Wettstein, H. i Uehling, T. (eds.) 1984.
- Berčić, B. 1995. *Znanost i istina* (Rijeka: Hrvatski kulturni dom).
- Block, N., Flanagan, O., Güzeldere, G. (eds.) 1997. *The Nature of Consciousness* (Cambridge, Mass.: A Bradford Book, MIT Press).
- Churchland, P. M. 1979. *Scientific Realism and the Plasticity of Mind* (Cambridge: Cambridge University Press).
- Churchland, P. M. 1981. Eliminative Materialism and the Propositional Attitudes. *Journal of Philosophy*, vol. 78. No. 2.
- Churchland, P. M. 1984. *Matter and Consciousness* (Cambridge, Mass.: A Bradford Book, MIT Press).
- Churchland, P. M. 1985. Reduction, Qualia, and the Direct Introspection of Brain States. *Journal of Philosophy*, vol. 83. No. 1.
- Churchland, P. M. 1989. *A Neurocomputational Perspective* (Cambridge, Mass.: A Bradford Book, MIT Press).
- Churchland, P. S. 1986. *Neurophilosophy* (Cambridge, Mass.: A Bradford Book, MIT Press).
- Churchland, P. S. 1997. Can Neurobiology Teach Us Anything about Consciousness. U: Block, N., Flanagan, O., Güzeldere, G., (eds.) 1997.
- Culicover, P. W. 1976. *Syntax* (New York: Academic Press).
- Davidson, R., Schwartz, G., Shapiro, D. (eds.) 1983. *Consciousness and Self-Regulation*, vol. 3. (New York: Plenum).
- Dennett, D. 1978a. *Intentional Systems*. U: Dennett 1978b.
- Dennett, D. 1978b. *Brainstorms* (Hassocks: Harvester Press).
- Devitt, M. i Sterelny, K. 1987. *Language and Reality* (Oxford: Blackwell).

² Ovaj rad se temelji na jednom dijelu moje doktorske disertacije. želio bih se zahvaliti Goranu Švobu, Juri Zovku, Tomislavu Janoviću, Zvonku Čuljku, Mari Grbiću, Ivanu Macanu, Lini Veljaku, Nenadu Mišćeviću, Marijanu Palmoviću, Slavku Brkiću, Dragani Sekulić i, *last but not least*, Anici Škrinjarić, Renati Pećnjak i Branku Pećnjaku na pruženim raznim oblicima pomoći.

- Flanagan, O. 1992. *Consciousness Reconsidered* (Cambridge, Mass.: A Bradford Book, MIT Press).
- Fodor, J. 1975. *The Language of Thought* (New York: Crowell).
- Foster, J. 1991. *The Immaterial Self* (London and New York: Routledge).
- Frankfurt, H. G. 1971. Freedom of the Will and the Concept of a Person. *Journal of Philosophy*, vol. 68. No. 1
- French, P., Wettstein, H. i Uehling, T. (eds.) 1984. *Midwest Studies in Philosophy IX* (Morris: University of Minnesota Press).
- Greenfield, S. 1995. *Journey to the Centres of the Mind* (New York: W. H. Freeman and Co).
- Grice, H. P. 1957. Meaning. *Philosophical Review*, vol. 66.
- Guyton, A. 1963. *Medicinska fiziologija*, preveli: Nikša Allegretti et al. (Zagreb: Medicinska knjiga).
- Hacking, I. 1983. *Representing and Intervening* (Cambridge: Cambridge University Press).
- Harman, G. 1973. *Thought* (Princeton: Princeton University Press).
- Horgan, T. and Woodward, J. 1990. Folk Psychology is Here to Stay. U: Lycan, W. (ed.) 1990.
- Levin M. 1984. Why We Believe in Other Minds?. *Philosophy and Phenomenological Research*, vol. 46. No. 3.
- Lycan, W. (ed.) 1990. *Mind and Cognition* (Oxford: Blackwell).
- Metzinger, T. (ed.) 1995. *Conscious Experience* (Paderborn: Ferdinand Schöningh Verlag).
- Miščević, N. 1990. *Uvod u filozofiju psihologije* (Zagreb: Grafički zavod Hrvatske).
- O'Hear, A. 1989. *An Introduction to the Philosophy of Science* (Oxford: Clarendon Press).
- Pinker, S. 1997. *How the Mind Works* (Harmondsworth: Penguin).
- Rey, G. 1983. A Reason for Doubting the Existence of Consciousness. U: Davidson, R., Schwartz, G., Shapiro, D. (eds.) 1983.
- Rey, G. 1995. Towards a Projectivist Account of Conscious Experience. U: Metzinger, T. (ed.) 1995.
- Rey, G. 1996. *Contemporary Philosophy of Mind* (Oxford: Blackwell).
- Rey, G. 1997. A Question About Consciousness. U: Block, N., Flanagan, O., Güzel-dere, G. (eds.) 1997.
- Rorty, R. 1965. Mind-Body Identity, Privacy, and Categories. *Review of Metaphysics*, vol. 19.
- Rorty, R. 1970. In Defense of Eliminative Materialism. *Review of Metaphysics*, vol. 24.
- Sekuler, R., Blake, R. 1994, 3rd ed. *Perception* (New York et al.: McGraw-Hill).
- Sesardić, N. 1986. U obranu znanstvenog realizma. *Filozofske studije*, vol. 18.

Stampe, D. 1977. Towards a Causal Theory of Linguistic Representation. U: Wettstein, H. i Uehling, T. (eds.) 1977.

Stich, S. 1983. *From Folk Psychology to Cognitive Science* (Cambridge, Mass.: A Bradford Book, MIT Press).

Wettstein, H. i Uehling, T. (eds.) 1977. *Midwest Studies in Philosophy II* (Morris: University of Minnesota Press).

Wilkes, K. V. 1984. Is Consciousness Important?. *British Journal for the Philosophy of Science*, vol. 35.

Wilkes, K. V. 1988. *Real People* (Oxford: Clarendon Press).

Elimination of eliminativisms

ABSTRACT: In this article, the author examines two kinds of eliminativisms in the philosophy of mind – eliminative materialism and functional eliminativism. He shows that mature neuroscience has to *explain* phenomena which are denoted by the concepts »perception«, »mind« or »consciousness« and that these concepts are not introduced as explanations of something. Consciousness, for example, is a factual phenomenon that should be explained and cannot be eliminated, by eliminative materialism or by functional eliminativism, as an *explanandum* and as a fact.

KEY WORDS: Eliminative materialism, functional eliminativism, mind, consciousness, reductionism, folk psychology, mature neuroscience, mind-machine analogy.