ETIKA I PRAKTIČNO RASUĐIVANJE: MORALNI RAZLOZI I PROBLEM MORALNE MOTIVACIJE

Sušnik, Matej

Doctoral thesis / Disertacija

2011

Degree Grantor / Ustanova koja je dodijelila akademski / stručni stupanj: **University of Zagreb, Department of Croatian Studies / Sveučilište u Zagrebu, Hrvatski studiji**

Permanent link / Trajna poveznica: https://urn.nsk.hr/urn:nbn:hr:111:354220

Rights / Prava: In copyright / Zaštićeno autorskim pravom.

Download date / Datum preuzimanja: **2024-04-25**

SVEUČILIŠTE U ZAGREBU
HRVATSKI STUDIJI
UNIVERSITY OF ZAGREB
CENTER FOR CROATIAN STUDIES

Repository / Repozitorij:

Repository of University of Zagreb, Centre for Croatian Studies

DIGITALNI AKADEMSKI ARHIVI I REPOZITORIJI

SVEUČILIŠTE U ZAGREBU
CENTAR ZA HRVATSKE STUDIJE

Matej Sušnik

ETIKA I PRAKTIČNO RASUĐIVANJE: MORALNI RAZLOZI I PROBLEM MORALNE MOTIVACIJE

DOKTORSKI RAD

Zagreb, 2010.

UNIVERSITY OF ZAGREB
CENTRE FOR CROATIAN STUDIES

Matej Sušnik

ETHICS AND PRACTICAL REASON: MORAL REASONS AND THE PROBLEM OF MORAL MOTIVATION

DOCTORAL THESIS

Zagreb, 2010

SVEUČILIŠTE U ZAGREBU
CENTAR ZA HRVATSKE STUDIJE

MATEJ SUŠNIK

ETIKA I PRAKTIČNO RASUĐIVANJE: MORALNI RAZLOZI I PROBLEM MORALNE MOTIVACIJE

DOKTORSKI RAD

Mentor: prof. dr. sc. Josip Talanga

Zagreb, 2010.

Sadržaj

1. Uvod ... 6
1.1. Kantovsko shvaćanje objektivnosti ... 8
1.2. Hjumovska vizija praktičnog rasuđivanja i moralne motivacije ... 13
1.3. Zašto biti moralan? ... 16
1.4. Želje kao motivi i razlozi ... 22

I. MOTIVI ... 25

2. Hjumovska teorija motivacije ... 26
2.1. Smjer odgovaranja ... 27
2.1.1. *Kontrafaktička ovisnost i status 'neobičnih' želja* ... 31
2.1.2. *Mentalno stanje s oba smjera odgovaranja?* ... 35
2.2. Želje i njihova priroda ... 38
2.2.1. *Divlje i tihe strasti* ... 38
2.2.2. *O još nekim aspektima želja* ... 42
2.3. Poteškoće hjumovskog modela ... 45
2.3.1. *Fenomenološki prigovor* ... 45
2.3.2. *Prudencijalnost, motivirane i nemotivirane želje* ... 51
2.3.3. Akrasia *i problem slobode volje* ... 59
2.3.4. *Eksplanatorna učinkovitost* ... 65

3. Internalizam i eksternalizam motivacije ... 68
3.1. Kognitivizam, nekognitivizam i motivacija ... 70
3.2. Varijabilna motivacija ... 73
3.3. Protuprimjeri ... 76
3.3.1. *Vojnik* ... 76
3.3.2. *Depresivac* ... 82
3.3.3. *Amoralist* ... 86
3.3.4. *Sotona ili o standardnoj upotrebi moralnih sudova* ... 90
3.4. Racionalistički (kantovski) internalizam motivacije ... 93
3.4.1. *Porijeklo moralnih zahtjeva* ... 94
3.4.2. *Poteškoće s formulacijom* ... 98
3.4.3. *Prihvaća li amoralist moralne tvrdnje?* ... 100
3.4.4. *Pružaju li moralne dužnosti razloge za djelovanje?* ... 103

II. RAZLOZI ... 106

4.Hjumovski internalizam razloga ... 107
4.1. Hume i instrumentalizam ... 107
4.2. Williamsov internalizam ... 113
4.2.1. *Argument iz krivnje* ... 122
4.2.2. *Argument iz apsurda* ... 125
4.2.3. *Argument iz 'tvrdih' etičkih pojmova* ... 131
4.2.4. *Argument iz konverzije* ... 135
4.3. Razlozi i vrijednosti ... 143
4.4. O instrumentalnoj normi ... 154

5.Kantovski internalizam razloga 159
5.1. Motivacijski i sadržajni skepticizam 162
5.2. Kantov argument 167
5.2.1. *Imperativi* 170
5.2.2. *Formule kategoričkog imperativa* 173
5.2.3. *Kategorički imperativ kao norma racionalnosti* 177
5.2.4. *Jedan problem za Kanta* 183
5.3. Smithov internalizam 184
5.3.1. *Koherencija želja* 191
5.3.2. *Protiv konvergencije* 192

6.Zaključak 201

Bibliografija 207

1. Uvod

Opisujući glavni cilj svoje knjige *Desire*, G. F. Schueler kaže: "… argumenti ove knjige usmjereni su protiv dva često zagovarana gledišta … prvo, grubo govoreći, jest da su želje na neki način esencijalne za racionalno opravdanje postupaka … drugo ih [gledište] vidi kao esencijalne za objašnjenje postupaka" (Schueler, 1995: 2–4). Problematika kojom se Schueler bavi u osnovnim se segmentima preklapa s temom ovog rada, ali s jednom ključnom razlikom: ovdje ne pokušavam argumentirati protiv, već *u prilog* navedenih gledišta. Konkretnije rečeno, u ovom radu nastojim obraniti poziciju koja svoje korijene vuče iz Humeove tradicije, a čije je centralno obilježje da želje predstavljaju neizostavan element kada je riječ kako o motivima tako i o razlozima za djelovanje. Naime, obično se smatra da je upravo ukazivanje na motive koji se nalaze u pozadini nekog postupka jedini način da se za taj postupak pruži adekvatno objašnjenje, dok je identificiranje razloga, s druge strane, konstitutivno za pružanje opravdanja tog istog postupka. Teza da su želje esencijalne za objašnjenje i opravdanje ljudskih postupaka zapravo nije ništa drugo nego teza da želje predstavljaju osnovu za razumijevanje prirode motiva i razloga.

Na distinkciju između motiva i razloga u literaturi se ponekad referira kao na distinkciju između "motivirajućih" i "normativnih" razloga.[1] Radi se o tome da razlozi koji motiviraju subjekta da učini neku radnju ne moraju biti dobri razlozi. Stoga i kažemo da su motivirajući razlozi oni koji objašnjavaju odnosno oni koji odgovaraju na pitanje *zašto* je netko nešto učinio, dok su normativni razlozi oni koji opravdavaju odnosno oni koji odgovaraju na pitanje zašto bi netko *trebao* nešto učiniti. Jedino u svjetlu te distinkcije ima smisla kazati da subjekt "može djelovati iz razloga koji nije razlog ili da nije postojao razlog

[1] O tome vidi Cullity i Gaut (1997), Dancy (2000), Darwall (1997), Lazović (1998), Schueler (1995) i Smith (1994, 2004a). U ovom ću radu, osim ako ne naznačim drukčije, pod razlogom uvijek podrazumijevati normativni razlog, dok ću o motivirajućim razlozima govoriti kao o motivima.

da … učini ono što je učinio premda je to učinio zbog određenog razloga" (Dancy 2000: 3). Kada bi subjekt djelovao na osnovi normativnih razloga, tada bi normativni razlozi za njega ujedno bili i oni motivirajući; oni bi tada ne samo opravdavali, već i objašnjavali njegov postupak. Međutim, to često nije slučaj. Ideja je da normativni i motivirajući razlozi sasvim lako mogu biti u raskoraku: kao što je moguće da osoba ima motivirajući, ali ne i normativni razlog za neko djelovanje, jednako tako je moguće da ona ima normativni razlog nešto učiniti, ali da je taj razlog ne uspijeva motivirati odnosno da on za nju ne predstavlja motivirajući razlog.

Imajući to na umu, nije teško razumjeti zašto mnogi ne nalaze atraktivnom tezu prema kojoj su želje esencijalne za razloge. S obzirom da vrše opravdavalačku funkciju, kod većine autora prevladava uvjerenje da razlozi ne mogu biti (i da nisu) ovisni o subjektovim psihološkim stanjima. Drugim riječima, čak i ako želje jesu esencijalne za motive (što također nije samo po sebi evidentno), to ni u kojem slučaju ne vrijedi za razloge – razlozi opravdavaju, a opravdanje jest nešto što izlazi izvan okvira subjektovih kontingentnih stanja i s tim stanjima ne može biti u izravnom kontaktu. Upravo iz potonjeg uvjerenja izrasta averzija prema hjumovskom modelu koji, kako je moguće naslutiti, između razloga i motiva ne čini tako oštru razliku, jer i jedno i drugo povezuje sa subjektovim željama.

Shvaćanje prema kojemu želje figuriraju kao esencijalni elementi ljudskog djelovanja ima značajne implikacije za moralnu filozofiju. Pretpostavka hjumovskog shvaćanja jest da moralno djelovanje ne predstavlja nekakvu posebnu sferu ljudskog djelovanja, sferu u kojoj vrijede sasvim drukčiji eksplanatorni i opravdavalački modeli. Upravo suprotno, zaključci o tome što je esencijalno za ljudske postupke vrijede za postupke općenito, što znači da moralni oblici postupanja pod takve zaključke jednako tako potpadaju. Oslanjajući se na ovu pretpostavku, hjumovci se suprotstavljaju onim autorima – prvenstveno kantovske orijentacije – čija je vizija moralnog ponašanja sasvim različita. Zasnovati moralno ponašanje na željama,

prema mišljenju mnogih kantovaca, može biti smisleno samo ako to za svoju posljedicu nema dokidanje objektivnosti i nužnosti morala ili zagovaranje neke vrste relativizma.

Već je na osnovi ovog kratkog i površnog uvoda moguće odrediti okvir unutar kojeg će se kretati moja rasprava. Budući da upravo u kantovski usmjerenim autorima vidim svoje glavne oponente, s jedne ću strane pokušati ocijeniti u kojoj mjeri njihovi prigovori predstavljaju stvarnu opasnost za hjumovski pothvat, dok ću, s druge strane, pokušati preciznije utvrditi kakvo shvaćanje *moralne* motivacije i *moralnih* razloga proizlazi iz prihvaćanja hjumovskog modela. No smjestimo za početak cijelu ovu diskusiju u nešto širi kontekst i pokušajmo jasnije utvrditi na koji se način naše razumijevanje prirode motiva i razloga reflektira na naše razumijevanje moralnosti.

1.1. Kantovsko shvaćanje objektivnosti

Pokazati da je etika zasnovana na objektivnim temeljima za mnoge filozofe morala predstavlja središnji zadatak i izazov. No što točno znači pokazati da je etika zasnovana na *objektivnim* temeljima? Vjerojatno jedan od najpoznatijih odgovora na to pitanje – koji je u povijesti etike poprimao različite oblike i nijanse – jest onaj realistički. Utvrditi objektivnost etike za moralne realiste znači utvrditi da moralne vrijednosti zaista postoje, da njihovo postojanje nije zavisno od naših vjerovanja i stavova te da moralne prosudbe, zahvaljujući njima, mogu biti istinite ili lažne. Zato je pitanje objektivnosti etike, prema viđenju moralnih realista, ekvivalentno pitanju postojanja objektivne moralne stvarnosti. Realistički pothvat, međutim, nailazi na ozbiljne prepreke i poteškoće koje se javljaju u obliku "argumenta iz čudnosti". U svom poznatom djelu *Ethics: Inventing Right and Wrong*, John Mackie tvrdi da je realističko stajalište osuđeno na postuliranje entiteta u ontologiju koji su "različiti od svega drugog u svemiru" te na zagovaranje "neke posebne sposobnosti moralne percepcije ili intuicije, sasvim različite od naših uobičajenih načina spoznavanja svega drugog" (Mackie,

1977: 38). Mackie argumentira da bi se misterioznost takvih entiteta manifestirala i u tome što bismo njihovim eventualnim spoznavanjem na neki krajnje zagonetan način bili pokrenuti na djelovanje. Ti bi entiteti u sebi imali usađenu neku specifičnu moć privlačenja koja bi se aktivirala u trenutku u kojem bi oni bili detektirani od strane subjekta. Mackieva je ideja vrlo jednostavna. Moralni sudovi imaju praktičnu prirodu: naše uvjerenje da je nešto ispravno, pogrešno ili da imamo određenu moralnu dužnost igra značajnu ulogu kada razmišljamo o tome kako postupiti; ako prosudimo da je nešto ispravno, onda ćemo u tom pravcu biti i motivirani djelovati, a ako prosudimo da je nešto pogrešno, onda ćemo biti motivirani izbjeći takvu vrstu postupanja. Međutim, praktičnost etike ne bi mogla biti objašnjena kada bi postojale objektivne moralne vrijednosti – time bi se izvor motivacijske snage prebacio iz unutarnje sfere u onu vanjsku. Realist bi *ili* morao priznati da se radi o entitetima koji su kauzalno potpuno inertni odnosno nesposobni da uzrokuju bilo kakav postupak *ili* bi pak morao priznati da oni imaju neku nevjerojatnu sposobnost privlačenja subjekta odnosno da je u njih nekako ugrađen "zahtjev za takvim i takvim postupkom" (Mackie, 1977: 40). Budući da nijedna od navedenih opcija ne zvuči zadovoljavajuće, mehanizam moralne motivacije u okviru realističkog pristupa mora ostati potpuno nerazjašnjen. Naravno, sve ovo vrijedi samo pod uvjetom da bismo entitete o kojima je riječ uopće mogli spoznati, a, kako smo vidjeli, čak i u tom segmentu postoje stanovite poteškoće jer ni izdaleka nije jasno u čemu bi se taj spoznajni proces mogao sastojati i na koji bi se način mogao odvijati. Kako onda sačuvati objektivnost etike, a istovremeno savladati metafizičke i epistemološke komplikacije na koje upozorava Mackie?

Iako je tom pitanju moguće pristupiti iz različitih uglova ovdje ću skicirati jedan odgovor koji će nas uvesti u samo središte ovog rada. Riječ je o odgovoru koji bi se otprilike mogao formulirati na sljedeći način: tvrdnja da postoje objektivne moralne vrijednosti nije tvrdnja da postoje opskurni metafizički entiteti za koje je nemoguće pronaći prostor u

znanstvenoj slici svijeta, već jednostavno tvrdnja da postoje objektivni *razlozi* za djelovanje. Ili kako to komentirajući Mackiev prigovor kaže Thomas Nagel: "Objektivna lošost boli, na primjer, nije neko misteriozno daljnje svojstvo koje sve boli imaju, već samo činjenica da za bilo koga sposobnog da sagleda svijet objektivno, postoji razlog da je poželi zaustaviti" (Nagel, 1986: 144). Prema ovoj liniji razmišljanja, kada smatramo da je neki postupak moralno ispravan, mi zapravo smatramo da postoji objektivan razlog da taj postupak učinimo, a kada smatramo da je neki postupak moralno pogrešan, tada smatramo da postoji objektivan razlog za njegovo izbjegavanje. Moralne činjenice, skraćeno rečeno, pružaju nam razloge za djelovanje. No netko bi na ovom mjestu odmah mogao prigovoriti da Mackiev argument i dalje ostaje na snazi i da se uvođenjem pojma razloga u cijelu diskusiju zapravo ništa ne mijenja. Uostalom, kako bi se drukčije mogla protumačiti tvrdnja da *postoje* objektivni razlozi za djelovanje? Kakva je priroda tih razloga i kako ih spoznajemo? Ako su ova pitanja na mjestu, izgleda da je objektivistička pozicija i dalje osuđena na zagovaranje postojanja intrinzično normativnih entiteta i da jednostavno ne postoji način na koji bi ona mogla izbjeći zadiranje u ontološku i epistemološku problematiku. Pogledajmo, međutim, zašto ovaj zaključak neki nalaze preuranjenim.

Esencijalno obilježje opisanog pristupa jest sasvim drukčije razumijevanje etičkog istraživanja: etika se više ne promatra kao teoretska disciplina u čijem se središtu nalaze metafizička pitanja (status objektivne moralne stvarnosti) i epistemološki problemi (je li moguća spoznaja te stvarnosti), već kao praktična disciplina kojoj je glavna svrha istražiti kako se moralna razmatranja reflektiraju (i trebaju reflektirati) u našem djelovanju ili, nešto preciznije, u našem razmišljanju o tome kako djelovati (usp. Korsgaard, 1996b: 44). Budući da se u samom središtu nalazi pitanje "Što imam razloga činiti?" odgovor treba potražiti u teoriji praktičnog rasuđivanja. Potrebno je stoga odustati od vizije prema kojoj zagovaranje istine u etici neminovno implicira zagovaranje objektivno postojećih entiteta i svoju pažnju

usmjeriti na mehanizam na koji se oslanjamo u donošenju moralnih (praktičnih) odluka; intencija je argumentirati da postoji ispravan način rasuđivanja o moralnim problemima koji nas na koncu može dovesti do istinitih odgovora na gore postavljeno pitanje.

Kantova filozofija morala predstavlja najpoznatiji pokušaj zasnivanja etike u tom pravcu. Grubo govoreći, Kant ne smatra da teza o postojanju objektivnih moralnih vrijednosti za sobom povlači suvišan metafizički teret, jer je zapravo riječ o tezi da postoje kategoričke norme rasuđivanja – norme čija se objektivnost sastoji u tome što vrijede podjednako za sve odnosno u tome što je njima podvrgnuto svako *racionalno* biće. A upravo te norme generiraju objektivne razloge za djelovanje. Objektivne moralne razloge, ako je to moguće tako formulirati, ne nalazimo u izvanjskom poretku stvari, već ih sami stvaramo; oni su jednostavno produkt naše racionalne prirode. Zato Kant i kaže da se moralni zakoni, "zato što imaju vrijediti za svako umno biće uopće, imaju izvesti iz općeg pojma umnog bića uopće" (*U* 78).[2] Etika se pokušava utemeljiti u praktičnom razumu, a ne u vanjskom stanju stvari; problem utemeljenja etike nema više ontološki karakter, već se težište interesa prebacuje na problem praktičnog rasuđivanja. Jednako tako, s obzirom da do zaključaka o tome što imamo razloga činiti dolazimo praktičnim rasuđivanjem, ne postoji opasnost od pozivanja na nekakvu izvanrednu epistemološku sposobnost.

Sada je lakše razumjeti i Korsgaardin komentar Mackieva "argumenta iz čudnosti": "Ako je etički dobar postupak jednostavno racionalan postupak, mi ne moramo postulirati posebna etička svojstva u svijet ili [posebne] sposobnosti u um da bismo etici priskrbili utemeljenje (Korsgaard, 1996a: 311). Zašto Korsgaard govori o *racionalnom* postupku? Zato što racionalan postupak nije ništa drugo nego postupak za kojeg postoje određeni razlozi. U tom smislu, ako činjenica da je neki postupak "etički dobar" pruža razlog da se taj postupak učini, onda slijedi da je taj postupak, pod jednakim uvjetima, racionalno učiniti. Racionalna bi

[2] U navođenju Kantova djela *Utemeljenje metafizike ćudoređa* koristim se prijevodom Josipa Talange (2001 [1785]) te se pritom služim skraćenicom *U*.

osoba trebala biti pokrenuta moralnim razlozima, a upravo se u tome i krije rješenje problema na koji je ukazao Mackie. Podsjetimo se da Mackie vjeruje kako nije moguće zastupati postojanje objektivnih moralnih vrijednosti a istovremeno ne biti primoran smatrati da te vrijednosti imaju misterioznu sposobnost privlačenja. Kantovci, međutim, smatraju da to nije točno te da je u njihovu načinu zasnivanja etike ostavljen prostor za adekvatno objašnjenje motivacijske prirode morala. Jer govor o razlozima na koncu ima smisla samo ako je moguće da nas ti razlozi motiviraju. Naravno, teško je za očekivati da će svatko uvijek biti motiviran razlogom koji ima za neko postupanje, ali to se može očekivati od onoga tko je racionalan. Dakle, ovaj objektivistički pristup, suprotno Mackievu mišljenju, nije osuđen na uvođenje entiteta koji bi nekim nejasnim načinom poticali na djelovanje, već je samo riječ o prilično nekontroverznoj tvrdnji da razlozi moraju biti u stanju motivirati. Doduše ne svakoga, već samo onoga tko je racionalan. Ukratko, ako netko prosuđuje da je nešto ispravno, kantovci će reći, on zapravo prosuđuje da ima razlog to učiniti. Ali ako prosuđuje da ima razlog to učiniti, onda će, ako je racionalan, biti i motiviran to učiniti.

Usmjeravajući svu svoju pozornost na pokušaj uspostavljanja bliske veze između morala i praktičnog razuma, a pritom smještajući metafizička i epistemološka pitanja u drugi plan, izgleda da kantovci ne samo da uspijevaju pokazati kako etika može biti zasnovana na objektivnim temeljima, već i kako objasniti praktičnu prirodu moralnosti – činjenicu da nas naša moralna uvjerenja obično motiviraju (ili da bi nas trebala motivirati) na djelovanje. No ostavljajući sada po strani kantovsko razumijevanje statusa moralnih razloga i moralne motivacije, pogledajmo u najosnovnijim crtama na koji način tim istim pitanjima pristupa zagovornik hjumovske pozicije.

1.2. Hjumovska vizija praktičnog rasuđivanja i moralne motivacije

Dobra teorija praktičnog rasuđivanja mora pružiti plauzibilno objašnjenje odnosa između razloga i motivacije, a veliko je pitanje uspijevaju li kantovci tom uvjetu doista udovoljiti. Poteškoća s kantovskim shvaćanjem, kako sam ga gore prikazao, jest u tome što odgovarajuća motivacija kod subjekta nastaje kao *rezultat* prepoznavanja moralnih razloga ili, slobodnije formulirano, pod pritiskom čistog razuma, a, kao što ćemo ubrzo vidjeti, nije sasvim očito da ovakvo rješenje izbjegava problem na koji upozorava Mackie. Hjumovci se zato slažu u procjeni da kantovci zapravo nisu sposobni objasniti prirodu te relacije na adekvatan način i vjeruju da iz te činjenice izviru gotovo sve poteškoće i slabosti njihova pothvata. (Spomenimo da kantovce prvenstveno zanima odnos između *moralnih* razloga i motivacije, dok hjumovci govore o razlozima općenito, jer ne misle da moralni razlozi predstavljaju izuzetak.) Zadržimo se sada na ovom problemu i pokušajmo detaljnije razjasniti o čemu se radi.

Već je rečeno da subjekt može djelovati na osnovi (normativnih) razloga odnosno da mora biti moguće da normativni razlozi ujedno budu i njegovi motivirajući razlozi. Jednostavnije rečeno, ako on ima razlog postupiti na određeni način, onda mora biti moguće da on taj postupak učini upravo zbog tog razloga. Uočimo da bez te pretpostavke praktično rasuđivanje zapravo ne bi imalo nikakvog smisla: svrha takvog rasuđivanja jest utvrditi kakve imamo razloge za djelovanje i stoga je razumno pretpostaviti da mora postojati veza između naših zaključaka o tome što imamo razloga činiti te naše motivacije da to činimo. Kako kaže Darwall: "Izgleda da je intrinzična veza između neke vrste praktičnog suda i motivacije pretpostavka ozbiljne deliberacije" (Darwall, 1997: 308). Jer zašto bismo uopće donosili nekakve praktične zaključke ako na osnovi tih zaključaka nikada ne bismo mogli postupati ili, nešto preciznije, ako nas ti praktični zaključci nikada ne bi mogli motivirati na postupanje?

Ali ako bolje razmislimo, odmah ćemo uočiti da stvari ipak nisu tako jednostavne. Poteškoća je u tome što na prvi pogled nije sasvim jasno na koji način objasniti činjenicu da

nas čisto formalna načela praktičnog razuma mogu motivirati? Ili možda bolje, kako objasniti činjenicu da nas jedan *isključivo* racionalan proces, kao što je proces rasuđivanja, može potaknuti na neko djelovanje? Nekako nam izgleda da se izvor motivacije mora nalaziti u našoj konativnoj (afektivnoj) sferi, a ne u čisto racionalnoj. Razum je, kako kaže Hume, po svojoj prirodi "potpuno inertan i … nikad ne može ni spriječiti ni proizvesti neku radnju ili čuvstvo" (*R* 391).[3] Problem pred kojim se stoga nalazimo jest sljedeći: iako je, s jedne strane, sasvim evidentno da proces rasuđivanja mora rezultirati formiranjem odgovarajuće motivacije, s druge nam strane izgleda da takva motivacija jednostavno ne može nastati čisto racionalnim putem odnosno da proces rasuđivanja – bez dodatne pomoći – takvu motivaciju nije u stanju generirati. Hjumovski odgovor na postavljenu dilemu ima otprilike sljedeću strukturu: da bi proces praktičnog rasuđivanja rezultirao stvaranjem odgovarajuće motivacije, nekakva prethodna motivacija već mora postojati; cijeli proces mora započeti iz nečega što nas je u stanju pokrenuti na djelovanje, a s obzirom da želje imaju takvu sposobnost, nameće se zaključak da upravo one moraju kontrolirati i nalaziti se u podlozi svakog deliberativnog procesa.[4]

Posljednjom opservacijom dolazimo do onoga što bi se moglo nazvati klasičnim argumentom u prilog *instrumentalizmu* – koncepciji praktičnog rasuđivanja koja se gotovo uvijek poistovjećuje s hjumovskim gledištem. Strukturu tog argumenta lijepo prikazuje Elijah Millgram: "Kako bi išta moglo biti razlog za djelovanje ako te ne bi motiviralo da doista *učiniš* nešto? A što bi te moglo motivirati da učiniš nešto osim neke od tvojih želja?" (Millgram, 1997: 4–5) Sličnu formulaciju nude Mark Schroeder i Stephen Finlay: "ako postojanje razloga povlači mogućnost motivacije, a mogućnost motivacije povlači postojanje želje, onda postojanje razloga povlači postojanje želje" (Finlay i Schroeder, 2008: 9). Prema

[3] U navođenju Humeova djela *Rasprava o ljudskoj prirodi* koristim se prijevodom Borivoja Nedića (1983 [1739-1740]) te se pritom služim skraćenicom *R*.

[4] Potrebno je napomenuti da ovdje govorim o *hjumovskoj*, a ne o *Humeovoj* teoriji praktičnog rasuđivanja. To je zato što sam Hume nije mislio da praktično rasuđivanje uopće postoji. O tome ću više govoriti u trećem poglavlju.

instrumentalističkoj koncepciji subjekt ima razlog činiti ono što na najbolji način zadovoljava njegove želje i preferencije. Formalnije rečeno, ako S želi *x* i vjeruje da je *y* najbolji način za ostvarenje *x*-a, onda S ima razlog učiniti *y*. Razlozi su uvijek izvedeni iz prethodno prisutnih želja i zato je hjumovcima na raspolaganju vrlo jednostavno objašnjenje načina na koji razlozi stječu svoju motivacijsku snagu – razlog može motivirati zato jer je motivacija (koja se javlja u obliku želje) preduvjet za njegovo prisustvo.

Nije teško naslutiti što se događa ako hjumovsku viziju praktičnog rasuđivanja prebacimo u moralni kontekst. U tom slučaju proizlazi da su i *moralni* razlozi ovisni o našim željama i subjektivnim ciljevima: netko tko nema želju koju bi mogao realizirati moralnim djelovanjem jednostavno nema ni razlog tako djelovati. Hjumovci stoga vrlo lako mogu objasniti motivirajuću snagu razloga, ali ipak ne mogu zahvatiti kantovsku intuiciju (a koju, po svemu sudeći, dijele i drugi) da postoje razlozi, poput moralnih, koji ne ovise o našim željama. S druge pak strane, kantovcima je izuzetno stalo do toga da pokažu kako barem neki razlozi imaju objektivan status, ali zapadaju u poteškoće pri pružanju objašnjenja motivirajuće snage takvih razloga. Jer kada bi postojali objektivni razlozi – razlozi koji bi bili potpuno nezavisni od naših unutarnjih izvora motivacije i čija bi objektivnost proizlazila isključivo iz naše racionalne prirode – i dalje bi bilo nejasno odakle bi oni mogli crpiti svoju motivacijsku energiju. Žele li se obraniti od ovog prigovora, kantovcima ne preostaje ništa drugo nego pokazati da je Humeova teza o motivacijskoj inerciji razuma sasvim pogrešna.

Oslonimo li se na tezu da je razum, kako to smatraju hjumovci, motivacijski inertan odnosno ako razum "ne može nikad neposredno spriječiti ni proizvesti neku radnju protiveći joj se ili odobravajući je, on ne može biti izvor moralno dobrog i lošeg o kojima nalazimo da imaju takav utjecaj" (*R* 391–392). Nije nepoznanica što nam Hume ovdje govori. On također ne dovodi u pitanje praktični utjecaj morala i za njega je evidentno da moral "utječe na naše strasti i radnje" (*R* 391). Ali ako moral ima praktični utjecaj, a razum nema, onda je sasvim

jasno da moral ne može biti utemeljen u razumu. Pružiti objašnjenje motivacijske prirode morala za hjumovce dakle nije moguće učiniti u okviru teorije praktičnog razuma, već samo referiranjem na nešto za što je već utvrđeno da nas je sposobno pokrenuti, a to su elementi koji čine konativnu stranu ljudske prirode.

1.3. Zašto biti moralan?

Kantovski pokušaj objektivnog zasnivanja etike i način rješavanja problema moralne motivacije isključivo zavise o tome je li istinito da moralni zahtjevi pružaju razloge za djelovanje. Ta se teza, koja je također poznata pod nazivom *moralni racionalizam*, nalazi u podlozi cjelokupnog kantovskog pothvata i zato se vrijedno na ovom mjestu na njoj nešto duže zadržati. Riječ je o gledištu prema kojemu postoji "nužna veza između moralnih dužnosti i izvrsnih razloga za djelovanje. Prema ovome gledištu, uvijek postoji vrlo dobar razlog da se čini kako to moral nalaže" (Shafer-Landau i Cuneo, 2007: 283). Kakav bi to razlog mogao biti? Uočimo da, prema mišljenju moralnih racionalista, uvijek imamo razloga ponašati se u skladu s moralnim normama, a veliko je pitanje na koji način to uvjerenje pomiriti s činjenicom da su one nerijetko u konfliktu s našim afinitetima i željama. Poteškoća je u tome što nam moral ponekad nalaže da činimo radnje koje su u sukobu s našim preferencijama te se čini da u situacijama takvog tipa nije sasvim bjelodano da imamo razlog podvrgnuti se moralnim zahtjevima. O tome možda i najbolje svjedoči priča koju nalazimo na uvodnim stranicama Platonove *Države.* Raspravljajući sa Sokratom o tome je li pravedan život bolji i korisniji od nepravednog, Glaukon se poziva na mit o pastiru Gigu koji, pronašavši čarobni prsten, počne upotrebljavati njegove moći u svrhu ostvarivanja svojih vlastitih ciljeva. Učinivši se nevidljivim, ubije kralja, zavede njegovu ženu i sebi prisvoji veliko bogatstvo i moć. Pouku koju bismo trebali izvući iz priče jest da "nitko nije svojevoljno nego pod silu pravedan, budući da pravednost pojedincu nije dobro; jer svatko

čini nepravdu, gdjegod misli, da će je moći činiti. Zaista svaki čovjek neoporecivo misli, da pojedincu mnogo više koristi nepravda od pravednosti" (Platon, 1997: 82).

Ima li smisla kazati da Gigo nije imao razloga učiniti ono što je učinio? Netko bi potvrdan odgovor na to pitanje mogao obrazložiti na sljedeći način: pravednost je – *dugoročno promatrajući* – korisnija od nepravde, što znači da moralni oblik djelovanja zapravo nije u protuslovlju s onime što nam je u interesu. Iako je nesporno da nemoralnim ponašanjem ponekad možemo profitirati, to zasigurno ne vrijedi na "duge staze" i zato je uvijek isplativije ponašati se u skladu s onime što moral od nas zahtijeva. Poteškoća ovakvog pristupa pitanju opravdanja morala jest u tome što unaprijed blokira mogućnost da netko može profitirati nemoralnim ponašanjem čak i u dugoročnom smislu. Na prvi se pogled čini da je to ipak moguće, a to znači da moral i vlastiti interes samo kontingentno mogu biti usklađeni. Ako je Gigo, kao što to uostalom i sugerira sama priča, znao da neće biti uhvaćen i da će njegove radnje proći bez posljedica, zašto onda ne učiniti to što je on učinio? Zašto bi se itko u takvim okolnostima ponašao moralno? Bilo kako bilo, opisana obrana moralnosti velikim dijelom proizlazi iz hobbesovske tradicije te ću je ovdje ostaviti po strani. Kao što sam ranije napomenuo, u ovom ću se radu orijentirati na razlike između hjumovaca i kantovaca u pogledu nekih središnjih problema moralne filozofije i zato će, između ostalog, moja pažnja biti usmjerena na kantovski pokušaj opravdanja morala prema kojemu moral pruža *kategoričke* razloge za djelovanje. Kantovci imaju namjeru pružiti opravdanje morala koje bi vrijedilo nužno za sve i očito je da se stoga moraju osloniti na nešto drukčije razumijevanje racionalnosti od onoga koje racionalnost poistovjećuje sa slijeđenjem vlastitog interesa.[5]

Moralni bi racionalisti dakle kazali da je Gigo imao razloga *ne* učiniti ono što je učinio te da taj razlog ni na koji način ne ovisi o njegovim željama. No opet, u tom slučaju nije lako

[5] Slijediti vlastiti interes *nije isto* što i činiti ono što želimo. Primjerice, netko može željeti činiti nešto što mu nije u interesu. Zato racionalnost shvaćena u prudencijalnom smislu nije isto što i racionalnost shvaćena u instrumentalnom smislu.

razumjeti u čemu bi se taj razlog mogao sastojati. Zašto bi netko imao razlog odabrati moralni način življenja ako mu to ne bi koristilo ili ako time ne bi mogao zadovoljiti nijednu svoju želju? (usp. Shafer-Landau i Cuneo, 2007: 282). Na taj je problem za moralni racionalizam možda na najbolji način ukazala Philippa Foot u svom sada već antologijskom članku "Morality as a System of Hypothetical Imperatives" (Foot, 2002 [1972]). Foot započinje ukazivanjem na poznatu distinkciju između hipotetičkih i kategoričnih imperativa te analizom izraza "trebati" koji se, sukladno toj distinkciji, obično koristi u dva različita smisla. Ponekad nešto trebamo učiniti zato što jedino tako možemo ostvariti svoje ciljeve. Primjerice, ako želim dobro igrati tenis, onda trebam puno trenirati. Takvi nas hipotetički imperativi uvijek obvezuju pod određenim uvjetom odnosno pod uvjetom postojanja nekog određenog cilja koji želimo realizirati. Hipotetična upotreba pojma trebati razlikuje se od kategorične upotrebe – od one kada taj pojam koristimo bezuvjetno odnosno neovisno o tome imamo li neku određenu preferenciju, sklonost ili želju. Kant je, kao što je poznato, smatrao da upravo moralni zahtjevi poprimaju oblik kategoričkih imperativa. Središnje pitanje koje Foot u ovom kontekstu postavlja moguće je formulirati na sljedeći način: u kojem su smislu moralni zahtjevi *specifični* u odnosu na neke druge zahtjeve da upravo oni poprimaju kategorički, a ne hipotetički oblik? Mi zaista mislimo da se moralne norme na nas odnose bezuvjetno, ali to nije karakteristika *isključivo* morala. Sigurno je da mi nećemo, kako to kaže Foot, "opozvati" moralni zahtjev da ljudi trebaju govoriti istinu ako shvatimo da ljudi nemaju želju to činiti ili ako im govorenje istine ne može poslužiti kao sredstvo za ostvarenje nekog vlastitog cilja. Oni će, čak i u takvim okolnostima, i dalje biti podvrgnuti navedenoj moralnoj normi. Ali iz toga se još uvijek ništa ne može zaključiti. Radi se o tome da postoje i drugi, izvanmoralni konteksti u kojima se također čini da se zahtjevi na nas odnose bezuvjetno, ali nećemo odmah biti spremni kazati da ti zahtjevi imaju nekakav privilegirani status. Foot navodi primjer normi koje reguliraju lijepo ponašanje. Takve se norme također na nas *odnose* neovisno o

tome kakve su naše sklonosti ili preferencije. Primjereno je, recimo, kazati da je on trebao doći prikladnije odjeven na dodjelu diplome iako to možda nije želio. Taj se zahtjev na dotičnu osobu odnosi bezuvjetno, ali teško da ćemo misliti da taj zahtjev ima poseban dignitet i nužnost kakvu Kant pripisuje moralnim zahtjevima.

U čemu se onda sastoji razlika između moralnih normi i normi lijepog ponašanja? Zašto moralne norme imaju takav poseban status, status kakav nemaju nijedne druge norme? Jedini odgovor koji se nameće jest da moralni zahtjevi, za razliku od zahtjeva lijepog ponašanja, pružaju *razloge* za djelovanje. Kantovac bi mogao tvrditi da dok kršenje normi lijepog ponašanja ne povlači da se ponašamo u suprotnosti s razlozima koje imamo, to svakako vrijedi u slučaju kršenja moralnih normi. Onaj tko ne poštuje moralne norme jest *iracionalan* zato jer su moralni zahtjevi zahtjevi o onome što imamo razloga činiti, a racionalno djelovanje jest ono djelovanje u kojemu osoba djeluje u skladu s razlozima. No poteškoća s ovim shvaćanjem, ističe Foot, jest u tome što razloge za djelovanje možemo imati samo onda kada imamo određeni cilj koji želimo ispuniti. Njezinim riječima: "Iracionalni postupci su oni u kojima osoba na neki način poništava svoje namjere, čineći ono što se smatra neprobitačnim ili ono što ugrožava njegove ciljeve. Nemoralnost nužno ne uključuje nešto takvoga" (Foot, [2002] 1972: 162). Nemoralno ponašanje ne mora biti ujedno i iracionalno ponašanje. Naprotiv, ponekad isključivo nemoralnim ponašanjem možemo zadovoljiti svoje ciljeve dok moralnim ponašanjem to ne možemo postići.

Ukratko, temeljna misao mita o Gigu nije daleko od ideje koju zastupa Philippa Foot. Ima li smisla kazati da je Gigo iracionalan zato što je u pogrešne svrhe odlučio koristiti prsten koji ga može učiniti nevidljivim? Upravo suprotno. Kao što Foot sugerira da moralni zahtjevi ne pružaju razloge za djelovanje osim ako slijeđenje tih zahtjeva nije način na koji možemo realizirati svoje ciljeve, tako i mit o Gigu, interpretirajući ga u suvremenom kontekstu, sugerira da subjekt ima razlog postupati na određeni način samo ako tim putem može ostvariti

svoje želje i ciljeve. Gigo je svoje ciljeve uspio ostvariti nemoralnim djelovanjem i zato se čini ispravnim kazati da je imao razlog tako postupati. Stoga se njegovo ponašanje, u suprotnosti s kantovskom intuicijom, i može opisati kao racionalno.

Sljedeći argument, koji se u neznatno drukčijem obliku može pronaći kod Jamesa Dreiera (1997: 91), izvrsno sažima razlike između instrumentalista i racionalista kada je riječ o opravdanju morala.

> (1) Moral može biti opravdan samo ako postoji razlog za moralno djelovanje koji nije ovisan o subjektovim željama i preferencijama.
>
> (2) Subjekt ima razlog učiniti *y* ako i samo ako želi učiniti *x* i vjeruje da je za ostvarenje *x*-a potrebno učiniti *y*.
>
> (3) Razlozi za djelovanje ovise o subjektovim željama i preferencijama.
>
> (4) Moral ne može biti opravdan.

Nije teško naslutiti da premisa (3), kojom se tvrdi da se u podlozi racionalnog djelovanja nalazi instrumentalno načelo, predstavlja ključni korak u argumentu. Moralni racionalisti upravu tu premisu nalaze prijepornom i vjeruju da jednostavno nije točno da su *svi* razlozi za djelovanje utemeljeni u subjektovim željama. To je, kako oni smatraju, ujedno i najveća slabost Footina napada na kantovsko gledište – ona samo pretpostavlja da je racionalno djelovanje ono djelovanje koje je probitačno i kojim se ostvaruju unaprijed zadani ciljevi, ali ne nudi nikakvu nezavisnu argumentaciju da bi taj stav potkrijepila (usp. Smith, 1994: 84). Pretpostavimo li instrumentalističku koncepciju racionalnosti, moralni racionalisti tvrde, onda zaista nećemo biti u stanju registrirati nešto što je posve evidentno – da postoje razlozi za djelovanje koji *ne* ovise o našim željama. To dobro ilustrira sljedeći primjer:

> Ako se osoba utapa ispred vaših očiju, a vi je lako možete spasiti, onda je to ono što trebate učiniti. To što vam je to nezgodno ne mijenja vašu dužnost. Niti iskorjenjuje vaš razlog da pomognete osobi čiji je život u vašim rukama. Vi joj ne želite pomoći. Time nećete ništa dobiti. Ali zasigurno postoji *neki* razlog za pružanje pomoći, zar ne postoji? (Shafer-Landau, 2004: 131, kurziv u originalu).

Zagovornici moralnog racionalizma se slažu da postoje objektivni razlozi za moralno ponašanje, ali razilaze se u pogledu pitanja o snazi tih razloga. Dok neki smatraju da su moralni razlozi *prima facie* razlozi, naime razlozi koji ne moraju uvijek prevladati ako se sukobe s drugim, izvanmoralnim razlozima, drugi vjeruju da moralni razlozi izvanmoralne razloge uvijek nadjačavaju. Međutim, nepostojanje konsenzusa kada je riječ o snazi moralnih razloga ne mijenja činjenicu da su oni posve nezavisni od subjektovih psiholoških stanja.

Vidljivo je da se glavni spor između instrumentalista i moralnih racionalista odvija oko toga kako shvatiti prirodu praktičnog rasuđivanja. Proces praktičnog rasuđivanja, kojim nastojimo utvrditi razloge za djelovanje, odvija se po određenim normama i jasno je da rezultat tog procesa uvelike zavisi od toga kakva je priroda tih normi. Na primjer, ako je instrumentalna norma *jedina* norma rasuđivanja, onda proizlazi da će naši razlozi uvijek biti ovisni o našim željama, što znači da je kantovski pothvat objektivnog zasnivanja etike već u samom početku osuđen na neuspjeh. Zato kantovcima preostaje pokazati da postoje i druge norme praktičnog rasuđivanja – kategoričke norme koje mogu generirati kategoričke razloge za djelovanje. Ali zašto je teret dokazivanja na kantovcima? Odgovor bi nam sada već trebao biti poznat: privilegirani status instrumentalističke koncepcije nije nezaslužen zato jer ona na vrlo elegantan način ispunjava uvjet koji vrijedi za sve razloge – uvjet da nas razlozi moraju biti u stanju motivirati. Utemeljujući sve razloge u željama, hjumovski instrumentalisti ostaju vjerni ideji da motivacija mora imati konativno porijeklo.

1.4. Želje kao motivi i razlozi

Iz cijelog gornjeg prikaza moguće je izdvojiti dvije tvrdnje koje su karakteristične za hjumovsku poziciju i oko kojih se vodi glavna rasprava: (a) izvor motivacije nalazi se u željama i (b) razlozi su ovisni o našim željama. Središnji cilj ovog rada jest obraniti ove dvije tvrdnje te jasnije utvrditi kakvo shvaćanje moralnosti proizlazi iz njihova prihvaćanja. Konkretnije rečeno, moja je intencija dvostruka: dok ću s jedne strane nastojati obraniti gornje tvrdnje odnosno tezu prema kojoj su motivi i razlozi ovisni o subjektovim željama, s druge ću strane razmotriti kakve implikacije ta teza ima specifično za naše razumijevanje prirode *moralne* motivacije te prirode *moralnih* razloga. Budući da je moje gledište hjumovsko, očito je da se najvećim dijelom suprotstavljam kantovski usmjerenim autorima koji, vidjeli smo – utemeljujući moralnost u praktičnom razumu – pružaju sasvim drukčije objašnjenje motivacijske prirode morala i statusa moralnih razloga.

Rad se sastoji od dvije cjeline od kojih je svaka podijeljena u dva poglavlja. Dok je prva cjelina posvećena motivaciji, u drugoj pokušavam razjasniti prirodu razloga. U sljedećem poglavlju branim hjumovsku teoriju motivacije – teoriju u čijem se središtu nalazi tvrdnja (a). Prvo prikazujem središnja obilježja ove teorije, a potom – upotrebljavajući neke ključne distinkcije koje nalazimo kod samoga Humea te ih smještajući u suvremeni kontekst – pokušavam pokazati kako one mogu poslužiti zastupniku ovog modela u obrani od nekih standardnih prigovora.

U trećem poglavlju razmatram kakvo razumijevanje *moralne* motivacije slijedi iz prihvaćanja hjumovske teorije. Pokazujući na koji je način ta teorija kompatibilna s *internalizmom* i s *eksternalizmom motivacije*, opredjeljujem se za prvu poziciju. Internalizam motivacije jest stajalište koje tvrdi da moralne prosudbe nužno motiviraju, dok eksternalisti relaciju između moralnih prosudbi i motivacije smatraju raskidivom i kontingentnom. U poglavlju tvrdim da neki od najutjecajnijih eksternalističkih prigovora predstavljaju opasnost

samo za, kako ću to ovdje nazvati, kantovski interpretiran internalizam motivacije, ali ne i za hjumovsku verziju tog gledišta. Osim toga, kantovsko tumačenje internalizma motivacije zasnovano je na, po mom mišljenju, vrlo kontroverznoj pretpostavci da moralne činjenice pružaju objektivne razloge za djelovanje. Oboružani tom pretpostavkom, kantovci s lakoćom mogu objasniti nužnost veze između moralnih prosudbi i motivacije. Budući da razlozi moraju biti u stanju motivirati, a moralni su sudovi sudovi o razlozima (jer moralne činjenice pružaju razloge), slijedi da moralni sudovi moraju biti u stanju motivirati. Dakle, ističući da je ova verzija internalizma motivacije uvjerljiva samo ako prihvatimo navedenu pretpostavku pružam neke površne naznake zašto to ipak ne bismo trebali učiniti, a sustavniju diskusiju o tom pitanju prolongiram za posljednja dva poglavlja.

U drugoj cjelini rada, koju tvore četvrto i peto poglavlje, analiziram prirodu razloga za djelovanje. U četvrtom poglavlju nastojim obraniti tvrdnju (b) odnosno shvaćanje razloga koje se u širem smislu može nazvati hjumovskim. Preciznije rečeno, branim gledište poznato pod nazivom *internalizam razloga*. Iako je ovdje opet riječ o *internalizmu*, izuzetno je važno razlikovati ove dvije internalističke pozicije. Dok je internalizam motivacije pozicija koja tvrdi da postoji nužna veza između subjektove *moralne prosudbe* i motivacije, internalizam razloga je pozicija koja tvrdi da subjekt *ima* razlog za djelovanje samo ako je motiviran na odgovarajući način ili, da se poslužim terminologijom Bernarda Williamsa, samo ako ima odgovarajući element u svom motivacijskom skupu. Razvijajući neohjumovsku poziciju, internalisti razloga smatraju da su razlozi utemeljeni u subjektovoj psihološkoj konstituciji odnosno da su razlozi uvijek unutarnji. Jedna od najvećih vrlina internalizma jest što se elementi motivacijskog skupa podvrgavaju određenoj korekciji, što znači da to koje razloge imamo ne ovisi o onome što smo *de facto* motivirani učiniti, već o onome što bismo bili motivirani učiniti u idealnim uvjetima. *Eksternalisti razloga*, s druge strane, nastoje pokazati da postoje razlozi za djelovanje koji su neovisni o tome je li subjekt motiviran (ili bi bio

motiviran) djelovati u skladu s njima (ovdje je prvenstveno riječ o moralnim i prudencijalnim razlozima). Ti su razlozi, oni smatraju, izvan nas te se stoga i nazivaju *vanjskim razlozima*. U ovom poglavlju nastojim eliminirati četiri eksternalistička prigovora.

Najveći izazov hjumovskom internalizmu razloga ponovno predstavlja kantovska interpretacija internalističkog stajališta. Glavna intencija takvog pristupa jest oslabiti razliku između internalizma i eksternalizma razloga odnosno potkopati pretpostavku da je *jedino* eksternalistima otvorena mogućnost za dokazivanje kategoričkih, ne-relativnih razloga. Kantovski orijentirani autori, drugim riječima, nastoje osporiti relativističke implikacije hjumovskog internalizma te žele pokazati da je unutar internalističkog okvira također ostavljen prostor za dokazivanje postojanja kategoričnih razloga. Napadu na kantovski internalizam razloga posvećeno je peto, posljednje poglavlje ovog rada.

I. MOTIVI

2. Hjumovska teorija motivacije

Prema jednom vrlo utjecajnom modelu motivacije, motivirajući su razlozi konstituirani od dva različita psihološka stanja: subjekt ima motivirajući razlog za djelovanje samo ako posjeduje određeno *vjerovanje* i odgovarajuću *želju*. Oba stanja obnašaju različite uloge u motivacijskom mehanizmu, ali je njihovo međusobno nadopunjavanje esencijalno za njegovo funkcioniranje. Vjerovanja su pasivna, motivacijski inertna i njihova je uloga uglavnom svedena na pružanje relevantnih informacija, dok su želje aktivne te predstavljaju ona mentalna stanja iz kojih proizlazi glavni motivacijski impuls. Za razliku od želja koje su "esencijalno motivirajuće", vjerovanja imaju "kontingentno motivirajuću" prirodu; vjerovanja svoju "sposobnost da motiviraju dobivaju iz drugog izvora – želja – dok želje motiviraju same po sebi" (Dancy, 1993: 2). Dakle, premda se motivacijski proces, kako nam kaže ovaj model, sastoji od dva potpuno nezavisna i po prirodi sasvim drukčija mentalna stanja (kognitivnog i nekognitivnog), subjekt ne može u potpunosti biti motiviran na djelovanje ako mu jedno od tih stanja nedostaje. Evo i jednog primjera. Želim li popiti sok od naranče, sama prisutnost te želje nije dovoljna da to učinim ako pritom nemam nikakvo vjerovanje o tome gdje se taj sok nalazi odnosno vjerovanje o tome kako do njega mogu doći. S druge strane, vjerujem li da se sok od naranče nalazi na stolu ispred mene, ta informacija na mene neće vršiti nikakav motivacijski učinak nemam li ujedno i želju popiti taj sok. Parafrazirajući poznatu Kantovu krilaticu, mogli bismo kazati da su želje bez vjerovanja slijepe, a vjerovanja bez želja prazna.

Iako neki osporavaju da se opisani model motivacije može izvorno pripisati Humeu, on je u literaturi ipak postao poznat pod nazivom "hjumovska teorija motivacije" (HTM). Hjumovski model je, mnogi su skloni tvrditi, "dogma u filozofskoj psihologiji", ali dogma koju bi svakako trebalo napustiti (Smith, 1994: 92). Upravo zbog tog razloga hjumovci najčešće zauzimaju defenzivnu strategiju, nastojeći tu "dogmu" obraniti od mnogobrojnih

kritika. Jačina njihove pozicije stoga najvećim dijelom ovisi o tome u kojoj se mjeri oni tim prigovorima uspijevaju othrvati, a znatno manje o tome uspijevaju li formulirati originalne i inventivne argumente u prilog svoje teorije. U ovom ću poglavlju detaljnije analizirati glavne elemente hjumovskog modela te pokušati ocijeniti ozbiljnost spomenutih prigovora. U prvom ću dijelu pažljivije promotriti zbog čega hjumovci upravo u željama nalaze glavni izvor motivacije te ću odbaciti neke primjedbe koje im se u tom kontekstu upućuju. Potom ću izdvojiti neka obilježja želja kao mentalnih stanja i sugerirati da ukazivanjem na ta obilježja hjumovac znatno može učvrstiti svoj položaj. Na koncu ću izdvojiti četiri prigovora HTM-e te na njih pokušati replicirati.

2.1. Smjer odgovaranja

U raspravi o razlozima i motivaciji, pod željom se nerijetko misli i na ona mentalna stanja koja se u strogom smislu riječi ne mogu podvesti pod tu kategoriju. Hjumovci u tome ne vide neki ozbiljniji problem jer ionako vjeruju da je za opstanak njihova modela bitno zadržati njegove središnje elemente – prisutnost dvaju mentalnih stanja koja mogu egzistirati neovisno jedno o drugome i čije su funkcije jasno definirane. Oni stoga i dozvoljavaju da pokretalačku ulogu koja je namijenjena želji vrše i druga konativna mentalna stanja koja sa željama nisu istovjetna (npr. htijenja i nade), ali ta činjenica, kako oni navode, nije dovoljna da se slomi "*duh* [hjumovskog] argumenta". Taj duh ostaje sačuvan jer je pojam želje moguće zamijeniti i pojmom "prostava" (usp. Smith, 1994: 117).

Antihjumovci ovakav pristup uopće ne nalaze uvjerljivim i naglašavaju da velika većina problema s kojima se suočava zastupnik HTM-a izvire iz toga što je priroda želje nedovoljno određena. Zapravo se iz cijele rasprave, što će uskoro postati vidljivo, stječe dojam da je hjumovac stavljen u prilično nezgodan položaj: ne odredi li jasnije što želje jesu, postoji velika opasnost da njegovo stajalište postane sasvim trivijalno, a ako to ipak pokuša

učiniti, onda postaje evidentno da njegovo stajalište nema čvrsto uporište. Potreba za konkretnijim određenjem želja javlja se iz vrlo jednostavnog razloga: hjumovci moraju pokazati što je toliko specifično za same želje da samo one mogu imati intrinzično motivirajuću snagu. Jer zašto bi netko bez dodatnog argumenta prihvatio tezu da su vjerovanja pasivna u motivacijskom smislu, dok su želje intrinzično motivirajuće? S lakoćom bi se moglo tvrditi da relevantna razlika između vjerovanja i želja nije utvrđena te da je, zbog te činjenice, vjerovanjima u sklopu hjumovskog modela neopravdano uskraćena motivacijska snaga. Obrana HTM-e dakle za sobom povlači i preciznije razlikovanje želja od drugih mentalnih stanja (u našem slučaju vjerovanja) – tvrdi li se da *samo* želje imaju pokretalačku funkciju, onda mora biti jasno čime one zaslužuju takav privilegirani status i u čemu je njihova *differentia specifica*.

U pružanju odgovora na postavljeni problem hjumovski orijentirani autori ne moraju daleko odmaknuti od onoga što tvrdi sam David Hume. U najosnovnijim crtama, teza o motivacijskoj inerciji vjerovanja izvor ima u Humeovoj slici razuma, dok je teza o želji kao "esencijalno motivirajućoj" potaknuta njegovim stavom da strast "ne sadrži nikakvu reprezentativnu kvalitetu koja je čini kopijom nekog drugog postojanja ili preinake" (*R* 355). Hume argumentira da je domena unutar koje djeluje razum ograničena na područje činjenica i ideja, što samo rasuđivanje čini ili demonstrativnim (ako su u pitanju ideje) ili empirijskim (ako su u pitanju činjenice). Budući da zaključci proizvedeni takvim tipom rasuđivanja ne mogu imati praktičan učinak, slijedi da je razum motivacijski inertan. Na primjer, baš kao što nekakav matematički zaključak ne može motivirati osobu da postupi na određeni način, to ne može učiniti niti nekakav empirijski zaključak kao što je, recimo, onaj da su svi gavrani crni. Slobodnije rečeno, razumu je nemoguće dodijeliti aktivnu motivacijsku ulogu zbog toga što je on [razum] zadužen za stvari koje su po samoj svojoj prirodi nesposobne za vršenje bilo kakvog praktičnog utjecaja. Premise Humeove argumentacije sigurno nisu nekontroverzne i

imune na kritiku, ali to u ovom kontekstu i nije od presudnog značaja. Važno je ne izgubiti iz vida da je razum, kako to smatra Hume, zapravo zadužen za stvari koje mogu biti istinite ili lažne, a "… istina ili laž sastoji se u slaganju ili neslaganju bilo sa *stvarnim* odnosom predstava bilo sa *stvarnim* postojanjem i stvarnom činjenicom" (Hume, *R* 391). Strasti, s druge strane, nisu "podložne takvom slaganju ili neslaganju" – one nisu reprezentacijska mentalna stanja i stoga ne mogu biti istinita ili lažna.

Humeovo razlikovanje razuma i strasti s obzirom na ovaj, kako bi ga se moglo označiti, "reprezentacijski kriterij" čini osnovu za tretiranje želja kao motivacijski aktivnih i vjerovanja kao motivacijski pasivnih. U današnjoj je raspravi sama ideja još uvijek prisutna, ali u nešto drukčijem obliku. Nadovezujući se na nju, razliku između vjerovanja i želja hjumovci nastoje produbiti pozivanjem na, kako se to u literaturi obično naziva, "smjer odgovaranja" (usp. Smith, 1994: 111–112; McNaughton, 2010 [1988]: 106; Dancy, 1993: 27–28). Budući da vjerovanja i želje stoje u različitim relacijama u odnosu na svijet, kaže se da ona imaju različite smjerove odgovaranja. Ta se različitost manifestira u činjenici što vjerovanja nisu imuna na utjecaj koje vrši vanjsko stanje stvari, dok vanjsko stanje stvari ponekad nije imuno na utjecaj koje vrše želje. Sasvim precizno, vjerovanja "pokušavaju" reprezentirati svijet kakav jest i zato, ovisno o tome koliko su ti pokušaji uspješni, kažemo da su ona istinita ili lažna. Ako su vjerovanja lažna, onda ih trebamo promijeniti kako bi odgovarala svijetu. Za želje, koje nemaju istinosnu vrijednost, vrijedi obrnuto. Želje "pokušavaju" biti ostvarene i zato – da bi u tome uspjele – potiču na mijenjanje svijeta ako on s njima nije usklađen. U sada već klasičnoj knjizi *Intention*, G. E. M. Anscombe na slikovit način objašnjava misao koja stoji u podlozi predloženog kriterija. Ona navodi primjer čovjeka koji, kupujući u trgovini namirnice služeći se listom koju ima kod sebe, biva praćen detektivom koji na svoju listu zapisuje svaki njegov potez. Anscombe potom piše: "…ako se lista i stvari koje čovjek zapravo kupuje ne podudaraju, […] onda pogreška nije u listi, već u

čovjekovoj izvedbi […] no ako se detektivova lista i ono što čovjek zapravo kupuje ne podudaraju, onda je pogreška u listi" (Anscombe, 1963 [1957]: 56).

Iako je sada razumljivo na koji način hjumovci podupiru svoju ideju da upravo želje trebaju biti klasificirane kao motivacijska stanja, mnogi prigovaraju da je kriterij smjera odgovaranja prilično nezgrapan, da zahtijeva daljnju artikulaciju te da ne podupire HTM. Prigovor je da se u određenju tog kriterija treba odmaknuti od metaforičkog načina izražavanja i da distinkcija između vjerovanja i želja mora počivati na sigurnijem temelju (usp. Dancy, 1993: 27; Smith, 1994: 112). Zato antihjumovac može izabrati dva načina kako da se suprotstavi svom protivniku. Prvo, on može tvrditi da kriterij smjera odgovaranja nije adekvatan kriterij za razlikovanje navedenih mentalnih stanja i u skladu s tim argumentirati da hjumovci ipak nisu u stanju pokazati da vjerovanja nemaju motivacijsku snagu. Priklanjajući se ovoj strategiji on ne mora ujedno i tvrditi da razlika između vjerovanja i želja uopće ne postoji, već samo to da ta distinkcija nije dovoljno precizirana, a što ostavlja prostor i da se kognitivnim stanjima pripiše motivirajuća uloga. Drugo, on može prihvatiti navedeni kriterij te pokazati da to nije u sukobu s gledištem prema kojemu *i* vjerovanja imaju aktivnu motivacijsku ulogu. U podlozi te strategije zapravo se nalazi teza da želje također imaju kognitivnu prirodu te da ozbiljno treba razmotriti mogućnost postojanja mentalnog stanja koje bi imalo oba smjera odgovaranja.[6]

U sljedećem ću dijelu razmotriti obje strategije. Mislim da su antihjumovci u pravu kada ukazuju na nedovoljno preciziranu formulaciju kriterija o kojemu je riječ, ali da to još uvijek ne znači da bi taj kriterij trebalo odbaciti kao nekoristan. Iako još uvijek nisu pružili precizan kriterij za razlikovanje vjerovanja i želja, hjumovci su ipak pružili kriterij koji uvelike zahvaća naše intuitivno poimanje te distinkcije. Pred njima je zasigurno još uvijek velika zadaća, što ne znači da su njihovi oponenti samim time u povoljnijem položaju. Jer, na

[6] Navedene strategije navodi McNaughton (2010 [1988]: 107). McNaughton je dobar primjer antihjumovca koji prihvaća kriterij smjera odgovaranja, ali koji smatra da time nije pokazano da vjerovanja ne mogu motivirati.

koncu konca, čak i ako ne postoji precizan kriterij za razlikovanje želja i vjerovanja, još je uvijek dalek put do zaključka da vjerovanja, kao kognitivna stanja, mogu motivirati: ta tvrdnja također ne dolazi sama od sebe, već je i za nju potrebno pružiti adekvatnu argumentaciju. O tome uspijevaju li antihjumovci to učiniti, bit će više riječi nešto kasnije.

2.1.1. *Kontrafaktička ovisnost i status 'neobičnih' želja*

Michael Smith kriterij smjera odgovaranja određuje na sljedeći način:

> Razlika između vjerovanja i želja s obzirom na smjer odgovaranja može se razumjeti kao razlika u funkcionalnim ulogama vjerovanja i želje. Vrlo grubo i pomalo pojednostavljeno, razlika je, *inter alia*, u kontrafaktičnoj ovisnosti vjerovanja da p i želje da p uslijed percepcije sa sadržajem da ne p: vjerovanje da p sklono je nestati u prisustvu percepcije sa sadržajem da ne p, dok je želja da p sklona potrajati, stvarajući kod subjekta u tom stanju dispoziciju da ostvari p (Smith, 1994: 115).

Razliku u kontrafaktičnoj ovisnosti ovih dvaju mentalnih stanja moguće je objasniti u obliku sljedećeg pitanja: što bi se dogodilo s mojim vjerovanjem da *p*, a što s mojom željom da *p* kada bih utvrdio da *p* nije slučaj? U slučaju da se radi o vjerovanju da *p,* to bih vjerovanje imao sklonost odbaciti; utvrdim li da *p* nije slučaj, to bi značilo da je moje vjerovanje takvo da ne odgovara stvarnom stanju stvari te da ga je potrebno zamijeniti istinitim vjerovanjem. S druge strane, ako želim da *p,* a utvrdim da *p* nije slučaj, onda neću odbaciti svoju želju da *p,* nego ću tu želju biti sklon zadržati te ću biti disponiran učiniti da *p* ipak bude slučaj.

Smithovo je određenje naišlo na udar različitih kritika (usp. Price, 1989; Humberstone, 1992; Sobel i Copp, 2001), a na ovom ću mjestu izdvojiti dvije za koje mislim da se ponajviše ističu te ću nastojati utvrditi u kojoj mjeri one ugrožavaju hjumovski položaj. Osvrnemo li se još jednom na gornji opis, možemo upitati sljedeće: što znači "*utvrditi* da *p* nije slučaj" ili,

koristeći točnu Smithovu formulaciju, što znači imati "percepciju sa sadržajem da ne p"? Ne smijemo zaboraviti da je intencija pružiti kriterij po kojemu se vjerovanja razlikuju od želja, što povlači da taj kriterij ne smije sadržavati bilo eksplicitno bilo implicitno pozivanje na neko od ta dva mentalna stanja (na taj bi se način samo pretpostavljalo ono što se tek treba dokazati). Međutim, izgleda da uvođenjem mentalnog stanja "percepcije" Smith upada upravo u tu zamku: stanje percepcije implicira prisutnost vjerovanja. Mogu li, na primjer, percipirati da je vani sunčan dan, a istovremeno ne formirati vjerovanje da je vani sunčan dan? Ako je odgovor negativan, onda je Smithovo određenje zapravo ekvivalentno tvrdnji da će subjektovo vjerovanje da *p* imati sklonost nestajanju pod utjecajem njegova *vjerovanja* da *ne p*. U slučaju da je tako, cijela formulacija postaje cirkularna jer objašnjenje prirode vjerovanja kao mentalnog stanja sadrži pozivanje na samo to mentalno stanje. Ipak, kako ispravno primjećuju Sobel i Copp (2001: 46–47), Smith bi mogao replicirati da je njegova formulacija u tom smislu neutralna odnosno da ostavlja prostor za razlikovanje percepcije od samog vjerovanja. Da upotrijebim jedan primjer, gledajući neku mađioničarsku predstavu, moguće je da imam percepciju sa sadržajem da čovjek lebdi iznad zemlje, a da pritom ne formiram vjerovanje da čovjek zaista lebdi iznad zemlje. Prema toj interpretaciji, iako imam percepciju sa sadržajem da *p*, još uvijek ne moram vjerovati da zaista *p*. Sobel i Copp, međutim, argumentiraju da to Smitha ipak ne spašava zapadanja u cirkularnost. Problem je u tome što percepcija sa sadržajem da *ne p* zasebno neće učiniti da vjerovanje da *p* ima sklonost nestati, već je, da bi do toga došlo, potrebno da subjekt prosudi da je ta percepcija pouzdana. Tek ako prosudi da je percepcija sa sadržajem da *ne p* pouzdana njezino će vjerovanje da *p* imati sklonost nestati. Ali sada smo opet na početku. Što znači *prosuditi* da je percepcija sa sadržajem da *ne p* pouzdana? Nameće se jedan mogući odgovor: to ne znači ništa drugo nego imati *vjerovanje* da je percepcija sa sadržajem da *ne p* pouzdana. Kako god interpretirali

Smithovo razumijevanje kriterija smjera odgovaranja, čini se da ono teško može izbjeći problem cirkularnosti.

No je li zaista tako? Cirkularni argument jest onaj u kojem se, kako je ranije navedeno, pretpostavlja ono što se tek treba dokazati. Optužba za cirkularnost može biti utemeljena samo ako argument počiva na pretpostavci koja jest predmet spora, tj. ako je ta pretpostavka ono oko čega se i vodi glavna rasprava. Međutim, nisam siguran da se to može kazati i u ovom slučaju. Da pojasnim. S obzirom da je namjera utvrditi kakva je *priroda* vjerovanja i želja ili, preciznije, u čemu se sastoji relevantna razlika između vjerovanja i želja *kao mentalnih stanja*, na prvi pogled zaista izgleda da predloženi kriterij ne smije sadržavati prethodno pozivanje na neko od tih stanja. Ali ovo je neprecizno. Naime, što se ovdje podrazumijeva pod "prethodnim pozivanjem" ili pretpostavljanjem nekog od navedenih stanja? Strogo rečeno, Smithovo prethodno pozivanje na vjerovanje nije pozivanje na vjerovanje *kao na neko mentalno stanje*, već na *sadržaj* tog vjerovanja. On naglasak ne stavlja na vjerovanje kao takvo i zato ostaje nerazjašnjeno na kojoj bi ga se osnovi moglo teretiti za cirkularnost. Dakle, čak i ako je točno, kako to sugeriraju Sobel i Copp, da "percepcija sa sadržajem da ne p" doista povlači prisutnost vjerovanja, time još uvijek nije pokazano da se ovdje pretpostavlja ono što tek treba dokazati. To bi se moglo zaključiti samo kada bi se pretpostavljalo nešto o samoj prirodi tog vjerovanja, a mislim da Smith ipak ne čini tu pogrešku.

Izuzev opisanog načina obrane, vjerujem da zagovorniku kriterija smjera odgovaranja ostaje još jedna mogućnost. On bi, suočen s gornjim optužbama, mogao učiniti sasvim neočekivani potez – priznati da su te optužbe potpuno osnovane i da ne postoji način na koji bi se izbjeglo pozivanje na jedno od mentalnih stanja čija priroda tek treba biti objašnjena, ali istovremeno zastupati stav da zapadanje u takvu vrstu cirkularnosti ne povlači potpuno odbacivanje kriterija koji je u pitanju. Uostalom, gornja kritika, on bi mogao reći, govori

samo to da još uvijek ne postoji precizna formulacija predloženog kriterija, ali da nepostojanje precizne formulacije ne znači da treba odbaciti cijelu ideju kao neutemeljenu. I, na koncu konca, sam Smith kaže da je njegova formulacija "gruba i pomalo pojednostavljena" čime eksplicitno daje do znanja da mu intencija nije pružiti takvu analizu koja bi bila oslobođena svih poteškoća, već samo površno skicirati zašto vjerovanja i želje imaju različite smjerove odgovaranja.

Još jedan problem na koji nailazi Smithovo određenje kriterija, a koji ovdje valja spomenuti, jest postojanje, kako ću ih nazvati, "neobičnih" želja. Poteškoća izvire iz činjenice što spektar nečijih želja zaista može biti raznolik i nije sasvim očito da svaka želja čini da je subjekt disponiran na odgovarajući način. Alfred Mele daje primjer čovjeka koji, dočekujući svog prijatelja na aerodromu, želi da je njegov avion krenuo na vrijeme.[7] Međutim, nema nikakvog smisla kazati da on sada ima dispoziciju to ostvariti jer je to ionako izvan domene njegova utjecaja (usp. Mele, 2003: 26). Moguću repliku na prigovor takvog tipa dao je sam Smith (1994: 208–209, fn.7) tvrdeći da je HTM teorija koja govori o motivirajućim razlozima, što znači da u njezin djelokrug, po definiciji, ne spadaju želje koje su usmjerene na prošlost ili one na čiji sadržaj subjekt nema nikakav utjecaj. Da bi se izbjegla spomenuta poteškoća dovoljno je ograničiti hjumovsku teoriju na onu vrstu želja čija realizacija zavisi isključivo od naših postupaka.[8] No to ne mora biti jedini način. Prigovor "neobičnih" želja, po mom mišljenju, zasnovan je na kontroverznoj pretpostavci da dispoziciju za ostvarenjem *p*-a možemo imati samo ako je za nas moguće da ostvarimo *p*. Teško je razumjeti zašto bismo tu pretpostavku smatrali opravdanom. Netko može imati dispoziciju ostvariti *p* a istovremeno

[7] Iako bi bilo ispravnije kazati da se čovjek *nada* da je avion krenuo na vrijeme, a ne da on *želi* da je avion krenuo na vrijeme, mislim da ta distinkcija nije od velike pomoći u obrani od Meleova prigovora. Jer prihvatimo li gledište nekih hjumovaca koji želje, kako sam već ranije napomenuo, supstituiraju "prostavovima", spomenuta distinkcija automatski gubi na značaju. Drugim riječima, nada je jedna vrsta prostava, a prostav u Meleovu primjeru opet kod subjekta ne stvara odgovarajuću dispoziciju (usp. Mele, 2003: 35, fn. 15). Nameće se dakle potreba za jasnijim određenjem želja. O tome će biti više riječi nešto kasnije.

[8] Zanimljivo je da Mele uopće ne spominje ovaj Smithov odgovor.

biti sasvim svjestan da ostvarenje *p*-a nije nešto na što on sam može utjecati. Iako on ništa ne može učiniti da ostvari *p*, još uvijek može biti istinito da on ima dispoziciju ostvariti *p*.[9]

2.1.2. *Mentalno stanje s oba smjera odgovaranja?*

Čak i ako je kriterij smjera odgovaranja pouzdan kriterij za razlikovanje vjerovanja i želja, to ne isključuje mogućnost postojanja mentalnog stanja koje bi imalo oba smjera odgovaranja. Specifičnost tog stanja bila bi u tome što bi ono s jedne strane bilo slično vjerovanju jer bi trebalo odgovarati svijetu, a s druge bi strane bilo slično želji jer bi svijet trebao odgovarati tom stanju. Prijedlog o kojemu je ovdje riječ hjumovci ne mogu uzeti ni u razmatranje budući da se kosi sa središnjim obilježjem njihove teorije – da je relacija između vjerovanja i želja isključivo kontingentne prirode. Kontingentnost te relacije znači da prisutnost nekog vjerovanja ne mora ujedno i povlačiti prisutnost neke konkretne želje, baš kao što ni prisutnost neke želje ne mora ujedno i povlačiti prisutnost nekog konkretnog vjerovanja. U svjetlu toga je lako razumjeti zašto bi egzistencija mentalnog stanja koje bi u sebi istovremeno sadržavalo elemente vjerovanja i želje bila posebno fatalna za hjumovce: antihjumovac bi time pokazao da kriterij smjera odgovaranja uopće ne potkrepljuje HTM.

Ime koje se najčešće spominje u kontekstu zagovaranja gornje strategije jest ono Johna McDowella (1988). Pomalo pojednostavljeno, McDowell argumentira da prisutnost odgovarajuće dispozicije za djelovanje čini neodvojiv dio sagledavanja neke situacije na pravilan način. Pravilno sagledavanje određene situacije sposobnost je koju ima moralna osoba – ona je uvijek potaknuta svojim "viđenjem" nekog stanja stvari i zbog toga potpuno objašnjenje njezina ponašanja ne mora uključivati pozivanje na neko dodatno motivacijsko

[9] U ovom kontekstu valja spomenuti i primjedbu koju zastupnicima smjera odgovaranja upućuje Berčić: "… želje smatramo ostvarivima ili neostvarivima, primjerenima ili neprimjerenima, itd. Meni se može sviđati ideja da budem ruski car, no ako bih svoje ponašanje doista podredio tom cilju, ne bi mi bila potrebna pomoć u realizaciji cilja već u odustajanju od njega. To pokazuje da i naše želje ipak trebaju odgovarati svijetu, ne samo da svijet treba odgovarati njima" (Berčić, 2009: 30, fn. 41). Međutim, kazati da su želje *ostvarive* nije isto što i kazati da one *odgovaraju svijetu*, a mislim da Berčić čini upravo tu pogrešku.

stanje. Ovdje je riječ o *kognitivnom* stanju, što znači da ono treba biti usklađeno sa svijetom, ali kognitivnom stanju koje nužno uključuje i prisutnost odgovarajućeg *poticaja*, što znači da i svijet treba biti usklađen s tim stanjem. Sagledava li subjekt pravilno određeno stanje stvari, on će imati i odgovarajući poticaj, a ako nema odgovarajući poticaj, onda ne sagledava određeno stanje stvari na pravilan način. McDowellovim riječima: "… stanje koje sebe predstavlja kao kognitivno, ali povlači konativno stanje, ipak mora biti samo djelomično kognitivno, te sadržavati konativno stanje kao dio" (McDowell, 1998: 82). McDowell dakle pokušava dokazati postojanje mentalnog stanja koje ima oba smjera odgovaranja.

Evo i nekih primjera koje McDowell spominje – doduše samo površno – da bi potkrijepio svoje razmišljanje. Mnogi ljudi jazz glazbu ne nalaze atraktivnom, zvuči im dosadno i u njoj ne nalaze neke pravilne melodijske linije. Uvjeriti takve ljude u suprotno značilo bi pokušati učiniti da oni jazz glazbu *vide* drukčije nego što je vide sada. U području morala stvari nisu nimalo drukčije. Moralna osoba određene okolnosti jednostavno sagledava drukčije nego ona koja nema razvijenu tu perceptivnu sposobnost; moralna će osoba činjenicu da je netko osjetljiv i sramežljiv vidjeti kao razlog da s njim postupa pažljivo, a činjenicu da je netko ozlijeđen vidjeti kao situaciju u kojoj ona (ili bilo tko drugi) nešto treba učiniti. A kako je ranije istaknuto, već je samo pozivanje na nečije pravilno sagledavanje određene situacije dovoljno za objašnjenje njegova djelovanja. Jer netko tko vidi da je nečija sramežljivost razlog da prema njemu postupa oprezno i pažljivo, samim time ima i poticaj prema njemu oprezno i pažljivo postupati te mu nikakva dodatna želja nije potrebna. Oni kod kojih taj poticaj izostaje nemaju potrebnu razinu osjetljivosti, a to ih po svemu sudeći i sprečava da navedene situacije vide u pravom svjetlu.

Unatoč privlačnosti opisane strategije, neke stvari ostaju nedorečene. Zašto bi bilo nemoguće zamisliti osobu koja bi određenu situaciju sagledavala u pravom svjetlu, ali kojoj bi

nedostajao odgovarajući poticaj?[10] Zašto bi, drugim riječima, posjedovanje perceptivne sposobnosti o kojoj priča McDowell nužno moralo značiti da je onaj koji tu sposobnost posjeduje ujedno i disponiran u određenim okolnostima postupati na određeni način? McDowellov opis moralne osobe na intuitivnoj razini zvuči prihvatljivo te zaista odgovara našem svakodnevnom shvaćanju motivacije, ali – čak i pod pretpostavkom da pravilno sagledavanje neke situacije u sebi uključuje kognitivni element – nije objašnjeno zašto se taj element ne bi mogao razdvojiti od onog konativnog. Michael Smith (1994: 120–123) neoboriv dokaz mogućnosti takvog razdvajanja vidi u slučajevima slabosti volje i "različitim oblicima depresije". Prema njegovu mišljenju, karakteristični učinak koji ti fenomeni imaju na subjekta jest da mu potpuno dokidaju motivaciju. Kada bi moralna osoba patila od *akrasije*, ona ne bi formirala poticaj djelovati na odgovarajući način iako bi njezino viđenje situacije i dalje bilo pravilno. No, za razliku od Smitha, mislim da pozivanje na fenomene poput depresije i slabosti volje uopće nije potrebno da bi se na uvjerljiv način zagovarala mogućnost razdvajanja kognitivnog i ne-kognitivnog momenta. Njih je moguće razdvajati i u regularnim okolnostima, naime onima kada nije riječ o depresiji i sličnim stanjima. Osim toga, ni izdaleka nije očito da fenomeni koje Smith spominje zaista imaju takav učinak kakav on misli da imaju.[11]

[10] Ovdje nalazim važnim spomenuti sljedeće. Iz svega rečenog lako je zaključiti da vidjeti situaciju u pravom svjetlu u osnovi znači vidjeti da nešto treba učiniti (da se prema nekome *treba* odnositi pažljivo, da nekome *treba* pomoći itd.). Riječ je dakle o vrijednosnim (moralnim) tvrdnjama. Imajući to u vidu, vidjeti da nešto treba učiniti za kognitivista znači *vjerovati* da nešto treba učiniti, a za nekognitivista znači *željeti* to učiniti (ovdje se misli na želju u širem smislu). Zato primjedbu da je moguće vidjeti situaciju u pravom svjetlu, a nemati odgovarajuću motivaciju upućujem *isključivo* kognitivističkoj interpretaciji, ali ne i onoj nekognitivističkoj. Jer uočimo da je za nekognitivista vidjeti situaciju u pravom svjetlu (vidjeti da nešto treba učiniti) isto što i imati odgovarajuću motivaciju. Nekognitivist, drugim riječima, ne misli da je *viđenje određene situacije u pravom svjetlu* kognitivno stanje. Na ovu ću se distinkciju ponovno vratiti u sljedećem poglavlju.

[11] O tome ću također kazati nešto više u sljedećem poglavlju.

2.2. Želje i njihova priroda

Iz prethodne bi rasprave trebalo biti vidljivo da čak ni hjumovci ne gaje iluziju da je kriterij smjera odgovaranja besprijekorno formuliran i imun na prigovore.[12] Ali unatoč tome, kako to lijepo kaže Dancy, teško da netko može osporavati "snagu intuicije koju gornja priča pokušava zahvatiti" (Dancy, 1993: 28). Poruka je jasna. Naše intuitivno razumijevanje razlike između želja i vjerovanja ponuđeni kriterij, usprkos svim svojim slabostima, ipak uspijeva zadržati. Čak ni antihjumovac to neće negirati. Uostalom, njegova primarna namjera nije kontrirati činjenici da želja ima izuzetno važnu ulogu u motivacijskom procesu, već dokazati da aktivnu ulogu u velikoj većini slučajeva može imati *i* vjerovanje. Jedino se radi o tome da u nepreciznom određenju kriterija on sada nalazi prostor za izgradnju svoje pozicije. Priroda želja, on smatra, još uvijek nije dovoljno dobro objašnjena da bi se sa sigurnošću moglo tvrditi da *samo one* mogu motivirati u pravom smislu te riječi. Zato je od presudnog značaja da hjumovac pruži nešto preciznije analizu tog mentalnog stanja, a koja će mu, kako ću kasnije tvrditi, pomoći da se ogradi od nekih prigovora i tako lakše diskvalificira antihjumovske pokušaje.

2.2.1. *Divlje i tihe strasti*

Budući da se model motivacije koji je u središtu mog razmatranja naziva *hjumovskim*, možda je primjereno i započeti osvrtanjem na ono što je sâm Hume imao kazati o pitanju koje nas ovdje zanima. Što Hume smatra pod željama odnosno strastima?[13] Odgovor nalazimo u II. knjizi *Rasprave o ljudskoj prirodi*: "Ono što obično razumijemo pod *strašću* jest snažna i osjetna emocija uma, kad je prisutno neko dobro ili zlo, ili neki predmet koji je, zbog originalnog ustroja naših duševnih moći, podešen da izazove sklonost" (*R* 372). Idejom o

[12] Treba spomenuti da pozivanje na taj kriterij nije specifičnost samo rasprave o motivaciji. Status tog kriterija igra važnu ulogu i u drugim filozofskim disciplinama kao što je filozofija jezika i epistemologija.
[13] Suvremene rasprave zanemaruju ovu distinkciju. Osim toga, želje za Humea potpadaju pod strasti (usp. Hume *R* 356).

strastima kao "osjetnim emocijama" Hume sugerira da strasti – kada ih imamo – nužno osjećamo i da smo uvijek svjesni njihove prisutnosti. Postavljajući to viđenje strasti u kontekst suvremenih rasprava, Humeu se ponekad pripisuje zastupanje koncepcije prema kojoj želje imaju određena fenomenološka svojstva, a koja čine da njihova prisutnost ne ostaje neregistrirana. Prisutnost određenog mentalnog stanja stoga nužno povlači i svjesnost o prisutnosti tog stanja.

Nije jednostavno razumjeti na koji se način ostali dijelovi Humeove *Rasprave* uklapaju u ovu fenomenološku koncepciju želja.[14] Hume naime također priznaje postojanje onih strasti koje "proizvode malo emocije u umu i [koje] su poznatije po svojim posljedicama nego po neposrednom osjećanju ili osjetu" (*R* 356). Tihe strasti "ne uzrokuju metež u duši" (*R* 356) i zato se razlikuju od "divljih" strasti koje intenzivnije osjećamo. Ključna razlika između ove dvije vrste strasti upravo je fenomenološka: dok "divlje" strasti osjećamo jačim intenzitetom, to nije slučaj kod "tihih" strasti kojih uopće ne moramo biti svjesni. Kao primjere tihih strasti Hume navodi "ljubav prema životu" i "dobrotu prema djeci". Te strasti ne moramo uvijek osjećati jakim intenzitetom, ali to ne znači da one ne postoje ili da naše djelovanje njima nije rukovođeno. Možda i ne moram u svakom trenutku biti svjestan svoje volje za životom, kao što je, primjerice, trenutak u kojem odlazim na rutinski zdravstveni pregled ili onaj u kojem pazim kako prelazim cestu. Na isti način ne moram u svakom trenutku biti svjestan ljubavi prema vlastitom djetetu, kao što je, primjerice, onaj kada ga vodim u šetnju. Vrlo je vjerojatno da ću i sam smatrati da sam u tim slučajevima motiviran "hladnim, naizgled nestrastvenim rasuđivanjem" (Shaw, 1989: 164–165). Moguće je da ću motiv za odlazak na pregled pronaći u vjerovanju da je odgovornost svake odrasle osobe da se brine za svoje zdravlje, dok ću ustanoviti da sam motiviran odvesti dijete u šetnju zbog mog vjerovanja da je za svako dijete dobro da što više vremena provodi na svježem zraku. Ali

[14] Izraz "fenomenološka koncepcija želja" posuđujem od Smith (1994: 105).

naravno, Hume ističe, takvo rasuđivanje samo naizgled nije popraćeno strastima. Strasti su uvijek osnovni pokretač i bez njihove prisutnosti djelovanje nije niti moguće.

Koji je odnos između intenziteta kojim osjećamo pojedinu strast i motivacijske moći te iste strasti? Slijedi li da su "divlje" strasti, s obzirom da ih osjećamo jačim intenzitetom, ujedno i motivacijski snažnije ili je možda tako da obje vrste strasti imaju podjednaku motivacijsku snagu? Kada bi "divlje" strasti ujedno bile i motivacijski snažnije, to bi značilo da mi ne bismo mogli ne djelovati na osnovi njih te da one "tihe" nikada ne bi mogle prevladati. No svakodnevno nam iskustvo govori da to nije točno: mi često djelujemo suprotno onim strastima koje intenzivnije osjećamo; unatoč tome što osjećamo da nas "divlje" strasti vuku u suprotnom smjeru, mi ih nerijetko uspijevamo nadvladati i djelovati suprotno njima. Kako bi to uopće bilo moguće kada bi "tiha" strast bila motivacijski slabija? Osim toga, Humeova pozicija ne bi bila konzistentna kada bi on veću motivacijsku snagu pripisivao "divljim" strastima jer bi to proturječilo gore navedenim primjerima iz kojih je vidljivo da nas nerijetko upravo "tihe" strasti pokreću na djelovanje. Proizlazi li onda da su "divlje" i "tihe" strasti na koncu ipak podjednake u motivacijskom smislu? Najbolje je pogledati što sâm Hume kaže o tom pitanju:

> (1) Očevidno je, strasti ne utječu na volju razmjerno njihovoj "divljini", ili pometnji koju izazivaju u raspoloženju… (*R* 357)

> (2) Općenito govoreći, "divlje" strasti imaju snažniji utjecaj na volju, iako se često vidi da ih one "tihe", kad ih podržava misao i potpomaže odlučnost, mogu kontrolirati u njihovim najbjesnijim nemirima. Ono što čitavu ovu stvar čini manje izvjesnom jest to što se "tiha" strast može lako pretvoriti u "divlju", bilo uslijed promjene raspoloženja bilo pak uslijed okolnosti i situacije objekta; kao pomoću posuđivanja snage od neke prateće strasti, pomoću navike, ili pomoću izazivanja mašte. (*R* 373)

Jesu li neke od Humeovih tvrdnji koje nalazimo u odlomcima (1) i (2) u kontradikciji? Hume, s jedne strane, argumentira da "strasti ne utječu na volju razmjerno njihovoj 'divljini'", dok, s druge strane, ipak tvrdi da "općenito govoreći, 'divlje' strasti imaju snažniji utjecaj na volju". Iako na prvi pogled doista izgleda da između ovih tvrdnji postoji određena tenzija, to ipak nije tako. Presudno je ne izgubiti iz vida da iskaz iz odlomka (2) vrijedi u generalnom smislu. Hume kaže da divlje strasti "*općenito govoreći* … imaju snažniji utjecaj na volju", a to ne proturječi tvrdnji da divlje strasti *ponekad* nemaju takav utjecaj. Iskaz u (1) je istinit baš zato što se ponekad takav utjecaj pripisuje upravo tihim strastima. Čini se da ova slika otvara prostor za razlikovanje fenomenološke komponente strasti od one motivacijske: stupanj intenziteta kojim osjećamo neku strast ne mora biti proporcionalan njezinoj motivacijskoj moći. Naprotiv, moguće je da "tihe" strasti – čije prisutnosti zbog nedostatka fenomenoloških svojstava ne moramo biti svjesni – i dalje zadržavaju svoju motivacijsku snagu.

Imajući u vidu navedenu distinkciju, moguće je i objasniti što se događa u onim slučajevima kada se "divlja" i "tiha" strast mijenjaju jedna u drugu. U slučaju promjene "tihe" strasti u onu "divlju", tiha strast dobiva na intenzitetu, ali ne nužno i na motivacijskoj snazi. Na sličan način, kada se "divlja" strast mijenja u "tihu", ona gubi svoje fenomenološke karakteristike, ali time ne gubi nužno i motivacijsku moć. Tako će se, da se poslužim Humeovim primjerom, moja ljutnja koju imam prema nekome intenzivirati ako mi taj čovjek nanese neko novo zlo, ali porast u intenzitetu ne mora *nužno* pratiti i porast u motivaciji. Na sličan način, ako mi neka djelatnost prema kojoj gajim veliku strast pređe u naviku, nije isključeno da će moja strast s vremenom oslabiti i postati manje intenzivna, ali smanjenje intenziteta ne implicira i smanjenje motivacije da se nastavim baviti tom djelatnošću. Jer, kako kaže Hume, "…navika pruža ne samo lakoću da se obavi neka radnja već isto tako i naklonost ili težnju ka toj radnji…" (*R* 361).

Kako je uopće moguće objasniti promjenu strasti iz jednog stanja u drugo? Hume navodi niz primjera kojima dokazuje tu mogućnost, a svima je zajednička misao da intenzitet strasti ponajviše ovisi o "živosti" njezina predmeta. Ako je predmet naše strasti "življi", time i sama strast time postaje intenzivnija, dok strasti postaju tiše i mirnije u slučaju da njihov predmet gubi na živosti. Na živost predmeta utječu različiti čimbenici poput upotrebe imaginacije, položaja tog predmeta te njegove udaljenosti u prostoru i vremenu. Neposredna prisutnost predmeta sam predmet čini jasnijim i življim, a sličan efekt polučuje i maštanje o tom predmetu, kao i zadovoljstvo koje osjećamo kada smo u neposrednom dodiru s tim predmetom.

Kao što će, nadam se, ubrzo postati jasno, pozivanjem na neke od opisanih karakteristika hjumovac lakše može odbaciti neke standardne prigovore koji mu se obično upućuju. Prije nego što na konkretnim primjerima pokažem što imam u vidu, ostanimo još malo na ovoj temi i pogledajmo postoje li poveznice između Humeova shvaćanja prirode strasti (želja) i onog kojeg nalazimo kod nekih suvremenih autora.

2.2.2. *O još nekim aspektima želja*

U svojoj recentnoj obrani HTM-e Neil Sinhababu (2009) korisno izdvaja ono što vidi kao središnja obilježja "želje" kao mentalnog stanja:

> Kada primjereno udružena s drugim mentalnim stanjima, želja motivira postupke, uzrokuje iskustva ugode i neugode, usmjeruje pažnju, te može postati divljija (kako bi rekao Hume) putem jasnih reprezentacija koje povezujemo s njezinim predmetom. Također se javlja u […] dva različita okusa – pozitivne želje i averzije" (Sinhababu, 2009: 468).

Na ovom je mjestu odmah potrebno navesti dvije stvari. Kao prvo, obilježja nekog mentalnog stanja moguće je identificirati isključivo na fenomenološkoj razini, a to, u kontekstu ove

diskusije, povlači da protivnici hjumovskog modela slobodno mogu negirati da želje zaista imaju gore navedene aspekte. Međutim, osporavati "zdravorazumsko" tumačenje svakodnevnog iskustva ne znači ništa drugo nego na sebe prebaciti teret dokazivanja. Kao drugo, čini mi se da i sam Hume prepoznaje gotovo sve aspekte želja o kojima je ovdje riječ. Iako je to u nekom smislu samo stvar interpretacije, vjerujem da neki odlomci iz njegove *Rasprave o ljudskoj prirodi* sugeriraju da i on dijeli slično razumijevanje tih mentalnih stanja. Osvrnimo se sada ukratko na citirani odlomak.

(1) Konstitutivni element HTM-e jest da "želja motivira postupke" ili, nešto preciznije, da želje imaju intrinzičnu motivacijsku snagu. Naravno, da bi motivacija rezultirala samim postupkom, potrebno je i odgovarajuće vjerovanje, ali vjerovanja nemaju samostalnu sposobnost motivacije. Već je nekoliko puta rečeno da antihjumovac upravo u ovoj posljednjoj tvrdnji nalazi glavnu slabost modela – on ne osporava (ili ne mora osporavati) da želja ima inherentnu sposobnost motivacije, nego samo osporava da *isključivo* želja ima takvu sposobnost. (2) Kada je riječ o odnosu želja, ugode i neugode, Sinhababu ispravno primjećuje da se ugoda kod subjekta javlja kada postane izvjesno da će njegova želja biti realizirana a da neugoda nastaje kao rezultat njegova uvida da svoju želju neće ostvariti ili da je to ostvarenje malo vjerojatno (usp. Sinhababu, 2009: 469). Taj aspekt, ako nisam u krivu, također prepoznaje i Hume: "Kao što živost ideje [predmeta strasti] pruža zadovoljstvo, tako i njezina izvjesnost sprečava nelagodu…" (*R* 385). Postoji mnogo primjera koji ilustriraju intuitivnost ovog aspekta. Recimo, teško da će se ugodno osjećati netko tko je, zbog fizičke ozljede, spriječen u svom pokušaju da postane profesionalni nogometaš. To je stoga što on spoznaje da sada postoji vrlo mala ili čak nikakva vjerojatnost da više ikada ostvari svoju želju. (3) Želja vrlo često "usmjeruje pažnju", tj. navodi nas da se usredotočimo na stvari koje su u bliskoj vezi sa željenim stanjem stvari. Ovdje je riječ o jednom aspektu želje koji je većina nas sposobna odmah prepoznati – imati želju da Φ ponekad znači potpuno

usmjeriti pozornost na svojstva Φ-a, a odvratiti pozornost od drugih stvari koje nisu u izravnom dodiru s Φ-om. Na ovo obilježje prvi ukazuje Thomas Scanlon: "Osoba ima želju u smislu usmjerenosti pažnje da P ako mu se pomisao na P iznova javlja u pozitivnom svjetlu, naime, ako je njegova pažnja uporno usmjerena na razloge koji govore u prilog P-u" (1998: 37–41). Scanlon navodi primjer gladi i kako želja za hranom može uzrokovati da svu svoju pažnju usmjerimo na njezino zadovoljenje. No možemo primijetiti da slično vrijedi i u slučaju nekih drugih, ne toliko biološki uzrokovanih želja. (4) Već sam naveo što je Hume mislio o razlici između tihih i divljih strasti te kako upravo *živost* slike predmeta naših strasti može proizvesti svakojake učinke. Jedan od tih učinaka jest upravo onaj fenomenološki: što je slika predmeta življa, strast intenzivnije osjećamo. Drugi učinak može, ali i ne mora biti sadržan u motivacijskom aspektu. Radi se o tome da porast u intenzitetu na koji osjećamo određene strasti *često* (ali ne uvijek) prati i porast u motivacijskoj snazi. Podsjetimo se također da na živost slike predmeta mogu utjecati različiti čimbenici: sam predmet može biti neposredno prisutan, ali i mi o tom predmetu možemo maštati ili intenzivnije razmišljati te ga na taj način u određenom smislu približiti. Kako to lijepo sažima Hume: "Što *manje* koraka činimo da bismo došli do predmeta, i što je put *glatkiji*, to se umanjenje živosti manje osjeća … (*R* 365, kurziv u originalu). (5) I na koncu, Hume na jednom mjestu ističe da se izvor motivacijske snage ne nalazi samo u "naklonjenosti" i spremnosti da "prigrlimo" predmet naših želja, već i u averziji da izbjegnemo ono što nam pruža nezadovoljstvo: "Očito je, kad je izgledno da neki objekt izazove u nama bol ili ugodu, da osjetimo posljedičnu emociju *averzije* i naklonjenosti, i poneseni smo da *izbjegnemo* ili prigrlimo ono što nam daje neugodu ili zadovoljstvo" (*R* 354, moj kurziv). No upravo je to još jedan aspekt želje koji navodi Sinhababu. Ponekad možda i ne možemo identificirati pozitivnu želju da nešto učinimo, ali zato možemo identificirati averziju koju imamo ako to *ne učinimo*.

2.3. Poteškoće hjumovskog modela

Hjumovac svoje razumijevanje motivacije ne mora braniti oslanjajući se *isključivo* na kriterij smjera odgovaranja koji – iako možda ne sasvim precizno – ipak objašnjava zašto su upravo želje aktivna motivacijska stanja. Ne ostajući samo na tome, on svoj položaj također može dodatno učvrstiti identificiranjem nekih dodatnih svojstava koja se obično pripisuju želji – svojstava koje ni njegov protivnik ne bi trebao smatrati kontroverznim. Stoga, oboružani intuitivnim razumijevanjem želja, a koje je znatnim dijelom prisutno i u Humeovoj slici, zastupnici HTM-e nalaze se u dobroj startnoj poziciji za suočavanje s nekim prigovorima koji dolaze iz antihjumovskog tabora. U narednom ću dijelu – analizirajući četiri problema za HTM – između ostalog pokušati pokazati zašto i na koji način neki od gornjih zaključaka hjumovcu ponekad mogu biti od pomoći.

2.3.1. *Fenomenološki prigovor*

Bit fenomenološkog prigovora jest da HTM ne uspijeva zahvatiti česte situacije iz svakodnevne ljudske prakse. O čemu se radi? Nije sporno da ljudi nerijetko djeluju na određeni način iako nemaju nikakvu želju tako djelovati, a jedino uvjerljivo objašnjenje tog fenomena jest da se u podlozi ljudskih postupaka, umjesto želja, ponekad nalaze i stanovite dužnosti. Da navedem samo jedan primjer, razlog zašto pomažem čovjeku u nevolji ne mora biti zato što to želim, već zato što vjerujem da bih to trebao učiniti, tj. zato što vjerujem da je to moja moralna dužnost. Ovakvih slučajeva postoji mnogo (ne moraju biti moralne prirode) te ih je suvišno nabrajati. Njihova je poruka, međutim, sasvim nedvosmislena: ljudsko ponašanje nije uvijek moguće objasniti u kategorijama konativnih stanja (što je esencijalna karakteristika HTM-e), već je ponekad plauzibilnije prikloniti se drukčijem tipu objašnjenja – da subjekt u onome što čini vidi neku vrijednost, da ga na takvo postupanje tjera moralna

dužnost, a za koju se ipak ne može reći da je u izravnom dodiru s njegovim afinitetima ili željama. Evo i tri različite formulacije ovog prigovora:

> (a) *Svakodnevno činimo stvari upravo zato što smatramo da ih trebamo činiti;* ni zato što ih želimo činiti, ni zato što imamo prirodnu sklonost da ih činimo, ni zato što smo ih navikli činiti, već upravo zato što smatramo da smo ih dužni učiniti sviđale nam se one ili ne (Berčić, 2009: 33).

> (b) Ne postoji ništa problematično ili misteriozno u tome da ljudi čine stvari koje nemaju želju činiti, stvari koje *uopće* ne *žele* (Schueler, 1995: 29).

> (c) Fenomenologija motivacijskog iskustva favorizira antihjumovsko gledište. Ponekad se želja uopće ne pojavljuje na sceni – 'Vidio sam da je to bilo potrebno učiniti te sam to učinio. Ne sjećam se da sam išta želio u tom trenutku' (Shafer-Landau, 2005: 123).

Zagovornici hjumovske teorije ne misle da su ozbiljnije ugroženi gornjim prigovorom jer ionako vjeruju da je postojanje spomenutih okolnosti u savršenoj harmoniji s njihovom vizijom ljudskog djelovanja. Mislim da je već na osnovi dosadašnje rasprave lako naslutiti u čemu bi se njihova obrana mogla sastojati: oni bi mogli argumentirati da se u postupcima poput navedenih – čak i ako to nije vidljivo na prvi pogled – moraju skrivati nekakve želje jer bez njih ti postupci ne bi bili niti mogući. Mnogi će prema ovoj liniji obrane odmah zauzeti skeptični stav: naime, ako imamo svjedočanstvo samog subjekta da nema želju činiti Φ, zar onda nije krajnje neuvjerljivo ustrajati na tome da on ipak *ima* želju činiti Φ? Insistirati da je on u krivu značilo bi negirati princip privilegiranog pristupa koji je, ako ništa drugo, intuitivno vrlo uvjerljiv i koji predstavlja osnovu za razumijevanje unutarnjeg života drugih ljudi. Jedino sam subjekt, prema tom principu, ima privilegirani pristup svojim mentalnim

stanjima, a to ga čini neprikosnovenim autoritetom kada je riječ o njihovu statusu. U slučaju da on vjeruje da nema želju činiti Φ, onda je to doista znak da ta želja kod njega ne postoji i svaki je pokušaj dokazivanja suprotnog uzaludan i unaprijed osuđen na neuspjeh.

Razmislimo li malo bolje, vrlo brzo ćemo zaključiti da to ipak nije tako te da ne bismo trebali s takvim stupnjem sigurnosti prihvatiti pretpostavku o nepogrešivosti subjekta u pogledu svih psiholoških stanja. Imati privilegirani pristup svojim stanjima ne mora ujedno značiti da u pogledu tih stanja nikada ne možemo pogriješiti. Evo što o tome kaže Neven Sesardić:

> Očito je da se doktrina nepogrešivosti ne može braniti u odnosu na sva mentalna stanja. Ako u mentalna stanja, kao što je uobičajeno, ubrojimo strah, mržnju, ljubav, bijes i slično, rijetko bi se tko upustio da u tom kontekstu brani nepogrešivost subjekta i njegov povlašteni pristup tim stanjima. I psihoanaliza i svakidašnje socijalne situacije naučile su nas uvjerljivo da čovjek često nema najbolji uvid u svoja psihička stanja. Upravo zbog toga razuman čovjek neće propustiti priliku da o sebi nešto više dozna /čak i o svojim unutrašnjim stanjima/ iz odnosa drugih ljudi prema njemu. (Sesardić, 1984: 20)

Prihvatimo li da čovjek može biti u krivu u odnosu na znatan broj svojih psiholoških stanja, zašto smatrati da želje u tom smislu predstavljaju izuzetak? Zašto i njih ne uvrstiti u kategoriju onih stanja u pogledu kojih subjekt također može pogriješiti? Ako za njihovo neuvrštavanje u navedenu skupinu ne postoje adekvatni argumenti (a nisam siguran da postoje), mislim da to baca dodatno svjetlo na fenomenološki argument.

Da pojasnim. Ono što bi hjumovac u obrani od tog argumenta trebao učiniti jest razlikovati *imanje* želje da Φ i *osjećanje* želje da Φ. Gornje formulacije fenomenološkog prigovora tu distinkciju previđaju, a ona sasvim korespondira Humeovoj distinkciji između divljih i tihih strasti: tihe strasti imamo, ali nije tako da ih uvijek osjećamo. Subjekt teško da

može pogriješiti u pogledu toga osjeća li određenu želju, ali činjenica da netko *ne* osjeća određenu želju ponekad nije dovoljan dokaz da ona ne postoji. Kažem "ponekad" jer u mnogim situacijama to ne vrijedi. Primjerice, ja ne osjećam želju da skočim s krova jednostavno zato što tu želju nemam i svaki pokušaj uvjeravanja u suprotno bih u najmanju ruku smatrao čudnim. Postoje, međutim, situacije koje nedvojbeno pokazuju da je negiranje ove distinkcije u suprotnosti s našim svakodnevnim iskustvom. Već je ranije naveden Humeov primjer "dobrote prema djeci" i "ljubavi prema životu". Iako roditelji žele sve najbolje svojoj djeci, malo je vjerojatno da oni tu želju stalno osjećaju i da su u svakom trenutku svjesni njezine prisutnosti; primjerenije je kazati da se ta želja kod njih manifestira u različitim prilikama. Ili, da upotrijebim drugi Humeov primjer, svi ljudi obično žele živjeti, ali ta želja nema trajne fenomenološke karakteristike – tu želju također intenzivnije osjećamo u određenim situacijama, kao što su one u kojima je naš život na neki način ugrožen. Stoga bi se moglo reći da iako je osjećanje neke želje ujedno i evidencija o njezinoj prisutnosti, to ne znači da je neosjećanje neke želje ujedno i evidencija o njezinoj neprisutnosti.[15] A zašto ponekad nismo sposobni identificirati svoje želje moguće je obrazložiti stavom da "u tim slučajevima ne razmišljamo o tome i ne kontempliramo objekt naše želje u trenutku djelovanja; … u tom konkretnom trenutku ne pitamo se ozbiljno želimo li to ili ne želimo i stoga nismo … prema tome aktualizirali naše osjećaje" (Shaw, 1989: 179).

Pokušajmo sada iz obrane preći u napad i razmotrimo je li fenomenološki prigovor zaista tako uvjerljiv kako se to čini na prvi pogled. Je li uopće koherentno tvrditi da ponekad činimo stvari koje "uopće ne želimo", da smo nerijetko motivirani isključivo spoznajom da imamo moralnu dužnost i da se u tim slučajevima "želja uopće ne pojavljuje na sceni"?

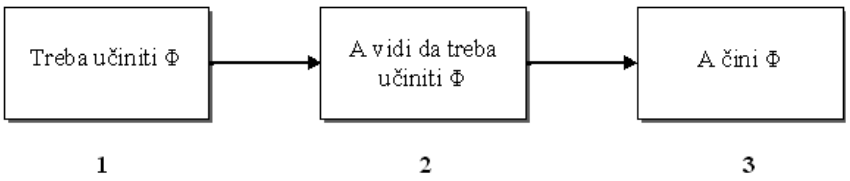
[15] Da zagovornici ovog prigovora doista ne čine spomenutu razliku vidljivo je iz primjera koji nudi Berčić: "… roditelji koji usprkos svim svojim obvezama i umoru tri puta tjedno vode dijete na sport u pravilu to čine zato što smatraju da to trebaju činiti a ne zato što *osjećaju želju* da čine (Berčić, 2009: 34, moj kurziv). Možda je i ne *osjećaju*, ali ne znači da je *nemaju*. Osim toga, primijetimo da nam se želja ne mora uvijek javljati u pozitivnom obliku, nego i u obliku averzije. Roditelji možda imaju averziju prema tome da im dijete *ne ide* tri puta tjedno na sport.

Nastojat ću pokazati da je odgovor na to pitanje niječan. Čini mi se da uvjerenje o postojanju moralne dužnosti nikada ne može imati samostalnu sposobnost motivacije – da sama misao da nešto *trebamo*, naime misao koja nije popraćena odgovarajućim željama, nije u stanju odigrati ulogu značajnog motivacijskog čimbenika. Jednostavno je teško zamisliti okolnosti u kojima je netko motiviran uvjerenjem da ima moralnu obvezu, a da istovremeno nema *baš nikakvu* želju koju bi izvršenjem te obveze mogao ostvariti.

Pretpostavimo da osoba A čini radnju Φ zato jer smatra da treba učiniti Φ, a ne zato jer želi učiniti Φ. Postoje dva objašnjenja zašto A može smatrati da treba učiniti Φ. Kao prvo, A može smatrati da treba učiniti Φ zato jer je činjenje Φ-a način na koji on može ostvariti *neku drugu* želju. No ovaj je opis sasvim sukladan hjumovskom modelu. Pobornik tog modela ne tvrdi da je motivaciju za činjenje neke radnje subjektu ispravno pripisati samo ako on ima neinstrumentalnu želju za činjenjem te konkretne radnje, već samo ako on ima bilo kakvu želju koja je s tom radnjom na neki način povezana. Jedino na čemu hjumovac mora ustrajati jest da nekakva želja u tom smislu mora postojati. Dakle, smatra li A da treba učiniti Φ zato jer time može realizirati neku drugu želju, onda fenomenološki argument nema nikakvu težinu: moguće je tvrditi da upravo ta želja predstavlja glavni izvor njegove motivacije i da A ipak ima želju učiniti Φ, ali želju koja je izvedena iz njegove druge želje. *Trebanje* za koje A smatra da se na njega odnosi stoga ima hipotetičku prirodu, što znači da pritisak koje ono vrši u konačnici proizlazi iz samih A-ovih želja. Nestankom tih želja automatski nestaje i normativni pritisak.

Ipak, fenomenološki prigovor možda i nije tako jednostavno odbaciti, a to je vidljivo ako se osvrnemo na drugo moguće objašnjenje A-ova uvjerenja da treba učiniti Φ. Kada A smatra da treba učiniti Φ, objašnjenje glasi, on to smatra *zato jer treba* učiniti Φ, a ne zato jer on ima neku želju koju bi činjenjem te radnje mogao ostvariti. A, drugim riječima, jednostavno prepoznaje normativni zahtjev koji se na njega bezuvjetno odnosi i stječe

motivaciju djelovati u skladu s tim zahtjevom. Ovdje ne postoji referiranje ni na kakvu želju, bilo izvedenu bilo neizvedenu. Stoga je, netko će reći, samo u duhu ovog objašnjenja ispravno tumačiti fenomenološki argument. Ideja je da ljudi ponekad čine stvari koje nemaju *apsolutno nikakvu* želju činiti, ali ih svejedno čine jer naprosto vide da ih *treba* učiniti. Priroda tog *trebanja* jest kategorička, što znači da je ono sasvim neovisno o našim željama i afinitetima. Međutim, ako je ovo način kako interpretirati gornji prigovor, onda mi se čini da se antihjumovci suočavaju sa sljedećom poteškoćom. Kako bismo tu poteškoću jasnije istaknuli, osvrnimo se na trenutak na sljedeći prikaz za koji vjerujem da u većoj ili manjoj mjeri zahvaća cijeli proces motivacije kako ga vide zagovornici fenomenološkog prigovora:



Slika 2.1. Jedno tumačenje fenomenološkog prigovora

Tvrdnju u (1), budući da se radi o kategoričkim normativnim zahtjevima, interpretiram kao tvrdnju prema kojoj za sve vrijedi da trebaju učiniti Φ ako se nalaze u takvim i takvim okolnostima. No ostavimo po strani (1) te se orijentirajmo na prijelaz između (2) i (3). Jasno je da vidjeti (smatrati) da treba učiniti Φ nije isto što i učiniti Φ. Ja mogu smatrati da treba učiniti Φ a da ipak ne učinim Φ. Hjumovac će zato argumentirati da ako ipak učinim Φ, da sam to učinio zato jer sam to na neki način želio učiniti. No pogledajmo što se događa ako prihvatimo sugestiju da se u ovom slučaju "želja *uopće* ne javlja na sceni". Prema toj interpretaciji, proizlazi da (2) mora *nužno* povlačiti (3). Jer uočimo da A, kako to proizlazi iz moje druge interpretacije fenomenološkog prigovora, nema *baš nikakvu* želju učiniti Φ. On

čini Φ samo zato jer vidi da treba učiniti Φ, ali prava je istina da on to *uopće ne želi*. Ali zar to, potpuno suprotno onome što misli Schueler u svojoj formulaciji argumenta, nije misteriozno? Ispada da normativni zahtjevi, prema tom shvaćanju, u sebi imaju ugrađenu neku moć privlačenja čija je priroda krajnje zagonetna i nedefinirana. Čim subjekt vidi ili prepozna normativni zahtjev (da treba učiniti Φ), taj ga zahtjev nekim misterioznim magnetizmom počinje privlačiti u željenom pravcu. Jer, podsjetimo se još jednom, subjekt prema ovom tumačenju nema nikakvu želju djelovati u skladu s tim normativnim zahtjevom, ali ipak tako djeluje. Kako je to moguće? Nekako izgleda da nije A onaj koji donosi odluku o tome hoće li se pokoriti normativnom zahtjevu, već normativni zahtjev, prema ovoj interpretaciji, pokorava njega. Stječe se dojam da svako djelovanje koje nije podložno hjumovskom tumačenju postaje misteriozno.[16]

2.3.2. *Prudencijalnost, motivirane i nemotivirane želje*

Jedan od argumenata na koji se pozivaju zastupnici HTM-e lijepo je formulirao David McNaughton: "…kada netko djeluje intencionalno, on u nekom smislu mora htjeti ili željeti to što čini" (McNaughton, 2010 [1988]: 106). Argument je vrlo jednostavan: njime se samo tvrdi da intencionalno djelovanje *per definitionem* povlači prisutnost želje ili htijenja kod djelatnika. Jednostavno je proturječno kazati da nemamo nikakvu želju činiti ono što činimo *namjerno*. Zato je, kako sam pokušao pokazati, snaga fenomenološkog prigovora samo prividna. Negiranje prisustva želje u motivacijskom procesu dovodi do stava da je intencionalno djelovanje nemoguće.

No ipak, suprotstavljajući se gornjoj obrani, antihjumovac bi mogao kazati da njegova namjera nije *sasvim* negirati prisutnost želje u motivaciji. Štoviše, želja u nekom minimalnom smislu mora postojati da bi subjekt bio pokrenut na djelovanje, ali sama prisutnost želje, on će

[16] Moj je posljednji odgovor na fenomenološki prigovor jednim dijelom inspiriran Searleovom raspravom o postojanju jaza između naših mentalnih stanja i samog djelovanja. Searle svoje opservacije, pomalo ironično, koristi upravo *protiv* hjumovskog modela. O tome ću govoriti u odsječku o slabosti volje.

tvrditi, nije evidencija da upravo ta želja i vrši esencijalnu motivacijsku funkciju. Ovo je otprilike linija razmišljanja koju razvija Thomas Nagel (1970). Iako priznaje da želja mora biti prisutna kod svakog intencionalnog djelovanja, Nagel argumentira da ta činjenica nije dovoljna za izvođenje zaključka prema kojemu je upravo želja ono *zbog čega* je subjekt i motiviran na to djelovanje. Konkretnije rečeno, iz toga što subjekt čini Φ slijedi da on ima želju činiti Φ (ili želju koja je s tome radnjom nekako povezana), ali ne slijedi (barem ne uvijek) da je on motiviran činiti Φ *zbog* želje koju ima. Očito je da se Nagel dijelom priklanja hjumovcima jer prihvaća njihov stav o nužnoj prisutnosti želje u procesu motivacije, ali isto je tako očito da se istovremeno udaljuje od njihove fundamentalne teze da *samo* želje igraju ulogu centralnog motivacijskog čimbenika. Tu ulogu, Nagel tvrdi, ponekad preuzima i *vjerovanje*. Nagelovim stopama ide i G. F. Schueler (1995) koji također smatra da se hjumovski argument u biti svodi na trivijalno istinitu tvrdnju te da stoga i nema duboki filozofski značaj. U jednom značenju pojma želje, smatra Schueler, konceptualno je istinito da svako intencionalno djelovanje implicira da je djelatnik *htio* učiniti upravo to što je učinio. Ali izraz "htjeti", prema njegovu mišljenju, ima vrlo široko značenje i ne mora nužno označavati određeno mentalno stanje kao što je želja. Schuelerovim riječima: "Ukratko, kada kažemo da je morao htjeti to učiniti budući da je to učinio intencionalno, izraz 'htjeti' kako je ovdje korišten, ako referira na bilo što, referira na *štogod* je vodilo djelatnika da izvrši tu radnju…" (Schueler, 1995: 34).

Nagel je vjerojatno jedan od najutjecajnijih kritičara HTM-e. Njegova kritika za mnoge predstavlja klasično mjesto u ovoj raspravi i teško da danas postoji ozbiljnija obrana hjumovskog stajališta koja ne sadrži barem nekakav odgovor na tu kritiku. Razmotrimo stoga pobliže što Nagel vidi kao glavnu slabost hjumovske teorije. Već je rečeno da on ne niječe prisutnost želje u motivaciji, ali ne negirati prisutnost želje u motivaciji ne znači ujedno i stati na stranu hjumovaca – za Nagela je, kako ćemo sada vidjeti, glavno pitanje na koji je način

određena želja formirana. Hjumovci jednostavno zanemaruju ovo pitanje, a upravo se u njemu krije glavni razlog zbog kojeg mogućnost kognitivne motivacije ne bi trebala biti isključena:

> Tvrdnja da želja leži u podlozi svakog postupka istinita je samo ako želje uključuju motivirane kao i nemotivirane želje te je istinita samo u smislu da *štogod* bila motivacija nečije intencionalne potjere za ciljem, na osnovi njegove potjere postaje *ipso facto* primjereno pripisati mu želju za tim ciljem. Ali ako je želja motivirana, onda će njezino objašnjenje biti isto kao i objašnjenje njegove potjere i, po svemu sudeći, nije očito da želja mora ući u to daljnje objašnjenje. […] Prema tome, može se priznati da je trivijalno da me, primjerice, razmišljanja o mojoj budućoj dobrobiti ili o interesu drugih ne mogu motivirati na djelovanje bez da je želja prisutna u vrijeme postupka. To da imam primjerenu želju jednostavno *slijedi* iz činjenice da me ta razmišljanja motiviraju (Nagel, 1970: 29).

Na početku odlomka Nagel uvodi distinkciju koju hjumovac ne prepoznaje, a koja će se kasnije pokazati esencijalnom za kognitivističko razumijevanje motivacije – distinkciju između motiviranih i nemotiviranih želja.[17] U čemu je razlika? Nemotivirane želje su, baš kao i neka vjerovanja, ona koja nas "jednostavno salijeću". Kao što ponekad formiramo vjerovanje da *p* premda nismo odlučili vjerovati da *p*, tako ponekad i formiramo želju učiniti Φ iako našoj želji da učinimo Φ nije prethodilo nikakvo rasuđivanje. Na primjer, moje vjerovanje da grmi formirao sam baš zbog toga što sam upravo čuo grom, a želju da nešto pojedem formirao sam jer sam upravo dobio "napad gladi". Za razliku od nemotiviranih želja, motiviranim željama ipak prethodi nekakva odluka ili određeni proces rasuđivanja. Pretpostavimo da kao rezultat mog vjerovanja da će me danas posjetiti neki prijatelji formiram želju da u obližnjoj trgovini kupim hranu i piće kojim ću ih ponuditi. Moja je želja u ovom slučaju motivirana, a ne motivirajuća, jer je za nju, baš kao i za sam čin odlaženja u

[17] Za raspravu o ovoj distinkciji također vidi Wallace (1990).

trgovinu, moguće pružiti *jedinstveno* "racionalno ili motivacijsko objašnjenje". Ali koja je onda uloga same želje u tom objašnjavalačkom procesu? Nikakva. Ako moju želju kao i sam postupak mogu objasniti svojim vjerovanjem da će me posjetiti prijatelji, onda se prirodno nameće zaključak da je samo to vjerovanje *dostatno* za motivaciju te da je pozivanje na želju u najmanju ruku suvišno.

Velik dio rasprave Nagel posvećuje prudencijalnoj motivaciji koja za njega predstavlja tipičan primjer kognitivne motivacije. U tom kontekstu formulira primjere kojima hjumovsko objašnjenje prudencijalnosti nastoji staviti pred izazov. Hjumovac, kako ga on ovdje prezentira, prudencijalnu motivaciju objašnjava djelatnikovom *željom* da čini ono što je u njegovu vlastitu interesu, a situacije u kojima je djelatnik motiviran svojim budućim vlastitim interesom objašnjava djelatnikovom trenutnom *željom* da čini ono što će zadovoljiti njegov budući interes (Nagel, 1970: 39). Nagel, kao kantovski usmjeren autor, opisano objašnjenje prudencijalnog djelovanja nalazi potpuno neutemeljenim: postuliranje želje za zadovoljenje budućih interesa krajnje je nepotrebno jer prudencijalnost predstavlja zahtjev racionalnosti, zahtjev kojemu ništa dodatno nije potrebno da bi imao motivirajuću snagu. Racionalna osoba jest ona koja je motivirana samom činjenicom da je nešto u njezinu budućem interesu, a želja za zadovoljenjem budućeg interesa može biti samo izvedena iz tog racionalnog zahtjeva. Ali ako je izvedena, onda je opet riječ o motiviranoj želji, a ne o onoj koja ima intrinzičnu motivirajuću snagu. Pokušajmo preciznije odrediti prirodu Nagelova prigovora. Pretpostavimo da A sutra ujutro odlazi s prijateljima na kampiranje. Budući da zna da će tijekom kampiranja biti gladan, odluči otići u trgovinu (iako trenutno nema nikakvu želju za hranom). Pobornici kognitivističkog modela motivacije pri ruci imaju vrlo elegantno rješenje kada je riječ o slučajevima ovakvog tipa: ono što me motivira nije želja, već moje vjerovanje

o razlogu koji ću imati u budućnosti. Ili, kako bi se to također moglo formulirati, motivira me čisti razum.[18]

Premda Nagelov prigovor ima jaku intuitivnu snagu, hjumovac se sigurno ne nalazi u bezizlaznoj situaciji. Kao prvo, zašto insistirati da su neke od naših želja *motivirane* vjerovanjima, a ne, recimo, radije ostati na tome da vjerovanja ponekad samo *objašnjavaju* nastanak nekih od naših želja? Kao što kaže Michael Smith: "Pretpostavimo da vjerovanja mogu racionalno objasniti želje, ali da, kada to čine, ona ih ne motiviraju. To bi ostavilo hjumovsku teoriju *motivirajućih stanja* netaknutom. Jer želje ipak mogu biti ono što motivira potonje radnje" (Smith, 2004a: 148). Nagel možda dobro primjećuje da postoji izvjesna razlika u tome kako pojedine želje nastaju, ali ipak griješi kada na tu razliku referira kao na razliku između *motiviranih* i *nemotiviranih* želja. Jer hjumovac, kako upozorava Smith, jednostavno ne nalazi uvjerljivom ideju da želja može biti motivirana nekim psihološkim stanjem koje nije želja.

Kao drugo, neće se svi složiti da prudencijalnost, kako to sugerira Nagel, predstavlja *zahtjev racionalnosti*. Shvatimo li prudencijalnost kao ono što bi osobi trebalo biti u interesu, tj. kao ono što je za osobu, generalno promatrajući, dobro, onda je teško shvatiti tvrdnju da bismo *trebali* (u racionalnom smislu) biti motivirani prudencijalno djelovati. Naime, shvatiti prudencijalnost na način na koji to čini Nagel znači tvrditi da prudencijalna motivacija nastaje pod pritiskom razuma te da bismo svakoga tko ne bi imao želju činiti stvari koje su njegovu vlastitom interesu morali smatrati iracionalnim. Nagel je uvjeren da je hjumovsko razumijevanje prudencijalnosti pogrešno i kontraintuitivno, a kao primjer, kako smo vidjeli, navodi slučajeve u kojima subjekt djeluje imajući u vidu ono što je u njegovu *budućem* interesu. Kada su takvi slučajevi u pitanju, hjumovac je, kako to kaže Nagel, obvezan prihvatiti mogućnost da "…u nedostatku … relevantne želje u sadašnjosti, ne moram imati

[18] Primjer modificiram prema Shaw (1989: 169–170). U svom odgovoru na Nagelov prigovor također ću se mjestimično oslanjati na navedeni članak.

razlog pripremiti se za ono što znam da ću imati razlog učiniti sutra" (Nagel, 1970: 40). Međutim, kako on dalje navodi, to je krajnje neprihvatljivo jer "sustav s posljedicama poput ovih ne samo da ne uspijeva zahtijevati najelementarniju konzistenciju u ponašanju kroz vrijeme, nego ustvari izoštrava mogućnosti konflikta ugrađujući pojedinčevu urotu protiv svog budućeg jastva u aparat racionalnosti (Nagel, 1970: 40). Kao što će, vjerujem, postati jasno u trećem poglavlju ovog rada, za hjumovca je Nagelovo gledište sasvim neprihvatljivo – on vjeruje da razum nikada ne može vršiti *pritisak* na želje jer bi time, slobodnije rečeno, prekoračio svoje ovlasti. Promatrajući iz perspektive hjumovske pozicije, prudencijalno djelovanje ovisi samo od toga kakve želje imamo, a ne kakve bismo želje trebali imati: ako subjekt nije motiviran ponašati se prudencijalno (ako mu nedostaje odgovarajuća želja), onda za njega ni ne postoji razlog da se tako ponaša. No ako za njega ne postoji takav razlog, onda ne postoji ni temelj za optužbu iracionalnosti.

Kao treće, je li zaista tako da je Nagelovo objašnjenje prudencijalnog djelovanja elegantnije? Mislim ne samo da je odgovor na to pitanje niječan, već da se to objašnjenje, za razliku od onog hjumovskog, suočava s nekim ozbiljnim poteškoćama. Dok je uloga želje u okviru hjumovske teorije precizno određena, to ipak ne vrijedi u Nagelovu slučaju. On nalazi trivijalno istinitom tezu da želja mora biti prisutna u motivacijskom procesu, a istovremeno argumentira da ta želja ne doprinosi "motivacijskog djelotvornosti" vjerovanja (usp. Nagel, 1970: 30). Zastupajući ovakvo stajalište, on kao da želi pronaći srednji put između HTM-e i nečega što bismo, slijedeći Dancya, mogli nazvati "čistom kognitivnom teorijom" (usp. Dancy, 1993: 18). Međutim, nisam siguran da u tome uspijeva: stječe se dojam da se u jednom trenutku ipak mora suočiti s izborom između ove dvije pozicije.

No pogledajmo za početak nešto detaljnije u čemu se sastoji hjumovsko objašnjenje prudencijalnog djelovanja. Na prvi pogled zaista može izgledati da su prudencijalni oblici postupanja tipičan primjer situacija u kojima smo motivirani "hladnim rasuđivanjem", ali to

ipak nije tako – hjumovac smatra da subjekt ne može biti motiviran svojim budućim interesima ako mu nedostaje "relevantna želja u sadašnjosti". Razlog zašto obično smatramo da nisu želje ono što nas u tim okolnostima pokreće jest taj što te želje uvijek ne osjećamo. Ali opet, to što ih ne osjećamo, ne znači da one ne postoje. Hume na jednom mjestu kaže: "Pričajte čovjeku o njegovom stanju poslije trideset godina od sada, i on se neće ni osvrnuti na vas. Recite mu što će se dogoditi sutra, i on će vam pokloniti pažnju" (*R* 365). Iako je činjenica da je ovo stav koji ljudi obično zauzimaju kada je riječ o njihovim *budućim* interesima, zanimljivo je pokušati pružiti nekakvo obrazloženje te činjenice. Zašto ljudi uglavnom zanemaruju svoj budući interes ako se radi o dalekoj budućnosti, a posvećuju mu izuzetno veliku pažnju kada je ta budućnost relativno bliska? Humeovo je obrazloženje sljedeće. Velika prostorna i vremenska udaljenost slike predmeta naših želja čini blijedima, dok te slike postaju sve jasnije i življe kako se vremenska i prostorna udaljenost smanjuje, a upravo *živost* tih slika na naše želje ima ogroman utjecaj – naime, što su te slike jasnije, želje postaju sve jače i intenzivnije. Prostorna i vremenska blizina, međutim, nisu jedini načini na koji naše želje mogu poprimiti izraženije fenomenološke karakteristike. U tome može pripomoći i imaginacija, pamćenje kao i ozbiljnije razmišljanje o našim radnjama. Ljudi svoju pažnju nerijetko usmjeruju na zadovoljenje trenutnih potreba i zato njihova želja da zaštite i svoj budući interes ostaje zatomljena. Za ponovno prizivanje te želje ponekad može biti dovoljno samo zamisliti neke moguće okolnosti ili buduće situacije koje bi osoba htjela izbjeći.

Hjumovac stoga može argumentirati da subjekt, koristeći upravo spomenutu sposobnost imaginacije te ozbiljnijim razmišljanjem o budućnosti, vrlo lako može osvijestiti neke svoje želje i afinitete koje, uslijed različitih čimbenika, trenutno ne mora uopće osjećati. Pogrešno je ovakvu obranu protumačiti kao *ad hoc* rješenje i smatrati da se ovime kod subjekta nastoji postulirati želju koja zapravo ne egzistira. Upravo suprotno, hjumovac za cilj

ima pokazati da subjekt može postati svjestan svoje želje samo ako bolje promisli o onome što čini i zašto to čini. Želja je *uvijek* prisutna u slučajevima poput ovih i ideja je da razum samo pomaže da tu želju osvijestimo u određenim trenucima. Primijetimo da želje, u okviru ovog shvaćanja, nisu ništa drugo nego dispozicije, što znači da one "imaju fenomenološki sadržaj u mjeri u kojoj je imanje određenih osjećaja jedna od stvari koje one imaju dispoziciju proizvesti pod određenim uvjetima" (Smith, 1994: 114). One poprimaju fenomenološke karakteristike *pod određenim uvjetima*, a sposobnost stvaranja tih uvjeta vrlo često ima i sam subjekt – za njega je dovoljno da upotrijebi vlastitu sposobnost imaginacije ili pak ozbiljnije razmisli o svom postupku da bi se neke njegove želje aktivirale i proizvele "imanje određenih osjećaja". Primijenimo li ovu analizu na naš primjer, dolazimo do sljedećih rezultata. Iako A u trenutku odlaženja u trgovinu ne osjeća nikakvu relevantnu želju, on ipak ima sposobnost tu želju osvijestiti. Dovoljno je da si pokuša predočiti situaciju u budućnosti u kojoj će imati hranu te tu situaciju usporediti s onom u kojoj neće imati hranu. Za očekivati je da će u tom trenutku osjetiti da preferira prvu situaciju, dok će prema drugoj osjećati averziju (usp. Shaw, 1989: 171–172).

No što se događa ako prihvatimo Nagelovu tezu da kod prudencijalnog djelovanja želja nema motivirajuću ulogu? U tom bi slučaju naše djelovanje – slično mom ranijem zaključku – bilo potpuno neobjašnjivo. Nagel, napomenimo još jednom, tvrdi da nas u slučajevima prudencijalne motivacije motivira čisti razum, te da želja – *iako je prisutna* – ne figurira kao središnji motivator. Ali ako želja ne figurira kao središnji motivator, koja je uopće njezina funkcija? Ako je cijeli postupak moguće objasniti pozivanjem na čisti razum, koja je uloga same želje u tom objašnjenju? Uostalom, je li uopće koherentna tvrdnja da smo ponekad motivirani čistim razumom? Razmišljajući o tom pitanju, Shaw zamišlja situaciju kojom Nagelov primjer dovodi do apsurda. Pretpostavimo da ja A potpuno indiferentan prema tome hoće li imati hranu na sutrašnjem kampiranju ili neće. Drugim riječima, kada razmišlja o

tim dvama izgledima, on jednostavno ništa ne osjeća. No pretpostavimo da, unatoč tome – motiviran čistim razumom – on ipak otiđe u trgovinu i kupi sve što mu je potrebno. Takav bi postupak bio misteriozan. Jer ako je A-u apsolutno svejedno, tj. ako nema ni pozitivne ni negativne osjećaje kada razmišlja o tome što će se dogoditi u budućnosti, kako je moguće da ipak čini odgovarajuće korake da tu budućnost dočeka spreman?

2.3.3. Akrasia *i problem slobode volje*

Postojanje *akrasije* predstavlja velik izazov za naše razumijevanje ljudskog djelovanja. Evo i zašto. Sasvim je očito da objašnjenje nečijeg ponašanja mora sadržavati pozivanje na razloge: tek kada smo upoznati s razlozima koji se nalaze u podlozi nekog postupka, taj nam postupak postaje razumljiv. Međutim, kako onda učiniti razumljivim postupak za koji sama osoba smatra da ga nema razloga činiti, ali ga ipak čini? Ovdje se dakle ne radi o tome da *mi* smatramo da subjekt nema razloga činiti to što čini, već je to nešto što *on sam* smatra, ali ipak izabire djelovati suprotno tom uvjerenju. Najkraće rečeno, ono što kod akratičnog djelovanja posebno zbunjuje jest činjenica da su razlozi na osnovi kojih subjekt djeluje oni razlozi za koje je, kada je uzeo sve u obzir, sam prosudio da su slabiji.

Pokušajmo sada još preciznije odrediti prirodu ovog fenomena. Prosuđujem li da je *x* bolje od *y*, onda prosuđujem da imam jače razloge učiniti *x*. Ali ako prosuđujem da imam jače razloge učiniti *x*, onda ću i učiniti *x*. No kako objašnjavamo situaciju u kojoj prosuđujem da imam jače razloge učiniti *x*, ali ipak učinim *y*? Je li takva situacija uopće moguća? Naše nam svakodnevno iskustvo sugerira da je odgovor na to pitanje potvrdan. Pretpostavimo da razmišljam o tome hoću li otići u kino ili ostati kod kuće ponavljati za sutrašnji ispit. Dilema pred kojoj se nalazim je sljedeća. Odlučim li se za kino, znam da kasnije neću imati vremena adekvatno ponoviti gradivo, a ako ipak ostanem kod kuće, znam da neću uspjeti pogledati film koji već dugo želim pogledati (pretpostavimo da se film neće više prikazivati). Nakon što

sam odvagnuo razloge *za* i *protiv*, prosuđujem da je za mene bolje ostati kod kuće i to vrijeme iskoristiti za pripremanje ispita koji me sutradan očekuje. Međutim, usprkos svojoj prosudbi, odlazim u kino. Što se ovdje zapravo dogodilo? Rekli bismo da je moj postupak paradigmatičan slučaj iracionalnog ponašanja. Jer kako drukčije objasniti da činim ono što sam smatram da nemam dovoljno razloga činiti? Poteškoća se, napomenimo još jednom, sastoji u tome da izabirem onu mogućnost za koju sam – razmotrivši sve razloge koji su mi bili dostupni – *sam* prosudio da je lošija. Također je presudno ne izgubiti iz vida da u ovakvim i sličnim situacijama subjekta nitko ne prisiljava da djeluje suprotno svojoj boljoj prosudbi, već on to čini slobodno i namjerno, a upravo ta činjenica i otežava objašnjenje ovakvih postupaka. Kako je moguće da svojevoljno i bez ikakve prisile biram opciju za koju sam ustanovio da je za mene lošija?

Neki su skloni vjerovati da fenomen *akrasije* hjumovci ne mogu na jednostavan način uklopiti u svoju teoriju. Konkretnije rečeno, jedini način na koji se hjumovac može nositi s tom pojavom jest da negira njeno postojanje, a to je, mnogi će smatrati, prilično kontroverzno rješenje. Snaga naših razloga, hjumovac vjeruje, zavisna je o snazi naših želja. Naime, ako S smatra da ima jače razloge učiniti *x* nego *y*, onda za hjumovca to ujedno znači da S jače želi učiniti *x*. Ali ako S jače želi učiniti *x*, onda će, pod jednakim uvjetima, i učiniti *x*. Međutim, ovakva vizija ljudskog djelovanja nije kompatibilna s postojanjem *akrasije* – kada bi hjumovac prihvatio postojanje slabosti volje bio bi obvezan tvrditi da je moguća situacija u kojoj osoba intencionalno i slobodno djeluje suprotno svojoj najjačoj želji. Hjumovac je – s obzirom na njegov stav da djelatnik uvijek čini ono što najjače želi činiti – prisiljen smatrati da u slučajevima *akrasije* osoba ipak griješi kada smatra da ima jače razloge učiniti *x*. Činjenica da on čini *y* ipak je pokazatelj da on ima jače razloge činiti *y*, a samim time da ima i jaču želju činiti *y*. Opisano hjumovsko rješenje problema *akrasije* u većoj ili manjoj mjeri

odgovara onome koje je pružio Donald Davidson (1980b).[19] Zagovarajući kauzalno objašnjenje naših postupaka, Davidson tvrdi da kada prosuđujem da je bolje učiniti *x* nego *y*, ali ipak učinim *y*, da su upravo oni razlozi koje sam prosudio slabijim zapravo bili *uzročno najjači*. Radi se o tome da su razlozi naših postupaka, prema Davidsonovu mišljenju, ujedno i *uzroci* mojih postupaka (usp. Davidson, 1980a). Vraćajući se na moj primjer, rekli bismo da je razlog koji sam prosudio slabijim (da neću uspjeti pogledati film jer se neće više prikazivati) na koncu ipak *uzrokovao* moj odlazak u kino. Iz svega proizlazi da je jednostavno nemoguće da subjekt intencionalno djeluje na osnovi onoga što su za njega slabiji razlozi, što znači da fenomen *akrasije*, kako glasi ovaj prigovor, za hjumovca zapravo nije niti moguć.[20] Ipak, kako to zgodno primjećuje Searle, rješenje koje bi nijekalo postojanje slabosti volje bilo bi sasvim kontraintuitivno jer "… *akrasia* je kod racionalnih bića uobičajena kao vino u Francuskoj. Svatko tko je ikada pokušao prestati pušiti, smršaviti ili manje piti na velikim zabavama znat će o čemu govorim" (Searle, 2001: 10).

U potonjem se pristupu skriva *još jedna* poteškoća za hjumovski model. Ako je točno da postoji kauzalno nužna veza između mojih "psiholoških antecedenta" (želje i vjerovanja) te samog intencionalnog djelovanja, onda ispada da se postupci na neki način odvijaju sami od sebe i da je sam subjekt u cijeloj priči potpuno irelevantan. Ovo je, u glavnim obrisima, primjedba koju Davidsonovoj argumentaciji također upućuje Searle (2001). Jer ako je tako da naša psihološka stanja, osim što objašnjavaju, ujedno i *uzrokuju* naše postupke, to znači da ti postupci nužno moraju slijediti kada za njih postoje kauzalno dostatni uvjeti. Ili, kako to na svoj prepoznatljiv način kaže Searle: "Kada bih stvarno mislio da su vjerovanja i želje dostatni da uzrokuju radnju, tada bih se samo mogao zavaliti i gledati radnju kako se

[19] Suvišno je isticati da je Davidsonovo rješenje u tom smislu mnogo kompleksnije i da gornji opis samo površno ocrtava smjer hjumovskog rješenja. Tako u prikazu recimo uopće ne spominjem distinkciju između *bezuvjetnih* i *uvjetnih* prosudbi koja čini osnovu za razumijevanje Davidsonova pristupa. Davidson vjeruje da akratična osoba nikada ne djeluje suprotno svojoj bezuvjetnoj prosudbi, već samo onoj uvjetnoj. Razlog tome jest taj što je intencionalno djelovanje za njega povezano samo s bezuvjetnim prosudbama. Dakle, za Davidsona je nemoguće intencionalno djelovati u suprotnosti sa svojom bezuvjetnom prosudbom.

[20] McNaughton ([2010] 1988: 120) razmatra ovo hjumovsko rješenje, ali ne izvodi (barem ne eksplicitno) neke od gornjih zaključaka.

odvija…" (Searle, 2001: 71). No ovakav je zaključak, suvišno je i reći, potpuno neprihvatljiv i apsurdan. Slijedi da je praktično rasuđivanje za hjumovce nešto sasvim suvišno: budući da su moja psihološka stanja sasvim dostatna da proizvedu postupak, onda ja o tim psihološkim stanjima ne mogu ni promišljati baš kao što, kada uzmem tabletu za glavobolju, ne moram "dodati neki daljnji psihološki trud na učinak tablete. Vjerujem li zaista da je tableta dovoljna, onda je racionalno da se opustim i pustim da ostvari učinak" (Searle, 2001: 72). U osnovnim crtama, da bi naše rasuđivanje o tome što ćemo učiniti uopće imalo smisla, Searle pokušava pokazati, mi moramo pretpostaviti postojanje jaza između "'uzroka' postupka u obliku vjerovanja i želja te 'učinka' u obliku postupka" (Searle, 2001: 13). Upravo u tom jazu subjekt nalazi svoj prostor za prakticiranje slobode volje, a očito je da zagovornici hjumovske teorije postojanje ovog jaza uopće ne priznaju.

(1) Mislim da kritiku prema kojoj hjumovci negiraju postojanje slabosti volje treba shvatiti s oprezom. Davidson recimo eksplicitno tvrdi da akratične radnje postoje i u tom smislu nastoji iznaći način na koji bi takve radnje pomirio s nekim načelima koji se nalaze u podlozi ljudskog djelovanja (usp. Davidson, 1980b). Potrebno je imati na umu da se Davidsonovo rješenje zasniva na razlikovanju između (a) razloga koje sam subjekt procjenjuje kao najjače i (b) razloga koji uistinu jesu najjači. On ne negira da mi često djelujemo suprotno onim razlozima koje smo *mi sami* prosudili najjačima, a to, na koncu, i jest središnje obilježje akratičnog djelovanja. Jedino što Davidson negira jest da subjekt može djelovati suprotno razlozima koji *zaista jesu* najjači, ali to nije isto što i negiranje fenomena *akrasije*.

(2) Nije teško primijetiti da je Searleova cijela argumentacijska konstrukcija utemeljena u (navodnom) postojanju jaza koji za njega nije ništa drugo nego prostor u kojemu subjekt može prakticirati svoju slobodnu volju. Oslanjanje na ovaj jaz Searleu i omogućuje da na vrlo jednostavan i efektan način dovede u pitanje cijelu hjumovsku tradiciju, a samim time

i hjumovski način tretiranja različitih filozofskih problema. Zato mislim da je vrijedno ukratko zadržati se na ovom ključnom mjestu te razmotriti u kojoj je mjeri istinita njegova tvrdnja da oslanjanjem na kauzalnost te negiranjem spomenutog jaza hjumovac svoju poziciju dovodi do apsurdnih konzekvenci. Fenomenološki promatrajući, nama se čini da nas razlozi na osnovi kojih djelujemo ne prisiljavaju na postupke i da uvijek imamo mogućnost izbora kada je riječ o tome što učiniti. Searle je tu sigurno u pravu. Međutim, nije odmah jasno zašto bi iz toga moralo slijediti da mi *zaista imamo* mogućnost izbora odnosno da moramo pretpostaviti postojanje jaza te da moramo odbaciti ideju prema kojoj su psihološka stanja dostatna da uzrokuju naše djelovanje. Searle će vjerojatno replicirati da tu postoji određena nekonzistencija: ako smo zaista sigurni da su naša psihološka stanja kauzalno dostatna za postupke (kao što smo sigurni da je tableta uzročno dovoljna za smirivanje glavobolje), onda bismo se u skladu s tim uvjerenjem trebali i ponašati (kao što jednostavno čekamo da tableta počne djelovati). Ali problem je u tome, Searle ističe, da se u slučaju naših postupaka jednostavno ne možemo ponašati u skladu s tim uvjerenjem jer je očito da se naši postupci ne mogu odvijati sami od sebe. Potrebno je *odlučiti* djelovati na osnovi želje koju imamo. Ali zašto smatrati da teza o kauzalnoj dostatnosti naših psiholoških stanja nužno mora isključivati moment odluke? Nije samo po sebi razumljivo da su te dvije tvrdnje u kontradikciji. Jer zastupnik hjumovskog modela nije obvezan argumentirati da psihološka stanja moraju uzrokovati naše postupke "direktno, neovisno o rasuđivanju i donošenju odluka" (Mele, 2002: 906). Što to znači? Hjumovac se, s jedne strane, sasvim lako može složiti sa Searleom da se nijedan postupak neće odviti sam od sebe ako subjekt ne *odluči* djelovati na osnovi želje koju ima, ali, s druge strane, može argumentirati da su sve subjektove odluke i izbori već unaprijed određeni i da jaz o kojemu Searle govori zapravo uopće ne postoji. On ne vidi nikakvu tenziju između ove dvije tvrdnje.

(3) Još jedan način na koji bi hjumovac vrlo lako mogao izbjeći Searleovu kritiku jest da pokaže kako HTM uopće nije osuđena na *kauzalno* objašnjenje djelovanja. Upravo je tim putem, ali u nešto drukčijem kontekstu, krenuo Smith (1994: 102–104). Kako on primjećuje, kauzalno shvaćanje razloga nije nešto što je specifično *samo* za hjumovsku teoriju, jer tezu o psihološkim stanjima kao uzrocima može prihvatiti i zagovornik kognitivističkog modela objašnjenja djelovanja, ali on će smatrati, suprotno hjumovcima, da su upravo vjerovanja ona stanja koja mogu imati uzročnu moć. Zato je pitanje mogu li razlozi za djelovanje – koji su konstituirani od psiholoških stanja – figurirati i kao uzroci sasvim neovisno od pitanja trebamo li prihvatiti hjumovski model motivacije. Smith zaključuje da hjumovce ništa ne sprečava da svoju teoriju interpretiraju na teleološki način.

(4) Vratimo se sada problemu slabosti volje. Neki hjumovci (Sinhababu, 2009) argumentiraju da upravo oni na raspolaganju imaju najuvjerljivije objašnjenje tog fenomena. Problem akrasije i jest problem prvenstveno zbog toga što je potrebno objasniti kako je moguće da osoba *zadržava svoju prosudbu* da je za nju bolje učiniti *x* iako intencionalno čini *y*. Prosudivši da je za mene bolje ostati kod kuće i učiti za sutrašnji ispit, ipak sam intencionalno i slobodno otišao u kino. No kada bi me netko naknadno upitao smatram li i dalje da je za mene bilo bolje ostati kod kuće, i dalje bih ostao pri svojoj izvornoj prosudbi. Kako to objasniti? Primjenjujući karakteristike koje želje imaju na pitanje akratičnog djelovanja, Sinhababu navodi da ovdje zapravo dolazi do raskoraka između motivacijske snage neke želje te činjenice da nas želje mogu natjerati da skrenemo pozornost na njihov predmet. Rečeno u kontekstu mog primjera, budući da je želja za odlaskom u kino bila jača od želje da ostanem kod kuće, na koncu sam i otišao u kino. Ali iako moja želja da ostanem kod kuće nije bila dovoljno jaka, ne znači da je time pala u drugi plan – ona me u nekom smislu nastavlja "tjerati" da usmjerim pozornost na njezin predmet, naime na sutrašnji ispit. To je osnovni razlog zašto zadržavam svoju prosudbu o tome što trebam učiniti.

2.3.4. *Eksplanatorna učinkovitost*

David McNaughton je formulirao jedan izuzetno zanimljiv argument protiv HTM-e. On smatra kako ponekad pozivanje na odgovarajuću želju uopće ne pomaže da određeno ponašanje učinimo razumljivim te da kognitivističko objašnjenje tog istog ponašanja u mnogim slučajevima polučuje mnogo bolje rezultate. Naime, HTM, gore je već rečeno, temelji se na tvrdnji da su vjerovanja i želje *odvojena* mentalna stanja, što znači da nema nikakvog proturječja u tome da jedno od tih mentalnih stanja egzistira bez drugoga. Vjerovanje, primjerice, da vani padi kiša ne mora biti popraćeno mojom željom da uzmem kišobran kao što ni moja želja da uzmem kišobran ne mora biti nužno povezana s mojim vjerovanjem da pada kiša. Ne postoji unaprijed određena kombinacija vjerovanja i želja koju bismo u određenoj situaciji trebali imati, a što u konačnici znači da nije isključena mogućnost postojanja vrlo neuobičajenih pa čak i nesuvislih želja. A upravo se u tome, prema McNaughtonovu mišljenju, i skriva osnovna poteškoća.

Čak i ako uvidom u nečije želje i možemo ponekad određeno ponašanje učiniti inteligibilnim, to ne vrijedi u slučajevima kada su upravo želje ono što zahtijeva objašnjenje. Ili, nešto drukčijim riječima, nemamo li objašnjenje želja koje se nalaze u podlozi nečijeg djelovanja, onda pozivanje na same te želje ne može biti od velike pomoći u razumijevanju tog istog djelovanja. Nama će to djelovanje i dalje izgledati potpuno nerazumljivo i neshvatljivo sve dok, kako kaže McNaughton, nemamo objašnjenje načina na koji "djelatnik poima željeno stanje stvari". Upravo je to "poimanje" *kognitivno* stanje, ali ono kognitivno stanje čija se uloga ne može shvatiti samo u instrumentalnom smislu: "Uloga djelatnikova poimanja kakav svijet jest nije ograničena na davanje informacije koja djelatniku omogućuje da bilo koju želju koju slučajno ima kanalizira u pravom smjeru. Djelatnikovo poimanje situacije ima važniju ulogu: da njegovu želju učini suvislom" (McNaughton, 2010 [1988]: 110). Moje poimanje neke situacije nije samo poimanje toga koji je najbolji način da ostvarim

svoje već unaprijed zacrtane želje, već uključuje mnogo više od toga. Tek kada shvatimo kako osoba u pitanju poima ili razumije predmet svoje želje, možemo sasvim shvatiti *i* njezino ponašanje. Bit argumenta jest da je kognitivističko objašnjenje ljudskog ponašanja bolje i potpunije od hjumovskog. Da upotrijebim McNaughtonov primjer, razlog zašto se netko penje po stijeni u nekom minimalnom smislu možda i možemo razumjeti ako mu pripišemo želju da osjeti opasnost i vjerovanje da to može postići penjanjem po stijeni. Međutim, takvo hjumovsko objašnjenje nije dovoljno jer još uvijek ostaje nerazjašnjeno zašto bi uopće netko imao takvu želju. Većina ljudi penjanje po stijeni smatra opasnim pothvatom u kojeg se nikada ne bi upustili i stoga im je neshvatljivo zašto bi itko takvu aktivnost smatrao privlačnom. No sagledamo li tu aktivnost iz perspektive onoga tko iskazuje takvu želju i tako na neki način dobijemo uvid u *njegovo poimanje* te aktivnosti, onda će nam i sam njegov postupak postati puno bliži i jasniji. Isti argument protiv hjumovskog modela pruža i Dancy kada piše: "Neke želje, naravno, ne mogu biti objašnjene. Ali ako ne mogu biti objašnjene, ne može niti postupak koji smo, želeći to što želimo, motivirani učiniti" (Dancy, 2000: 85). McNaughton i Dancy sugeriraju sljedeću mogućnost: ono što objašnjava određeni postupak nije želja, već ono što se nalazi u podlozi te želje odnosno razlozi zbog kojih te želje uopće imamo. Stoga i znanje o tome u kakvom svjetlu djelatnik vidi svoj postupak jest nužan, ali i dovoljan uvjet za potpuno objašnjenje tog istog postupka.

Mislim da opisani argument ima ograničenu snagu jer ne predstavlja opasnost *samo* za hjumovsku teoriju. Pomalo iznenađujuće, argument ne samo da dovodi u pitanje hjumovsku teoriju, već i samu teoriju za koju se Dancy i McNaughton zalažu. Da objasnim. Argument tvrdi da subjektove želje ponekad mogu postati razumljive samo ako zauzmemo njegovo poimanje situacije. U tom smislu, ako subjektovo poimanje situacije njegove želje može nama učiniti razumljivima, onda to poimanje može učiniti razumljivim i *postupak* koji je motiviran tim željama. Pozivanje na same želje apsolutno je nepotrebno za samo objašnjenje postupka.

Međutim, problem leži u sljedećem. Složimo li se s Dancyem i McNaughtonom da postoje želje koje uslijed svoje nesuvislosti i čudnovatosti ne mogu biti objašnjene, zašto ne pretpostaviti da postoje i subjektova *poimanja* određenih situacija koja isto tako ne mogu biti objašnjena zbog svoje nesuvislosti i čudnovatosti? Nadalje, ako neka subjektova poimanja ne mogu biti objašnjena, baš kao što to ne mogu biti i neke želje, kako onda ona mogu poslužiti kao objašnjenje onoga što on čini? Mogli bismo inzistirati na tome da bi sam postupak, bez dodatnog objašnjenja, i dalje ostao nerazjašnjen. McNaughton doduše smatra da subjektovu "poimanju ne treba dodati ništa kako bi se objašnjenje upotpunilo" (McNaughton, 2010[1988]: 111), ali, kako vidimo, to ipak nije točno. Ako potpuno objašnjenje nekog postupka ne može stati, kao što to hjumovci tvrde, na uvidu u želje koje se nalaze u njegovoj podlozi, već zahtijeva i razloge za same želje, zašto onda ne tražiti i objašnjenje tih razloga itd.? Problem beskonačnog regresa teško da se može izbjeći.

Dancyu i McNaughtonu je moguće odgovoriti i na nešto drukčiji način. Oni, ako ne griješim, ne tvrde da je za svaku želju potrebno pružiti objašnjenje, već da je za neke želje potrebno pružiti objašnjenje: HTM uspijeva postupak učiniti inteligibilnim u mnogim situacijama, ali ne u svim. No ako to ne može učiniti u *svim* situacijama, znači li to da taj model treba napustiti? Mislim da ne. Jer pogledajmo na primjer što se događa ako prihvatimo alternativu koju nude Dancy i McNaughton. Za njih razumljivost postupka ovisi samo o tome imamo li uvid u način na koji subjekt poima željeno stanje stvari. Ipak, pretpostavimo li da osoba *ne zna* objasniti svoju želju, onda A-ovo poimanje njegove želje zapravo ne postoji i stoga ne može ni figurirati kao objašnjenje samog postupka. Jedino što nam u tom slučaju preostaje jest prikloniti se hjumovskom modelu objašnjenja ponašanja.

3. Internalizam i eksternalizam motivacije

Moralne prosudbe imaju motivirajuću snagu: prosudba da je neka radnja moralno ispravna za sobom povlači i motivaciju da se ta radnja izvrši. Nepostojanje motivacije kod nekoga da djeluje u skladu sa svojim moralnim uvjerenjima, često ćemo pomisliti, dovoljan je razlog da toj osobi prestanemo pripisivati ta uvjerenja. I ne samo to. Ponekad smo skloni otići i korak dalje te takve ljude etiketirati kao neiskrene i licemjerne. Jer izgleda da i nema metode putem koje bi bilo moguće ustanoviti u kojoj mjeri netko stoji iza svojih vlastitih uvjerenja, a koja pritom ne bi sadržavala obraćanje pozornosti na njegove postupke i ponašanje općenito. Pretpostavimo da tvrdim kako je moralno ispravno učiniti radnju Φ okolnostima C. Zar ne bi bilo opravdano dovesti u pitanje to moje moralno uvjerenje kada bi se pokazalo da nikada nisam imao ni najmanju namjeru učiniti Φ kada sam se nalazio u okolnostima C? Odgovorimo li na to pitanje potvrdno, približavamo se poimanju odnosa između moralnih prosudbi i motivacije koje je prilično rasprostranjeno i koje predstavlja okosnicu pozicije poznate kao *internalizam motivacije* – pozicije prema kojoj je prisutnost odgovarajuće motivacije inherentna moralnim prosudbama.

Jednu od prvih formulacija internalizma nalazimo kod W. D. Falka koji kao središnje obilježje tog stajališta vidi tezu da "moralnosti nije potrebna vanjska sankcija" te da je "sama misao da mi moralno trebamo učiniti neku radnju dostatna, bez referiranja na neki skriveni motiv, da nas opskrbi razlogom da je učinimo …" (Falk, 1952 [1947-48]: 494). Ne pridajući na ovom mjestu značajniju pozornost Falkovoj nejasnoj upotrebi pojmova "razlog" i "motiv", evidentno je da naglasak leži na riječi "dostatno". Falk želi istaknuti da za stjecanje motivacije činiti ono što nam nalaže moralna dužnost nije potrebno ništa drugo izuzev same misli da smo toj dužnosti podvrgnuti. Nadovezujući se na Falkovu diskusiju, William K. Frankena internalizam vidi kao gledište prema kojemu prepoznavanje dužnosti "povlači postojanje

motivacije…" i prema kojemu je "paradoksalno i logički nemoguće" da subjekt prepozna moralnu dužnost a da pritom nema motivaciju tu dužnost ispuniti. Time se internalisti suprotstavljaju eksternalistima koji "inzistiraju da motivacija nije dio analize moralnih sudova" (Frankena, 1997 [1958]: 399–404). Falkova i Frankenina definicija nisu toliko različite. Pitanje o tome uključuje li analiza moralnog suda referiranje na motivaciju ekvivalentno je pitanju je li uopće ispravno nekome pripisati moralni sud ako on nije motiviran na odgovarajući način. Internalisti smatraju da to nije moguće budući da je motivacija već "ugrađena" u moralni sud, a što u konačnici znači da je moralni sud "dostatan" za motivaciju.

Eksternalisti svoju poziciju grade na pretpostavci da moralna prosudba nije dostatna za motivaciju te da je, da ponovno upotrijebim Falkov termin, moralnosti ipak potrebna "vanjska sankcija". Eksternalistička intuicija jest da nema nikakve proturječnosti u tome da kod subjekta izostane motivacija činiti ono što prosuđuje da bi (moralno) trebao, jer je relacija između moralne prosudbe i motivacije kontingentna te je za generiranje adekvatne motivacije potrebno – uz sam moralni sud – imati dodatno psihološko stanje. Budući da je eksternalistička pozicija rijetko kada formulirana u pozitivnom obliku, tj. budući da se ona ponajviše razvija u opreci s internalizmom, priroda dodatnog psihološkog stanja o kojemu je riječ nije uvijek jasno precizirana. U literaturi oko tog pitanja ne postoji konsenzus tako da različiti autori koji naginju eksternalizmu to stanje razumiju na različite načine: neki ga definiranju kao *želju za moralnim ponašanjem*, neki kao *strah od mogućih sankcija*, a neki kao *simpatiju* koja je zajednička većini ljudi. Zašto je subjekt motiviran djelovati u skladu s prosudbom da je Φ moralno ispravno (da treba učiniti Φ) dakle nije zato što ga sama ta prosudba uspijeva motivirati na činjenje Φ-a, već zato što se boji društvene osude, ili zato što ima želju činiti ono što je moralno ispravno, ili pak zato što mu stanje simpatije koje ima prema drugima omogućuje da njegove želje ionako uvijek budu usklađene s onime što mu

nalaže dužnost.[21] Eksternalistička strategija ponajviše se svodi na osmišljavanje različitih protuprimjera čiji je cilj pokazati da je sasvim koherentno ostati potpuno nepokrenut moralnim sudovima koje izričemo. Internalistima, na koje je time (navodno) prebačen teret dokazivanja, tako preostaje pokazati zašto ti primjeri nisu uvjerljivi i zašto njima internalizam ipak nije ozbiljnije uzdrman. Cilj ovog poglavlja jest argumentirati da se internalist uspješno može nositi s navedenim poteškoćama. No to se, kao što ću tvrditi, odnosi samo na nekognitivističkog (hjumovskog) internalista (vidi tablicu 3.1.).

3.1. Kognitivizam, nekognitivizam i motivacija

Iako je problem moralne motivacije, strogo govoreći, nezavisan od pitanja imaju li moralne prosudbe kognitivni ili nekognitivni sadržaj, nije slučajno da upravo nekognitivisti u metaetici najčešće zastupaju internalizam.[22] Razlog tome prvenstveno treba tražiti u hjumovskom modelu objašnjenja ponašanja – modelu koji sam nastojao obraniti u prethodnom poglavlju i koji za mnoge još uvijek predstavlja "standardnu sliku ljudske psihologije" (Smith, 1994: 7). Hjumovski model, podsjetimo se, tvrdi da su upravo želje ona mentalna stanja koja su intrinzično motivirajuća, dok su vjerovanja pasivna i motivacijski inertna. Oslonimo li se na to shvaćanje u kontekstu rasprave između kognitivista i nekognitivista, dolazimo do sljedećih rezultata. Kognitivisti drže da su moralne prosudbe nalik vjerovanjima (mentalnim stanjima koja reprezentiraju objektivno stanje stvari i koja, uslijed toga, mogu biti istinita ili lažna), a takvo ih gledište stavlja u nezgodan položaj žele li istovremeno prihvatiti i internalističku ideju prema kojoj su moralne prosudbe nužno povezane s motivacijom. Budući da su za

[21] Eksternalisti koji dodatno konativno stanje definiraju kao želju za moralnim ponašanjem su Prichard (1968 [1928]: 225), Ross (2002 [1930]: 164), Svavarsdóttir (1999: 170, 2006: 166) te Zangwill (2007: 244). O sankcijama kao pravim motivatorima moralnog djelovanja govori Mill (2003 [1861]: 74), dok eksternalizam u vidu postojanja simpatije kao psihološke činjenice zagovara Brink (1996: 49).

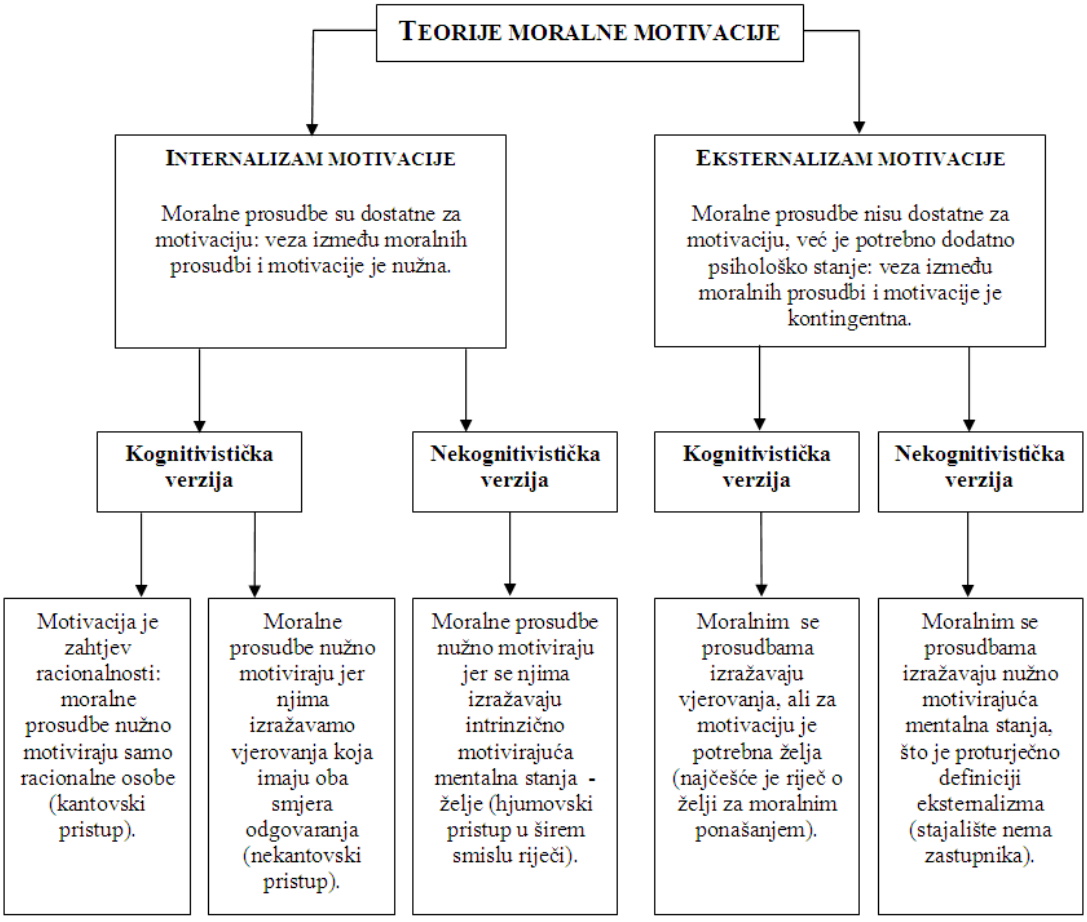
[22] Nekognitivističku verziju internalizma u različitim interpretacijama brane Stevenson (1952 [1937], Hare (2000 [1963]) i Blackburn (2000). Unatoč tome što su nekognitivisti najčešće internalisti u pogledu motivacije, možda treba naglasiti da je u okviru nekognitivističkog pristupa također moguće braniti i eksternalizam. Međutim, takva pozicija nije uobičajena jer bi nekognitivist u tom slučaju morao pokazati da moralnim sudovima izražavamo nekognitivna stanja koja nisu intrinzično motivirajuća. Na neplauzibilnost nekognitivističkog eksternalizma ukazuje Zangwill (2009).

kognitivista moralne prosudbe vjerovanja, a vjerovanja nemaju samostalnu motivirajuću snagu (kako to kaže hjumovski model), slijedi da moralne prosudbe nisu dostatne za motivaciju. Nekognitivist se, s druge strane, nalazi u znatno boljoj poziciji želi li smjestiti internalizam u okvir svoje teorije. Osnovu nekognitivističkog pristupa čini teza da se moralnim sudovima izražavaju nekognitivna mentalna stanja – stanja koja jednostavno ne mogu postojati bez da motiviraju. Iz tog shvaćanja internalistička misao proizlazi kao prirodna posljedica: nije moguće izreći moralni sud a istovremeno ne biti motiviran.[23]

Krenemo li od pretpostavke da internalizam pruža plauzibilno objašnjenje moralne motivacije, moglo bi se čak tvrditi da to predstavlja neovisan argument za prihvaćanje nekognitivizma. Naime, internalizam mnogi nalaze intuitivno vrlo privlačnim jer to stajalište velikim dijelom zahvaća ono što u svakodnevnom životu ionako mislimo o moralu i moralnom ponašanju, a nekognitivizam tu intuiciju uspješno uspijeva zadržati. Na koncu konca, argumentaciju koja ide u tom pravcu razvija i sam Hume kada piše sljedeće: "Moralnost izaziva strasti i proizvodi ili sprječava radnje. Razum je, sam po sebi, potpuno nemoćan u toj pojedinosti. Prema tome, pravila etike nisu zaključci našeg razuma" (*R* 391). Prva premisa Humeova argumenta nije ništa drugo nego nešto drukčija formulacija internalističkog stava, dok je u drugoj premisi sadržana bazična misao hjumovske teorije motivacije. Konkluzija da "pravila etike nisu zaključci našeg razuma" slobodno može biti protumačena kao teza da moralnost nema kognitivnu osnovu. Način na koji će kognitivist pristupiti izloženom argumentu zavisi od toga koji su njegovi ciljevi. On *ili* može negirati drugu premisu i braniti mogućnost postojanja kognitivne motivacije (tvrditi da vjerovanja imaju motivacijsku moć) te tako svoju poziciju učiniti kompatibilnom s internalizmom *ili* se pak prikloniti eksternalizmu – poziciji prema kojoj je za motivaciju potrebno još neko

[23] U suvremenoj literaturi nekognitivizam je nerijetko definiran kao pozicija prema kojoj se moralnim sudovima izražavaju *želje*. O tome, kao i o bliskoj vezi nekognitivizma i motivacije, također vidi Dancy (1993: 3), Smith (2004a: 363), McNaughton (2010: 31) i Miller (2003: 3–7).

dodatno psihološko stanje izuzev same moralne prosudbe.[24] U narednom ću se dijelu prvo orijentirati na kritiku druge kognitivističke strategije. Karakteristika te strategije jest negiranje intuitivnosti internalizma i dokazivanje tvrdnje da eksternalizam adekvatnije objašnjava prirodu moralne motivacije. U posljednjem ću dijelu ukazati na neke poteškoće kognitivističkog (kantovskog) internalizma. (Kritici pozicije koju sam u tablici 3.1. označio kao ne-kantovski internalizam već sam posvetio određeni dio prošlog poglavlja.)



Tablica 3.1. Teorije moralne motivacije

[24] Ovim dvjema strategijama nisu iscrpljene sve mogućnosti. Kognitivist također može argumentirati da njegova pozicija uopće nije u neskladu s internalizmom i hjumovskom teorijom motivacije. Najpoznatiji zagovornik tog pristupa je Smith (1994). Također treba spomenuti da zagovaranje eksternalizma ne znači uvijek i zagovaranje hjumovskog modela. Verziju kognitivističkog internalizma koja isključuje hjumovski model zastupaju McDowell (1998) i Nagel (1970).

3.2. Varijabilna motivacija

Ima li eksternalist pravo kada smatra da svakodnevna moralna praksa uopće ne podržava internalističke intuicije? Je li moguće da veza između moralnih sudova i motivacije zapravo nije tako bliska kako nam se obično čini? Nick Zangwill vjeruje da je odgovor na ta pitanja potvrdan i da, pomalo ironično, fenomenologija moralne motivacije više favorizira kognitivizam u odnosu na nekognitivizam. Kao glavnu slabost internalistâ Zangwill vidi njihovu nemogućnost da objasne ono što on naziva fenomenom "varijabilne motivacije": internalistu nekognitivne orijentacije nije jednostavno pružiti uvjerljivo objašnjenje činjenice da "isti moralni sudovi pokreću različite ljude na različite načine i [da] oni različito pokreću istu osobu u različitim okolnostima", kao što, primjerice, "misao da je moja moralna i profesionalna dužnost ocijeniti [studentske] radove na mene ima različiti motivacijski učinak prije i poslije ispijanja moje jutarnje kave" (Zangwill, 2009: 416). Fenomen varijabilne motivacije, kako napominje Zangwill, govori u prilog kognitivističkom eksternalizmu zato jer je taj fenomen najbolje objasniti ukazivanjem na različitu snagu želja koje su eksternalne moralnim sudovima. Pojednostavljeno rečeno, kognitivistički eksternalist vjeruje da motivacija ne proizlazi iz samih moralnih sudova, već iz dodatnih psiholoških stanja (najčešće želja), a to mu omogućuje da različite motivacijske učinke moralnih sudova o kojima govori Zangwill pripiše različitoj snazi samih tih želja (jer nema razloga pretpostaviti da te želje kod svih imaju jednaku snagu ili pak da kod iste osobe imaju jednaku snagu neovisno o okolnostima). Slično objašnjenje nije dostupno internalistu. Kada bi motivacija bila inherentna moralnim prosudbama kao što to vjeruje internalist, tada bi moralne prosudbe nužno motivirale, ali s različitim motivacijskim učincima. Postojanje varijabilne motivacije u tom bi slučaju ostalo nerazjašnjeno. Srž Zangwillova argumenta jest da je, suprotno općeprihvaćenom uvjerenju, potpuno pogrešno smatrati internalizam intuitivno privlačnim. Intuitivnost te pozicije ugrožena je fenomenom varijabilne motivacije – fenomenom koji ipak

najbolje objašnjava eksternalist. Samim time je, naravno, oslabljen i nekognitivizam: jer ako moralni sudovi "nisu intrinzično motivirajući, tada se čini da oni moraju biti kognitivni" (Zangwill, 2009: 417).

Koje objašnjenje varijabilne motivacije internalist (nekognitivist) može ponuditi da bi se obranio od Zangwillova prigovora? Mislim da ovako postavljeno pitanje navodi na krivi put i da se internalist uopće ne mora boriti na tom terenu. Naime, da bi se internalist uopće upustio u pokušaj objašnjenja fenomena varijabilne motivacije, on bi trebao imati mnogo jasniju sliku o tome što varijabilna motivacija uopće jest, a čini mi se da Zangwillova argumentacija u tom smislu ne pruža dovoljno. To ćemo najbolje registrirati osvrnemo li se ponovno na eksternalističko objašnjenje kojemu se Zangwill priklanja. On napominje kako je eksternalist u boljoj poziciji od internalista jer može tvrditi kako "varijacija u motivacijskoj snazi proizlazi iz varirajućih snaga određenih želja" (Zangwill, 2009: 417). Ovdje mi nije namjera negirati tvrdnju da snaga motivacije korespondira snazi želje, ali mi je namjera izraziti skepsu prema jednom tumačenju te tvrdnje koje, kako ćemo vidjeti, Zangwillu omogućuje da u prvom redu i postavi problem varijabilne motivacije. Tumačenje koje imam u vidu zasnovano je na, po mom mišljenju, sasvim pogrešnoj pretpostavci da snaga motivacije (a time i želje) ovisi o tome kojim intenzitetom želju *osjećamo*. Zangwillovu opservaciju da moralni sudovi na njega vrše "različiti motivacijski učinak prije i poslije ispijanja jutarnje kave" moguće je čitati upravo kao dokaz njegova prihvaćanja navedenog fenomenološkog kriterija. Ako nisam u krivu, u tom je primjeru sadržano uvjerenje kako jedino sam subjekt može biti konačni arbitar u procjeni snage svojih motivacijskih stanja. Jer kako on zna da moralni sudovi na njega vrše različiti motivacijski učinak? Jedini odgovor koji se nameće jest: zato što *osjeća* da mu jačina želje za činjenjem onoga što treba činiti varira ovisno o okolnostima. Ali ako on *osjeća* da mu je motivacija (želja) jača poslije ispijanja jutarnje kave, onda mu je, Zangwill zaključuje, motivacija (želja) zaista jača poslije ispijanja jutarnje kave.

Posljednja je pretpostavka kontroverzna. Prisjetimo li se Humeove distinkcije između "tihih" i "divljih" strasti koju sam analizirao u prošlom poglavlju, odmah postaje jasno zašto bismo sa zadrškom trebali pristupiti tezi da uvijek postoji korelacija između intenziteta kojim neku želju osjećamo i motivacijske snage te želje. Za ilustraciju ću se poslužiti primjerom prudencijalne motivacije ili, da upotrijebim Humeove riječi, primjerom "ljubavi prema životu". S obzirom da mi je u trenutku dok ovo tipkam pažnja usmjerena na neke druge stvari a ne na moju dobrobit kao takvu, nije čudno da želju za životom *ne osjećam* nekim izrazito jakim intenzitetom. No je li iz toga opravdano zaključiti da mi je u ovom trenutku želja (motivacija) za životom slabija nego inače? Pokazuje li taj primjer da moja motivacija za životom varira ovisno o okolnostima? Teško mi je prihvatiti da je odgovor na to pitanje potvrdan. Gotovo je sigurno da bi fenomenološke karakteristike moje želje za životom postale izraženije u okolnostima u kojima bi mi, recimo, život bio ugrožen, ali sasvim je pogrešno to upotrijebiti kao dokaz za tezu da bi mi u tim okolnostima sama želja za životom bila jača. Do istog ćemo rezultata doći promotrimo li primjer koji se tiče odnosa moralnog suda i motivacije. Pretpostavimo da prosuđujem da je moralno ispravno govoriti istinu u okolnostima C. Internalist vjeruje da sam u tom slučaju nužno motiviran govoriti istinu u okolnostima C. No znači li to da će moja motivacija biti najjača kada se zaista nađem u okolnostima C? Ne vidim zašto bismo to trebali pretpostaviti. Sasvim je moguće da ću želju za govorenjem istine u tim okolnostima osjećati jače nego uobičajeno, ali iz toga ne slijedi nikakav zaključak o samoj snazi te želje. Nije pokazano zašto ne bismo mogli smatrati da je moja motivacija u ovom slučaju podjednako jaka sve dok zadržavam svoju moralnu prosudbu. Slabost Zangwillova argumenta leži u njegovu mišljenju da snaga želja uvijek korespondira njihovim fenomenološkim obilježjima.[25]

[25] Garrard i McNaughton, čini mi se, dolaze do vrlo sličnog zaključka koji ovdje nastojim obraniti iako se pri tome ne služe Humeovom distinkcijom između "divljih" i "tihih" strasti. Razmišljajući o pitanju što motivacija jest, oni navode kako djelatnikov izvještaj o tome da nije motiviran na neko djelovanje zapravo nema veliki značaj "osim ako ne pretpostavimo da je biti motiviran u osnovi fenomenološko stanje oko kojeg djelatnik ima

3.3. Protuprimjeri

U prethodnom sam odsječku pokušao pokazati da fenomen "varijabilne motivacije" ne predstavlja ozbiljniju prijetnju nekognitivističkoj verziji internalizma, a u narednom ću se dijelu orijentirati na tipične prigovore koji se upućuju internalističkom kampu. Ti prigovori najčešće poprimaju oblik protuprimjera za koje eksternalisti vjeruju da nedvojbeno dokazuju kako je veza između moralne prosudbe i motivacije raskidiva i kontingentna. Kada je riječ o protuprimjerima internalizmu, ne treba izgubiti iz vida sljedeću činjenicu. Protuprimjer internalizmu nije situacija u kojoj netko ima duboka i sofisticirana moralna uvjerenja, ali se u skladu s tim uvjerenjima ne ponaša. Jer internalizam tvrdi da postoji nužna veza između moralne prosudbe i *motivacije*, a ne da postoji nužna veza između moralne prosudbe i samog *postupka*. Čini se plauzibilnim pretpostaviti da subjekt može biti motiviran učiniti Φ, ali da ipak ne učini Φ. Mjestimično se pozivajući i na ovu distinkciju, tvrdit ću da su eksternalisti u krivu i da njihovi protuprimjeri ne uspijevaju ostvariti zadani cilj.

3.3.1. *Vojnik*

Russ Shafer-Landau pretpostavlja sljedeću situaciju:

> Zamislite vojnika za vrijeme ratnog vremena koji vjeruje da je njegova dužnost boriti se za svoju zemlju. On zna da će u periodu od dva tjedna biti pozvan na front. Kako dani prolaze, on razmišlja o strahotama s kojima se može suočiti [...] Poslije nekoliko dana on ustanovi da ima potpuno odsustvo motivacije za borbom. [...] Kada bi prigrlio pacifizam ili se barem moralno protivio trenutnom ratu, tada bi mirne savjesti mogao opravdati svoju bezvoljnost za borbom. Ali on to ne može učiniti. On sam sebe smatra kukavicom baš zbog nedostatka motivacije da

neosporan autoritet", a što je, kako dalje pišu, "jako neplauzibilna tvrdnja" (Garrard i McNaughton, 1998: 49–50).

učini svoju zapaženu dužnost. Iz svoje vlastite perspektive, on je potpuno nepokrenut onime što moralno treba učiniti. (Shafer-Landau, 2003: 150)[26]

Pogledajmo koje su opcije internalistu otvorene u njegovu pokušaju da se obrani od Shafer-Landauova protuprimjera. On bi mogao inzistirati na tome da vojnikovo moralno uvjerenje nije *iskreno* moralno uvjerenje. Internalist, prema ovoj liniji razmišljanja, nema razloga negirati da ponekad dolazi do raskida između moralnog uvjerenja i motivacije, ali ne dozvoljava mogućnost raskida između iskrenog moralnog uvjerenja i motivacije. Da je vojnik iskreno prihvatio svoju dužnost boriti se u ratu, tada bi bio nužno motiviran to i učiniti. Budući da nije motiviran, slijedi da tu dužnost nije iskreno prihvatio. Nisam siguran da bi u ovom slučaju priklanjanje opisanoj strategiji bilo najbolji izbor za internalista. Kao prvo, netko bi mogao prigovoriti da bi se na taj način internalizam samo pretpostavljao, ali ne i dokazivao, a, kao drugo, mislim da je sam primjer konstruiran tako da unaprijed blokira mogućnost ovakvog izlaza.[27] Naime, problem za vojnika nastaje *baš zato* što on zadržava svoje moralno uvjerenje i zato što ne želi prihvatiti drukčiji moralni svjetonazor. Stoga mi se čini da bi se internalist trebao poslužiti drukčijom strategijom – argumentirati da nedostatak motivacije kod vojnika nije dokazan. Mnogima će na prvi pogled i ta opcija zvučati neuvjerljivo. Jer zašto ustrajati na tome da je vojnik ipak motiviran kada sve upućuje na to da on ustvari nema nikakvu motivaciju? Uostalom, zar nije tako da on sam priznaje da nije motiviran i da sam sebe smatra kukavicom baš zbog toga što je izgubio motivaciju?

Mislim da se internalistički odgovor krije u samim pitanjima. Zaključci o fenomenu varijabilne motivacije sugeriraju da ne bismo bez zadrške trebali prihvatiti tezu da zadnju riječ u pogledu svog motivacijskog statusa ima sam subjekt. Nije očito da postojanje motivacije uvijek korelira s fenomenološkim karakteristikama određene želje i zato činjenica da je vojnik

[26] Shafer-Landauov primjer pretpostavlja kognitivističko čitanje jer govori o vojniku koji *vjeruje* da ima dužnost. Primjer bi bio metaetički neutralan kada bi Landau govorio o vojnikovu *uvjerenju* da ima dužnost.
[27] Na prigovor cirkularnosti ću se detaljnije osvrnuti nešto kasnije.

utvrdio kako nije motiviran ne povlači da on zaista nije motiviran. Sasvim je moguće da on trenutno ne osjeća nikakvu želju za odlaskom na front, ali to nije dovoljno za zaključak da ta želja ne postoji odnosno da on nije motiviran to učiniti. Evo dodatnog razloga zašto bismo u ovom slučaju potonje obrazloženje zaista trebali i prihvatiti. Shafer-Landau tvrdi da vojnik "sam sebe smatra kukavicom baš zbog nedostatka motivacije", što je pomalo zbunjujuće. Sama činjenica da je izgubio motivaciju (ako je to uopće činjenica) sigurno ne objašnjava zašto on sam sebe smatra *kukavicom*, već objašnjenje toga, po svemu sudeći, mora ležati u nečem drugom – njegovu strahu za vlastiti život. Dakle, precizniji opis cijelog slučaja bi glasio da on sam sebe smatra kukavicom, ne zato što nije motiviran ići u rat, već zato što se boji za vlastiti život. Ovo nije zanemarivo jer se time ukazuje da zaključak o odsustvu motivacije još uvijek nije legitimno donijeti.

Shafer-Landau bi mogao replicirati da situacija time nije bitno promijenjena jer ono što *objašnjava* zašto vojniku nedostaje relevantna motivacija jest činjenica da on ima drugu i jaču motivaciju – motivaciju da sačuva vlastiti život. Budući da je upravo ta motivacija *potkopala* njegovu motivaciju za borbom, odgovor bi vjerojatno glasio, sasvim je u redu kazati da se vojnik smatra kukavicom zbog nedostatka motivacije. Ne vidim, međutim, zašto bismo se morali prikloniti takvom opisu situacije? Zašto je u ovom slučaju razumnije pretpostaviti da je jedna motivacija *potkopala* drugu, a ne da je jedna motivacija samo *nadjačala* drugu? Uočimo da internalist nema nikakvih poteškoća u prihvaćanju drugog objašnjenja. U tom bi slučaju motivacija za odlaskom u rat kod vojnika i dalje bila prisutna, ali ne bi bila kauzalno djelotvorna, dok bi prvo objašnjenje impliciralo da je vojnikova ne-moralna motivacija toliko jaka da je njegova moralna motivacija jednostavno prestala postojati.[28] Za Shafer-Landaua u nekom smislu ovaj spor uopće ne postoji jer, kako on kaže,

[28] I sam Shafer-Landau ukratko razmatra slučajeve u kojima postoji motivacijski *defeater*. U slučaju kada je motivacija nadjačana, on govori o *pro tanto* motivaciji, dok u slučaju potkopavanja govori o *prima facie* motivaciji. Njegov primjer vojnika zamišljen je kao primjer u kojemu je motivacija *prima facie* (Shafer-Landau, 2003: 148).

vojnik "priznaje odsustvo motivacije" (Shafer-Landau, 2003: 152). No ako vojnikovo priznanje ovdje ima presudnu ulogu, onda je svaka daljnja diskusija bespredmetna i internalist odmah može priznati poraz. Naravno, internalist to neće učiniti jer odbija pretpostavku da su nečiji introspektivni izvještaji u svakom trenutku pouzdan indikator prisutnosti motivacije te svoju tezu, između ostalog, gradi na uvjerenju da subjekt nije nepogrešiv u pogledu prirode svojih mentalnih stanja. Kako onda utvrditi tko je u pravu? Kako utvrditi je li vojnikova motivacija za odlaskom u rat *potkopana* ili, s čime bi se internalist sasvim lako mogao složiti, samo *nadjačana* njegovom željom za preživljavanjem?

Čini mi se da bi se jedan od načina mogao sastojati u primjeni kontrafaktičkog testa koji, u nešto drukčijem kontekstu, formulira Parfit: "Mi možemo biti motivirani nešto učiniti bez da smo pokrenuti to učiniti. Ali, da bismo bili motivirani, mora biti istinito da bismo to i učinili kada bi nam bila dana prilika i kada bi suprotne i konkurirajuće motivacije bile odsutne" (Parfit, 1997: 100). Distinkcijom između "biti motiviran" i "biti pokrenut" Parfit sugerira da određene motivacije koje imamo ne moraju biti dovoljno jake da nas pokrenu na djelovanje; moguće je da subjekt istovremeno ima različite motivacije za različitim radnjama iako neće biti pokrenut svim tim motivacijama, već samo onom najjačom. Internalist, kako sam ga ovdje prikazao, pretpostavlja da se slično događa i u Shafer-Landauovu primjeru – vojnik je ustvari motiviran ići na front, ali nije pokrenut to činiti jer mu je ta motivacija nadjačana konkurirajućom motivacijom. No kako to dokazati? Prijedlog je sljedeći. Želimo li utvrditi da je subjekt motiviran učiniti neku radnju, dovoljno je upitati se bi li on učinio tu radnju kada kod njega ne bi postojala konkurirajuća motivacija. Primijenimo li taj test na naš slučaj, čini se da prisutnost motivacije više nije tako jednostavno negirati: u prilog tome da je vojnik motiviran govori činjenica da bi se on pridružio ratnim zbivanjima kada se ne bi u tolikoj mjeri bojao za vlastiti život.

Naravno, internalist u ovom slučaju odnosi pobjedu samo pod uvjetom da Parfitova kontrafaktička analiza zaista može poslužiti kao mjerodavan test za prisutnost motivacije, a to, kako je moguće naslutiti, ipak nije nešto što će svi bezrezervno prihvatiti. Shafer-Landau bi se možda i složio da bi vojnik otišao na front u odsustvu konkurirajuće motivacije, ali bi vjerojatno i dalje tvrdio da iz toga ne slijedi ništa o vojnikovu sadašnju stanju: činjenica da bi on učinio Φ kada se ne bi bojao ne znači da je sada, kada se boji, motiviran učiniti Φ. To ne slijedi, on bi mogao tvrditi, zato jer Parfitova analiza govori samo o nužnim, ali ne i dovoljnim uvjetima za prisutnost motivacije. Međutim, kada bi Shafer-Landau svoju poziciju branio na spomenuti način, mislim da bi teret dokazivanja bio na njemu: on bi u tom slučaju trebao pokazati koji bi to još uvjeti trebali biti ispunjeni da bismo subjektu pripisali motivaciju za djelovanje te dodatno pojasniti zašto vjeruje da je u slučaju vojnika smislenije govoriti o *prima facie,* radije nego *pro tanto* motivaciji.

Osvrnut ću se sada na jedan prigovor koji Parfitovoj kontrafaktičkoj analizi u nešto drukčijem kontekstu upućuju Garrard i McNaughton (1998: 50). Oni navode dva primjera u kojima je subjekt motiviran učiniti Φ, ali za koga nije istinito kazati da bi on učinio Φ kada ne bi imao suprotnu i konkurirajuću motivaciju. Ako bi se ti primjeri pokazali uspješnima, slijedilo bi da činjenje Φ-a u kontrafaktičkoj situaciji nije ni nužan uvjet za prisutnost motivacije. U prvom je primjeru S motiviran imati seksualni odnos sa suprugom svog najboljeg prijatelja, ali za njega nije istinito kazati da bi – u slučaju da nije riječ o supruzi njegova najboljeg prijatelja – s njom zaista i imao seksualni odnos: razlog tome jest taj što njega privlače samo žene koje su udane za njegove prijatelje. U drugom je primjeru S motiviran kupiti jaknu, ali to ipak ne učini jer je jakna jako skupa. No opet nije istinito kazati da bi on tu jaknu kupio kada ne bi bila skupa, jer ono što ga na toj jakni, između ostalog, i privlači jest činjenica da toliko košta.

Nijedan od navedenih primjera, po mom mišljenju, ne dovodi u opasnost Parfitovu analizu. Najjednostavnijim riječima, shvatimo li konkurirajuću motivaciju kao motivaciju koja, uslijed svoje snage, subjektu onemogućuje da učini neke druge stvari za koje je također motiviran, onda nije jasno u čemu se ta konkurirajuća motivacija uopće sastoji kada je riječ o situacijama koje prezentiraju Garrard i McNaughton. Da bih to pojasnio, osvrnut ću se na drugi primjer. Ono što S-a sprečava da kupi jaknu jest činjenica da je ona skupa. No, strogo promatrajući, to nije točno. To bi značilo konkurirajuću motivaciju shvatiti kao određenu *činjenicu* o vanjskom svijetu, a ne kao određeno *psihološko stanje*.[29] Jasnije rečeno, puka činjenica da je jakna skupa ne može predstavljati suprotnu i konkurirajuću motivaciju ako osoba prema toj činjenici ne zauzme određeni stav. Dakle, objašnjenje zašto S ipak ne kupuje jaknu jest to da *ne želi* potrošiti puno novaca (psihološko stanje), a ne to da je jakna skupa (činjenica). Zato je potrebno upitati se bi li S tu jaknu kupio kada bi to psihološko stanje bilo odsutno (kada, recimo, želja da uštedi novac kod njega ne bi postojala), a ne, kako to pretpostavljaju Garrard i McNaughton, bi li S tu jaknu kupio kada bi određena činjenica bila odsutna (kada jakna ne bi bila skupa). Jer ako su kontrafaktičke okolnosti one u kojima je odsutna konkurirajuća motivacija te ako je konkurirajuća motivacija psihološko stanje, onda su kontrafaktičke okolnosti one u kojima je odsutno psihološko stanje. No ono po čemu je ovaj slučaj specifičan, netko bi mogao reći, jest to da je S motiviran kupiti jaknu *baš zato* što je ona skupa te u odsustvu te činjenice S jaknu ne bi kupio. No opet je teško razumjeti zašto bi to dovodilo u pitanje predloženu analizu. Prema toj analizi, ono što moramo učiniti jest zamisliti okolnosti u kojoj S-a ništa ne onemogućuje da učini radnju koju je navodno motiviran učiniti te utvrditi bi li S*, u takvim okolnostima*, tu radnju i učinio. Ali ako to ne znači zamisliti okolnosti u kojima kod S-a nedostaje određeno psihološko stanje (želja da uštedi novac i tome slično), onda nije jasno koju bismo drugu konkurirajuću motivaciju

[29] Ova distinkcija možda jasnije dolazi do izražaja kada je riječ o prvom primjeru.

trebali zamisliti kao odsutnu. Pojednostavljeno rečeno, ako odbacimo ideju da S-a u realizaciji njegove želje onemogućuje njegova druga i, po svemu sudeći, jača želja, što ga onda onemogućuje? Sigurno ne činjenica da jakna nije skupa, jer ona to jest. Imajući to na umu, složimo li se da je S-ova konkurirajuća motivacija psihološko stanje, onda je kontrafaktička situacija ona u kojoj upravo to psihološko stanje nedostaje (a ne neka činjenica), a ako S-ova konkurirajuća motivacija nije psihološko stanje, onda proizlazi da S konkurirajuću motivaciju zapravo uopće nema. Primjeri koje nude Garrard i McNaughton ne uspijevaju uzdrmati Parfitov kontrafaktički test.

3.3.2. *Depresivac*

Primjer depresivca, mnogi vjeruju, ozbiljno potkopava internalističke ambicije. Tako već ranije spomenuti Dancy otvoreno priznaje da se eksternalist, kada je riječ o ovom primjeru, nalazi "u jačoj poziciji" (Dancy, 1993: 6). Ideja je da depresija može izvršiti jako negativan učinak na nečiji motivacijski sustav u smislu da subjekta njegova moralna uvjerenja više ne uspijevaju motivirati ni u najmanjoj mjeri. Ne radi se o tome da osoba koja pati od depresije gubi svoja moralna uvjerenja, nego o tome da je ona sada prema njima sasvim indiferentna. Drugim riječima, ona ne mijenja svoje poglede o tome što je moralno ispravno, a što moralno pogrešno odnosno o tome što bi trebala, a što ne bi trebala, već središnji problem nastaje u tome što su ta uvjerenja, gledajući iz perspektive njihova praktičnog utjecaja, za nju sada potpuno beznačajna. Alfred Mele nudi sljedeći primjer:

> Razmotrimo nesretnu osobu … koja pati od kliničke depresije zbog nedavnih tragičnih smrti njezina muža i djece u zrakoplovnoj nesreći. *Naizgled* možemo zamisliti da ona zadržava neka od svojih vjerovanja da ima moralnu obvezu učiniti određene stvari … dok je sasvim lišena motivacije da shodno tome djeluje. … Ona je godinama pomagala svom bolesnom ujaku,

> vjerujući da ima moralnu obvezu to činiti. Moguće je da i dalje to vjeruje ali sada je potpuno nemotivirana da mu pomaže. (Mele, 2003: 111)

Ako je Meleov scenarij plauzibilan, onda postojanje depresivne osobe zaista ugrožava internaliste. Jer za opovrgavanje internalizma dovoljno je skicirati samo jednu situaciju u kojoj dolazi do raskida između moralnih sudova i motivacije, a ostaje dojam da je taj uvjet zadovoljen gornjim protuprimjerom. Ipak, nastojat ću pokazati da je riječ samo o dojmu i da je internalistu i ovaj put nekoliko izlaza na raspolaganju. No prije toga je vrijedno primijetiti da Meleov misaoni eksperiment, slično kao i u prethodnom slučaju, pretpostavlja kognitivističko čitanje internalizma: depresivna osoba, Mele eksplicitno naglašava, zadržava svoja *vjerovanja* o tome što bi bilo ispravno učiniti, a ne, što bi vjerojatno bilo metaetički neutralnije, *uvjerenja* ili *prosudbe* o tome što bi bilo ispravno učiniti. Imajući u vidu navedeni opis, ispada da nekognitivistički internalist svoju poziciju od Meleova primjera uopće ne mora braniti, jer on ionako ne smatra da se moralnim sudovima izražavaju vjerovanja.[30] Međutim, gornji je primjer sasvim lako protumačiti na način da obje internalističke interpretacije budu ugrožene. Naime, kada bi jedan od učinaka depresije zaista bilo stvaranje jaza između moralnih prosudbi (bez obzira na pitanje jesu li te prosudbe vjerovanja ili želje) i odgovarajuće motivacije, tada bi to zasigurno bacalo i negativno svjetlo na kombinaciju nekognitivizma i internalizma. U tom bi slučaju bilo potrebno ili napustiti stav prema kojemu se moralnim sudovima izražavaju neka nekognitivna stanja ili se, s druge strane, prikloniti eksternalističkom shvaćanju moralne motivacije.

Postoji nekoliko načina kako da se internalist suprotstavi Meleovu protuprimjeru, a neki od njih bi nam već trebali biti dobro poznati. Kao prvo, već je ranije dovedena u sumnju

[30] Podsjetimo se da internalist tog tipa smatra kako moralnim prosudbama izražavamo neka nekognitivna stanja koja su, sama po sebi, već motivacijska. Zato nije moguće nemati motivaciju učiniti Φ, a istovremeno zadržati moralnu prosudbu da je ispravno (da treba) učiniti Φ. Zanimljivo je da i sam Mele to izričito spominje te piše kako "nekognitivistički internalizam nije meta [njegove] kritike…" (Mele, 2003: 112; također vidi 132, f.12).

pretpostavka prema kojoj uvijek imamo nepogrešiv pristup svojim mentalnim stanjima. Nadovezujući se na tu sumnju, moglo bi se argumentirati da depresivac sigurno nije mjerodavan u pogledu procjene svog motivacijskog statusa. Ili, kako to kaže Björnsson: "…introspekcija očito nije nepogrešivi vodič … za prisutnost stanja koja upravljaju našim ponašanjem i razmišljanjem, pogotovo ne pod abnormalnim psihološkim uvjetima. Stoga, čak i ako [depresivac] ne misli da je u određenom motivacijskom stanju, on to može biti" (Björnsson, 2002: 336). Kao drugo, prisutnost motivacije kod depresivne osobe mogla bi se dokazivati na intuitivnoj razini. Prisjetimo li se samo da je, prema glavnim obilježjima hjumovskog modela, upravo želja ono intrinzično motivirajuće mentalno stanje, razumno je pretpostaviti da je osoba prije svoje depresivne faze bila motivirana pomagati svom bolesnom ujaku zato jer je to na neki način željela. No znači li to da sada, kada pati od teške depresije, zapravo ne želi više pomagati svom bolesnom ujaku? Vjerujem da nije sasvim očito kako je odgovor na to pitanje potvrdan. Teško je osporiti da je njezina želja, fenomenološki promatrajući, uslijed nastanka depresije sasvim utihnula, ali zašto inzistirati na mnogo jačem zaključku prema kojemu ta želja više uopće ne postoji? Plauzibilnije je pretpostaviti da je ona još uvijek prisutna, ali da je, slično kao i u prethodnom primjeru, jednostavno nadjačana učincima depresije. Kao treće, to što ona zadržava svoja moralna uvjerenja može biti pokazatelj da njezine vrijednosti ostaju netaknute, a ako su, kako to zgodno kaže Blackurn, "…vrijednosti stvarno tamo, očekivat ćemo da se manifestiraju negdje drugdje: u žaljenju ili kajanju, ili u krivnji ili sramu" (Blackburn, 2000: 61). Primjenjujući tu misao na naš primjer, za očekivati je da bi ona osjećala nešto od navedenoga kada bi, uslijed njezine neangažiranosti, bolest njezina ujaka uzela maha. Složit ćemo se da bi to bio ipak nekakav znak njezine brige prema ujaku. Jer zašto bi netko osjećao žaljenje ili sram zbog nečega što nije učinio ako mu je sasvim svejedno što to nije učinio? Pitanje je, naravno, smisleno samo pod pretpostavkom da bi ona u tom slučaju zaista osjećala žaljenje ili sram, a nije

samoočigledno da bismo tu pretpostavku trebali prihvatiti. Možda i nije, ali Mele bi nam za odbacivanje te pretpostavke ipak trebao pružiti adekvatne razloge. I na koncu, zanimljiva je Meleova opservacija da njegov primjer nije zasnovan na empirijskim postavkama u smislu da depresija u nekim slučajevima zaista ugrožava vezu između moralnih prosudbi i motivacije. Mele ne odbacuje mogućnost da su depresivne osobe u stvarnosti uvijek motivirane svojim moralnim prosudbama, ali, čak i ako je to tako, on smatra, to još uvijek ne isključuje ideju da "stvarna ljudska bića nisu *u biti* imuna na depresiju…" kakva je opisana u njegovu scenariju (Mele, 2003: 123). Iako može biti tako da depresija *de facto* ne funkcionira na način kako je to pretpostavljeno u njegovu primjeru, Mele tvrdi, to nije u suprotnosti s idejom da je moguće da depresija kod ljudi ipak ostvaruje takav učinak. Mele je ovdje sigurno u pravu. Jer podsjetimo se da eksternalisti svoju poziciju ne moraju braniti ukazivanjem na stvarne situacije u kojima dolazi do raskida između moralnih uvjerenja i motivacije, već im je u tu svrhu sasvim dovoljno utvrditi da je takav raskid logički moguć. Uostalom, empirijska evidencija ovdje može ići na štetu samo internalistima: kada bi se empirijski utvrdilo da depresija dokida motivaciju, opovrgavanje Meleova primjera bilo bi sasvim uzaludno i svaka daljna filozofska rasprava o tom pitanju bila bi besmislena. No ipak, izlazeći izvan okvira onoga što se zaista događa u slučaju depresije i pozivajući se na puku mogućnost da depresija dokine nečiju motivaciju, Mele više pretpostavlja eksternalističku poziciju nego što je potkrepljuje nezavisnim argumentima. Konkretnije rečeno, argumentirajući da je moguća situacija u kojoj subjekt nije motiviran svojim moralnim uvjerenjima, eksternalist samo ponavlja svoju osnovnu tezu koja i jest glavni predmet spora – naime, tezu da je veza između morala i motivacije kontingentna. To je, između ostalog, i jedna od poteškoća amoralističkog izazova kojemu ću se sada posvetiti.

3.3.3. *Amoralist*

U eksternalističkoj galeriji likova za koje se vjeruje da predstavljaju prijetnju internalizmu lik amoralista nedvojbeno je najpoznatiji. Za razliku od prethodnih primjera, amoralist nije zamišljen kao osoba kod koje motivacija izostaje u jednoj konkretnoj situaciji i kod koje postoji suprotstavljena i konkurirajuća motivacija, već je riječ o osobi čija je nemotiviranost vlastitim moralnim uvjerenjima jedno od njezinih središnjih karakternih obilježja. Amoralist ima vrlo čvrsta uvjerenja o tome što je moralno ispravno a što moralno pogrešno, ali takva ga uvjerenja jednostavno ne uspijevaju motivirati na djelovanje. David Brink vjeruje da se poteškoća za internalizam sastoji u tome što "ne shvaća amoralistički izazov dovoljno ozbiljno" te kasnije nadodaje da je "takva osoba moguća" i da "moramo pokušati objasniti zašto bi amoralista trebalo biti briga za moral" (Brink, 1996: 47–48). Iz Brinkovih je riječi vidljivo da se "amoralistički izazov" ne svodi na tvrdnju da amoralist *de facto* postoji, već da je moguće da postoji. Eksternalist je zato uvjeren da mogućnost postojanja amoralista njegovog protivnika stavlja u vrlo nezgodan položaj: internalist bi trebao objasniti zašto je takva mogućnost isključena ali na način koji ne bi pretpostavljao njegovu poziciju. Drugim riječima, internalist ne može negirati postojanje amoralista navodeći da je veza između moralne prosudbe i motivacije nužna jer bi time učinio logičku pogrešku *petitio principii*. Međutim, mnogi internalisti odbijaju pretpostavku da je na njima teret dokazivanja. Jer ako bi internalist trebao argumentirati nezavisno od svog stajališta da amoralist nije moguć, zašto onda eksternalist isto tako ne bi trebao argumentirati nezavisno od svog stajališta da amoralist jest moguć? Zar intuicija da postoji osoba koja nikada nije motivirana svojim moralnim uvjerenjima nije fundirana u prethodno prihvaćenoj pretpostavci da je veza između moralnih uvjerenja i motivacije kontingentne prirode? Premda nije sklon internalizmu, slično razmišljanje dijeli i Jonathan Dancy: "Ali teško da može biti predstavljeno kao činjenica ili da postoje [amoralisti] ili da bi mogli postojati. Da je takva osoba moguća sigurno je više

pretpostavka eksternalističke pozicije nego nezavisna potpora za nju" (Dancy, 1993: 5). Internalist, moglo bi se kazati, ne shvaća amoralistički izazov dovoljno ozbiljno zato jer amoralistički izazov za njega uopće nije – izazov. S obzirom da obje strane u sporu imaju sasvim suprotne intuicije o amoralističkom prigovoru, proizlazi da se rasprava o ovom pitanju uopće ne može pokrenuti s mrtve točke. Zato ću, prije nego što se upustim u nešto dublju analizu ovog eksternalističkog protuprimjera, na ovom mjestu pokušati ocijeniti jesu li eksternalisti ipak u pravu kada smatraju da teret dokazivanja u ovom slučaju nije jednak i da je internalist onaj koji mora braniti svoju poziciju. Pokušaj argumentiranja u tom pravcu nalazimo kod Sigrun Svavarsdóttir (1999).

Svavarsdóttir prezentira slučaj osobe (Patrik) koja se slaže sa svojom sugovornicom u moralnoj procjeni da bi bilo ispravno pomoći politički progonjenom neznancu, ali koja, za razliku od svoje sugovornice, ne pokazuje ni najmanju sklonost to učiniti. Iako pružanje pomoći neznancu za Patrika ni u kojem pogledu nije riskantno te od njega zahtijeva samo minimalnu dozu napora, on jednostavno priznaje da se ne želi opterećivati moralnim pitanjima i da problemi takve vrste na njega ne vrše nikakav učinak. Virginia, Patrikova sugovornica, prikazana je kao njegova sušta suprotnost: u nastojanju da pokrene Patrika na djelovanje, ona se poziva na "eksplicitne moralne tvrdnje: u ovom slučaju, pomaganje neznancu jest njegova moralna dužnost i stvar suzbijanja goleme nepravde" (Svavarsdóttir, 1999: 176). Unatoč njezinim naporima, Patrik i dalje ostaje pri svom izvornom stavu te izvještava da ga nije moguće dovesti u stanje u kojemu bi ga bilo briga za moralna pitanja. No činjenicu da Patrika nije briga za moralna pitanja također ne treba pogrešno shvatiti. Time on samo daje do znanja da moralni propisi nemaju *praktični* utjecaj u njegovu svakodnevnom životu, ali to ne govori ništa o njegovu razumijevanju samih tih pitanja. Štoviše, dodatne informacije o Patriku sugeriraju da on vrlo kompetentno i inteligentno može sudjelovati u

raspravama o moralno osjetljivim temama i da je, kada je potrebno, uvijek sposoban izraziti vlastito mišljenje u pogledu moralnog statusa neke radnje ili osobe.

Svavarsdóttir je vrlo oprezna da cijeli slučaj predstavi na način koji bi bio zadovoljavajuć kako za internaliste tako i za eksternaliste – fokus je na Patrikovu opažljivom ponašanju, a ne na njegovim mentalnim stanjima. Sve što je o Patriku rečeno vidljivo je bilo kojem vanjskom promatraču, a takvim opisom Svavarsdóttir već u samom startu blokira moguću optužbu da njezin scenarij pretpostavlja neku od sukobljenih pozicija. Pitanje na koje sada preostaje odgovoriti jest, kojim ćemo se hipotezama prikloniti pri pokušaju objašnjenja Patrikova ponašanja? Svavarsdóttir zaključuje da će eksternalist prihvatiti sve predložene hipoteze, dok će internalist uvesti ograničenje – kao potencijalno objašnjenje Patrikova ponašanja internalist sigurno neće prihvatiti ono prema kojemu Patrik nije motiviran moralnim sudom koji izriče. Naravno, potonje je objašnjenje za eksternalista ujedno i najprihvatljivije, ali za Svavarsdóttir to nije od presudne važnosti. Internalističko odbacivanje jedne od hipoteza u Patrikovu slučaju ona navodi u svrhu pozivanja na sljedeći metodološki princip: "…kada postoji konflikt intuicija (među inteligentnim i razumnim ljudima) u pogledu toga koje su hipoteze u opticaju kao objašnjenje nekog opažljivog fenomena, teret dokazivanja je na onima koji inzistiraju na ograničenijem skupu objašnjenja" (Svavarsdóttir, 1999: 179).

Nisam siguran da je spomenuti metodološki princip prihvatljivo upotrijebiti protiv internalista. Zašto je Svavarsdóttir uvjerena da će internalist na predložene hipoteze uvesti ograničenje, tj. da će smatrati kako eksternalističko objašnjenje Patrikova ponašanja nije u opticaju? Odgovor je sljedeći: zato jer je, gledano iz internalističke perspektive, "konceptualno nekoherentno" ne biti motiviran moralnim sudovima koje izričemo. Međutim, još uvijek ne vidim zašto bi teza o konceptualnoj nekoherentnosti ujedno obvezivala internalista da izbaci eksternalističko objašnjenje *iz opticaja*. Takvim bi pristupom, kao što se

ranije dalo naslutiti, cijela diskusija postala trivijalna i filozofski potpuno nezanimljiva. Drugim riječima, internalist zaista smatra da je postojanje amoralista konceptualno nekoherentno, ali on još uvijek može biti u krivu; on za svoju tezu tek treba pružiti adekvatne argumente i baš zbog toga rasprava o moralnoj motivaciji ima smisla. Nema proturječja u tome da internalist prvo prihvati eksternalističko objašnjenje kao jedno od objašnjenja koja su u opticaju, a tek onda pokuša pokazati kako to objašnjenje ipak nije *najbolje objašnjenje* Patrikova ponašanja.

Pogledajmo sada u čemu bi se taj internalistički pokušaj mogao sastojati? Koju bi on strategiju mogao primijeniti pri pokušaju diskvalificiranja eksternalističke hipoteze? Zanimljive opservacije u pogledu tog pitanja nudi Robert Johnson (2001).[31] Johnson primjećuje da su, suprotno Svavarsdóttirinim namjerama, u njezinu primjeru zapravo sadržane internalističke intuicije, a to je vidljivo usmjeri li se pažnja na Patrikovu sugovornicu, Virginiju. Podsjetimo se da Virginia pokušava učiniti da Patrik ipak počne brinuti o sudbini politički progonjenog neznanca, a metoda koju u tu svrhu koristi jest pozivanje na "eksplicitne moralne tvrdnje". No njezino pozivanje na "eksplicitne moralne tvrdnje" u nastojanju da preobrati Patrika sugerira da ona sama vjeruje da su moralne tvrdnje dovoljne za stvaranje motivacije – njezino je ponašanje, bez te pretpostavke, sasvim neobjašnjivo. Jer ako je vjerovati Patrikovu izvještaju da se "davno riješio bilo kakve težnje da živi u skladu s moralnim standardima", zašto se i dalje pozivati na te moralne standarde u namjeri da ga se pokrene na djelovanje? Zašto Patrika i dalje uvjeravati naglašavanjem da je "pomaganje neznancu njegova moralna dužnost i stvar suzbijanja goleme nepravde"? Patrik eksplicitno

[31] Johnson također kritizira Svavarsdóttirino shvaćanje metodološkog principa, ali na nešto drukčiji način. Kako on kaže: "U empirijskom istraživanju, hipoteze nisu evaluirane s obzirom na njihovu sposobnost da objasne pojedinačnu pojavu. One su evaluirane s obzirom na njihovu sposobnost da objasne *obrazac* pojava. Ako se pokazalo da jedna hipoteza pristaje obrascu bolje nego alternativne [hipoteze], izrada jedinstvenog primjera *ne bi trebala* ponovno uvesti alternativnu hipotezu kao kandidata" (2001). Dakle, Johnsonova je ideja da postojanje jednog izoliranog slučaja još uvijek nije dovoljan razlog da napustimo određeni eksplanatorni model. Ograničavanje hipoteza koje su u opticaju za internalista dakle može biti sasvim opravdano pod uvjetom da se objašnjenje koje on sam nudi najbolje uklapa u uzorak. Patrikov slučaj jest slučaj koji se doduše ne uklapa u uzorak, ali to ne znači da zbog toga treba uvesti alternativno objašnjenje i napustiti internalizam.

navodi da moralna razmatranja na njega ne vrše nikakav utjecaj, a Virginiju to ipak ne sputava da se njima nastavi koristiti kao svojim glavnim oružjem. Ne preostaje nijedno drugo objašnjenje njezina ponašanja nego ono internalističko: ona očito mora smatrati da Patrik nije dovoljno dobro uvidio značaj tih moralnih tvrdnji jer bi u suprotnom bio motiviran nešto poduzeti.[32] Ako je to točno, onda to u nekom smislu implicira da moralni sudovi u sebi zaista imaju "ugrađen" motivacijski element te da oni koji tim sudovima nisu motivirani zapravo ne predstavljaju nikakvu prijetnju internalizmu: najjednostavnije rečeno, oni ne izriču *moralne sudove*. Posljednja nas opaska dovodi do vjerojatno najpopularnije internalističke strategije u obrani protiv eksternalističkih izazova – strategije čija je svrha pokazati da amoralist (ali ne samo on) moralne sudove ne upotrebljava na pravi način, već u "navodnicima".

3.3.4. *Sotona ili o standardnoj upotrebi moralnih sudova*

Funkcija moralnog jezika za nekognitivista ne sastoji se u opisivanju, nego u evaluiranju.[33] Iako je moralne izraze (a time i sudove u kojemu se ti izrazi pojavljuju) moguće koristiti u deskriptivnom značenju, evaluativno značenje u većini je slučajeva primarno, a time je definiran i način njihove normalne upotrebe – onaj koji ne koristi moralne izraze u evaluativnom značenju te izraze ne koristi na uobičajen način. Upravo tu dijagnozu internalist pripisuje amoralistu: kada amoralist prosuđuje da je nešto moralno ispravno, on ustvari samo opisuje kakav je vrijednosni stav drugih ljudi prema toj stvari, ali sam se s tim vrijednosnim stavom ne poistovjećuje. Upotrebljavajući terminologiju Richarda Harea (1998 [1952]: 140), za amoralista bi se stoga moglo kazati da moralne sudove upotrebljava u "navodnicima": nije tako da on nije motiviran onim što smatra ispravnim, već onim što smatra "ispravnim". S

[32] Svavarsdóttir (2001) pokušava ponuditi odgovor na Johnsonov prigovor, ali ostaje dojam da to nije lako učiniti. To ilustrira i njezin komentar da je Johnson nije uvjerio, ali da "prihvaća izazov".

[33] Budući da pitanje primarne funkcije moralnog jezika, strogo govoreći, spada u domenu moralne semantike a ne moralne epistemologije, možda bi, umjesto o *kognitivistima* i *nekognitivistima*, bilo ispravnije govoriti o *deskriptivistima* i *nedeskriptivistima*. U našem kontekstu, međutim, to nije presudno. Uostalom, u samim počecima metaetičkog istraživanja – kada su upravo semantička pitanja bila u središtu interesa – termini kao što su kognitivizam i deskriptivizam koristili su se naizmjenično.

druge strane, također postoji mogućnost da netko upotrebljava vrijednosne sudove, kako to naziva Blackburn, iz *inercije*: "Ako smo, iz nekog razloga, izgubili naš okus za vino i više ne uživamo u njegovu kušanju, i dalje bismo mogli nazivati neko vino dobrim ako znamo da je jednom zaslužilo tu etiketu" (Blackburn, 2000: 60). Iako nije uvijek jednostavno razlučiti u kojemu od ova dva navedena načina netko upotrebljava vrijednosne termine, važno je primijetiti da je u njima jasno istaknuta misao da vrijednosti neće imati praktični utjecaj u onim situacijama u kojima subjekt u odnosu na te vrijednosti drži stanoviti odmak. Primjenjujući ove opservacije na debatu o moralnoj motivaciji internalist svoj položaj znatno učvršćuje. On dozvoljava mogućnost pojavljivanja raskida između prividno vrijednosnih sudova i motivacije i sam jasno specificira okolnosti u kojima se to *de facto* događa. Ali naravno, postojanje takvih specifičnih okolnosti ni na koji način ne oštećuje njegovu poziciju jer još uvijek nije pokazano da je taj raskid moguć kada se vrijednosne sudove upotrebljava u njihovu standardnom značenju.

Opisana strategija internalistu služi i u borbi s još jednom potencijalno opasnom figurom iz eksternalističkog tabora – Sotonom. Njegova pojava predstavlja opasnost za internalizam prvenstveno zbog toga što Sotona ima uvid u ono što je dobro i ispravno, ali je motiviran djelovati u sasvim suprotnom pravcu. Zato je pozivanje na različitost u načinima na koje se moralni sudovi mogu upotrebljavati za ovaj slučaj iznimno značajno. Kako kaže McNaughton: "…nekognitivist sada može pojasniti naizgled paradoksalne primjedbe poput one koju Milton pripisuje Sotoni: 'Zlo, budi ti moje dobro.' Termin 'zlo' ne koristi se evaluativno, već samo radi opisivanja one vrste djelovanja koju većina ljudi smatra zlom" (McNaughton, 2010 [1988]: 131). Nekognitivističko rješenje mnogima neće zvučati zadovoljavajuće, a o tome svjedoči i McNaughtonova opaska nekoliko redaka dalje da je nekognitivist time "došao do jeftine pobjede". Tom primjedbom on želi istaknuti da nekognitivistički internalist blokira svaku daljnju raspravu jer svojim razumijevanjem

moralnog diskursa svoj položaj čini neopovrgljivim.[34] Meni se čini da McNaughton ipak nije u pravu. Naime, za očekivati je da će i eksternalist priznati kako postoje okolnosti poput navedenih u kojima se moralni sudovi upotrebljavaju s izvjesnom distancom. Drugim riječima, eksternalist ne bi trebao negirati da postoji stanovita razlika između, recimo, ironičnog korištenja nekog vrijednosnog izraza, njegova korištenja u svrhu izražavanja vlastitog iskrenog stava (vjerovanja) i korištenja tog istog izraza da bi se opisao stav (vjerovanje) drugih ljudi. No ako prihvati tu pretpostavku (a ne vidim zašto to ne bi učinio), onda je upravo on taj koji treba pokazati zašto je internalist u krivu kada spomenute razlike koristi kao oružje protiv amoralista i Sotone. Internalist, s druge strane, uopće ne mora misliti da je ovime izborio pobjedu: njegova se argumentacija svodi samo na to da se neki eksternalistički protuprimjeri daju *najbolje objasniti* upotrebom navedenih distinkcija, a na eksternalistu je da sada pokaže što s tom argumentacijom nije u redu. Međutim, ne znam ni za jedan drugi način na koji bi on to mogao učiniti osim da nastavi inzistirati kako nema ničeg nekoherentnog u tome da amoralist moralne sudove koristi u njihovu pravom značenju, ali da, unatoč tome, ostaje nemotiviran. Ali ako je to jedina strategija koju eksternalist ima na

[34] Radi izbjegavanja pogrešnog dojma, treba napomenuti da rješenje problema Sotone koje ovdje nastojim obraniti i koje McNaughton analizira nije *jedino* nekognitivističko rješenje. Simon Blackburn, s čijim se imenom najčešće poistovjećuje suvremeni nekognitivizam, primjerice, argumentira da "eksternalisti mogu pobijediti pojedinačne bitke … ali [da] internalizam dobiva rat" (Blackburn, 2000: 61, 65). Slučaj Sotone, kao i neki drugi eksternalistički slučajevi, on tvrdi, predstavljaju netipične situacije, a netipične situacije imaju smisla samo ako pretpostavimo postojanje tipičnih (normalnih) situacija u kojima je veza između etike i motivacije nužna. Blackburn dakle smatra da moralne prosudbe *nužno* motiviraju u *normalnim* okolnostima (usp. Zangwill, 2009: 418). Na primjer, iako zaljubljenost u normalnim okolnostima znači da želimo biti u društvu osobe u koju smo zaljubljeni, ipak je moguća situacija da tu osobu, u svrhu izbjegavanja boli, ne želimo vidjeti ili da je zbog ljubomore čak želimo povrijediti. Na sličan način, roditelj svom djetetu može zabraniti da nešto učini a da pritom nema namjeru sankcionirati ga ukoliko tu zabranu prekrši, nego mu, sasvim suprotno, ima namjeru oprostiti. Međutim, kao što prvi slučaj ne opovrgava činjenicu da stanje zaljubljenosti u normalnim okolnostima nužno povlači takva i takva stanja, drugi izolirani slučaj ne opovrgava tezu da u tipičnim situacijama postoji "nužna veza između čina zabrane i namjere za primjenom sankcija" (usp. Blackburn, 2000: 61–62). Priznajući da eksternalist dobiva "pojedinačne bitke", Blackburn tako oslabljuje internalizam koji bi se sada, čini mi se, mogao formulirati na sljedeći način: "Ako A prosuđuje da je Φ moralno ispravno, onda je A *pod jednakim uvjetima* motiviran učiniti Φ." No, za razliku od Blackburna, mislim da iz njegovih analogija nije vidljivo da eksternalist dobiva "pojedinačne bitke" te da oslabljivanje internalizma ovdje uopće nije potrebno. Što se tiče njegove prve analogije, to što zaljubljena osoba ne želi vidjeti onoga u koga je zaljubljena ne pokazuje da ona nema sve one želje (motivacije) koje se obično pripisuju zaljubljenim ljudima. Upravo suprotno, razumno je pretpostaviti da je ne želi vidjeti baš zato što ima sve te želje, a zna da ih neće moći ispuniti. Poteškoća druge analogije jest u tome što izricanje zabrane s namjerom opraštanja nije *iskreno* izricanje zabrane, a prisjetimo se da internalist tvrdi da je motivacija nužan rezultat *iskrene* moralne prosudbe. Eksternalist, prema mom mišljenju, ne dobiva čak ni "pojedinačne bitke".

raspolaganju, onda bi ga se lako moglo optužiti za cirkularnost jer time ne bi dokazivao svoju poziciju, već bi je samo pretpostavljao.[35]

3.4. Racionalistički (kantovski) internalizam motivacije

Do sada sam internalističku poziciju razmatrao u njezinoj nekognitivističkoj interpretaciji te sam nastojao pokazati na koji se način ona može nositi s različitim prigovorima koji joj se obično upućuju. Internalističku tezu, međutim, moguće je braniti i u njezinoj kognitivističkoj verziji – verziji prema kojoj je moralna motivacija zahtjev racionalnosti. Kognitivistički aspekt te internalističke verzije sadržan je u ideji da moralnim sudovima izražavamo *vjerovanja*, a da je također riječ o racionalističkom pristupu vidljivo je iz toga što su to vjerovanja o tome što imamo *razloga* činiti. U narednom ću dijelu o ovom obliku internalizma govoriti kao o *racionalističkom internalizmu motivacije* (RIM). Kao što će, nadam se, ubrzo postati jasno, RIM je zasnovan na kontroverznoj pretpostavci koju bez dodatne argumentacije ne bismo trebali prihvatiti, naime, da moralne činjenice zaista jesu činjenice o tome što imamo razloga činiti. Razrađenijem napadu na tu pretpostavku, koja je još poznata pod nazivom *moralni racionalizam*, bit će posvećena naredna poglavlja, dok ću u ostatku ovog

[35] S mojim se zaključkom neće složiti ni svi internalisti. Michael Smith (1994: 60–76) tako smatra da internalisti i eksternalisti podjednako dijele teret dokazivanja kada je riječ o fenomenu amoralista. Za njega je, kako izričito navodi, cijela diskusija analogna diskusiji o tome je li vizualno iskustvo konstitutivno za izricanje sudova o bojama. Preciznije rečeno, kao što postoji razlika u intucijama u pogledu pitanja da li osoba koja je slijepa od rođenja uopće izriče *sudove o bojama*, tako postoji razlika u intucijama u pogledu pitanja da li osoba koja nikada nije motivirana svojim moralnim sudovima uopće izriče *moralne sudove*. S obzirom da Smith smatra kako je dilemu jedino moguće riješiti pružanjem nezavisnog argumenta – argumenta koji ne bi pretpostavljao nijednu od pozicija u raspravi – on upravo takav argument nastoji formulirati. U najosnovnijim crtama, cilj argumenta jest pokazati da eksternalizam ne zahvaća pravu prirodu motivacije kod moralno dobre osobe jer su moralno dobre osobe motivirane *izravno* samim stvarima, a ne *neizravno* kako to proizlazi iz eksternalističkog stajališta. Da upotrijebim jedan primjer, mehanizam moralne motivacije za eksternaliste otprilike ima sljedeći oblik: (1) A prosuđuje da je moralno ispravno pomagati svojim prijateljima. (2) A želi (ima motivaciju) činiti ono što je moralno ispravno. (3) A želi (ima motivaciju) pomagati svojim prijateljima. Budući da je A-ova želja (motivacija) za pomaganjem prijateljima *izvedena* iz njegove želje (motivacije) da čini ono što je moralno ispravno, slijedi da je ona isključivo instrumentalne prirode. Međutim, to nije u skladu s našom slikom moralno dobre osobe za koju pretpostavljamo da ne želi pomagati svojim prijateljima u instrumentalnom smislu, već ne-instrumentalno. Eksternalizam bi stoga, Smith tvrdi, trebalo napustiti te se – kada je riječ o dilemi spomenutoj na početku – prikloniti onom mišljenju prema kojemu je prisutnost motivacije ipak konstitutivna za izricanje *moralnih sudova*. Dakle, amoralist ne izriče moralne sudove. Smithov argument bio je izložen ogromnom broju kritika od kojih treba izdvojiti Brink (1997), Miller (1996, 2003), Sadler (2003), Shafer-Landau, R. (1998), Svavarsdóttir (1999) te Zangwill (2007).

poglavlja nastojati ukazati na neke druge čimbenike zbog kojih RIM ne bismo trebali smatrati atraktivnom pozicijom te ću samo površno naznačiti odnos između moralnosti i razloga za djelovanje. No prvo ću nešto sustavnije prikazati središnja obilježja ovog, kako bi ga se također moglo smatrati, kantovski orijentiranog internalizma.

3.4.1. *Porijeklo moralnih zahtjeva*

Priroda morala je takva da na nas postavlja vrlo jake normativne zahtjeve: mi imamo dojam da se trebamo ponašati onako kako nam moral nalaže i da tim normama ni na koji način ne možemo izbjeći. Moralne se norme, barem nam se tako čini, zato i razlikuju od drugih normi kojima možemo biti podvrgnuti. Imamo osjećaj da neke druge norme, poput, primjerice, onih koje reguliraju lijepo ponašanje ili onih koje reguliraju ponašanje u prometu, nemaju takvu obvezujuću snagu. Objasniti porijeklo tog osjećaja, kako sam već kazao u poglavlju o hjumovskoj motivaciji, predstavlja velik izazov svakom filozofu morala; objasniti porijeklo tog osjećaja u nekom smislu ne znači ništa drugo nego objasniti porijeklo samih moralnih zahtjeva. Odakle ti moralni zahtjevi uopće dolaze? Zašto su moralne norme toliko specifične i zašto one imaju takvu snagu kakvu ne nalazimo ni kod jedne druge norme?

Jedan prilično raširen odgovor u etičkoj literaturi jest i onaj koji nas vodi natrag do samog Kanta. Osjećaj o neizbježnosti moralnih zahtjeva karakterističan je samo za racionalna bića jer zahtjevi morala nisu ništa drugo nego zahtjevi racionalnosti. Kao što normama racionalnosti ne možemo izbjeći jer smo racionalna bića, tako ne možemo izbjeći niti moralnim normama, a koje u konačnici nisu ništa drugo nego norme racionalnosti. Svođenjem morala na racionalnost, poseban status moralnih zahtjeva tako prestaje biti misterijom: budući da zahtjevi racionalnosti nisu misteriozni, nisu misteriozni niti zahtjevi morala.

Kako se u opisanu sliku može uklopiti internalistički model moralne motivacije? Vrlo jednostavno. Moralne norme na nas se odnose samim time što smo racionalna bića, a to znači

da njihovo porijeklo ne treba tražiti "izvan" nas, već se one "imaju izvesti iz općeg pojma umnog bića uopće" (*U* 78). Racionalno biće jednostavno "ne sluša nikakav drugi zakon osim onoga koji ono ujedno sâmo daje" (*U* 94). No ako je naša priroda kao racionalnih bića takva da mi *sami sebi* postavljamo zahtjeve koje ćemo slijediti, zar onda nije razumno pretpostaviti da nam, pored toga, nije potreban nikakav *dodatan* motivacijski impuls da te zahtjeve zaista i slijedimo? Dodatan motivacijski impuls možda bi i bio potreban u onim situacijama u kojima smo podvrgnuti nekim zahtjevima čije porijeklo nije u nama samima ili, Kantovom terminologijom, u situacijama kada je naša volja heteronomna. Na primjer, ako mi netko, suprotno mojoj volji, naredi da svaki dan trčim par kilometara, vrlo je vjerojatno da će mi biti potrebna dodatna motivacija da tu radnju svaki dan i izvršim. Dodatna motivacija u tom slučaju može poprimiti različite forme, poput one da ću biti podvrgnut nekoj sankciji ukoliko taj zahtjev ne ispunim, ili one da ću ispunjenjem tog zahtjeva na neki način profitirati. Stvari, međutim, poprimaju sasvim drukčije značenje kada je porijeklo zahtjeva u nama samima. Obvežem li sam sebe na neku radnju, onda je taj čin obvezivanja dovoljan da me na tu radnju i motivira, a porijeklo moralnih zahtjeva, kantovci argumentiraju, upravo je takvo: oni proizlaze iz nas samih te nam nisu nametnuti "izvana" i zato nam nije potrebna "vanjska sankcija" da bismo stekli motivaciju te zahtjeve doista i slijediti. Netko tko ne slijedi zahtjeve koje sam sebi postavlja može biti samo iracionalan, pogotovo ako ti zahtjevi i nisu ništa drugo nego zahtjevi racionalnosti. U opisanom je pristupu duh internalističke pozicije u potpunosti sačuvan.

Ideja da normativnost morala izvire iz racionalnosti sadrži jednu implikaciju koju je izuzetno važno ne izgubiti iz vida. S obzirom na činjenicu da racionalnost i racionalno ponašanje vežemo uz razloge, moralna razmatranja, prema ovom shvaćanju, zapravo pružaju razloge za djelovanje. Kazati da je moral normativan stoga ne znači ništa drugo nego kazati da moralna razmatranja nezaobilazno ulaze u sferu našeg praktičnog zaključivanja, tj. da

predstavljaju bitan čimbenik kada razmišljamo o tome što činiti. Uočiti normativnost morala, najkraće rečeno, isto je što i uočiti da moral pruža razloge za djelovanje. Pod pretpostavkom da područje morala zaista jest područje koje generira razloge za djelovanje, nekome tko je sklon braniti gore opisano shvaćanje moralne motivacije još samo preostaje preciznije odrediti odnos između razloga kao takvih i same motivacije. Autori koji naginju internalizmu taj odnos sasvim jasno preciziraju: *samo su racionalne osobe* motivirane djelovati u skladu s razlozima koje imaju odnosno veza između razloga i motivacije nužna je samo kod racionalnih osoba. Nadalje, ako nešto vrijedi za razloge općenito, tada vrijedi i za *moralne* razloge: dakle, samo su racionalne osobe motivirane djelovati u skladu s moralnim razlozima odnosno veza između *moralnih* razloga i motivacije nužna je samo kod racionalnih osoba. Moralna je motivacija stoga kognitivne prirode jer se u konačnici radi o tome da nas čisti praktični razum može motivirati. U najosnovnijim crtama, prosudba da je moralno ispravno (da trebamo) postupiti na neki način ekvivalentna je prosudbi da imamo razloga tako postupiti, a netko tko ostaje nemotiviran razlozima koji se na njega odnose jednostavno nije racionalan.

Nužnost veze između moralnih sudova i motivacije osigurana je samo kod racionalnih osoba, što znači da moralna motivacija, najjednostavnije rečeno, nastaje pod pritiskom racionalnosti. Jednu varijantu ove pozicije koja nam može poslužiti kao polazište u raspravi i koju valja nešto detaljnije razmotriti razvija Wallace (2006). Veza između moralnih sudova i motivacije nužna je, prema Wallaceu, samo kod onih koji *prihvaćaju* da moralna razmatranja pružaju razloge za djelovanje. Moralna razmatranja doduše svima pružaju razloge za djelovanje, ali subjekt, uslijed pogrešnog rasuđivanja, to ne mora prihvatiti kao činjenicu – on ne mora prihvatiti da moralna razmatranja imaju normativni značaj, a u tom bi ga slučaju bilo pogrešno proglasiti iracionalnim zbog nedostatka odgovarajuće motivacije. Jer netko tko *uopće ne prihvaća* da se nalazi pod normativnim pritiskom da nešto učini ne može biti

iracionalan zbog nedostatka motivacije da to učini. Ali unatoč tome što bi bilo pogrešno smatrati ga iracionalnim, smatra Wallace, taj ipak pogrešno rasuđuje jer ne prihvaća ono što bi trebao prihvatiti.

A što je s onima koji prihvaćaju da imaju razlog, a ipak nisu motivirani? Zar ti slučajevi nisu izravan dokaz tome da je veza između moralnih sudova i motivacije na koncu ipak samo kontingentna? Ne. Poznato je da, primjerice, slučajevi poput slabosti volje mogu raskinuti vezu između normativnosti i motivacije, a takvi slučajevi predstavljaju klasične primjere iracionalnog ponašanja. Bilo kako bilo, Wallaceova pozicija uspijeva blokirati dvije vrste slučaja kojima se internalizam motivacije obično nastoji osporiti. Prema njegovu, kako ga on naziva, "motivacijskom zahtjevu", oni koji nisu motivirani moralnim sudovima *ili* zbog pogrešnog rasuđivanja ne uviđaju normativnu snagu tih sudova (ne uviđaju da im ti sudovi pružaju razloge za djelovanje) *ili* potpadaju pod neke od klasičnih primjera praktične iracionalnosti. Dakle, netko tko, kao što je to slučaj s amoralistom, nije motiviran moralnim sudovima sigurno spada u jednu od dvije navedene skupine.

Velik broj kantovski usmjerenih autora moralnu motivaciju shvaća u većoj ili manjoj mjeri na gore prikazani način. Zato ću u dijelu koji slijedi pružiti formulaciju RIM-a koju bi, unatoč njihovim manjim razlikama, većina kantovaca mogla prihvatiti. Kao što će ubrzo postati vidljivo, već se sama ta formulacija suočava s određenim poteškoćama za koje vjerujem da bi ih svaki pobornik RIM-a trebao prevladati ako želi da se njegova pozicija shvati ozbiljno. Potom ću, oslanjajući se na Svavarsdóttir, ukazati na neke nedostatke Wallaceove verzije tog stajališta te ću tvrditi kako Svavarsdóttirina kritika Wallacea zapravo ide u prilog nekognitivističkom internalizmu, a ne eksternalizmu. Naposljetku ću nešto kazati o moralnom racionalizmu – pretpostavci koja stoji u podlozi RIM-a.

3.4.2. *Poteškoće s formulacijom*

Kantovci se u osnovi slažu da formulacija RIM-a ima sljedeći oblik:

> (RIM): Ako A prosuđuje da je moralno ispravno učiniti Φ ili ako prosuđuje da ima moralnu dužnost učiniti Φ, onda će, *ako je praktično racionalan*, biti motiviran učiniti Φ.

Budući da je nužnost veze između moralnih prosudbi i motivacije zajamčena *samo* kod racionalnih osoba, jasno je da RIM u nekom smislu predstavlja slabiju varijantu internalizma. Jaka varijanta internalizma motivacije (JIM), koju sam u njezinu nekognitivističkom obliku nastojao obraniti u prvom dijelu, obično je formulirana na sljedeći način:

> (JIM): Ako A prosuđuje da je moralno ispravno učiniti Φ ili ako prosuđuje da ima moralnu dužnost učiniti Φ, onda će biti motiviran učiniti Φ.[36]

Svi zastupnici RIM-a (ali ne samo oni) smatraju da je JIM izuzetno teško braniti. Problem je u tome što JIM ne uzima u obzir različite fenomene koji očito dovode u opasnost praktični utjecaj moralnosti. O kakvim je točno fenomenima riječ vidljivo je iz sljedećih komentara. Smith, primjerice, smatra da nas JIM stavlja u položaj da moramo negirati kako "praktične iracionalnosti različitih vrsta – različite vrste 'depresije' kako ih naziva Stocker – mogu ostaviti nečije evaluativno gledište netaknutim a istovremeno odstraniti njegovu sveukupnu motivaciju" (Smith, 1994: 121). Isto tvrdi i Sadler kada piše da je JIM "neplauzibilan zbog slučajeva slabosti volje" (Sadler, 2003: 64). I Svavarsdottir, koja naginje eksternalizmu, također napominje da JIM zahtijeva modifikaciju budući da "pod uvjetima duboke depresije, jakih slučajeva slabosti volje i drugih bolesti duše, veza između moralnog suda i motivacije često je slomljena…" (Svavarsdottir, 1999: 164). Na koncu i Wallace, čiju sam razradu RIM-

[36] Vidi Smith (1994: 61).

a uzeo kao polazište u ovoj diskusiji, prisjetimo se, vjeruje da je "u slučajevima slabosti volje i nekim oblicima samoobmane" subjekt nemotiviran svojim moralnim prosudbama, ali da je za takve subjekte "prirodno kazati" da su iracionalni (Wallace, 2006: 187).

Veliko je pitanje je li uopće točno da *akrasia* vrši takav utjecaj kao što se to pretpostavlja u citiranim tvrdnjama. Konkretnije rečeno, slabost volje najčešće pripisujemo onima koji se ne *ponašaju* onako kako sami misle da bi trebali, a u gornjim se komentarima pretpostavlja nešto sasvim drugo, naime, da slabost volje potkopava *motivaciju* da činimo ono što mislimo da bismo trebali. Neki će reći da je za fenomen akrasije karakteristično to da subjekt ne djeluje u skladu sa svojom boljom prosudbom, a ne da subjekt nije motiviran djelovati u skladu sa svojom boljom prosudbom (usp. Dancy, 1993: 16, fn.13; Garrard i McNaughton, 1998: 49). No, po mom mišljenju, puno ozbiljniji problem za formulaciju RIM-a izvire iz činjenice što nije samo *akrasia* ono što navodno raskida vezu između moralnosti i motivacije, već to, prema nekima, čine i stanja "duboke depresije". Prisjetimo se da se na stanja duboke depresije u debati s internalistom poziva eksternalist. On vjeruje da su takva stanja izravan dokaz neistinitosti internalizma. Ranije sam razmotrio jedan takav prigovor te sam pokušao pokazati da za nekognitivističkog internalista ipak postoje načini da s njim izađe "na kraj". Međutim, ono što je ovdje bitno jest da zastupnici RIM-a, s druge strane, protuprimjer tog tipa diskvalificiraju u samom početku. Dopuštajući da depresija nečije moralne sudove može lišiti motivacijskih učinaka, oni se od tog prigovora ograđuju time što ga anticipiraju u samoj formulaciji svog gledišta: depresija zapravo nije ništa drugo nego jedan oblik praktične iracionalnosti. Na prvi pogled to izgleda kao dobra strategija. Jer zašto protivniku uopće dopustiti da artikulira svoje prigovore ako već postoji način da se to izbjegne? Možda to doista i jest dobra strategija, ali opasnost je u tome što RIM na taj način počinje poprimati oblike sasvim trivijalne pozicije (usp. Miller, 2003: 221). Oslanjajući se na svoju formulaciju, zastupnik RIM-a dolazi do izvrsnog položaja budući da je sve potencijalno

opasne eksternalističke slučajeve u mogućnosti okarakterizirati kao slučajeve praktične iracionalnosti.

3.4.3. *Prihvaća li amoralist moralne tvrdnje?*

Vratimo se sada na Wallaceovu verziju RIM-a te ukažimo na još neke nedostatke njegove pozicije. Svavarsdóttir (2006) najveću pozornost pridaje Wallaceovoj tezi da je uvjet racionalnosti biti motiviran nečime čiji normativni značaj *prihvaćamo*. Za nju je uvjerljivost te teze u velikoj mjeri ovisna o shvaćanju izraza "prihvatiti", ali i o shvaćanju racionalnosti uopće. Podsjetimo još jednom da Wallace ne zagovara postojanje nužne veze između normativnih zahtjeva i motivacije, već postojanje nužne veze između *prihvaćanja* normativnih zahtjeva i motivacije. Jer, kako on ističe, iako nije upitno da motivacija za činjenjem Φ-a može izostati kod nekoga tko ne prepoznaje da treba učiniti Φ, ta je mogućnost isključena kod onih koji *prihvaćaju* da treba učiniti Φ. Normativni zahtjevi usmjeravaju naše djelovanje i zato ih je moguće shvatiti kao vodiče/putokaze kojima se uvijek možemo rukovoditi, a moralni sudovi, po svemu sudeći, imaju upravo takvu ulogu: prihvatiti da je određeni smjer djelovanja moralno ispravan stoga je isto što i prihvatiti da tim smjerom treba krenuti. Imajući to u vidu, što kazati o onome tko prihvaća da treba krenuti određenim smjerom, ali to nije motiviran učiniti? Za Wallacea, kako smo vidjeli, postoje samo dva moguća objašnjenja: ta osoba *ili*, suprotno onome što sama tvrdi, ipak ne prihvaća moralne sudove kao vodiče u djelovanju (ne prihvaća njihovu normativnu prirodu) *ili* nije racionalna.

Svavarsdóttir za cilj ima pokazati da nijedno od objašnjenja ne predstavlja ispravan opis onoga što se događa kada kod osobe izostaje odgovarajuća motivacija. Što se prvog objašnjenja tiče, osnovno je učiniti distinkciju između prihvaćanja nečega kao vodiča u djelovanju i prihvaćanja nečega *kao svog* vodiča u djelovanju. Sasvim je moguće da netko prihvati moralne sudove kao vodič u djelovanju, ali da ostane nemotiviran djelovati onako

kako taj vodič nalaže. Ta je mogućnost otvorena prvenstveno zbog toga što on ne mora prihvatiti moralne sudove kao *svoj* vodič odnosno kao nešto čime se *on sam* rukovodi u svom ponašanju. Odbacivanjem moralnih sudova kao vlastitog vodiča subjekt ne mora ujedno negirati i normativni značaj tih sudova jer je sasvim konzistentno tvrditi da moralni sudovi imaju normativni značaj a u isto vrijeme ustrajati na tvrdnji da ti sudovi nemaju normativni značaj za neku konkretnu osobu. Uostalom, to je i glavna odlika cinika, kako ih naziva Svavarsdóttir, koji prihvaćaju činjenicu da moralni sudovi predstavljaju putokaze za moralno ponašanje, ali koji istodobno te sudove ne prihvaćaju kao *svoje* vodiče. Na ovom je mjestu moguće slijediti Svavarsdóttir i ovu poziciju još više razjasniti povlačenjem analogije s normama koje reguliraju ponašanje u prometu. Nema proturječnosti u tome da netko prihvaća normativnost prometnih pravila, ali da istodobno ta pravila ne prihvaća kao nešto što regulira *njegovo* ponašanje kao sudionika u prometu. On se čak može složiti s time da bi trebao, ukoliko uhvaćen u prekršaju, kazneno odgovarati, ali to ne znači da ga samo prihvaćanje tih normi mora i motivirati da u skladu s njima i djeluje.[37]

Sada se čini opravdanim preispitati i drugo objašnjenje koje Wallace nudi za one koji nisu motivirani činiti ono što im nalažu norme koje prihvaćaju. Je li Wallace zaista u pravu kada kaže da nedostatak odgovarajuće motivacije kod takvih osoba također može biti manifestacija njihove iracionalnosti? Nije očito da je odgovor potvrdan. Na primjer, pozovemo li se na gornju distinkciju između dva načina na koji neku normu možemo prihvatiti (prihvatiti neku normu nije isto što i prihvatiti tu normu *kao svoju* normu), Wallaceovoj optužbi za iracionalnost nije teško umaknuti. Netko tko prihvaća prometne norme, ali koji te iste norme ne prihvaća kao svoje norme, tj. kao nešto na što se on sam

[37] Josip Talanga spominje anegdotu koja, čini mi se, zgodno može poslužiti u ovom kontekstu budući da lijepo ilustrira procjep između normativnosti i motivacije: "Profesor Scheler je živio dosta u raskoraku s uobičajenim akademskim pravilima ponašanja i bio je poznat po svojim brojnim aferama sa studenticama. Zbog toga su ga neki upitali zašto živi u neskladu sa svojim lijepim teorijama o moralu. Scheler je na to odgovorio: 'Moji su spisi o moralu dobri putokazi za moralni život, ali, kao što znate, putokazi samo pokazuju kuda treba ići, ali sami tamo ne idu" (Talanga, 1999: 15).

oslanja kao sudionik u prometu ne mora nužno biti iracionalan ako tim normama nije motiviran. Kršenjem tih normi on možda pokazuje neke druge nedostatke. Možda bi za njega bilo prigodnije kazati da nema dovoljno senzibilnosti ili možda, oslanjajući se na Svavarsdóttir, da svojim ponašanjem samo pokazuje da nije uzoran građanin. Bilo kako bilo, nije jasno zašto je potrebno inzistirati na tome da je njegova nemotiviranost ujedno i evidencija njegove iracionalnosti.

Ako su prigovori koje Svavarsdóttir nudi uvjerljivi (a meni se čini da jesu), tada možda nije neistinito zaključiti da je RIM ozbiljno uzdrman. Jer iako su ti prigovori prvenstveno usmjereni protiv Wallacea, njegova je verzija internalizma, sklon sam vjerovati, nešto s čime bi se u znatnoj mjeri mogao složiti svaki simpatizer kantovskog shvaćanja motivacije. No znači li to da je Svavarsdóttir ipak u pravu kada smatra da eksternalizam adekvatno objašnjava prirodu moralne motivacije? Mislim da je preuranjeno kazati da je odgovor na to pitanje potvrdan. Čak i ako je RIM neplauzibilna pozicija, još uvijek smo prilično daleko od zaključka da je eksternalizam motivacije stajalište kojemu bi se trebalo prikloniti. Da skratim, vjerujem da Svavarsdóttirina kritika polučuje sasvim suprotan učinak od onog koji je namjeravan. Intuicija koja stoji u podlozi njezine distinkcije između prihvaćanja nečega kao vodiča u djelovanju i prihvaćanja nečega *kao svog* vodiča u djelovanju jest upravo ona intucija koju ima nekognitivist kada se poziva na Hareove opservacije o navodničkoj upotrebi moralnih sudova. Pojednostavljeno rečeno, nekognitivistički internalist će se s lakoćom složiti da netko može prihvatiti nešto kao vodič u djelovanju a pritom ne biti motiviran. Razlog tome jest što i on sam smatra da se to događa upravo u onim situacijama kada se moralni sudovi upotrebljavaju s izvjesnim odmakom, tj. kada se subjekt s vrijednostima koje su sadržane u tim sudovima ne poistovjećuje u potpunosti. Kazati da netko upotrebljava moralne sudove u "navodnicima" isto je što i kazati da on te moralne sudove prihvaća kao vodiče, ali ne da ih prihvaća *kao svoje* vodiče. No kao

što nekognitivist smatra da nije moguće da netko ostane nemotiviran moralnim sudovima kada ih koristi u njihovu *pravom značenju* tako Svavarsdóttir prešutno tvrdi da nije moguće da netko ostane nemotiviran moralnim sudovima ako ih prihvati kao *svoje vodiče* u djelovanju. Nekognitivist možda ne upotrebljava Svavarsdóttirinu terminologiju da bi opisao svoju intuiciju, ali to ionako nije od presudne važnosti. Važno je samo to da je ta intucija kod obje strane potpuno identična. Možda zvuči pomalo paradoksalno, ali proizlazi da čak i neki eksternalisti prešutno prihvaćaju internalizam.

3.4.4. *Pružaju li moralne dužnosti razloge za djelovanje?*

Fundamentalna premisa na kojoj počiva RIM jest teza *moralnog racionalizma*:

(MR): Ako je Φ moralno ispravno, tada je to razlog da se učini Φ.

Moralni racionalisti odbijaju instrumentalnu koncepciju racionalnosti jer smatraju da moralne dužnosti svima pružaju razloge za djelovanje i da je stoga uvijek racionalno postupati u skladu s onim što nam moral nalaže. Kada bi racionalnost bila instrumentalna, tada moralne činjenice ne bi *svima* pružale razloge za djelovanje, nego samo onima koji bi imali odgovarajuće želje i ciljeve. Drugim riječima, to da je neka radnja moralno ispravna, instrumentalist smatra, samo po sebi ne pruža nikakav razlog. Za instrumentalista sfera morala ne predstavlja neku posebnu sferu u kojoj vrijede drukčija pravila i zakonitosti: ako su razlozi za djelovanje ovisni o subjektovim preferencijama i željama, onda su *svi* razlozi za djelovanje ovisni o subjektovim preferencijama i željama. Stoga će, prema ovome gledištu, moralna ispravnost neke radnje nekome predstavljati razlog za djelovanje samo ako, recimo, on ima želju činiti ono što je moralno ispravno. U nedostatku takve želje o postojanju razloga ne može biti govora. Za moralne racionaliste to nije opcija: moralni su razlozi prisutni

neovisno o tome koje su subjektove preferencije, a pretpostavka tog stajališta jest da su moralni zahtjevi zahtjevi praktične racionalnosti. Svako racionalno biće ima *razlog* ponašati se moralno samim time što je racionalno.

Dok ću u sljedećim poglavljima tvrditi da instrumentalisti u ovom sporu odnose pobjedu te da se još uvijek ne naziru argumenti koji bi nas mogli uvjeriti da doista postoje kategoričke norme racionalnosti, tj. norme koje generiraju kategoričke razloge za djelovanje, ovdje ću samo ukratko razmotriti Wallaceov argument u prilog moralnom racionalizmu. Wallace ne pokušava pružiti konkluzivan argument za moralni racionalizam te ne isključuje mogućnost da nekome moral možda i ne pruža razloge za djelovanje, ali ipak navodi neke činjenice za koje vjeruje da najbolje mogu biti objašnjene u sklopu racionalističkog stajališta:

> Ono što se možda može kazati o odnosu morala i motivacije jest da moralna razmatranja *teže* normativnom statusu razloga za djelovanje. … U raspravama s drugim ljudima mi tipično navodimo moralna razmatranja kao čimbenike koji imaju direktan normativni značaj, koji govore za ili protiv opcija koje su pod aktivnim razmatranjem. Mi, na primjer, podižemo svoju djecu da se prema moralnim razmatranjima odnose na taj način. … Ti aspekti naše moralne prakse … sugeriraju da se na moralnost široko gleda kao na normativnu domenu, u onoj mjeri u kojoj moralna razmatranja igraju ulogu u našim individualnim i skupnim rasuđivanjima koja se obično pripisuje *pravim razlozima za djelovanje*. (Wallace, 2006: 192, moj kurziv)

Wallaceova pretpostavka da fenomenološke karakteristike moralne prakse najbolje mogu biti objašnjene u racionalističkom okviru, međutim, nije osnovana. Oslanjanje na moralnu praksu, čini se, nije dovoljno za izgradnju stajališta prema kojemu moral generira *objektivne* razloge za djelovanje. Činjenica da se mi u raspravama često pozivamo na moralna razmatranja kao normativna i da nam je, na koncu, stalo da se naša djeca prema tim razmatranjima odnose na takav način također je kompatibilna s instrumentalističkim pogledom na racionalnost. U našoj

moralnoj praksi mi jednostavno pretpostavljamo da je drugim ljudima stalo do morala i da nemaju ciljeve čiji je status moralno upitan. Pristajući na tu prešutnu pretpostavku, tvrdimo da i oni imaju razlog slijediti moralne zahtjeve jer vjerujemo da time neće nužno ugroziti te ciljeve, nego ih možda čak i lakše realizirati. No ako su naši sugovornici sasvim indiferentni prema moralnim pitanjima i ako su njihovi ciljevi u suprotnosti sa zahtjevima morala, tada možda i nema smisla ustrajati na tvrdnji da moralna razmatranja i za njih imaju "direktan normativni značaj" ili da oni ipak imaju *razlog* ponašati se moralno. Oni su svakako moralno neosjetljivi, ali ne nužno iracionalni.

II. RAZLOZI

4. Hjumovski internalizam razloga

U prvom dijelu ovog rada u središtu mog interesa bili su motivirajući razlozi – razlozi koji objašnjavaju naše postupke. Tako sam prvo nastojao pokazati da su želje (u kombinaciji s odgovarajućim vjerovanjima) ona psihološka stanja koja nas motiviraju te sam potom argumentirao da upravo nekognitivističko shvaćanje *moralne* motivacije proizlazi iz prihvaćanja ove tvrdnje. No što je s normativnim razlozima? Jesu li i normativni razlozi – razlozi koji opravdavaju – također zavisni o našim željama? Mnogi nisu spremni ići tako daleko i kazati da je odgovor na to pitanje potvrdan. Takvi razlozi, pomislit ćemo, moraju vrijediti objektivno tako da njihovo postojanje ne može (i ne bi smjelo) biti zavisno o čimbenicima koji su sasvim subjektivni i arbitrarni. Uostalom, zar se nećemo složiti da to posebno dolazi do izražaja kada je riječ o *moralnim* razlozima? Zar nije očito da su moralni razlozi potpuno neovisni o kontingentnim činjenicama ljudske prirode? Jedna od mojih glavnih zadaća u narednim poglavljima jest negirati ovo prevladavajuće uvjerenje.

4.1. Hume i instrumentalizam

Poznat je status razuma u Humeovoj filozofiji: razum je u potpunosti podređen strastima i ne postoji nijedan način na koji im se može suprotstaviti. Taj stav na prvi pogled zaista zvuči radikalno i zato je vrijedno na njemu se zadržati i promotriti što stoji u njegovoj pozadini. Razum se, prema Humeu, bavi idejama i činjenicama, tj. stvarima koje mogu biti istinite ili lažne, a strasti, s druge strane, nemaju istinosnu vrijednost budući da nisu reprezentacijska mentalna stanja. Domena unutar koje djeluje razum i domena unutar koje djeluju strasti jasno su odijeljene i svaki je sukob između njih, već po samoj prirodi stvari, a priori isključen. Humeova poznata krilatica o razumu kao robu strasti samo je konzekvenca njegova razumijevanja prirode "posla" kojim se razum bavi i nije motivirana, kao što bi se moglo

naslutiti, stavom da strasti imaju važniji status u okviru njegove cjelokupne filozofije. Koristeći termin "rob", Hume samo želi jasno dati do znanja da razum, metaforički rečeno, nema ovlasti strastima "kazati" kakve bi one trebale biti – s obzirom da im ne može ukazati na njihovu "neistinitost", razum strasti nije u mogućnosti kritizirati ni na jedan drugi način jer drugi način kritike jednostavno ne postoji. Nemogućnost kritike strasti od strane razuma, međutim, ne povlači da razum na strasti ni na koji način ne može *utjecati*. Upravo suprotno, Hume smatra da postoje dva načina na koja se to događa. Njegovim riječima:

> […] um, u strogom i filozofskom smislu, može na naše vladanje utjecati samo na *dva* načina: bilo tako da pobuđuje strast obavještavajući nas o postojanju onog što je pravi predmet te strasti, bilo tako da nam omogućuje sredstvo zadovoljenja te strasti otkrivajući nam vezu uzroka i posljedica. To su jedine vrste prosudbi koje mogu pratiti naše radnje i o kojima se može reći da ih na bilo koji način proizvode; a uz to moramo dopustiti da su te prosudbe često lažne i pogrešne (*R* 45).

Koristeći suvremeniju terminologiju, Hume napominje da se uloga razuma, u kontekstu njegova mogućeg utjecaja na strasti, sastoji u davanju relevantnih informacija o predmetu strasti te davanja informacija o sredstvima putem kojih strasti mogu biti zadovoljene. Strasti su ponekad zasnovane na pogrešnim prosudbama. Pretpostavimo li, na primjer, da imam želju pojesti jabuku koja se nalazi na stolu ispred mene, ali da istodobno ne znam da to nije prava jabuka, vrlo je vjerojatno da bih, kada bih znao tu informaciju, izgubio želju za jedenjem te konkretne jabuke. Isto tako, razum nam daje informaciju o tome kako da ostvarimo svoje želje. Želim li pročitati današnje novine, svoju želju neću moći realizirati ako nemam informaciju o tome kako do njih mogu doći: moja želja može postati efektivna tek kada dobijem tu informaciju. Presudno je ne izgubiti iz vida da navedeni primjeri ne pokazuju kako razum željama ipak može "kazati" kakav bi *trebao* biti njihov predmet. To što bismo mi

trebali željeti ili koji bi trebali biti predmeti naših želja, nešto je nad čim razum nema ingerenciju i zato se razum, slikovito rečeno, javlja na sceni tek nakon što su predmeti naših želja jasno utvrđeni. Jedino što razum može učiniti jest ukazati na to da su naše želje rezultat pogrešnog shvaćanja stanja stvari ili nas pak može usmjeriti na pravi put u njihovoj realizaciji. Želje mogu biti nerazumne samo u minimalnom pogledu, a to je u onome kada nastaju kao rezultat pogrešnih prosudbi i u onome kada ne mogu biti zadovoljene uslijed biranja pogrešnih sredstava, ali, "čak ni tada, strogo govoreći, nije strast nerazumna, nego sud" (*R* 356). Uloga je razuma, prema Humeovoj slici, dakle prilično jasno definirana. U kontekstu rečenoga treba razumjeti sljedeći Humeov odlomak:

> Nije protivno razumu pretpostaviti uništenje cijelog svijeta ogrebotini mog prsta. Nije protivno razumu da izaberem svoju potpunu propast da bih spriječio najmanju neugodnost nekog Indijca, ili osobe meni potpuno nepoznate. Isto tako, malo se protivi razumu da pretpostavim čak svoje priznato manje dobro svojemu većem dobru, i da vatrenije volim ono prvo negoli ovo drugo (*R* 355).

Hume ovdje govori o nekoliko stvari. Prvi dio odlomka ima moralne implikacije, dok je drugi više orijentiran na ono što se obično naziva prudencijalnošću. Netko tko uistinu preferira uništenje cijelog svijeta ogrebotini na prstu ima preferenciju koja sigurno nije imuna na *moralnu* kritiku. Prema svemu sudeći rekli bismo da se radi o osobi koja ima teški moralni poremećaj, ali ne bi bilo istinito kazati, barem ne u sklopu Humeova shvaćanja, da se radi o osobi koja je nerazumna. Slično vrijedi i za slučaj promicanja vlastite dobrobiti. Što reći o onome tko ima želju žrtvovati se za nekoga koga uopće ne poznaje i koga nikada nije vidio, i to samo zato da bi ga poštedio neke bezazlene neugodnosti? Onaj koga bi ta želja zaista pokrenula na postupak izlazio bi daleko izvan okvira onoga što bi moral od njega mogao zahtijevati (vjerojatno bi se radilo o paradigmatičnom primjeru supererogacije), ali Hume ne

misli da bi to bilo iracionalno. Imam li strast činiti nešto što je za mene, sveukupno promatrajući, lošije, još uvijek ne mora biti tako da želim nešto što se protivi načelima razuma. Jer želje su izvan domene racionalne kritike i Hume namjerno bira ovako radikalne primjere samo da bi jasnije istaknuo tu središnju misao.

Moralnost i prudencijalnost, gornji primjeri također ilustriraju, nisu zahtjevi racionalnosti. Kazati da je neki pravac djelovanja zahtjev racionalnosti ima smisla samo ako bi taj pravac djelovanja predstavljao adekvatno sredstvo za ostvarenje određene želje. Humeova misao da se utjecaj koji razum vrši na strasti iscrpljuje u pružanju odgovarajuće informacije o načinu njihova zadovoljenja na koncu je i dovela do toga da se njemu obično pripisuje instrumentalizam u pogledu praktičnog zaključivanja. Praktično zaključivanje, shvaćeno na instrumentalni način, svodi se na biranje sredstava za ostvarenje određenih ciljeva ili, formulirano u kategorijama razloga za djelovanje, instrumentalno načelo zaključivanja tvrdi da imamo razloga poduzeti ona sredstva putem kojih na najbolji način možemo zadovoljiti naše želje ili preferencije. Za instrumentaliste se praktično zaključivanje dakle svodi na razmišljanje o tome kako realizirati ono što već želimo.

Pitanje je, međutim, je li Humea uopće opravdano svrstati među instrumentaliste. Upravo suprotno općeprihvaćenom stajalištu, neki argumentiraju da je to posve neutemeljeno jer za njega praktično rasuđivanje zapravo ne postoji (Millgram, 2005: 198–217; Smith, 2004b: 75–92). Kada bi postojalo praktično rasuđivanje, to bi značilo da razum ipak može biti u konfliktu s motivacijskim stanjima, a vidjeli smo da Hume takvu mogućnost ne dozvoljava. Humeovo prepoznavanje načina na koji razum može utjecati na strasti nedvojbeno je razlog zašto je dugo vremena vladala tendencija da se korijeni instrumentalizma traže upravo u njegovoj filozofiji, ali, strogo promatrajući, ne postoji nikakva osnova za mišljenje da je Hume nešto takvo imao u vidu.

Slika koju nam Hume nudi otprilike izgleda ovako: strasti, metaforički rečeno, šalju razum po informaciju o načinu kako mogu biti zadovoljene. No jednom kada su stekle tu informaciju, strasti s njom mogu raspolagati kako žele; razum, budući da je njihov rob, nema ovlasti strastima *narediti* što da rade s tom informacijom. Evo jednog primjera. Pretpostavimo da imam želju stići na određeno mjesto i da imam informaciju o tome kako tu želju mogu realizirati na najbolji i najjednostavniji način. Na primjer, otiđem li na brod koji isplovljava jednom dnevno, na željeno mjesto ću stići u najkraćem mogućem roku. Vrlo je vjerojatno da će ta informacija na mene izvršiti stanoviti utjecaj u smislu da ću na koncu ipak formirati želju otići na taj brod. Međutim, ovdje ne smije postojati *pritisak razuma* za formiranjem takve želje. Postojanje takvog pritiska povlačilo bi da razum prekoračuje svoje ovlasti koje ima kao rob. Formiram li želju da na odredište krenem plivajući, to bi samo značilo da biram put koji je teži i duži, ali ne i to da se ponašam iracionalno. Sve je ovo, međutim, nekonzistentno s instrumentalizmom. Instrumentalističko načelo, suprotno Humeu, tvrdi da je racionalno formirati želju otići na brod i da postoji stanoviti pritisak razuma na moja motivacijska stanja. Humeova je vizija nešto drukčija: jednom kada dobijemo informaciju da je naša želja zasnovana na pogrešnoj prosudbi ili informaciju o tome na koji način tu želju možemo najlakše realizirati, to će gotovo sigurno imati odgovarajući utjecaj na naše daljnje ponašanje. Netočno je međutim zaključiti da je tu riječ o nekakvom normativnom pritisku na naše želje, već se prije radi o načinu na koji smo, kao ljudska bića, konstituirani – oni koji korigiraju svoje želje u svjetlu relevantnih informacija jednostavno bolje prolaze od onih koji to ne čine. Gotovo će svatko, na primjer, prestati željeti šetati na lokaciji gdje postoji opasnost od miniranja, ali nema temelja za tvrdnju, barem ne iz onoga što sam Hume kaže, da bi raspolaganje tom informacijom *trebalo* rezultirati nestankom te želje. Nekonzistentno je biti instrumentalist u pogledu praktičnih razloga, a istovremeno tvrditi da je "razum rob strasti".

Pod pretpostavkom da je gornja interpretacija Humeova viđenja prihvatljiva odnosno ako je uistinu tako da je Humeu pogrešno pripisati zastupanje instrumentalnog načela, ta činjenica ne mora imati značajne posljedice za suvremenu raspravu. Kao što je na jednom mjestu rekao Dancy, "moj je pravi interes za poziciju inspiriranu Humeom, iako ne dostupnu samom Humeu, da je IN [instrumentalno načelo] jedini racionalni zahtjev" (Dancy, 2000: 47 fn12). Mislim da je ovdje lako složiti se s Dancyem. Čak i ako je samoga Humea kao povijesnu figuru pogrešno svrstati u zastupnike instrumentalizma, nije prijeporno da je instrumentalizam glavno obilježje hjumovske pozicije. Današnji hjumovci uglavnom smatraju da je instrumentalističko shvaćanje praktične racionalnosti osnovno, ako ne i jedino shvaćanje praktične racionalnosti. Instrumentalizam zato predstavlja *received view* i rasprava se uglavnom odvija na način da protivnici ovog načela nastoje dokazati njegovu neistinitost. Instrumentalno se načelo smatra istinitim sve dok se ne dokaže suprotno i zato strategija hjumovaca uglavnom ima defenzivan oblik.

Hjumovski instrumentalisti slažu se u stavu da razlozi za djelovanje ne mogu biti prisutni bez odgovarajuće želje. U narednom ću dijelu prikazati i pokušati obraniti najpoznatiju artikulaciju te ideje koju je pružio Bernard Williams (1981, 1995a, 1995b, 2001, 2006) – poziciju poznatu kao *internalizam razloga*. No prije nego što to učinim, evo jedne preliminarne napomene. U literaturi ne postoji konsenzus oko *preciznog* određenja odnosa između instrumentalizma i internalizma. Neki, na primjer, smatraju da su "[i]nstrumentalizam i internalizam razloga povezani … kao vrsta i rod. Instrumentalizam je najpopularniji oblik internalizma. Instrumentalizam navodi nečije želje kao jedini relevantni internalni izvor razloga" (Shafer-Landau, R. i Cuneo, T., 2007: 283). Nešto drukčije viđenje tog odnosa nalazimo u tvrdnji da "internalizam pretpostavlja instrumentalizam" te u tvrdnji da je "internalizam … sofisticiraniji nasljednik instrumentalizma" (Millgram, 2001: 6–7). Kao što će ubrzo postati vidljivo, Williams se udaljuje od instrumentalizma utoliko što smatra da

postoje i neki drugi oblici praktičnog zaključivanja koji – pored "sredstvo-cilj" zaključivanja – generiraju razloge za djelovanje. Međutim, kako to na jednom mjestu zgodno formulira Korsgaard, ta načela rasuđivanja nisu ništa drugo nego "prirodne ekstenzije" instrumentalizma. U čemu god da se sastojali procesi rasuđivanja, oni, Williams će reći, moraju biti pod kontrolom nečega što nas je u stanju motivirati. Na tom mjestu leži poveznica s instrumentalizmom (usp Korsgaard, 1996a: 326–328).[38]

4.2. Williamsov internalizam

Williamsova distinkcija između tzv. *unutarnjih* (internalnih) i *vanjskih* (eksternalnih) razloga predstavlja nezaobilazno mjesto u diskusiji o praktičnom rasuđivanju. McDowell je svojevremeno možda sasvim opravdano naznačio da je Williamsova ideja "nedostatno diskutirana", ali takvu dijagnozu danas nije moguće postaviti (McDowell 1995: 68). U posljednje je vrijeme pojam razloga došao u samo središte etičkih rasprava, a tako se i Williamsova obrana internalizma našla na samom udaru kritika. Internalizam razloga je, u vrlo grubim crtama, pozicija prema kojoj su svi razlozi za djelovanje unutarnje (internalne) prirode, što znači da njihovo porijeklo treba tražiti u subjektovu motivacijskom ustrojstvu. Prvo ću prikazati osnovna obilježja internalizma, a potom ću izdvojiti četiri, po mom sudu, najvažnija prigovora koja su toj poziciji bila upućena te ukazati na koji bi se način internalist od njih mogao obraniti.

U svrhu boljeg razjašnjenja internalizma, mislim da je potrebno prvo odgovoriti na sljedeće pitanje: koji je odnos između razloga i motiva? Razmišljajući o ovom pitanju, vrlo brzo možemo uočiti sljedeće. Kada prosuđujemo kako postupiti u određenoj situaciji, naše se rasuđivanje svodi na "vaganje" razloga te se u većini slučajeva (izuzev slučajeva *akrasije*) opredjeljujemo za one razloge za koje smo utvrdili da su dobri. Drugim riječima, nakon što

[38] Za razliku od instrumentalista, Williams pojam želje shvaća u nešto širem značenju. O tome će ubrzo biti više riječi.

smo rasuđivanjem utvrdili da imamo dobar razlog učiniti neku radnju, taj bi nas razlog – izuzmemo li slučaj *akrasije* – trebao motivirati. Ovu činjenicu ističe i sâm Williams tvrdnjom da ukoliko je R dobar razlog da A učini Φ, tada ima smisla reći da A može učiniti Φ upravo zbog razloga R: "Ako je istinito da A ima razlog učiniti Φ, tada mora biti moguće da bi on trebao učiniti Φ iz tog razloga; te ako djeluje iz tog razloga, tada će taj razlog biti objašnjenje njegova djelovanja (Williams 1995a: 39). Ili kako to, nadovezujući se na Nagela, ističe Christine Korsgaard: "… kada prepoznavanje tvrdnje o razlogu ne bi uključivalo prepoznavanje motiva, tada bi netko tko ima razlog za djelovanje mogao pitati: zašto učiniti ono što imam razloga učiniti?" (Korsgaard: 1996a: 317).

Ako postoji razlog da postupimo na određeni način, onda je sasvim smisleno reći da ponekad tako postupamo upravo iz *tog* razloga. Ideja se svodi na to da bi (normativni) razlozi također trebali figurirati u objašnjenju naših postupaka. Ovo je zaista prihvatljiva tvrdnja koju je prilično teško dovesti u pitanje. Poslužimo se jednim primjerom. Pretpostavimo da imam razlog napustiti prostoriju u kojoj se nalazim (recimo da je u njoj bomba). Pretpostavimo da na koncu i napustim tu prostoriju. Koji je razlog *zašto* sam to učinio? Što me motiviralo da to učinim? Izgleda razumnim pretpostaviti da sam prostoriju napustio zbog toga što je u prostoriji bomba. Činjenica da je u prostoriji bomba ne samo da *objašnjava* zašto sam prostoriju napustio, već ta ista činjenica moj postupak i *opravdava*. Premda se razlozi i motivi u mnogim slučajevima ne poklapaju, ovaj primjer samo pokazuje da se *barem ponekad* poklapaju odnosno da se u mnogim situacijama naše djelovanje može ne samo opravdati, već i objasniti (normativnim) razlozima.

Plauzibilno je pretpostaviti da nas razlozi moraju biti u stanju pokrenuti na djelovanje, a svaka teorija razloga mora biti u stanju pružiti detaljnije objašnjenje te pretpostavke. Najjednostavnijim riječima, svaka teorija razloga – da bi se uopće mogla smatrati dobrom teorijom – mora biti u stanju objasniti na koji nas točno način normativni razlozi uopće mogu

motivirati.[39] No zašto tom zahtjevu i nije toliko jednostavno udovoljiti, možda možemo bolje shvatiti ako se još malo zadržimo na gornjem primjeru. Činjenica da je u prostoriji bomba jest razlog zašto bih tu prostoriju trebao napustiti, ali, strogo gledajući, ono što objašnjava zašto sam tu prostoriju na koncu napustio nije *činjenica* da je u prostoriji bomba, već moje *vjerovanje* da je u prostoriji bomba. Ali opet, samo vjerovanje da je u prostoriji bomba nije dostatno da me motivira da napustim prostoriju ukoliko pritom nemam i nekakvu *želju* koju mogu realizirati napuštanjem prostorije (primjerice, želju da ostanem neozlijeđen). Normativni razlozi su, ako prihvatimo ovu sliku, *činjenice u svijetu*, dok razlozi koji objašnjavaju naše postupke (tj. motivirajući razlozi) nisu činjenice, već *psihološka stanja*. Sada je sasvim jasno gdje se skriva problem: dok se s jedne strane čini potpuno neuvjerljivim dovoditi u pitanje pretpostavku da nas normativni razlozi mogu motivirati, s druge strane proizlazi da to *nikada* ne može biti slučaj jer su normativni i motivirajući razlozi, da se poslužim Dancyevim riječima, "metafizički različite zvijeri" (Dancy 2000: 106).[40]

Svjesni ove poteškoće, neki su autori došli do zaključka da razlog – *da bi se uopće smatrao razlogom* – mora biti u stanju motivirati onoga na koga se odnosi. O razlozima za djelovanje, prema toj liniji razmišljanja, nema smisla govoriti prije nego što ih osoba ne prepozna kao svoje razloge: nije li osoba motivirana učiniti nešto što ima razloga učiniti, to znači da ona zapravo i nema razlog to učiniti. U tom smislu Williams, razvijajući neohjumovsku poziciju, nastoji pokazati da su razlozi za djelovanje uvijek unutarnji odnosno da je subjektova motivacija uvjet za prisustvo razloga. Svi vanjski razlozi, on tvrdi, u konačnici kolabiraju u one "unutarnje".

Williams započinje analizom rečenice "A ima razlog učiniti Φ". Tu rečenicu možemo interpretirati na dva različita načina – internalistički i eksternalistički. Prema internalističkom

[39] Jonathan Dancy (2000: 101) to je nazvao "objašnjavalačkim ograničenjem" ("explanatory constraint").
[40] Moglo bi se tvrditi da su psihološka stanja isto tako činjenice. Ovdje je, međutim, prvenstveno riječ o različitoj *prirodi* tih činjenica. Želi se sugerirati da su normativni razlozi činjenice koje su neovisne o našoj psihološkoj konstituciji, dok su motivirajući razlozi činjenice o našim psihološkim stanjima.

čitanju rečenica "A ima razlog učiniti Φ" istinita je ako i samo ako A ima motiv učiniti Φ. Jedino kada je istinito za nekoga reći da ima razlog djelovati na određeni način, Williams tvrdi, jest u onim okolnostima u kojima je on motiviran djelovati na taj način. Stoga je, prema internalističkoj interpretaciji, rečenica "A ima razlog učiniti Φ" lažna ako A nije motiviran učiniti Φ. Takva interpretacija pretpostavlja da su razlozi utemeljeni u subjektovoj psihološkoj konstituciji odnosno da su razlozi uvijek *internalni*. Eksternalističko shvaćanje negira da istinosna vrijednost te rečenice ovisi o prisutnosti motivacije: čak i ako subjekt nije motiviran učiniti Φ, to ne mora povlačiti da on nema razlog učiniti Φ. Normativni razlozi se, prema eksternalistima, nalaze izvan subjekta te se zato i nazivaju *vanjskim* razlozima.[41]

Jedna od prednosti internalističke pozicije jest u tome što – zbog svog shvaćanja ontološkog statusa normativnih razloga – na jednostavan i elegantan način udovoljava objašnjavalačkom zahtjevu o kojem je prethodno bilo riječi. Činjenica da razlozi mogu motivirati lako može biti objašnjena internalističkim čitanjem jer razlozi, kako ih internalisti shvaćaju, nisu *samo* činjenice u svijetu. Sama činjenica da je u prostoriji bomba, internalist će reći, za osobu još uvijek ne predstavlja dobar razlog da tu prostoriju napusti. Razlog za to može postojati samo ako ona ima odgovarajuću motivaciju (na primjer, želju da ostane neozlijeđena). Najjednostavnije rečeno, razlozi mogu motivirati zato što je motivacija *uvjet* za njihovu prisutnost. Da bi nešto za subjekta predstavljalo razlog, on taj razlog mora internalizirati odnosno uvidjeti da je to razlog za njega samog. Problem s eksternalističkim čitanjem, s druge strane, jest u tome što relaciju između razloga i motiva ostavlja nerazjašnjenom. Eksternalisti su primorani pružiti objašnjenje načina na koji nas razlozi uopće mogu motivirati odnosno moraju pokazati kako se, unutar te pozicije, može udovoljiti objašnjavalačkom ograničenju.

[41] Stajalište da postoje vanjski razlozi može se pripisati Kantu. No Kanta je moguće interpretirati i kao internalista. O tome će biti više riječi u sljedećem poglavlju.

Internalistička se pozicija već na prvi pogled suočava s određenim poteškoćama. Jer što ako je netko motiviran nešto učiniti, a da istovremeno nismo skloni kazati da za to postoji razlog? Pretpostavimo, koristeći Williamsov primjer, da želim popiti džin-tonik te da vjerujem da je u boci, koja se nalazi ispred mene, džin. Prema internalističkoj interpretaciji slijedi da imam razlog pomiješati sadržaj boce s tonikom. No pretpostavimo da je moje vjerovanje o sadržaju boce lažno (u boci se ne nalazi džin, već benzin). Imam li tada razlog pomiješati sadržaj boce s tonikom? Opis internalističke pozicije sugerira da je internalist na ovo pitanje primoran odgovoriti potvrdno.[42] Ipak, Williams pruža negativan odgovor te smatra kako ovaj prigovor ne može uzdrmati internalističku poziciju: normativni razlog za djelovanje ne ovisi o tome jesmo li trenutno motivirani na neko postupanje, već o tome bismo li bili motivirani na to postupanje kada bismo imali više saznanja o svijetu te kada bismo rasuđivali na ispravan način. Moje bi se želje u svjetlu novih informacija vrlo vjerojatno promijenile te, kada bih znao da se u boci nalazi benzin, tada više ne bih bio motiviran pomiješati sadržaj boce s tonikom.

Svaka osoba posjeduje nešto što Williams naziva subjektivnim motivacijskim skupom (S). Formalni izraz koji on koristi da bi okarakterizirao elemente S-a jest "želja", ali pojam želje ovdje je shvaćen u prilično širokom smislu jer S "može sadržavati stvari kao što su dispozicije za evaluaciju, uzorke emocionalnih reakcija, osobne privrženosti i, kao što se može apstraktno nazvati, različite projekte, koji utjelovljuju djelatnikove obveze" (Williams 1981: 105). Elemente S-a podvrgavamo određenoj korekciji, a motivacijski skup, koji proizlazi kao rezultat te korekcije, usko je povezan s motivacijskim skupom s kojim je rasuđivanje započelo. Korekcija motivacijskog skupa sastoji se u formiranju istinitih vjerovanja, odbacivanju lažnih vjerovanja te ispravnom rasuđivanju. Williamsovim riječima:

[42] Jonathan Dancy, primjerice, tvrdi da, promatrajući iz perspektive djelatnika, imam opravdan razlog pomiješati tonik s benzinom. Premda je moje vjerovanje o svijetu lažno, u ovoj situaciji nisam svjestan te činjenice i moj se postupak stoga može opravdati. Dancy nadalje napominje kako je stoga bolje koristiti termin "normativni razlog" umjesto "opravdani razlog". Naime, premda je moje vjerovanje lažno, ono me u ovom slučaju opravdava, ali to ne znači da imam normativni razlog pomiješati tonik s benzinom (Dancy, 2000: 107).

> Kao rezultat takvih procesa, djelatnik može vidjeti da ima razlog učiniti nešto što nije uopće vidio kao razlog za djelovanje. Na taj način, proces rasuđivanja može dodati nove radnje za koje postoje internalni razlozi, baš kao što može dodati nove internalne razloge za dane radnje. Proces rasuđivanja također može oduzeti elemente iz S-a. Promišljanje može dovesti djelatnika da vidi da je neko vjerovanje lažno i da tako shvati da ustvari nema razlog učiniti nešto za što je mislio da ima razlog. Još preciznije, on bi mogao pomisliti da ima razlog promicati neki razvoj stanja stvari zato što nije dovoljno upotrijebio maštu o tome kako bi bilo kada bi se to dogodilo. … mašta može stvoriti nove mogućnosti i želje. … Procesi rasuđivanja mogu imati svakojake učinke na S, i to je činjenica koju bi teorija internalnih razloga trebala s veseljem prihvatiti (Williams 1981: 104–105).

Procesi rasuđivanja ne uključuju samo instrumentalno rasuđivanje odnosno rasuđivanje o tome kako na najbolji način realizirati svoje želje, već uključuju i upotrebu mašte koja može proizvesti nove želje, ali i odbaciti neke želje koje trenutno imamo. Osim toga, ukoliko iz nekog razloga ne možemo realizirati neke naše želje i projekte, promišljanjem možemo utvrditi kojim željama i projektima pridajemo veću težinu te se pokušati usredotočiti na njihovu realizaciju. Isto tako, uspijemo li odbaciti lažna vjerovanja koja imamo o svijetu, tada je razumno za očekivati da će se neke naše želje, koje su na tim vjerovanjima zasnovane, također promijeniti. No formiranje novih želja pod utjecajem rasuđivanja ne nastaje *ex nihilo*, jer su nove želje na neki način već usađene u postojeći motivacijski skup. Ukoliko, upotrebljavajući maštu, formiram želju učiniti Φ, to znači da sam tu želju razvio iz već postojećih motiva koje sam imao prije nego što je proces rasuđivanja uopće započeo. Rasuđivanjem možemo samo "osvijestiti" naše već postojeće motive ili pak shvatiti da možda ipak nismo motivirani učiniti nešto za što smo prethodno bili uvjereni da nam ne nedostaje motivacije. Proces rasuđivanja mora započeti iz našeg postojećeg S-a odnosno iz elemenata

koji su nas u stanju motivirati na neko djelovanje jer taj proces, sam po sebi, ne može proizvesti nove motive koji u nekom smislu već ranije nisu bili prisutni u motivacijskom skupu.[43] Opisano shvaćanje internalnih razloga mnogo je uvjerljivije te mnogo teže ostavlja prostor za kritiku.[44]

Glavna karakteristika Williamsova internalizma očituje se u stajalištu da nikada nije istinito reći da netko ima razlog za djelovanje ukoliko u njegovu subjektivnom motivacijskom skupu ne postoji nešto što bi ga na to djelovanje moglo potaknuti. No, ako je to točno, kako onda objasniti porijeklo eksternalističke intuicije da je ponekad sasvim primjereno govoriti o razlozima za djelovanje čak i ako nedostaje bilo kakva motivacija? Mnogi su skloni reći da su razlozi takve vrste upravo prudencijalni i moralni razlozi te da se oni na nas odnose neovisno o tome jesmo li motivirani (ili pak bismo li mogli postati motivirani) postupati u skladu s njima. Postojanje prudencijalnih i moralnih razloga, prema ovoj liniji razmišljanja, nije ovisno o elementima nečijeg S-a, što znači da se svi imamo razloga brinuti o vlastitoj dobrobiti, baš kao što svi imamo i razlog za moralno ponašanje.

[43] Ovo je, kako ćemo kasnije vidjeti, središnja tvrdnja Williamsova internalizma oko koje se danas vode žustre rasprave i čije porijeklo nalazimo u Humeovoj tezi o "razumu kao robu strasti". Kantovci, o kojima će kasnije biti riječi, dovode u pitanje upravo ovu pretpostavku. Za kritiku Williamsova shvaćanja racionalne deliberacije također vidi Hooker, B. (2001) i Cohon, R (1986).

[44] Imajući to na umu, moram priznati da mi nisu jasne neke interpretacije prema kojima se Williamsa svrstava u skupinu onih autora koji zastupaju mogućnost kognitivne motivacije. Tako Shafer-Landau, na primjer, piše kako je pomalo iznenađujuće da Williams u "utjecajnom članku koji je inače posvećen obrani hjumovskog poimanja praktičnih razloga naglašava plauzibilnost tvrdnje da vjerovanja mogu konstituirati motivaciju za djelovanje" (Shafer-Landau, 2003: 122). Sličan stav zastupa i John Skorupski koji smatra kako se Williams slaže da "A-ovo vjerovanje o razlozima – na primjer njegovo vjerovanje da nagaziti nečiju nogu jest razlog za ispriku – može 'proizvesti ili uistinu konstituirati motivaciju za djelovanje'. Dopuštajući ovo, odnosno uključujući vjerovanje u A-ov S, čini se da se on udaljava od hjumovstva u pogledu motivacije" (Skorupski, 2007: 82). Iako neki pasusi u Williamsovu članku zaista sugeriraju da vjerovanja mogu biti dio nečijeg motivacijskog skupa, takva je interpretacija, meni se čini, daleko od istinite. Radi se o tome da na prvom mjestu uopće ne bih formirao vjerovanje da *imam razlog* učiniti Φ kada prethodno ne bih imao odgovarajuću motivaciju. Ovdje je prvenstveno riječ o redoslijedu kojim se proces motivacije odvija. Nije tako da prvo formiram vjerovanje o postojanju razloga, a onda postajem motiviran tim razlogom, već upravo suprotno. Stoga što jesam motiviran *uopće imam razlog*, ali, uslijed nedostatka relevantnih informacija, ne moram odmah vjerovati da je taj razlog prisutan. Kada jednom formiram vjerovanje da imam razlog, to znači da sam postao *svjestan* činjenice da sam *već* motiviran, a ne da me *tek* to vjerovanje može motivirati. Na osnovi cjelokupne Williamsove rasprave, mislim da upravo takvu interpretaciju on ima u vidu kada kaže da "vjerovanje može proizvesti motivaciju". Williams bi zastupao mogućnost kognitivne motivacije jedino u slučaju kada bi tvrdio da vjerovanje o postojanju razloga, *samo po sebi*, može proizvesti motivaciju, a u tekstu ne postoji nikakva naznaka da bi se njemu takav stav mogao pripisati.

Poslužimo se jednom ilustracijom. Zamislimo da se A-ov prijatelj nalazi u velikoj nevolji i da mu jedino A može pružiti pomoć. Pretpostavimo da A nije motiviran to učiniti jer pružanje pomoći njegovu prijatelju ne zadovoljava nijednu njegovu preferenciju. Pomaganje prijatelju jednostavno nije kompatibilno s njegovim ciljevima i projektima. Ima li A, unatoč tome, razlog pružiti pomoć? Eksternalist će na ovo pitanje odgovoriti potvrdno i isticati da je razlog o kojemu je ovdje riječ *moralni* razlog – razlog koji postoji neovisno o nečijem motivacijskom skupu. Williams se, s druge strane, nije spreman složiti s takvim gledištem te ističe kako ne možemo sa sigurnošću tvrditi da je u S-u svake osobe sadržana motivacija za prudencijalno i moralno djelovanje:

> Netko može reći da je svaki racionalni prosuditelj vezan ograničenjima morala … Ali ako je to tako, onda su ograničenja morala dio svačijeg S-a i svaki ispravan moralni razlog bit će internalni razlog. Ali mora postojati argument za takav zaključak. Netko tko tvrdi da su ograničenja morala sama po sebi ugrađena u pojam onoga što znači biti racionalnim prosuditeljem ne može izvesti takav zaključak tek tako. Jednostavno ukazujući da imamo pojam dobrog razloga koji uključuje prudencijalne i moralne tvrdnje … ne utječe ni na što (Williams 1995a: 37).

Premda u svakodnevnom životu (pogotovo u moralnim situacijama) često rabimo pojam razloga u eksternalističkom smislu – naime, pretpostavljamo da su razlozi odvojeni od naše psihološke domene – takva pretpostavka ipak nije utemeljena. Svi se prividno vanjski razlozi, uključujući prudencijalne i moralne, na koncu svode na one internalne. Na primjer, moje inzistiranje na tome da neka osoba ima razlog uzeti lijek koji joj je potreban, unatoč činjenici da ona pokazuje nezainteresiranost za očuvanje vlastitog zdravlja, može biti pokazatelj mojeg uvjerenja kako je taj razlog "izvan" samog subjekta. No takvo je uvjerenje pogrešno. Uvjeravanjem nekoga da mu je lijek potreban ne želimo da on prepozna razlog koji postoji

neovisno o njegovoj motivaciji, već se nadamo da on "na nekoj razini *mora* željeti biti dobro" (Williams 1981: 106). Kada bi on na koncu pristao uzeti lijek, to ne bi pokazalo da je cijelo vrijeme bilo istinito da je on imao razlog to učiniti. Upravo suprotno, internalist će tvrditi da on taj razlog nije imao sve do trenutka dok nije stvorio motivaciju to učiniti. Slično će, prema internalizmu, vrijediti i za moralne razloge. Moralni razlozi nemaju veći "autoritet" od drugih razloga samo zato što su moralni; kao što je za prisutnost bilo kojeg razloga potrebna odgovarajuća motivacija, tako je i za prisutnost moralnog razloga potrebno da subjekt u svom motivacijskom skupu ima želju koja može biti zadovoljena moralnim ponašanjem. Ako takav element nije prisutan, ne može biti prisutan ni moralni razlog.[45]

Implikacija internalističke pozicije sasvim je jasna. Budući da "vanjski" razlozi jednostavno ne postoje, nije jasno zašto bi moralni i prudencijalni razlozi u tom smislu predstavljali izuzetak. Stoga možda i ne bi trebalo čuditi da Williamsovo stajalište nije naišlo na odobravanje mnogih autora koji su skloni kantovskoj tradiciji. Kant je, kao što je poznato, nastojao dokazati da moralni zahtjevi obvezuju sva racionalna bića te da nam moralnost postavlja zahtjeve kojima ne možemo izbjeći jednostavno zato što udovoljavanje tim zahtjevima ponekad nije usklađeno s našim željama. Kada bi internalizam Williamsova tipa bio točan, kantovci su uvjereni, o moralnim zahtjevima zapravo ne bi imalo smisla govoriti.

U svakom slučaju, navedena razmatranja otvaraju brojne poteškoće i pitanja na koja ću se vratiti u sljedećem poglavlju kada ću govoriti o kantovskom internalizmu. Prije toga ću izdvojiti četiri prigovora koji su bili upućeni Williamsovoj poziciji te ću nastojati pokazati da nijedan od njih ne uspijeva uzdrmati njegovo stajalište. Također vjerujem da kritičko

[45] Ovdje je važno istaknuti Parfitovu distinkciju između "supstantivne" i "proceduralne" racionalnosti. Biti supstantivno racionalan znači "brinuti [se] o određenim stvarima, kao što je naša vlastita dobrobit" dok proceduralna racionalnost znači "rasuđivati na određene načine", ali ne nalaziti se pod "[zahtjevom] za imanjem bilo kakvih pojedinačnih želja ili ciljeva, kao što je briga za našu vlastitu dobrobit" (Parfit, 1997: 101). Parfit proceduralizam pripisuje internalistima, a supstantivizam eksternalistima. Distinkciju je moguće opisati i na sljedeći način. Internalisti vjeruju da imamo moralne i prudencijalne razloge *ako* smo motivirani ponašati se moralno i prudencijalno (kao rezultat proceduralnog rasuđivanja). Eksternalisti, s druge strane, tvrde da smo, ukoliko supstantivno racionalni, motivirani ponašati se moralno i prudencijalno *zato* jer za to imamo razloga. Eksternalizam, prikazan u ovom obliku, ponekad se naziva i kantovskim internalizmom. Kritici te pozicije posvetio sam sljedeće poglavlje.

razmatranje tih prigovora može poslužiti u dodatnom rasvjetljavanju glavnih karakteristika internalizma odnosno jasnijem formuliranju ove pozicije.

4.2.1. *Argument iz krivnje*

Internalističko stajalište svodi se na tezu da su razlozi za djelovanje subjektivni i da njihovo porijeklo treba tražiti u elementima koji čine nečiji motivacijski skup. Na prvi pogled čini se razumljivim zašto je takvo stajalište teško prihvatiti. Ti se elementi, po svemu sudeći, razlikuju od osobe do osobe, što znači da razlozi u krajnjoj instanci ovise o nečemu sasvim arbitrarnom i da takvo shvaćanje vodi u neku vrstu relativizma. Kada bi postojanje razloga bilo ovisno o našim motivima, tada bi se svatko mogao ponašati u skladu s vlastitim preferencijama i ciljevima; za svaki bi se postupak, koliko god bio moralno neprihvatljiv, mogao pronaći dobar razlog jer bi samo bilo dovoljno pozvati se na motive onoga koji je tako postupio. No ljudsko je ponašanje podložno etičkoj procjeni, pa tako uvijek ima smisla reći da za neke postupke jednostavno ne postoje dobri razlozi odnosno da postoje dobri razlozi *ne* postupati na određeni način. Osim toga, pripisivanje krivnje u nekom smislu pretpostavlja postojanje vanjskih razloga – zašto osobu na koncu i smatramo krivom jest zato što vjerujemo da ona nije imala dobar razlog postupiti onako kako je postupila odnosno da je imala razlog izbjeći određenu vrstu djelovanja.

Pod pretpostavkom prihvaćanja internalizma, praksa pripisivanja krivnje očito bi izgubila svoj smisao i opravdanje. Mi jednostavno ne bismo mogli tvrditi da netko – neovisno o tome koje je njegovo motivacijsko stanje – nema razlog činiti ono što čini: prisutnost motivacije da Φ povlači i prisutnost razloga da Φ i zato, ako je internalizam istinit, ne postoji osnova za tvrdnju da razlog za Φ ipak ne postoji. Na tragu ovog razmišljanja vrlo zanimljivu kritiku internalistima je uputio Russ Shafer-Landau (Shafer-Landau, 2003: 187–189). Kako on ističe, mi nećemo nekog mizantropa ili pak nekoga tko se ponaša na najbrutalniji mogući

način prestati smatrati krivim zato jer mu nedostaje motivacija za promjenom takvog ponašanja. Internalist, prema njegovu mišljenju, stoga može birati između dvije, ne baš tako obećavajuće, strategije: s jedne strane, on mora ili revidirati standarde na osnovi kojih se krivnja pripisuje ili, s druge strane, on mora potpuno izuzeti od krivnje neke najgore oblike ljudskog ponašanja. Priklanjanje prvoj opciji impliciralo bi da internalist neku osobu može smatrati krivom za činjenje Φ-a premda ona u danoj situaciji nije imala razlog ne činiti Φ. Priklanjanje toj opciji nije privlačno jer bi ostalo sasvim nerazumljivo na čemu bi pripisivanje krivnje u tom slučaju uopće bilo zasnovano. Prihvaćanje druge opcije impliciralo bi da bi najgori zločinci mogli biti oslobođeni, ne samo krivnje, već i kazne jer kazniti nekoga ima smisla jedino ukoliko je prethodno utvrđena krivnja.

Vjerujem da bi internalisti ovaj prigovor mogli izbjeći tako da negiraju samu njegovu pretpostavku, naime da tvrde kako opcije pred koje ih eksternalisti stavljaju i nisu toliko kontroverzne kao što to može izgledati. Takav bi odgovor zahtijevao opširniju analizu uvjeta i okolnosti pod kojima je krivnju primjereno pripisivati i iz koje bi moglo slijediti da internalisti tu praksu mogu bolje objasniti od svojih protivnika. Na ovom se mjestu u takvu *opširniju* analizu ne mogu upuštati, ali ću se u narednom dijelu ipak upustiti u nagađanje o mogućim rezultatima takve analize. Kako i sam Williams napominje, u većini slučajeva krivnju ipak pripisujemo ljudima koji imaju odgovorajući element u svom motivacijskom skupu. U tim je slučajevima takva praksa opravdana budući da osoba možda nije dovoljno dobro rasuđivala odnosno, prije odluke o tome kako postupiti, nije razmotrila sve relevantne činjenice koji bi mogli znatno utjecati na ishod odluke. Takve situacije pokazuju da je osoba ipak imala razlog postupiti drukčije iako u danim okolnostima možda nije bila svjesna da u svom S-u ima motivaciju koja bi bila zadovoljena činjenjem neke druge radnje. S druge strane, osoba možda stvarno nije imala razlog postupiti drukčije jer joj motivacija da to učini tada nije bila najjača. Svrha okrivljavanja i tada ima uporište jer se time želi postići da ljudi u

nekoj idućoj sličnoj situaciji razmotre stavove koji su sadržani u samoj toj osudi te koji bi ih stoga mogli nagnati na drukčije djelovanje. Osim toga, moguće su i situacije u kojima netko uistinu nema nikakvu direktnu motivaciju koja bi bila zadovoljena činjenjem neke radnje, ali zato ima općenitiju motivaciju da izbjegne osudu i kritiku te da s ljudima koje cijeni ostvari dobar i prijateljski odnos. Osuđivanje i pripisivanje krivnje u tim je okolnostima također primjereno jer se time želi ukazati da je nekakav razlog za drukčije ponašanje ipak postojao (Williams, 1995a).

Internalističko shvaćanje krivnje, prema mojem mišljenju, ne bi smjelo *a priori* zanemariti. Tim se shvaćanjem pretpostavlja da krivnju u većini situacija pripisujemo ljudima za koje vjerujemo da će to na njih polučiti određeni *učinak*.[46] Ukazivanjem na pogreške u ponašanju i na moralno pogrešne postupke vjerujemo da je većina ljudi sposobna shvatiti da njihovo ponašanje u određenim okolnostima možda nije bilo primjereno i da bi u budućnosti trebali izbjegavati donošenje sličnih odluka. Svrha pripisivanja krivnje velikim dijelom sastoji se upravo u tome. No što je s ljudima kod kojih se takav učinak ne može postići? Što je s onima koji u svojem motivacijskom skupu ne posjeduju nijedan element koji bi ih mogao potaknuti na promjenu ponašanja te koji su nesposobni shvatiti da njihovi postupci nisu društveno prihvatljivi? Znači li to, kako tvrdi Shafer-Landau, da su takvi ljudi izuzeti od krivnje i kazne? Na to je pitanje teško pružiti odgovor kojim se pritom ne bi pretpostavljala neka od brojnih filozofskih teorija kazne i moralne odgovornosti, ali možda vrijedi spomenuti sljedeće.[47] Čini mi se da je ovdje potrebno razlikovati značenja u kojima se pojam krivnje može pojavljivati. Vjerujem da bi bilo pogrešno tvrditi da je krivnja isključivo *moralna* kategorija budući da krivnju ima smisla pripisivati čak i u izvanmoralnim situacijama, tj. u

[46] Uočimo da takvo shvaćanje krivnje eksternalisti zapravo i ne mogu u potpunosti prihvatiti. Netko bi mogao zaista ustanoviti da ima razlog postupati drukčije nego što postupa, ali to bi ga i dalje moglo ostaviti nemotiviranim da promijeni svoje ponašanje.

[47] Shafer-Landauov argument između ostalog počiva na pretpostavci da je postojanje krivnje preduvjet kažnjavanja. Iako je tu pretpostavku izuzetno teško dovesti u pitanje, sigurno nije nemoguće. Na primjer, pod pretpostavkom prihvaćanja utilitarističkog opravdanja kazne, ponekad bi bilo sasvim opravdano kazniti nevinu osobu ako bi se time mogao postići željeni učinak.

situacijama u kojima osobu ne tretiramo kao *moralni* subjekt. Stoga mi se čini da Shafer-Landau u svom argumentu protiv internalizma ne koristi pojam krivnje u istom značenju. On smatra da zastupanje internalizma povlači da bi oni "najgori od najgorih … tako bili imuni na krivnju … a vjerojatno i na kaznu" (Shafer-Landau, 2003: 187). Ali to ne mora biti sasvim točno. "Najgori od najgorih" mogli bi biti imuni na *moralnu* krivnju, ali ne i krivnju shvaćenu u drugom, kauzalnom smislu. Kažnjavanje bi u takvim situacijama imalo osnovu jer svrha kažnjavanja ne mora biti ista kao i svrha pripisivanja krivnje. Kažnjavanje ponekad može biti opravdano jer se time zaštićuje društvo od mogućih postupaka osobe koja je kažnjena te kažnjavanje ne mora ujedno povlačiti i *moralnu* krivnju kažnjene osobe. Naravno, postoji mnogo graničnih slučajeva u kojima je teško procijeniti ima li pripisivanje moralne krivnje uopće smisla. U takvim situacijama može biti relevantno pokazuje li osoba žaljenje ili kajanje jer time daje do znanja da bi moralna kritika mogla u budućnosti polučiti određene rezultate. No možda ne bi bilo čudno smatrati da određeni broj ljudi ipak ne bi trebao biti predmetom moralne osude, već da bismo ih, umjesto toga, radije trebali smatrati, kako kaže Williams, "beznadežnim ili opasnim likovima" (Williams, 1995a: 43).

4.2.2. *Argument iz apsurda*

U situacijama u kojima A smatra da B ima razlog učiniti Φ, a B smatra da to nije istina, A, prema internalizmu, samo vjeruje da negdje u B-ovu S-u ipak postoji element koji bi ga mogao navesti da učini Φ. Ali što ako u B-ovu S-u nedostaje potrebni element? Što ako B ne promijeni svoju motivaciju nakon što ju je podvrgnuo korekciji ispravnog rasuđivanja? Ima li tada i dalje smisla tvrditi da B ima razlog učiniti Φ? Internalisti i eksternalisti, vidjeli smo, pružaju različite odgovore na to pitanje. Eksternalist smatra da stajalište prema kojem su svi razlozi u konačnici povezani sa subjektovim motivima zvuči kontraintuitivno i neistinito. Dovoljno je, on će kazati, samo prisjetiti se različitih situacija koje nedvosmisleno sugeriraju

da razlozi postoje "izvan nas" odnosno nezavisno o našim subjektivnim psihološkim kategorijama. Sličnu liniju razmišljanja prihvaća i John Searle, koji ukazuje na jednu takvu situaciju i pritom ističe apsurdne posljedice internalizma:

> Pretpostavite da uđete u bar i naručite pivo. Konobar vam donese pivo i vi ga popijete. Tada vam konobar donese račun, a vi mu kažete: 'Pogledao sam u svoj motivacijski skup i ne nalazim unutarnji razlog da platim ovo pivo. Platiti pivo nije nešto što želim niti je to sredstvo za neki cilj koji imam u svojemu motivacijskom skupu. Pročitao sam profesora Williamsa i također sam pročitao Humea u pogledu ovog pitanja, te sam pažljivo pogledao u svoj motivacijski skup, ali jednostavno ne mogu pronaći želju da platim ovo pivo! Jednostavno ne mogu!' (Searle, 2001: 27).

Nije čudno da Searle opisanu situaciju smatra apsurdnom. Tvrdnja da su svi razlozi za djelovanje vezani za subjektivni motivacijski skup internalista dovodi u prilično nezahvalnu poziciju, a to Searle i namjerava pokazati gornjim primjerom. Budući da osoba nema želju platiti pivo, slijedi da nema niti razlog platiti pivo, što je, tvrdi Searle, apsurdno upravo zbog toga što ste narudžbom "intencionalno *stvorili* razlog koji je nezavisan od vaših želja, razlog da učinite nešto bez obzira na ono što se nalazilo u vašem motivacijskom skupu kada je došlo vrijeme da to učinite" (Searle, 2001: 28).

(a) Jedna strategija koju internalist ima na raspolaganju jest argumentirati da gornja situacija zapravo uopće nije toliko apsurdna kakvom je Searle namjerava prikazati. On je sigurno u pravu kada vjeruje da postoje mnoge situacije i okolnosti u kojima nam se čini da je pripisivanje razloga za djelovanje neovisno o motivaciji. Ali ako nam se čini da u nekim situacijama i okolnostima razlozi nisu ovisni o motivaciji, to ne znači da to uistinu *jest* tako. Meni se čini da se Searle ovdje opasno približava onome što Williams naziva "blefom": on zamišlja okolnosti u kojima je pripisivanje eksternalnih razloga intuitivno očito, ali takva je intuicija samo blef jer su razlozi o kojima je riječ *prividno* vanjski. Da bih to pokazao, ostat ću

u okviru njegova primjera. Ako A iskreno tvrdi da plaćanje narudžbe nije vezano uz njegove preferencije ili želje, da ne samo da mu je svejedno hoće li zbog toga biti kažnjen, već da je spreman prihvatiti sve moguće posljedice koje mogu proizići iz njegova neplaćanja itd., zašto i dalje inzistirati na tome da on *ima razlog* platiti ono što je naručio? Kakav bi to razlog uopće mogao biti? Mi i dalje možemo kritizirati A zato što u svom S-u nema elemente za koje vjerujemo da bi bilo bolje da ih ima, ali ako je zaista tako da A putem racionalne deliberacije ne može doći do nekog motivacijskog impulsa za radnju, tvrdnja o vanjskom razlogu u najmanju ruku počinje zvučati misteriozno. Postavimo li stvari na ovaj način, čini mi se da ono što je u početku bila intuicija za eksternalizam sada postaje intuicija za internalizam. Pretpostavimo da u nekom trenutku A ipak stvori motivaciju to učiniti. Tek bi u tom trenutku bilo istinito kazati da A ima razlog jer bi to bio pokazatelj da je A ipak uspio pronaći neki element u svom motivacijskom skupu koji bi plaćanjem mogao biti zadovoljen. Williams dakle ne negira da se u motivacijskom skupu može dogoditi promjena, već samo negira eksternalističku tvrdnju da osoba može uvidjeti kako je cijelo vrijeme bilo istinito da je ona imala razlog učiniti nešto što ona nije htjela učiniti. Osnovna Williamsova tvrdnja jest da to ne može biti istinito sve dok osoba ne usvoji motivaciju da djeluje na odgovarajući način.

(b) Čak i ako prethodna strategija ne donosi povoljne rezultate, to još uvijek ne znači da je to jedina strategija kojom je Searleov argument moguće oslabiti. Drugi internalistički pristup tom argumentu mogao bi se sastojati u tome da se dovede u pitanje ekvivalentnost tvrdnji "A treba učiniti Φ" i "A ima razlog učiniti Φ". Ta bi strategija uspješno objasnila zašto Searleov primjer pobuđuje intuiciju koja ide u prilog eksternalističkom kampu, ali bi istovremeno ostavila internalizam netaknutim. Internalistička bi pozicija tom strategijom bila obranjena, a vjerovanje da postoje dužnosti koje se na nas odnose bezuvjetno ostalo bi sačuvano. Tako bi se, primjerice, moglo tvrditi da Searleov argument pokazuje samo to da A *treba* platiti ono što je naručio, ali ne i to da A ima *razlog* platiti ono što je naručio.

Zanimljivo je da sam Williams u jednom drugom kontekstu razmatra mogućnost razdvajanja ova dva normativna pojma:

> Kada kažemo da je netko trebao djelovati na neki zahtijevan ili poželjan način, a na koji nije djelovao, ponekad kažemo da je *postojao razlog* za njega da djeluje na taj način – obećao je, na primjer, ili to što je stvarno učinio povrijedilo je nečija prava. Premda to možemo reći, ne čini se da je to povezano ni na koji čvrst način s idejom da je *on imao razlog* djelovati na taj način. Možda uopće nije imao razlog. Kršenjem dužnosti, nije se nužno ponašao iracionalno ili nerazborito, već loše (Williams, 1985: 192).

Kazati da netko nešto treba učiniti ponekad može biti sasvim opravdano iako to ne mora uvijek značiti da on ima i razlog to učiniti. Prisjetimo li se rasprave Philippe Foot o kategoričkim i hipotetičkim imperativima, moguće je njezinu glavnu misao primijeniti i kao odgovor na Searleov, kako sam ga nazvao, argument iz apsurda. Ideja je da postoje situacije u kojima se imperativi na nas *odnose* u kategoričkom smislu iako ne mora biti istinito da nam ti isti imperativi pružaju razlog da u skladu s njima djelujemo. Searleov primjer govori upravo o jednoj takvoj situaciji. Možda na prvi pogled doista zvuči nekonzistentno kažemo li nešto poput "Razumijem da trebam učiniti Φ, ali ne vidim da imam razlog učiniti Φ", ali mislim da se nelogičnost takve formulacije gubi ako tu formulaciju shvatimo na sljedeći način: "Razumijem da postoji zakon koji mi govori da trebam platiti ono što sam naručio, ali ne mislim da imam razlog taj zakon slijediti". Ta je tvrdnja sasvim konzistentna. Nekako nam izgleda da legalne norme konzistentno možemo dovoditi u pitanje na način da tražimo razloge koji bi opravdali njihovo slijeđenje. Opravdanje neke norme mora biti opravdanje *za mene*, opravdanje koje mogu prihvatiti i tek bi tada imalo smisla kazati da imam razlog postupati onako kako mi ta norma nalaže.

Iz Searleove se diskusije lako da iščitati njegova odbojnost prema činjenju spomenute distinkcije. Upotrebljavajući na jednom mjestu primjer obećanja, on dolazi do sasvim

suprotnih zaključaka: "*U slučaju obećanja, želja je izvedena iz prepoznavanja razloga koji je neovisan o željama, naime, dužnosti*" (Searle, 2001: 170, kurziv u originalu). Searle dakle sugerira da ne postoji nikakva razlika između *dužnosti* izvršiti obećanje i *razloga* izvršiti obećanje. Nitko ne misli da postoji imperativ koji nam nalaže da obećanja trebamo davati, ali ako smo dali obećanje, tada smo na sebe preuzeli obvezu da to obećanje ispunimo, a preuzimanje te obveze zapravo nije ništa drugo nego intencionalno stvaranje razloga koji sada postoji neovisno o našem motivacijskom skupu.

Prije svega je potrebno navesti da se Searle pozivanjem na primjer obećanja nadovezuje na svoju poznatu argumentaciju o premošćivanju procjepa između "jest" i "treba" (Searle, 1969). U vrlo kratkim crtama, njegova je namjera pokazati da je skok iz "čisto" deskriptivnih premisa na normativnu konkluziju ponekad sasvim opravdan. Takav skok je moguć u okviru onoga što on naziva *institucionalnim činjenicama*. Ne ulazeći detaljnije u slijed njegove argumentacije, u ovom je kontekstu sasvim dovoljno kazati sljedeće. Kao što u okviru svake institucije vrijede određena konstitutivna pravila, tj. pravila bez kojih ta institucija ne bi postojala, tako i institucija obećanja ne bi postojala kada davanje obećanja ne bi ujedno impliciralo *obvezu* da se to obećanje izvrši. Primjerice, kao što se, kada igram šah, obvezujem da ću igrati po pravilima koja reguliraju igranje šaha, na isti se način, kada dajem obećanje, obvezujem se da ću to obećanje i izvršiti. Međutim, prihvatimo li sugestiju prema kojoj "imati obvezu učiniti Φ" i "imati razlog za Φ" nisu ekvivalentne tvrdnje, tada bismo primjer obećanja mogli promotriti u nešto drukčijem svjetlu. Ima li smisla kazati da se obećanja trebaju izvršavati, ali ne i da ih imamo razloga izvršavati? Ne vidim zašto bismo na to pitanje morali odgovoriti negativno.

Neke prigovore koji su Searleu bili upućeni u kontekstu njegove rasprave o jazu "jest-treba" (Flew, 1969; Hare, 1969; McClellan i Komisar, 1969; Thomson i Thomson, 1969) možemo sasvim lako spomenuti i na ovom mjestu. Searleovi su kritičari ukazali da on ipak

nije uspio pokazati da davanje obećanja implicira kategoričko trebanje, već samo hipotetičko. Obećanja trebamo izvršavati jedino ukoliko želimo da nam drugi ljudi u budućnosti vjeruju ili pak ukoliko želimo izbjeći osudu. Budući da nam je u interesu da nam drugi ljudi vjeruju, proizlazi da obećanja trebamo ispunjavati. Moj je prijedlog da bismo taj prigovor Searleu mogli preformulirati na sljedeći način. Čak i pod pretpostavkom da je Searle zaista pokazao da obećanja *kategorički trebamo* izvršavati, on nije pokazao da imamo *kategoričke razloge* obećanja izvršavati. Norma koja proizlazi iz davanja obećanja jest društvena norma koja se na nas *odnosi* kategorički. Upravo zato što se ta norma odnosi na sve ljude neovisno o tome koji su njihovi individualni ciljevi, imamo pravo kritizirati one koji tu normu ne slijede. Time, međutim, nije pokazano da svi imaju i razlog tu normu slijediti. Razlog za to imaju samo oni čije bi želje slijeđenjem te norme mogle biti zadovoljene. Razlozi za izvršavanje obećanja samo su hipotetičke prirode jer ih imamo jedino ako, na primjer, želimo izbjeći društvenu kritiku i zadržati povjerenje ljudi do kojih nam je stalo. Searle je dakle u krivu kada smatra da su razlozi za izvršavanje obećanja kategorički jer postojanje kategoričkih razloga ne bi smjelo sadržavati pozivanje na naš motivacijski skup, a primjer obećanja taj uvjet ipak ne zadovoljava. Oboružani tom distinkcijom, sada možemo i uvidjeti zašto Searle griješi u sljedećem razmišljanju. Kada bi razlozi za održavanje obećanja bili hipotetički, tada bi, on vjeruje, slijedilo da ne bismo imali nikakvu obvezu izvršavati obećanja u situacijama u kojima bismo bili uvjereni da možemo proći nezapaženo. Davanje obećanja umirućoj osobi ne bi povlačilo postojanje nikakve obveze jer za to obećanje nitko nikada ne bi morao saznati (Searle 2001: 195–196). Internalist se, međutim, bez imalo poteškoća može složiti sa Searleom i tvrditi da bi davanje obećanja doista povlačilo postojanje obveze čak i u okolnostima u kojima nitko za to obećanje ne bi znao. Jedino što internalist mora negirati jest da bi davanje takvog obećanja pružalo razlog za njegovo izvršavanje, razlog koji bi bio neovisan o sadržaju S-a.

4.2.3. *Argument iz 'tvrdih' etičkih pojmova*

Zanimljiva je posljednja Williamsova opaska u ranije citiranom odlomku da kršenje dužnosti nije uvijek iracionalan postupak, već da je to ponekad samo loš postupak. To nas dovodi do još jednog obilježja internalizma koje sam već ranije spomenuo: prihvaćanje te pozicije ni na koji način ne povlači da elementi nečijeg S-a ne mogu biti predmetom kritičke prosudbe i vrednovanja. Opet ostajući vjerni Searleovu primjeru, internalist onoga koji ne plaća svoje narudžbe ili onoga koji ne drži svoja obećanja slobodno može nazvati moralno lošim, ali misliti da je netko moralno loš nije isto što i misliti da je on iracionalan. Tu misao Williams razvija razmišljajući o tome što bi bilo prigodno reći nekome tko se loše odnosi prema svojoj supruzi i koji istovremeno nema nijedan element u svom S-u koji bi mogao biti ostvaren promjenom takvog ponašanja: "Postoje mnoge stvari koje mogu reći o [toj osobi] ili toj osobi: da je nezahvalan, bezobziran, tvrd, seksist, nepristojan, sebičan, brutalan i mnoge druge nepovoljne stvari" (Williams, 1995: 39). Williamsova je ideja da je, pored svih navedenih kvalifikacija, sasvim suvišno kazati još i to da on ima *razlog* bolje tretirati svoju suprugu. Govor o razlozima ne bi pridodao ništa novo, a što već nije iscrpljeno spomenutim pojmovima. Osim toga, inzistira li na tome da on ima razlog tretirati svoju suprugu bolje nego što to trenutno čini, eksternalist time sugerira da je ta osoba iracionalna budući da taj razlog ne prepoznaje. Ali optužba za iracionalnost u ovom se kontekstu ipak čini krivo usmjerenom. Nekako smo prije skloni reći da se radi o nekome tko svojim ponašanjem manifestira stanoviti moralni nedostatak, a ne o nekome tko je iracionalan.

U svojoj raspravi o razlozima, Scanlon određeni prostor posvećuje upravo navedenom Williamsovu razmišljanju. Polazeći od određenja iracionalnosti prema kojemu se iracionalnost javlja u slučajevima "kada osoba prepoznaje nešto kao razlog ali njime ne biva pogođena na jedan od relevantnih načina", Scanlon predlaže jednu drukčiju verziju eksternalizma. On se u potpunosti slaže s Williamsom da bi nekoga tko se loše ponaša prema

svojoj supruzi bilo neopravdano proglašavati iracionalnim, ali također navodi da eksternalist u prvom redu to uopće nije obvezan tvrditi. Jedino što je eksternalist obvezan tvrditi jest to da takva osoba ima praktični razlog koji je neovisan o njegovoj motivaciji. Za osobu iz primjera bi stoga bilo primjerenije kazati da je neosjetljiva na postojanje razloga za djelovanje, a ne da je iracionalna jer ne djeluje u skladu s tim razlogom. Jer, kao što je rečeno, iracionalnost se javlja samo onda kada nismo pokrenuti onime što sami *prepoznajemo* kao razlog za djelovanje, ali ne i onda kada uopće ne uviđamo da imamo razlog za djelovanje.[48] Onaj koji se ne ponaša primjereno prema svojoj supruzi uopće ne uviđa da ima razlog za promjenu svog ponašanja i to ga, prema Scanlonu, oslobađa optužbe za iracionalnost. Međutim, to što je on oslobođen optužbe za iracionalnost, Scanlon vjeruje, ne znači da je s njim sve u redu. Sama činjenica da on ne uspijeva uvidjeti da ima razlog dokaz je njegova karakternog nedostatka. Pitanje koje se sada nameće jest: kako Scanlon dokazuje da vanjski razlog za drukčije djelovanje ipak postoji u ovom slučaju? Njegova je ideja da dokaz za postojanje vanjskog razloga, pomalo ironično, proizlazi iz epiteta čiju upotrebu dozvoljava sam Williams:

> On je, međutim, vrsta osobe za koju bi nam Williams dopustio da kažemo da je neobazriva, okrutna, neosjetljiva i tako dalje. Ove kritike uključuju optuživanje [te osobe] za vrstu nedostatka, naime neuspjeh da bude pokrenut određenim razmatranjima koje smatramo razlozima. … Ako je nedostatak za osobu da ne uspijeva vidjeti ta razmatranja kao razloge, čini se da ona moraju biti razlozi za njega. (Ako nisu, kako može biti njegov nedostatak da ih ne uspijeva prepoznati?) (Scanlon, 1998: 367).

[48] Scanlon se dakle oslanja na ono što naziva "usko protumačenom" racionalnošću (Scanlon, 1998: 25). Ideju sličnu Scanlonovoj razvijaju Hooker i Streumer (2004). Oni nude primjer osobe koja ima mogućnost znatno umanjiti patnju nevinih ljudi jednostavnim pritiskom gumba, ali koja nema apsolutno nikakvu želju to učiniti. Hooker i Streumer se slažu da bi takvog čovjeka bilo neosnovano prozvati iracionalnim. Jer to što on nema odgovarajuću motivaciju nije u proturječju s činjenicom da on može biti vrlo hladan i proračunat u svom ponašanju. Ipak, ostaje dojam da s takvim čovjekom nešto nije u redu. Unatoč njegovoj nemotiviranosti, Hooker i Streumer argumentiraju, on je podložan racionalnoj kritici zato jer ipak ima razlog pritisnuti gumb.

Argumentom se ističe da upotreba gore navedenih pojmova pretpostavlja eksternalizam. Kritiziramo li nečije ponašanje kao okrutno ili neobazrivo, samim time sugeriramo da je onaj koga kritiziramo imun na određene činjenice koje mi smatramo razlozima protiv takvog postupanja. U nešto drukčijim terminima, nekoga koga smatramo okrutnim ili neobazrivim zapravo optužujemo za određeni moralni nedostatak – nemogućnost da činjenice koje su uključene u brutalno ili neobazrivo ponašanje registrira *kao* razloge protiv takvog ponašanja. Budući da je njegov moralni nedostatak u tome što određene činjenice ne uspijeva prepoznati kao razloge, to znači da te činjenice zaista *jesu* razlozi. U slučaju da nisu, ne bi bilo jasno u čemu bi se točno sastojao njegov moralni nedostatak.

(a) Opravdanost posljednjeg prijelaza je upitna. Ukoliko je moj *nedostatak* da neke činjenice ne prepoznajem kao razloge, slijedi li nužno iz toga da te činjenice, kako Scanlon smatra, zaista *jesu* razlozi za mene? Je li nekonzistentno tvrditi da je A-ova nemogućnost da određene činjenice prepozna kao razloge zaista njegov moralni nedostatak, ali da te činjenice istovremeno nisu razlozi za A? Internalist misli da tu nema nekonzistencije i vjerujem da tu i leži glavni spor između dva tabora. Williams napominje da, upotrebom gore navedenih pojmova, internalist samo želi reći da bi za subjekta bilo *bolje* da te činjenice prepozna kao razloge, a ne da on ima *razlog* te činjenice prepoznati kao razloge. U potonjem slučaju opet ne bi bilo jasno odakle bi taj razlog mogao dolaziti i kako bi on mogao biti povezan sa subjektovom motivacijom (usp. Williams, 2001: 96).

Williams, međutim, ne daje dodatno objašnjenje svog prijedloga. Pretpostavimo da A postavi sljedeće pitanje: "zašto bi za mene bilo *bolje* da činjenice koje su sadržane u okrutnom ponašanju tumačim kao razloge protiv takvog ponašanja kada te činjenice, budući da nemam apsolutno nikakvu motivaciju prestati se ponašati okrutno, za mene jednostavno *nisu* razlozi"? Internalist je sada, koliko mi se čini, u poziciji tvrditi da je upravo A-ova nemogućnost da odgovori na to pitanje ujedno i A-ov *moralni nedostatak*. S obzirom da A nije u stanju steći

motivaciju za promjenu ponašanja, istinito je da on nema razlog za promjenu ponašanja, ali njegov je moralni nedostatak baš u tome što on ne vidi da je za njega bolje tu motivaciju ipak imati. Na nekoga tko je svjestan svojih okrutnih ili neobazrivih postupaka i tko, povrh svega, nije motiviran ništa učiniti da bi to promijenio možda i nema smisla primjenjivati moralne kategorije. Oslonimo li se na internalističku interpretaciju, za njega je također netočno reći da ima razlog promijeniti svoje ponašanje, ali nije isključeno da bi za njega samog, sveukupno gledajući, bilo bolje kada bi to učinio. Pod pretpostavkom da bi tada i sam shvatio da mu je život kvalitetniji i potpuniji, ima smisla reći da je za njega *bolje* da svoje ponašanje i promijeni.

(b) Nešto drukčiji odgovor na Scanlonovu kritiku zahtijeva kratak osvrt na pojmove o čijoj je upotrebi ovdje riječ – tzv. "tvrde" etičke pojmove. Kao što je poznato, taj naziv potječe od samog Williamsa koji njime želi istaknuti ono što je za te pojmove karakteristično, a to je "jedinstvo činjenice i vrijednosti" (Williams, 1985: 129). Radi se o tome da – za razliku od "mekih" (npr. "loš", "dobar", "ispravan") – "tvrdi" etički pojmovi, pored evaluativne, imaju izraženu *i* deskriptivnu komponentu. Pravilna upotreba tih pojmova regulirana je objektivnim stanjem stvari u svijetu, što ujedno omogućuje tezu o postizanju moralnog znanja: uvjeti pod kojima je ispravno upotrijebiti pojam "sebičan", "brutalan" ili "okrutan" nešto je oko čega ipak postoji stanoviti konsenzus i zato je moguće, barem prema Williamsu, govoriti o *istinitosti* moralnih tvrdnji u kojima se takvi pojmovi pojavljuju. Osim toga, postojanje takvog konsenzusa ujedno postavlja i granice unutar kojih je pojam moguće upotrebljavati. Međutim, kako ističe Cowley (2005: 365–368), nisu rijetke situacije u kojima je ipak moguće neslaganje oko toga je li pravilno upotrijebiti "tvrdi" etički pojam, a da istovremeno postojanje takvog neslaganja *ne implicira* da netko od sukobljenih strana ne razumije taj pojam u potpunosti. Osoba koja je optužena za okrutno ponašanje prema svojoj supruzi tako može reći da njegovi postupci samo odaju dojam okrutnosti, ali da u stvarnosti

nije tako jer se ti postupci ne smiju promatrati izolirano od njihova odnosa u cjelini. Što nam to govori? Ne slaže li se on da je njegovo ponašanje opravdano etiketirati kao "okrutno", time ujedno i negira da neke činjenice za njega zaista predstavljaju razloge. Te činjenice nisu razlozi za njega samog zato što on odbija optužbu da je njegovo ponašanje okrutno. U slučaju da prva strana i dalje nastavi ustrajati na tome da je pojam ispravno upotrijebiti, prezentirana nam je situacija u kojoj upotrebom "tvrdog" etičkog pojma jedna strana sugerira da osoba ne prepoznaje određene činjenice kao razloge, ali to nužno ne znači da te činjenice zaista jesu razlozi za tu osobu. Upotreba tvrdih etičkih pojmova, suprotno onome što vjeruje Scanlon, ipak ne pretpostavlja eksternalizam.

Ideja da činjenice o okrutnom ili neobazrivom ponašanju u normalnim okolnostima zaista predstavljaju razloge kompatibilna je s internalizmom. Sam Williams napominje da su tvrdi etički pojmovi "karakteristično povezani s razlozima za djelovanje" i da "često pružaju nekome razlog za djelovanje iako taj razlog ne mora biti presudan i može biti nadjačan drugim razlozima…". (Williams, 1985: 140). Naglasak je na riječi "često". Upotrebom pojma kao što je okrutan, govornik ukazuje na činjenice koje ljudi *često* tumače kao razloge protiv takvog ponašanja jer ipak imaju motivaciju koja ne proturječi moralnim zahtjevima. Internalist nema problem složiti se sa Scanlonovom tvrdnjom da se korištenjem pojma okrutan ukazuje na nečiju neosjetljivost da stanovite činjenice prepozna *kao razloge*.[49] Nije sporno da neke činjenice često smatramo razlozima, ali sporno je da te činjenice za svakoga zaista *jesu* razlozi. Nisu za one koji nemaju adekvatnu motivaciju u svom S-u.

4.2.4. *Argument iz konverzije*

Prema eksternalističkoj interpretaciji tvrdnju o postojanju razloga nije moguće falsificirati utvrđivanjem da relevantna motivacija nije prisutna. Vanjski razlozi su oni koji postoje

[49] U svom posljednjem kratkom osvrtu na problem unutarnjih i vanjskih razloga Williams ovo eksplicitno priznaje (usp. Williams, 2001: 95).

neovisno o subjektovoj motivaciji tako da nedostatak odgovarajućeg motivacijskog elementa kod subjekta nije nešto zbog čega bi bilo potrebno ustručavati se pripisati mu razlog za djelovanje. Ne priznavajući motivaciju kao uvjet za postojanje razloga, eksternalisti razlozima pripisuju objektivan status, a time ujedno postavljaju i nezavisan kriterij za evaluaciju ljudskih postupaka. To je sigurno glavna eksternalistička prednost, ali ostaje preispitati je li cijena koju su oni spremni platiti za zadržavanje te prednosti možda ipak prevelika. Uvjet prema kojemu razlozi moraju motivirati vjerojatno predstavlja najveću prepreku potpunom izgrađivanju teze o postojanju vanjskih razloga: eksternalisti tu prepreku nisu u stanju prevladati i svaki njihov pokušaj u tom pravcu na koncu uvijek kolabira u internalizam. Pogledajmo zašto. Za razliku od internalista, eksternalisti vjeruju da bi "subjekt trebao zadobiti motivaciju *zato* jer počinje vjerovati u iskaz o razlogu, a potonje bi uostalom trebao učiniti zato jer, na neki način, stvari razmatra pravilno" (Williams, 1981: 108–109). Subjektovo prepoznavanje vanjskog razloga dakle mora polučiti stvaranje odgovarajuće motivacije, ali upravo ta veza – između prepoznavanja razloga i stvaranja motivacije – u sklopu eksternalizma ostaje misteriozna i neprecizirana. Subjekt bi, nakon racionalne deliberacije, trebao prepoznati da ima razlog koji je cijelo vrijeme postojao i steći motivaciju djelovati u skladu s tim razlogom.

Poteškoća s tim shvaćanjem jest u tome što ono obvezuje eksternalista na stav prema kojemu se motivacija stječe čisto racionalnim putem, neovisno o tome što se nalazi u subjektovu S-u prije deliberacije. Racionalna deliberacija u internalističkom slučaju uvijek kreće iz subjektova S-a i zato stvaranje nove motivacije nije *ex nihilo*, dok eksternalisti racionalnu deliberaciju nisu u mogućnosti shvatiti na isti način jer bi time samo priznali da su razlozi internalne prirode. Racionalne procese eksternalisti moraju shvatiti kao procese koji su samostalno sposobni stvoriti motivaciju, a upravo to i jest glavni problem. Budući da "ne postoji motivacija *iz* koje [djelatnik] rasuđuje da bi dostigao novu motivaciju", nije objašnjeno kako novu motivaciju subjekt uopće može dosegnuti.

Eksternalisti vjeruju da osoba putem racionalne deliberacije prepoznaje razlog koji je uvijek bio prezentan i da kao rezultat tog prepoznavanja stvara odgovarajuću motivaciju. Međutim, jesu li oni doista osuđeni pristati na takav prikaz njihove pozicije? Eksternalist koji na to pitanje daje negativan odgovor jest McDowell (1995). On napada pretpostavku prema kojoj se prijelaz iz stanja u kojem osoba nije motivirana razlogom za djelovanje u stanje u kojem ona jest tako motivirana mora odvijati pod utjecajem *racionalne deliberacije*. McDowell je svjestan da bi upravo prihvaćanje te pretpostavke eksternalista dovelo u nezgodan položaj jer bi tada tu tranziciju, baš kako naslućuje Williams, mogao objasniti samo ukoliko bi prethodno dokazao mogućnost nastajanja motivacije čisto racionalnim putem. McDowell stoga sugerira da za eksternaliste postoje alternativni načini objašnjenja tranzicije iz stanja nemotiviranosti u stanje motiviranosti. Ne mora uvijek postojati racionalni put koji vodi iz jednog stanja u drugo, već do stvaranja motivacije ponekad dolazi ne-racionalnim načinom kao što je konverzija. McDowellova je ideja da racionalna deliberacija u nekim slučajevima ima smisla samo ako prvo dođe do radikalne promjene u motivacijskom ustrojstvu jer subjekt na osnovi racionalne deliberacije uopće ne mora biti sposoban zahvatiti razlog koji postoji neovisno o njemu. Netko može "stvari razmotriti pravilno" jedino pod uvjetom da prethodno promijeni svoj motivacijski skup na način koji ne uključuje ispravno rasuđivanje: ponekad je potreban "preokret u motivacijskom usmjerenju" da bi se spoznao vanjski razlog za djelovanje jer taj učinak nije moguće ostvariti racionalnom deliberacijom (McDowell, 1998: 102). McDowell prvenstveno ima u vidu nekoga tko nije imao adekvatan odgoj i tko zbog toga nije razvio dovoljnu senzibilnost za pravilno sagledavanje neke situacije odnosno za prepoznavanje nekih čimbenika koji mogu figurirati kao razlozi. Čak i da taj podvrgne elemente svog S-a korekciji putem racionalne deliberacije, on neće doseći stanje u kojemu prepoznaje vanjski razlog. To se neće dogoditi upravo zato što su elementi njegova S-a takvi kakvi jesu.

Pretpostavka McDowellove kritike jest da je do spoznaje razloga ponekad moguće doći nekim ne-racionalnim putem, a ne isključivo oslanjanjem na racionalnu deliberaciju. Nedostaje li motivacija za činjenje Φ-a nakon racionalne deliberacije, za Williamsa je to dovoljan dokaz da subjekt nema razlog učiniti Φ. McDowell smatra da je taj zaključak preuranjen: subjekt i dalje može imati razlog učiniti Φ i zato racionalna deliberacija nije jedini način putem kojeg se taj razlog može spoznati. Budući da na McDowellovu kritiku u literaturi već postoji određeni broj replika, u narednom ću je dijelu ostaviti po strani.[50] Umjesto toga, orijentirat ću se na kritiku koju je internalističkom shvaćanju razloga uputio Shafer-Landau (2003: 185–186) i koja, kako mi se čini, s McDowellovim pristupom dijeli barem jedno slično obilježje. I Shafer-Landau i McDowell slažu se u ideji da racionalna deliberacija nije jedini način za spoznavanje razloga, a to je sasvim dovoljno da te kritike ne promatram izolirano jednu od druge. Zato mislim da će moje razilaženje sa Shafer-Landauom također baciti i dodatno svjetlo na McDowellov argument.

Shafer-Landau zamišlja melankoličnu osobu (A) koja ne vidi nikakvu mogućnost za realizaciju vlastite sreće. A jednostavno nema nikakvu motivaciju za bilo kakvom vrstom društvenog angažmana, nikakvu želju – jednostavno ništa u njegovu motivacijskom skupu ne postoji što bi ga moglo natjerati da se "pokrene" i pokuša ostvariti sreću. Smatrajući da ne postoji nikakva djelatnost u kojoj bi pronašao zadovoljstvo, on se u potpunosti izolira iz društvenog života. No vrlo je izvjesno da bi A ostvario životnu sreću samo kada bi promijenio svoj pogled na svijet. Ključno je ne izgubiti iz vida da bi *on sam*, kada bi barem malo promijenio svoje shvaćanje svijeta, otkrio da je pogriješio i da postoje mnoga vrijedna iskustva koja bi ga mogla učiniti sretnim i koja bi mu život mogla učiniti boljim i kvalitetnijim. Budući da se nalazi u stanju u kojem se nalazi, on to nije u mogućnosti shvatiti. Da bi to shvatio, potrebno je da svijet promotri "drugim očima". No ako je to točno, zar nema

[50] Sam Williams (1995b) je replicirao na McDowellov argument. Za nešto drukčiju kritiku McDowella također vidi Cowley (2005) i Pettit i Smith (2006). McDowell (2006) također odgovara na argumentaciju Pettita i Smitha te dodatno pojašnjava svoje razilaženje s Williamsom.

smisla reći da on ima *razlog* dovesti se u stanje u kojemu svijet vidi "drugim očima"? Budući da je jedini način da to učini taj da stvarno "izađe u svijet", "otvori" se drugim ljudima i počne baviti se stvarima čiju će vrijednost sam kasnije spoznati, zar nije točno kazati da on ima razlog to činiti premda to ni na koji način nije povezano s njegovim S-om? Postoji razlog, ali on do tog razloga ne može doći rasuđivanjem, već iskustvom.[51]

Moguću strategiju koju bi internalist mogao koristiti jest ustrajati na tome da je pripisivanje razloga u ovoj situaciji sasvim konzistentno s njegovom pozicijom. Činjenica da bi *sam A* naknadno uvidio da ipak pronalazi zadovoljstvo u određenim aktivnostima i da je mnogo sretniji nego što je bio prije promjene jest indikator da u danim okolnostima jednostavno ne rasuđuje na dovoljno dobar način o tome kako bi svoju dobrobit mogao ostvariti. Nesporno je da A u svom S-u ima motivaciju da ostvarenjem dobrobiti, ali rasuđivanjem ne uspijeva pronaći način na koji bi to mogao realizirati. Zato navedeni primjer, internalist bi mogao tvrditi, samo pokazuje da A, ako se tako može reći, svoje rasuđivalačke kapacitete ne koristi u maksimalnom smislu. Da A rasuđuje kako treba, tada bi utvrdio da je za ostvarenje njegove sreće, primjerice, potrebno "vratiti se" u društveni život te bi pronašao i motivaciju to učiniti. Ukratko, razlog nije eksternalni, već internalni.

Mislim da opisani način obrane protiv Shafer-Landauova primjera ipak ne bi bio uvjerljiv. Pretpostavimo li analogiju s McDowellovim argumentom, centralna pretpostavka jest da racionalna deliberacija ne mora uvijek biti od pomoći. Ljudi – čak i uz besprijekorno rasuđivanje – nisu uvijek u stanju pronaći motivaciju za djelovanje. Ponekad je potrebno zauzeti drukčiju perspektivu jer je jedino tako moguće ustanoviti da je naše shvaćanje svijeta zasnovano na pogrešnim premisama. Zauzimanje drukčije perspektive "često zahtijeva promjenu karaktera. Prije te promjene, izgledi novog života nisu privlačni zato jer nisu racionalno povezani s nečijim trenutnim gledištem. Ali netko može imati razlog promijeniti to

[51] Važnost iskustva u spoznavanju razloga posebno ističe Millgram (1996).

gledište" (Shafer-Landau 2003: 186). Promjena karaktera o kojoj govori Shafer-Landau sasvim korespondira onome što je McDowell nazvao konverzijom. Racionalna deliberacija, ukoliko je pod kontrolom motivacijskih svojstava koje A trenutno ima, jednostavno neće polučiti nikakav efekt: A neće biti u stanju spoznati da ima razlog učiniti Φ.

Pokušat ću Shafer-Landauovu kritiku dodatno objasniti koristeći jednu analogiju. Ljudi koji žive u Platonovoj pećini vjeruju da je njihovo viđenje svijeta ispravno i da su sjene koje promatraju na zidovima pravi predmeti. Jedini način na koji oni mogu ustanoviti da to nije slučaj jest izlaskom iz pećine. Sa sigurnošću se može tvrditi da bi oni sami uvidjeli da je svijet mnogo savršeniji samo kada bi se natjerali da izađu, ali problem je u tome što čak i kada bi ih netko uporno uvjeravao da se pravi svijet nalazi izvan pećine, još uvijek ne bi postojalo jamstvo da bi to uvjeravanje na koncu rezultiralo njihovim izlaskom. I sam Platon o tome govori kada opisuje čovjeka koji se, nakon što je iskusio blagodati vanjskog svijeta, vraća u pećinu po druge s namjerom da ih povede sa sobom. Što bi se njemu dogodilo, Platon pretpostavlja u sljedećem odlomku:

> …zar mu se ne bi ismijali, i zar se ne bi o njemu govorilo, da se gore uspeo samo zato, da se vrati s pokvarenim očima, te da nije vrijedno ni kušati gore ići? I ako bi kako mogli dobiti u ruke onoga, koji bi ih htio otkrivati i gore voditi, zar ga ne bi i ubili? (Platon, 1997: 270).

Kao što stanovnici pećine, slušajući o blagodatima vanjskog svijeta, ne moraju nužno pronaći motivaciju da izađu iz pećine, tako i melankolična osoba, deliberativnim rasuđivanjem, ne mora nužno pronaći motivaciju da "izađe u svijet". Ali kao što se čini da stanovnici pećine ipak imaju razlog izaći iz pećine premda za to ne moraju biti u stanju pronaći motivaciju, tako se čini da melankolična osoba ipak ima razlog svijet sagledati na drukčiji način iako to ni na koji način ne mora biti povezano s njezinim S-om. Pripisivanje razloga ovdje ima smisla zato što bi sami stanovnici pećine naknadno shvatili da su živjeli u "iluziji", baš kao što bi i

melankolična osoba sama naknadno shvatila da je njezino shvaćanje svijeta bilo zasnovano na potpuno pogrešnim temeljima.

Pogledajmo sada koje su slabosti "argumenta iz konverzije". Kao prvo, njegovo sporno mjesto sigurno jest teza da spoznaja nekih vrijednih iskustava ponekad iziskuje promjenu karaktera. Promijeni li osoba svoja karakterna obilježja i tako počne uživati u stvarima u kojima ne bi uživala da do te promjene nije došlo, pitanje je bi li još uvijek bilo riječ o *potpuno* istoj osobi. S tom je opservacijom usko povezana i sljedeća poteškoća. Nečije karakterne crte uvelike određuju kakve razloge netko uopće ima u danim okolnostima. Plauzibilno je naime pretpostaviti da promjena karakternih osobina uzrokuje promjenu samih okolnosti i, naposljetku, promjenu razloga koje netko ima za djelovanje. Kažemo li da A ima razlog činiti Φ, tada je taj razlog velikim dijelom determiniran i A-ovim okolnostima, a A-ovo psihološko stanje zasigurno "ulazi" u opis tih okolnosti i možda baš zato što se A nalazi u stanju u kojem se nalazi vjerujemo da on ima razlog djelovati na ovaj, a ne neki drugi način. Shafer-Landau, vidjeli smo, pretpostavlja da u nekim situacijama jednostavno ne možemo ustanoviti da imamo razlog za djelovanje ukoliko stvari ne sagledamo iz drukčijeg kuta. Čak i ako rasuđujemo na savršen način, taj proces ne mora uvijek rezultirati spoznajom razloga. To je stoga što se to rasuđivanje uvijek odvija *unutar prethodno prihvaćene slike svijeta* i zato tu sliku ponekad imamo razloga promijeniti. Pretpostavimo da promjena prihvaćene slike svijeta u ovom slučaju doista iziskuje "izlazak u svijet" i "otvaranje" drugim ljudima. Izlazak iz stanja melankolije moguć je samo pod uvjetom da se A natjera baviti stvarima za koje trenutno nije motiviran, ali koje će s vremenom na njega ipak polučiti pozitivan učinak. Pokazuje li to da A *sada* ima razlog baviti se tim stvarima? Nije evidentno da je odgovor potvrdan. Baš zato što se nalazi u stanju melankolije, on ne mora imati iste razloge koje bi imao kada se ne bi nalazio u tom stanju. On rasuđivanjem možda i nije u stanju spoznati da bi

ga određene djelatnosti učinile sretnim kada bi se u njih upustio, ali isto je tako moguće da bavljenje tim stvarima u melankoličnom stanju može polučiti sasvim suprotan učinak.

Kao drugo, zašto u prvom redu i prihvatiti pretpostavku da A ima sposobnost utjecati na svoje trenutno stanje? Pitanje je, naravno, kako izaći izvan okvira prethodno prihvaćene slike svijeta? Kako učiniti da ljudi izađu van iz "pećine" u kojoj se cijeli život nalaze? Budući da će A, jednom kada izađe iz stanja melankolije, sam uvidjeti da je imao razlog promijeniti svoje stanje melankolije, to znači da i *sada*, kada se nalazi u stanju melankolije, on ima razlog promijeniti svoje stanje melankolije. No ako je to točno, onda je razumno pretpostaviti da on zaista *može* promijeniti svoje stanje melankolije. Otvoreno je pitanje u kojoj mjeri je to realno. Poteškoća se skriva u tome što se pripisivanje razloga u ovom slučaju odvija retroaktivno. Složimo li se da su neki naši izbori determinirani našim psihološkim ustrojstvom, ima li smisla kazati da su ti izbori mogli biti drukčiji nego što su bili? Na to je pitanje teško pružiti konkluzivan odgovor, ali to ne znači da takvo razmatranje treba odmah odbaciti. Svakodnevne životne situacije također mogu biti dobar pokazatelj. Razmišljajući o nekim ranijim izborima i odlukama, često kažemo "Da sam barem…" time implicirajući da smo možda i mogli izabrati alternativan put djelovanja. Ali takvo razmišljanje možda i nema osnovu. Zašto bi bilo netočno reći da jednostavno nismo mogli izabrati alternativan pravac djelovanja jer se nismo nalazili u okolnostima u kojima smo to mogli učiniti?

Kao treće, Shafer-Landauov argument zasnovan je na pretpostavci da *vrijednost* Φ-a povlači *razlog* da se čini Φ. Budući da bi A, kada ne bi bio melankoličan, utvrdio kako su za njega mnogobrojna iskustva ipak vrijedna i da njegov život mogu učiniti boljim, to znači da on, sada kada je melankoličan, ima *razlog* "izaći iz svog oklopa". Zašto onda ne bi bilo bolje kazati da je za A *dobro* da izađe iz stanja melankolije radije nego da A ima *razlog* izaći iz stanja melankolije? Takav bi opis Shafer-Landauovog primjera bio kompatibilan s internalizmom jer bi pripisivanje razloga i dalje bi bilo moguće samo u situacijama u kojima

je prisutna odgovarajuća motivacija. S druge strane, tvrdnja da vrijednosti pružaju razloge za djelovanje, kao što ću nastojati pokazati u sljedećem odsječku, nije sasvim očita i zasigurno nailazi na mnoge poteškoće.

Nadam se da sam u prethodnoj diskusiji uspio pokazati da je Williamsov internalizam imun barem na neke prigovore. U najosnovnijim crtama, mislim da internalističko gledište ne odbacuje svakodnevno razumijevanje pojma "krivnje", da ne dovodi do apsurdnih posljedica, da internalist konzistentno može koristiti "tvrde" etičke pojmove i da mogućnost konverzije ne povlači postojanje vanjskih razloga. Rasprava o internalizmu, naravno, time nije završena. Dok ću u sljedećem poglavlju promotriti jednu drukčiju verziju internalističke pozicije koja predstavlja "srednji put" između internalizma Williamsova tipa i eksternalizma, u ostatku ovog ću se orijentirati na još neke poteškoće koje se obično povezuju s ovim hjumovskim stajalištem.

4.3. Razlozi i vrijednosti

Internalist tvrdi da razlog za djelovanje postoji samo *ako* je prisutna odgovarajuća želja. Riječ je dakle o pokušaju da se specifiraju *uvjeti* pod kojima je istinito kazati da netko ima praktični razlog.[52] Međutim, kazati da je razlog ispravno pripisati samo ako je kod subjekta prisutna odgovarajuća želja nije isto što i kazati da su želje izvor tih razloga ili da želje pružaju razloge. Jednostavnije rečeno, to da imamo razlog samo *ako* imamo odgovarajuću želju ne znači da imamo razlog *zato* jer imamo odgovarajuću želju (usp. Finlay i Schroeder, 2008: 7). Potonja je relacija, kako primjećuje Dancy, "…različita, jača i mnogo interesantnija od mnogo slabije relacije 'biti nužan uvjet za postojanje'" (Dancy, 2000: 17). Oslonimo li se na tu distinkciju, slijedi da je internalističko stajalište pogrešno poistovjećivati s jačom tezom te da ga nije moguće osporavati argumentiranjem da razlozi nisu zasnovani na željama. Ali mnogi

[52] Williams eksplicitno tvrdi da je srž internalističkog stajališta u tome da je želja nužan uvjet za pripisivanje razloga te da pitanje o tome je ona ujedno i dostatan uvjet ostaje otvoreno (Williams, 2001: 91).

tu grešku ipak čine. Tako Ruth Chang (2004: 56), primjerice, piše da teoriju prema kojoj su "praktični razlozi *utemeljeni* u željama" brani Williams, dok Derek Parfit smatra da su "internalističke teorije *zasnovane* na željama" te da zbog toga i "promašuju većinu onoga što je istinito" (Parfit, 1997: 128, moj kurziv). Budući da je jedna od mojih primarnih namjera u ovom dijelu rada obraniti hjumovski internalizam razloga, tj. gledište prema kojoj su želje *uvjet* za prisustvo razloga, smatram da nisam primoran braniti jaču poziciju prema kojoj su želje ujedno i *izvor* svih razloga odnosno prema kojoj su svi razlozi za djelovanje ujedno i *zasnovani* na željama. No ipak, mislim da je vrijedno zadržati se na ovoj raspravi te promotriti argumente koji se navode protiv te teze. Razlog tome jest što hjumovac ovdje nema što za izgubiti: ako se zaista pokaže da želje ne mogu predstavljati temelj za praktične razloge, njegov internalizam i dalje ostaje sačuvan, a ako se pokaže da argumenti protiv jače teze nisu dovoljno uvjerljivi, to njegovu poziciju može samo dodatno učvrstiti.

Stajalište koje ovdje valja promotriti jest ono prema kojemu razlozi nisu zasnovani na željama, već na vrijednostima (usp. Dancy, 2000; Parfit, 1997; Raz, 1997; Scanlon, 1998). Zašto imam razlog, primjerice, pročitati određenu knjigu nije zato što je želim pročitati, već zato što je knjiga dobra. Među ovim autorima doduše ne vlada konsenzus oko pitanja na koji način vrijednosti pružaju razloge, tj. u čemu se točno sastoji odnos između vrijednosti i razloga. Je li naime tako da sama vrijednost knjige (da je knjiga *dobra*) pruža razlog za djelovanje ili nam takav razlog pružaju činjenice koje, uslijed toga, tu knjigu i čine dobrom? Imamo li, jednostavnije rečeno, razlog pročitati knjigu zato što je ona dobra ili je knjiga dobra zato što je mi imamo razloga pročitati? U prvom slučaju vrijednosti prethode razlozima, dok u drugom slučaju vrijedi obrnuto.[53] No koji god bio odgovor na postavljeno pitanje, jedna stvar ipak ostaje neupitna: razlozi ni u kojem slučaju nisu zasnovani na našim željama, već u onome *što* želimo. Razmotrimo to pobliže.

[53] Prema drugoj interpretaciji, također poznatoj kao *buck-passing account*, "kazati da je nešto vrijedno jest kazati da ono ima druga svojstva koja pružaju razloge za ponašanje na određene načine… (Scanlon, 1998: 96).

Karakteristično obilježje navedenog pristupa jest stav da nisu razlozi zasnovani na željama, već da su želje zasnovane na razlozima. Kada nešto želimo, onda to želimo zato jer u tome vidimo neku vrijednost, a budući da "vrijednost … stvari i opcija konstituira razlog da ih imamo ili da ih činimo, njihova *vrijednost* jest također *razlog* da ih *želimo* imati ili činiti (Raz, 1997: 115, moj kurziv). Želim se baviti sportom zato jer smatram da ću time održati svoju kondiciju i pozitivno utjecati na svoje zdravlje. Želim naučiti strane jezike zato jer ću na taj način moći lakše upoznavati ljude iz drugih kultura. Želim kupiti bolje računalo jer će mi time biti olakšano izvršavanje nekih zadataka. Kao što ovi primjeri ilustriraju, naše želje uvijek nastaju kao rezultat vjerovanja da je njihov predmet nešto što ima određenu vrijednost te da o njima uopće ne bi bilo smisleno govoriti kada one tim vjerovanjem ne bi bile popraćene. Želje ne moraju biti zasnovane na razlozima, ali ako to jest slučaj, onda to više nisu *naše* želje. Netko tko ne vidi nikakvo dobro u onome što želi, kako to smatra Raz, ima želju koju ne prihvaća i koja nije pod njegovom kontrolom. Njegovim riječima: "Želja je naša sve dok je imamo zbog vjerovanja u vrijednost njezina predmeta te bi nestala kada bismo se tog vjerovanja odrekli" (Raz, 1997: 116). Zato bi želje koje nisu pod našom kontrolom, Raz nastavlja, bilo ispravnije okarakterizirati kao "porive" te ih na taj način razlikovati od, kako ih on naziva, "filozofskih" želja. Iako i porivi ponekad mogu biti zasnovani na razlozima, oni nas u većini slučajeva samo prisiljavaju da ih zadovoljimo: mi ćemo djelovati u skladu s njima ne zato jer u tome vidimo neku vrijednost, nego zato što nemamo izbora. Njihova je priroda uglavnom patološka te stoga i ne mogu poslužiti kao reprezentativni primjer onoga što obično podrazumijevamo pod željom. U tu je svrhu potrebno osvrnuti se na filozofske želje koje nemaju "osjećajnu kvalitetu" i koje se uvijek pojavljuju zato što vjerujemo da je njihov predmet u nekom smislu vrijedan.

S obzirom da su naše želje uvijek zasnovane na razlozima, sama želja u tom smislu ne predstavlja *dodatan* razlog. Drugim riječima, ako želim učiniti Φ, onda to već želim učiniti

zbog određenih razloga, što znači da sama činjenica da *želim* učiniti Φ-a ne može predstavljati razlog *više* da to učinim. Želim li, da ponovno upotrijebim isti primjer, učiti strane jezike zato jer vjerujem da ću time proširiti svoja poznanstva, bolje razumjeti vrijednosti drugih kultura i lakše pratiti zbivanja u drugim državama, onda mi sama činjenica da to želim ne daje *još jedan* razlog da to činim. Želja, kako to zgodno kaže Dancy, "ne utječe na zalihu razloga" (Dancy, 2000: 36). Raz doduše priznaje da se u rijetkim okolnostima to ipak može dogoditi. Riječ je o okolnostima u kojima bi postojali podjednaki razlozi na svakoj strani i u kojima bi stoga bilo iracionalno djelovati suprotno svojoj želji. Recimo da sam suočen s izborom između A i B te da je – zbog toga što su razlozi za obje opcije podjednako jaki – svejedno što ću izabrati. U tom bi slučaju činjenica da želim učiniti A predstavljala *razlog* da učinim A. Jer ako bih, unatoč tome što želim učiniti A, ipak učinio B, onda bih po svim standardima djelovao iracionalno. Za Raza je ovo jedina situacija u kojoj želje mogu biti razlozi, ali, kako navodi, "vrlo neobični razlozi" (Raz, 1997: 124). Dancy, čiji je smjer argumentacije u osnovnoj strukturi vrlo sličan Razovu, na ovom se mjestu s njime ipak razilazi. On ne samo da ne dijeli zaključak da želje u navedenim okolnostima mogu predstavljati razloge, nego smatra da je sam opis tih okolnosti sasvim nekoherentan. Jer kako je uopće moguće, on argumentira, da *više* želim učiniti A ako istodobno vjerujem da su razlozi za A i B podjednaki te da A ni na koji način nije bolji od B? (Dancy, 2000: 40). Stoga za Dancya ne postoje okolnosti u kojima bi želje predstavljale *dodatne* razloge u odnosu na one koji već postoje za činjenje neke radnje. To ujedno znači da želja za činjenjem neke radnje stoga ne može biti razlog ni u situacijama kada za tu radnju ne postoje nikakvi razlozi. Naime, "ako netko želi nešto za što ne postoji razlog, njegova mu želja ne daje neki razlog da to učini, razlog koji nije postojao ranije. Ako je radnja blesava, željeti je učiniti ne čini je manje blesavom" (Dancy, 2000: 36–37). I ne samo to. U slučaju da ne postoje razlozi za naše želje, onda ne mogu

postojati ni razlozi da učinimo stvari putem kojih bismo te želje mogli ostvariti; razlog za sredstvom možemo imati samo ako prethodno imamo razlog za ciljem.

Gornji prikaz, iako pojednostavljen i nepotpun, u velikoj mjeri ocrtava liniju argumentacije onih autora koji vjeruju da vrijednosti predstavljaju metafizički temelj praktičnih razloga. Ako je ta argumentacija uvjerljiva, već je rečeno, onda je hjumovsku poziciju moguće braniti samo u njezinoj slabijoj verziji. U tom slučaju hjumovac mora priznati da se priroda razloga za djelovanje ipak ne može *u potpunosti* objasniti referiranjem na naša konativna stanja. No je li ta argumentacija zaista uvjerljiva? To je pitanje na koje ću se sada osvrnuti.

Vratimo li se na Razovu diskusiju, čini se opravdanim izraziti skepsu prema njegovoj tezi da svaka moja želja mora biti popraćena vjerovanjem da je njezin predmet nešto dobro ili "vrijedno truda". Podsjetimo se, želja koja nije popraćena takvim vjerovanjem, prema Razovu mišljenju, nije *moja* u pravom smislu te riječi – tu je želju primjerenije kategorizirati kao poriv ili nagon, psihološko stanje koje me "napada" i koje ne mogu sasvim kontrolirati. Raz, međutim, previđa da postoje i "srednji" slučajevi. Izgleda da je ipak zamisliva situacija u kojoj subjekt ima *svoju* želju a istovremeno ne vjeruje da ima razlog željeti to što želi odnosno ne vjeruje da je predmet njegove želje na neki način vrijedan. Da bi to pokazala, Chang navodi primjere u kojima subjekt nema odgovarajuće vjerovanje zato jer ga, uslijed nekih čimbenika, jednostavno *nije u mogućnosti* imati. Dijete, koje još nema pojam razloga, može željeti jesti mrkvu a da pritom nema vjerovanje o razlogu. Ono nema odgovarajuće vjerovanje jer mu za to nedostaje "konceptualni aparat", ali to ne znači da samim time nema želju o kojoj je riječ. Izuzev ove "konceptualne nemogućnosti", moguć je i slučaj "racionalne nemogućnosti". Primjerice, dijabetičar koji želi komad čokoladne torte može vjerovati da za to nema nikakvog razloga jer zna da ne voli okus čokolade i da bi ostvarenjem te želje pao u dijabetičku komu. Također, postoji i slučaj "logičke nekoherentnosti" – recimo, kada

ekscentrični filozof želi samo one stvari za koje vjeruje da ne postoji razlog da ih ima (usp. Chang, 2004: 66–68). No možda i nije potrebno ići tako daleko i zamišljati ovako egzotične primjere. U namjeri da ojača svoje stajalište, hjumovac bi također mogao argumentirati da postoji određena kategorija želja – poput *hirovitih želja* – koje generiraju praktične razloge premda same nisu na njima zasnovane. Hjumovac doduše ne mora tvrditi da to vrijedi za *sve* hirove, već samo za one koji su do te mjere bezazleni da ne mogu biti u konfliktu s drugim željama koje subjekt ima (usp. Hubin: 2003: 324). Tako nije jednostavno razumjeti zašto mi, recimo, hirovita želja da pjevam iz sveg glasa, u okolnostima u kojima znam da me nitko ne može čuti, ne bi dala razlog da to činim. Nekako mi izgleda da je to primjer situacije u kojoj mi želja, premda nije zasnovana na razlogu, ipak pruža razlog – imam razlog pjevati iz sveg glasa baš *zato* što to želim.

U ovom bi se kontekstu moglo prigovoriti da je, generalno promatrajući, poziciju prema kojoj su praktični razlozi utemeljeni u vrijednostima uzaludno osporavati osmišljavanjem različitih primjera. Jer kako će oni biti protumačeni u krajnjoj instanci zavisi samo od toga kakve su nečije intuicije: netko tko vjeruje da su sve naše želje praćene vjerovanjem u vrijednost njihova predmeta prije da će u opisanim slučajevima vidjeti potvrdu te teze nego njenu falsifikaciju. Dakle, prigovor bi mogao glasiti, iluzorno je smatrati da bi navođenjem protuprimjera opisano stajalište moglo biti ozbiljnije ugroženo; za postizanje takvog učinka potrebno je argumentirati nezavisno od pretpostavki koje suprotna strana nema nikakvog razloga prihvatiti. No ima li hjumovac na raspolaganju bolju strategiju? Mislim da je odgovor na to pitanje potvrdan.

Samo da ukratko ponovimo, pozicija koja je predmet razmatranja jest ona prema kojoj su praktični razlozi zasnovani na vrijednostima. Naše želje, u okviru ovog prijedloga, samostalno ne generiraju nikakve razloge. Ako nešto želimo, onda to želimo zbog određenog razloga, a taj je razlog najčešće "neko dobro koje vidimo u predmetu koji želimo" (Dancy,

2000: 35). Dakle, "vrijednost … konstituira razlog", a subjekt može željeti samo ono za što vjeruje da ima neku vrijednost odnosno samo ono "što mu se doima dobrim ili vrijednim" (Raz, 1997: 115–117). Mislim da je glavnu poteškoću cijelog modela moguće iščitati upravo u ovim tvrdnjama. Želje su zasnovane na razlozima jer, kako ovi autori ističu, mi često "vidimo" neko dobro u predmetu koji želimo, "vjerujemo" u vrijednost onoga što želimo ili nam se predmet koji želimo "doima" vrijednim. Ali, strogo promatrajući, iz toga slijedi samo to da su želje zasnovane na onome za što mi *vjerujemo* da je vrijedno, a ne na onome što zaista *jest* vrijedno. Drugim riječima, iz rasprave je moguće zaključiti samo to da su naše želje zasnovane na onome što mi *vjerujemo* da su razlozi ili na onome što *vidimo* kao razloge, a ne na onome što zaista *jesu* razlozi. Ta distinkcija nije beznačajna jer se njome ukazuje na mogućnost situacije u kojoj subjekt želi Φ zato jer pogrešno vjeruje da Φ ima određenu vrijednost. No, ako on pogrešno vjeruje da Φ ima određenu vrijednost, onda on pogrešno vjeruje da postoji razlog željeti Φ, što znači da bi takvu situaciju bilo ispravno opisati kao onu u kojoj subjekt ima želju koja zapravo nije zasnovana na razlogu.

Dancy je, po svemu sudeći, svjestan ove poteškoće jer na jednom mjestu kaže da su sve želje zasnovane na razlozima, "stvarnim ili *prividnim*" (Dancy, 2000: 35, moj kurziv). Unatoč tome što subjektova želja nije zasnovana na stvarnom razlogu, ona je, Dancy bi rekao, ipak zasnovana na "prividnom" razlogu. Ali ako se tvrdnja da su želje zasnovane na razlozima ne odnosi samo na stvarne razloge, već i na one prividne, onda je to nešto s čime bi se svaki hjumovac lako mogao složiti. Najkraće rečeno, očita implikacija takve interpretacije jest da nijedna želja u konačnici ne može biti predmetom racionalne kritike – jer ako u razloge koji *opravdavaju* nečije želje ne ulaze samo oni razlozi koji *de facto* postoje, već i oni razlozi koji se subjektu "doimaju" da postoje, onda je teško razumjeti na kojoj bi osnovi nečije želje ikada mogle biti podvrgnute kritičkom preispitivanju praktičnog razuma. Naime, ako su sve želje zasnovane na razlozima, bilo stvarnim bilo prividnim, onda su sve želje opravdane, bilo

objektivno bilo subjektivno. Kako onda opisati situaciju u kojoj subjekt pogrešno vjeruje da postoji razlog za njegovu želju? Izgleda da postoje samo dva načina: *ili* je tako da subjekt u tom slučaju ima želju koja nije zasnovana na razlogu *ili* je tako da je njegova želja ipak zasnovana na razlogu, ali razlogu koji je samo "prividan". Ako je prvo tumačenje istinito, slijedi da nisu *sve* želje zasnovane na razlozima, a ako je drugo tumačenje istinito, onda se ovo stajalište opasno približava hjumovskim intuicijama. Nijedna od opcija nije zadovoljavajuća za Dancya i njegove istomišljenike.

Autori na koje sam se osvrnuo argumentiraju da želje, same po sebi, nikada ne mogu racionalizirati nečiji postupak te da izvor praktičnih razloga treba tražiti u vrijednostima na kojima su one zasnovane. Da upotrijebim često spominjani primjer Warrena Quinna, moja želja da upalim svaki radio koji vidim da je ugašen – pod pretpostavkom da to ne radim zato da bih čuo vijesti ili slušao muziku – ne čini moje postupanje smislenim. Ta neobična želja ili, Quinnovim riječima, "to bizarno funkcionalno stanje" možda i objašnjava moje ponašanje u kauzalnom smislu, ali zasigurno ne doprinosi njegovoj racionalizaciji (usp. Quinn, 2002: 189–190). Na sličan način, netko tko ima želju brojati vlati trave u svom vrtu ili bojati krumpire u zeleno nema nikakav razlog to činiti ako pritom ne vidi nikakvu vrijednost u tim radnjama (usp. Raz, 1997: 118). Predmet želje potrebno je vrednovati, a vrednovanje je ovdje shvaćeno kao kognitivno stanje – *vjerovanje* da je ono što želimo u nekom smislu dobro ili vrijedno truda. Jedino u svjetlu tog vjerovanja naše djelovanje može imati smisla, što znači da su vrijednosti, a ne želje, ono što djelovanje na koncu čini inteligibilnim.

Zanimljivo je da hjumovac – bez opasnosti da učini svoju poziciju nekoherentnom – ovaj prijedlog ne mora u potpunosti odbaciti. On se sasvim lako može složiti da nam neke želje ne pružaju razloge te priznati postojanje bliske veze između naših želja i vrijednosti. Njegova bi argumentacija u tom smislu mogla imati sljedeću strukturu. Evidentno je da između psiholoških stanja koje obično nazivamo "željama" postoje izvjesne razlike. Kao što

ispravno primjećuje Raz, neke želje koje imamo nisu *naše* želje zato jer se s njima ne poistovjećujemo, imamo dojam da nas "napadaju" i da im se ne možemo tek tako oduprijeti. U skupinu takvih želja, koje bi uostalom bilo ispravno okarakterizirati kao porive, mogli bismo uvrstiti i one poput gore navedenih – netko tko ima želju upaliti svaki radio, brojati vlati trave u svom vrtu ili bojati krumpire u zeleno nema želju koja mu pruža razlog za djelovanje. U njegovu je slučaju prije riječ o funkcionalnom stanju koje ga nagoni ili tjera na postupanje, postupanje koje ni on sam ne nalazi (ili ne mora nalaziti) vrijednim. Dakle, sasvim je ispravno tvrditi da je o našim željama primjereno govoriti samo onda kada su udružene s određenim vrijednostima. Ono što međutim nije ispravno kazati jest da su te vrijednosti *nešto različito* od samih želja. Vrednovati određeno stanje stvari ne znači *vjerovati* da je to stanje stvari vrijedno ili u nekom smislu dobro, već to znači *željeti* to stanje stvari. Vrijednosti su u nekom smislu inherentne našim željama. Ključni elementi prijedloga koji ovdje razmatram potječu od Hubina:

> Neohjumovac koji drži da postoje različite vrste konativnih stanja ne mora, govorim, držati da sva djelatnikova konativna stanja utemeljuju razloge za djelovanje. On je, kao temelje djelatnikovih razloga za djelovanje, slobodan identificirati neku posebnu vrstu konativnog stanja. Na primjer, pretpostavimo da među djelatnikovim konativnim stanjima možemo identificirati odgovarajući podskup stanja koji bismo mogli smatrati djelatnikovim *vrijednostima*. Teorija koja drži da je ono što djelatnik ima razloga učiniti (u smislu koji je relevantan za racionalnu opravdanost) određeno ne svim njegovim konativnim stanjima, već samo njegovim vrijednostima, i dalje bi bila neohjumovska teorija. (Hubin, 2003: 329)

Pretpostavimo li da svatko ima nešto što bismo mogli nazvati njegovim sustavom vrijednosti, onda je hjumovac u mogućnosti tvrditi da upravo taj sustav vrijednosti predstavlja pravi izvor praktičnih razloga. Presudno je registrirati da Hubin o vrijednostima govori kao o konativnim

stanjima, ali konativnim stanjima koja su "relativno trajna u vremenu, relativno stabilna pri refleksiji [i] relativno centralna za djelatnikovu konativnu psihologiju" (Hubin, 2003: 333). Zastupnik neohjumovske teorije je zato slobodan odrediti granicu između onih konativnih stanja koja generiraju praktične razloge i onih koja to ne čine. Samo ona psihološka stanja s kojima se subjekt poistovjećuje, koja jesu pod njegovom kontrolom i koja reflektiraju njegove prave vrijednosti mogu biti ona stanja na osnovi kojih on ima razloga djelovati. Hjumovac se ne bi trebao ustručavati prihvatiti prijedlog prema kojemu vrijednosti predstavljaju izvor praktičnih razloga baš zato što vrijednosti za njega i nisu ništa drugo nego konativna stanja.[54]

Bitna pretpostavka prijedloga jest da nam razloge ne pružaju bilo kakva konativna stanja, već samo ona s kojima se subjekt poistovjećuje. Upravo je zbog oslanjanja na tu pretpostavku hjumovac sposoban i eliminirati neke standardne prigovore koji bi mu se u ovom kontekstu mogli uputiti. Pogledajmo o čemu se radi. Vjerojatno središnji problem za ovakvu analizu vrednovanja – analizu prema kojoj je vrednovanje shvaćeno u kategorijama želja – predstavljaju primjeri čija je svrha ilustrirati da je rascjep između vrednovanja i želje ne samo moguć, već i prilično uobičajen (usp. Smith, 1994: 133–136). To što želim učiniti Φ još uvijek ne mora značiti da vrednujem činjenje Φ-a. Primjerice, "…ovisnik može željeti svoju euforičnu omamljenost, ali je ne [mora] vrednovati. … on se može mrziti zbog toga što želi nešto što uopće ne vrednuje. To je želja koje se on vrlo jako želi riješiti (Lewis, 2000: 70). Ideja je da ovisnik ne mora prihvatiti svoju želju da uzima drogu, već može preferirati stanje u kojemu tu želju nema. Zato bi u njegovu slučaju bilo sasvim netočno insistirati na tome da on vrednuje uzimanje droge. Nesporno je da naime ovisnik želi uzimati drogu, ali je isto tako nesporno da se on s tom željom ne poistovjećuje. Ali ako se on s tom željom ne poistovjećuje,

[54] Iako bi možda bilo ispravnije kazati da je *vrednovanje* konativno stanje, a ne da su *vrijednosti* konativna stanja, mislim da smisao ostaje isti. Jer vrijednosti su, kako to na jednom mjestu kaže David Lewis, "ono što imamo dispoziciju vrednovati" (Lewis, 2000: 68), a vrednovati nešto, prema prijedlogu koji ovdje razmatram, znači biti u određenom konativnom stanju – željeti nešto. Ideja je da su vrijednosti ovdje shvaćene isključivo kao subjektivna kategorija.

onda ta njegova želja ne reflektira njegove prave vrijednosti. Vrednovati neko stanje stvari dakle nije isto što i željeti to stanje stvari.

Uočimo da ovaj prigovor ni na koji način nije u koliziji s gornjim prijedlogom. Jer i sam hjumovac – kako je gore prikazan – prepoznaje našu sposobnost distanciranja od naših psiholoških stanja odnosno mogućnost da ih kritički preispitujemo i da o njima reflektiramo. U tome on vidi jedno od glavnih obilježja nas kao ljudskih bića. To mu i omogućuje da, kako je već rečeno, odredi granicu između onih psiholoških stanja koja mogu poslužiti kao izvor praktičnih razloga i onih kojima se ta uloga ne može dodijeliti. Ovdje ću samo ukratko razmotriti i komentirati dva načina na koja bi se takvo razlikovanje moglo provesti. Prvo, hjumovac bi u kontekstu ove rasprave mogao upotrijebiti distinkciju između želja prvog i drugog reda (usp. Frankfurt, 1988) te zagovarati tezu da samo one potonje reflektiraju prave vrijednosti i tako generiraju razloge za djelovanje. Ovisnik koji želi uzimati drogu, ali koji nema želju na osnovi te želje djelovati nema razlog uzimati drogu. To je stoga što se on ne poistovjećuje s onime što želi, nego s onime što želi da želi. Drugo, hjumovac bi mogao argumentirati da govor o željama koje nisu *naše* želje jednostavno nije govor o željama. Porivi, sklonosti i nagoni nisu želje u pravom smislu te riječi i zato je netočno ovisnika opisati kao onoga koji *želi* uzimati drogu. Slučaj ovisnika, najjednostavnijim riječima, nije slučaj u kojemu dolazi do rascjepa između subjektovih želja i vrijednosti, nego onaj u kojemu se sukobljavaju subjektove vrijednosti i nagoni. Ono od čega se subjekt ovdje distancira nisu njegove želje, već njegovi porivi.

Spomenuta rješenja nemaju jednaku težinu. Iako činjenje distinkcije između želja prvog i drugog reda na prvi pogled zvuči vrlo uvjerljivo, meni se čini da to ipak nije put kojim bi hjumovac trebao krenuti. Slabost tog rješenja jest u tome što nije jednostavno obrazložiti zašto bi upravo želje drugog reda trebale predstavljati pravi izvor razloga za djelovanje, a ne, recimo, želje trećeg ili četvrtog reda. Ugraditi vrijednosti u želje drugog reda u tom smislu

izgleda sasvim proizvoljno i teško je ne priznati da bi netko tko se protivi zasnivanju praktičnih razloga u željama bio sasvim u pravu kada bi ukazivao na tu činjenicu. Osim toga, opredijeliti se za tu strategiju značilo bi svom oponentu dozvoliti mnogo više nego što je to potrebno. Naime, hjumovska teorija razloga često je pod udarom brojnih kritika upravo zbog naširoko prihvaćene pretpostavke da je simpatizer te teorije obvezan smatrati da *sva* konativna stanja – neovisno o tome kakve su njihove međusobne razlike – pružaju razloge za djelovanje. Međutim, kako to napominje i Hubin u gore citiranom odlomku, to nije točno. Hjumovac sasvim konzistentno može braniti svoje gledište čak i ako neka od tih stanja proglasi nelegitimnim izvorom praktičnih razloga. Primjer ovisnika predstavljao bi prijetnju hjumovskom programu samo kada bi hjumovac smatrao da razlike među konativnim stanjima nisu relevantne, a moram priznati da ne vidim zašto bi on bio osuđen zagovarati takvo gledište. Zato mislim da bi druga, gore predložena strategija ipak mogla polučiti bolje rezultate. Vrijednosti su, prema toj strategiji, konstitutivni element želja, što znači da hjumovac s lakoćom može priznati da upravo one generiraju razloge za djelovanje.

4.4. O instrumentalnoj normi

Za Williamsa procesi praktičnog rasuđivanja nisu iscrpljeni instrumentalističkim načelom, već on prepoznaje i druge racionalne procese (kao što je, recimo, mašta) koji također vrše značajnu ulogu u korekciji motivacijskog skupa. Ipak, kao što je već ranije navedeno, Williamsa je opravdano svrstati među zastupnike instrumentalizma budući da deliberativni procesi o kojima on govori imaju instrumentalni duh: u čemu god da se sastojali i kakva god bila njihova ostala obilježja, ti procesi moraju biti pod kontrolom postojeće motivacije. Budući da su želje te koje nas motiviraju, slijedi da upravo one moraju biti u podlozi svakog rasuđivalačkog procesa. Hjumovski internalizam i instrumentalizam stoga predstavljaju dvije strane istog novčića: obje su pozicije utemeljene u pretpostavci da razlozi za djelovanje ne

mogu postojati bez prisutnosti odgovarajućih želja. Instrumentalistička se koncepcija međutim u novije vrijeme nalazi pod udarom brojnih kritika i stječe se dojam da ona zapravo uopće ne zaslužuje status kakav ima u ovoj diskusiji. Prema toj koncepciji, ono što imamo razlog učiniti jest poduzeti ona sredstva kojima možemo ostvariti svoje prethodno određene ciljeve: želim li učiniti *x* i vjerujem li da je *y* način da ostvarim *x*, onda imam razlog učiniti *y*. Iako se ovaj princip na prvi pogled čini nekontroverznim, mnogi autori, kao što se moglo naslutiti i u prethodnoj diskusiji, osporavaju njegovu istinitost. Postoji jako puno slučajeva – barem tako glasi glavni prigovor – u kojima subjekt jednostavno nema razlog poduzeti sredstva koja su mu potrebna za realizaciju vlastitih želja. Ovdje ću takve slučajeve, radi jednostavnosti i lakšeg razumijevanja ovog prigovora, podijeliti u dvije kategorije: razlog za poduzimanjem sredstva ne postoji u (i) slučajevima u kojima ne postoji razlog za sam cilj (za samu želju) te (ii) u slučajevima u kojima je poduzimanje sredstva bilo u sukobu s nekim drugim ciljevima (željama) koje subjekt može imati, bilo u sukobu s nekim drugim normama kojima je subjekt isto tako podvrgnut.

Što se tiče prve grupe prigovora, mnogi se autori slažu da mi jednostavno ne možemo *stvoriti* razlog za ostvarenjem nekog cilja samo zato što smo taj cilj učinili *svojim* ciljem (Broome, 1999, 2004, 2005; Dancy 2000; Raz, 2005). Evo i nekoliko relevantnih formulacija tog prigovora:

> [Ž]elja za koju ne postoji razlog ne može stvoriti razlog za činjenje onoga što bi joj bilo od koristi (Dancy, 2000: 39).
>
> Nije izvjesno da ste, samo zato što ste usvojili neki cilj, učinili da imate razlog ostvariti ga. Možda se ništa ne može kazati u prilog vašeg cilja; može biti besmislen, glup ili štetan (Broome, 2005: 1).

> Ljudi mogu imati ciljeve koje …. ne bi trebali imati, ciljeve koji su bezvrijedni ili čak zli. Mogu li oni imati razlog omogućiti njihovu realizaciju samo zato jer su to njihovi ciljevi? (Raz, 2005: 10).

Problem s ovim prigovorom – čiju snagu svakako ne treba zanemariti – jest u tome što polazi od sasvim drukčijeg shvaćanja praktičnih razloga, shvaćanja koje hjumovski usmjereni autori, da se tako izrazim, nemaju nikakvog razloga prihvatiti. Dok hjumovac polazi od pretpostavke da su razlozi za djelovanje *subjektivni*, u gornjim je formulacijama sadržan objektivistički pristup – pretpostavlja se da za određene ciljeve jednostavno ne postoje dobri razlozi. Drugim riječima, u prigovoru se pretpostavlja da za neke želje jednostavno "ne postoji razlog", da su neki ciljevi "glupi", "štetni" ili "bezvrijedni", a pritom ostaje neodgovoreno na čemu su te procjene u prvom redu zasnovane. Nekako se čini da je vrlo teško (ako ne i nemoguće) postići konsenzus oko kriterija kojim bismo se trebali ravnati pri vrednovanju određenih ciljeva, tj. izgleda da ne postoji neko objektivno mjerilo na koje se možemo pozvati pri klasificiranju nekog cilja kao "besmislenog" ili "bezvrijednog". Vjerujem da upravo to Kolodny ima na umu kada, razmatrajući ovo gledište, pita: "Kako možemo riješiti pitanja o tome što je razlog za što? Što se tiče pukog pojma razloga, čini se da bilo što može biti razlog za bilo što drugo" (Kolodny, 2005: 511). Hjumovac, s druge strane, nema ovu poteškoću. Kazati da je neki cilj "besmislen" ili "bezvrijedan" ima smisla samo ako je to rečeno *u odnosu* na neke druge ciljeve (odnosno želje) koje subjekt ima. Neki cilj nikada nije štetan ili bezvrijedan sam po sebi, već uvijek samo *za nekoga*.[55]

[55] Ne treba zaključiti kako mislim da ova opservacija potkopava stajalište koje zastupaju Broome, Raz i Dancy. Bitno je imati na umu da to stajalište počiva na *potpuno drukčijim pretpostavkama* nego što je to slučaj kod hjumovskih instrumentalista i moralnih racionalista. Glavna razlika je u tome što su spomenuti autori *neredukcionisti* u pogledu razloga. Dok hjumovci i moralni racionalisti razlog definiraju u kategorijama racionalnosti (ono što subjekt ima razlog učiniti jest ono što bi učinio kada bi bio racionalan), ovi autori "negiraju da razlozi za djelovanje … mogu biti objašnjeni u terminima racionalnosti, ili u nekim drugim terminima" (Kolodny, 2005: 511). (Za eksplicitno razdvajanje razloga i racionalnosti posebno vidi Broome (1999)). Budući da su polazne točke u raspravi sasvim drukčije, u ovom ću radu navedenu neredukcionističku kritiku instrumentalizma ostaviti po strani.

Usmjerimo sada pažnju na drugu kategoriju prigovora instrumentalnom shvaćanju razloga. Karakteristično obilježje ove vrste prigovora jest u tome što se njegovi zastupnici oslanjaju na naše intuitivno razumijevanje praktičnih razloga – oni pokušavaju konstruirati situacije u kojima je intuitivno očito da osoba, unatoč tome što ima određenu želju, nema razlog poduzeti sredstvo kojim bi tu želju mogla ostvariti. Pogledajmo primjer koji nudi Peter Railton: "Osoba koja želi proizvesti masovnu paniku može prosuditi … da je najbolje raspoloživo sredstvo za taj cilj pustiti otrovni plin u gradsku podzemnu željeznicu, ali moglo bi se razumno sumnjati da ona (ili bilo tko) ima ikakav *dobar razlog* to učiniti (Railton, 2006: 272–273, moj kurziv). Nije teško razumjeti zašto Railton smatra da bi bilo "razumno sumnjati" da osoba u ovom slučaju ima razlog. Imati dobar razlog za neku radnju znači imati razlog koji *opravdava*, a – intuitivno promatrajući – malo tko će smatrati da je opravdano pustiti otrovni plin u podzemnu željeznicu unatoč tome što to može biti jedini način za realizaciju vlastitog cilja. Zanimljiv primjer nudi i Dancy: "Ostavljen sam u pustinji, očajnički želim hladno pivo; Vrag mi ponudi pivo u zamjenu za moje dijete. Imam li nekakav razlog dati moje dijete Vragu? Mislim da nemam" (Dancy, 1993: 181). Zanimljivo je da Dancy ne obrazlaže svoje uvjerenje, već ga samo navodi. Možda zato što smatra da je svako obrazlaganje u ovom slučaju sasvim suvišno? Naime, zar nije *evidentno* da osoba nema nikakav razlog dati svoje dijete Vragu?

Odgovor na to pitanje, čini mi se, opet ovisi samo o tome što podrazumijevamo pod pojmom razloga. Imati razlog za neku radnju znači da je ta radnja – promatrajući iz perspektive racionalnosti – opravdana. Ali ako je neka radnja opravdana iz perspektive racionalnosti, to ne znači da je ona ujedno opravdana iz perspektive moralnosti. Vratimo li se na Railtonov primjer, zašto se mnogima čini da čovjek nema razlog pustiti plin u podzemnu željeznicu jest taj što je za takvu radnju teško pronaći moralno opravdanje. Ali uočimo da u takvoj radnji ne bi bilo ništa iracionalno. Slično vrijedi i za Dancyev primjer. Iako se na prvi

pogled može doista činiti da osoba nema *nikakav* razlog predati svoje dijete Vragu, meni izgleda da je takav zaključak ipak preuranjen. *Neki* razlog, koliko god bio minimalan, u ovom slučaju ipak postoji, a hoće li taj razlog ujedno biti i onaj na osnovi kojeg će osoba djelovati ovisi isključivo o tome koje su njezine druge želje i sklonosti.[56] Kada bi netko doista i djelovao na osnovi tog razloga, nitko ne dvoji da bi takav čin bio podložan svakoj moralnoj osudi i da bismo onoga tko bi to doista i učinio smatrali moralno monstruoznom osobom, ali iz takvih moralnih kvalifikacija ne bi slijedio nikakav zaključak o (i)racionalnosti takvog postupka. Naravno, ovaj odgovor vrijedi samo pod pretpostavkom da moralni zahtjevi nisu zahtjevi same racionalnosti. Netko bi naime mogao primijetiti da je neosjetljivost na moralne razloge u konačnici jedan oblik iracionalnosti. Kako na jednom mjestu Hubin: [N]eki kritičari hjumovstva traže palicu kojom bi potukli one koji se ne obaziru na moralne razloge. Oni su ih već potukli palicom nemoralnosti – bez učinka. Oni žele veću palicu; oni žele palicu iracionalnosti" (Hubin, 1999: 40). Kritičari o kojima Hubin govori su oni koji svoju inspiraciju nalaze u Kantu – oni kritičari koji moralnost nastoje reducirati na racionalnost, ali racionalnost shvaćenu u kategoričkom, a ne instrumentalnom smislu. Njihov je pristup tema sljedećeg poglavlja.

[56] Ovdje se djelomično oslanjam na Hubina koji također kritički razmatra jednu varijantu Dancyeva slučaja. Hubin zaključuje da za osobu ovdje "doista postoji razlog predati dušu djeteta, ali da ona ima toliko jači razlog to ne učiniti da je ovaj razlog teško vidjeti … " (Hubin, 1999: 34).

5. Kantovski internalizam razloga

Konzekvence hjumovskog internalizma sasvim su jasne: ako je prisustvo razloga ovisno o elementima naših subjektivnih motivacijskih skupova, onda takvo stajalište ne ostavlja prostor *kategoričnim* razlozima. Razlog je uvijek razlog *za nekoga* te govor o razlozima zapravo i nema smisla ako prethodno nije utvrđena odgovarajuća motivacija kod određenog subjekta. Nije teško uvidjeti da tim shvaćanjem internalistička pozicija vrlo brzo zapada u jedan oblik relativizma. Relativistički aspekt sastoji se u tome što elementi nečijeg motivacijskog skupa, nakon što bivaju korigirani procesom rasuđivanja, uvijek reflektiraju elemente s kojima je takvo rasuđivanje započelo, a jasno je da su motivacijski skupovi različiti od osobe do osobe i teško je za očekivati da proces rasuđivanja na tu činjenicu može značajnije utjecati. Taj proces mora krenuti od želja koje *de facto* imamo u svojemu motivacijskom skupu te stoga nije realno za pretpostaviti da hipotetičke želje – koje nastaju kao rezultat tog procesa – konvergiraju u jedan zajednički skup.[57] Nezgodna implikacija takvog oblika relativizma jest u tome što razlozi za djelovanje u krajnjoj instanci mogu proizlaziti iz moralno dvojbenih želja: kada bi moralno neprihvatljive želje nekim slučajem "preživjele" proces rasuđivanja, internalist bi bio obvezan tvrditi da netko tko ima takve želje ima i *razlog* na osnovi njih djelovati.

Zašto se mnogima ta ideja činila u tolikoj mjeri nevjerojatnom i neodrživom, zgodno je ilustrirao Parfit koristeći poznati Humeov primjer. Pretpostavimo da netko – premda *jasno prosuđuje* i *posjeduje sva relevantna vjerovanja* (dakle, netko tko je već korigirao svoj motivacijski skup) – preferira uništenje cijelog svijeta ogrebotini na svom prstu. Parfit nalazi očitim da takva preferencija – iako je "preživjela" proces rasuđivanja – *ne* pruža razlog za djelovanje onome tko je ima. Ili, možda bolje, ne samo da netko s takvom preferencijom

[57] O mogućnosti konvergencije će kasnije biti više riječi.

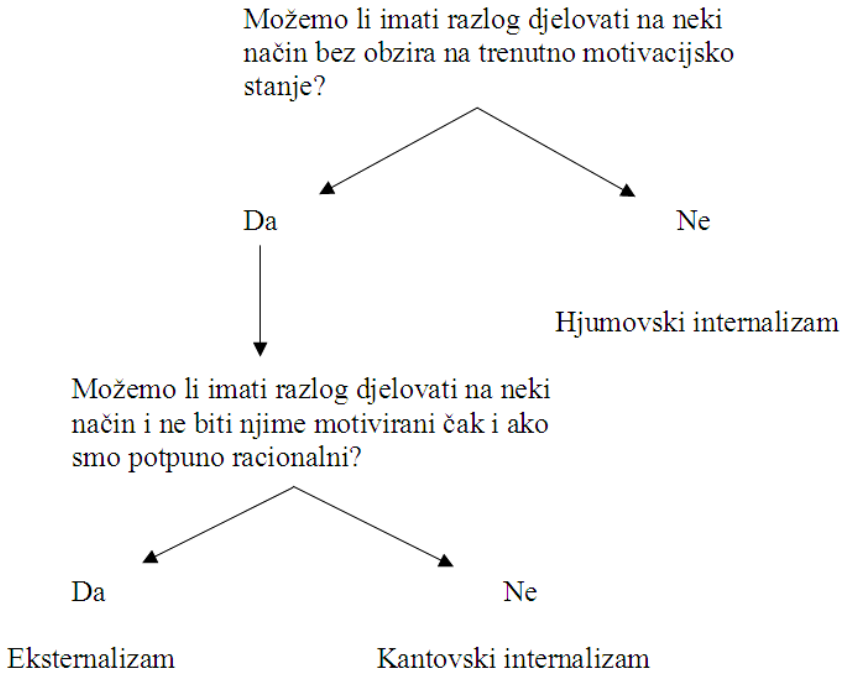
nema razlog na osnovi nje *djelovati*, već se čini, smatra Parfit, da ne postoji razlog zašto bi netko takvu preferenciju uopće imao. Može li internalist pronaći izlaz iz ove poteškoće? Parfit sugerira mogući odgovor, ali potom odmah pokazuje zašto bi internalist takvim odgovorom značajno rasklimao svoju poziciju. Moguća strategija kojom se internalist može poslužiti jest da tvrdi da "netko tko ima takve želje ne 'prosuđuje jasno'. Ali ovime zapravo prihvaća prigovor. Definirajući što znači 'jasno prosuđivanje", [internalist] bi trebao referirati na želje koje su u pitanju. Trebao bi odlučiti koje su želje *intrinzično iracionalne*" (Parfit, 1984: 118, moj kurziv).

Parfitovo odbacivanje mogućeg internalističkog odgovora nije teško razumjeti. Budući da, prema hjumovskom gledištu, naše želje ne mogu biti podvrgnute racionalnoj kritici, internalist bi – smatrajući iracionalnom preferenciju za uništenjem cijelog svijeta – svoju poziciju učinio nekoherentnom. Na taj bi način u znatnoj mjeri odstupio od njezine temeljne karakteristike – da razlozi za djelovanje proizlaze iz želja *kakve god da one jesu*. Navodna nemogućnost internalista da pruže uvjerljiv odgovor na taj problem za neke je bila sasvim dovoljna da to stajalište proglase krajnje neuvjerljivim. Međutim, postoje i oni koji vjeruju da priklanjanje eksternalizmu nije jedina opcija te da je napuštanje internalizma ipak preuranjeno. Izbjeći relativističke implikacije, oni tvrde, nije moguće samo putem eksternalističke strategije, već je objektivnost razloga moguće dokazivati unutar same internalističke pozicije. Internalizam razloga ne mora nužno biti osuđen na hjumovsku interpretaciju te je, kako oni smatraju, moguće izgraditi takvu internalističku poziciju koja bi sadržavala središnje elemente Kantove etike. Ti bi elementi doduše mogli biti prisutni samo pod pretpostavkom napuštanja fundamentalne premise Williamsova argumenta – premise da je praktična racionalnost u osnovi instrumentalne prirode.

Moguće je navesti barem jedan razlog zašto prihvaćanje instrumentalne koncepcije ne ide u prilog Kantovoj koncepciji morala. Kant je, kao što ćemo ubrzo i vidjeti, smatrao da je

moralno djelovanje racionalno djelovanje. Ali ako se praktična racionalnost ne svodi na ništa drugo nego na biranje najboljih sredstava za zadovoljenje naših želja, onda neminovno proizlazi da moral nije nužno, već samo *kontingentno* utemeljen u racionalnosti. Moral bi u tom slučaju bio sustav hipotetičkih imperativa, što bi značilo da bi moralni razlozi bili ovisni o našim proizvoljnim ciljevima. Taj je rezultat nešto što kantovci pod svaku cijenu žele izbjeći. Njihovo je gledište da moralni razlozi nisu samo individualni, već univerzalni te da je njihova priroda kategorička odnosno da mi te razloge imamo neovisno o našim arbitranim željama ili drugim psihološkim stanjima.

Kantovski internalizam je pozicija koja, slobodnije rečeno, nastoji pomiriti hjumovsku verziju internalizma i eksternalističko stajalište. Dok zagovornici hjumovskog internalizma uspješno udovoljavaju objašnjavalačkom ograničenju prema kojemu razlozi moraju biti u stanju motivirati, intuiciji da postoje objektivni razlozi ta pozicija ne uspijeva udovoljiti, a upravo u tome mnogi i vide njezinu glavnu slabost. S druge strane, eksternalistički pokušaji udovoljavanja objašnjavalačkom ograničenju suočavaju se s velikim poteškoćama, ali zato je njihovo stajalište u skladu s nekim uobičajenim i svakodnevnim intuicijama o postojanju objektivnih razloga. Drugim riječima, relacija razloga i motivacije u okviru eksternalizma ostaje neprecizirana, ali zato se o razlozima ipak govori kao o objektivnim. Kantovski je internalizam moguće promatrati kao pokušaj objedinjavanja ovih dviju suprotstavljenih pozicija odnosno kao pokušaj argumentiranja u prilog postojanju objektivnih razloga, ali objektivnih razloga koji bi nas trebali pokretati na djelovanje. Za razliku od hjumovskih internalista, oni kantovske orijentacije ne smatraju da je motivacija preduvjet za pripisivanje razloga, već da motivacija djelovati u skladu s objektivnim razlozima nastaje pod pritiskom razuma. Odnos triju opisanih pozicija shematski se može ovako prikazati:



Tablica 5.1. Taksonomija internalizama (preuzeto iz Gert, 2004: 172)

5.1. Motivacijski i sadržajni skepticizam

Tvrdnja da procesi rasuđivanja ne mogu sami po sebi generirati motivaciju vjerojatno je jedna od najosnovnijih pretpostavki hjumovskog internalizma. Sva motivacija mora biti izvedena iz već postojeće prethodne motivacije. Tu misao, podsjetimo se, nalazimo u Humeovoj tvrdnji o "razumu kao robu strasti" ili, drukčije formulirano, u tezi da "nijedna pojedinačna temeljna motivacija [ne može biti] racionalno zahtijevana ili zabranjivana" (Gert, 2004: 41). Mi ne možemo, takoreći, stvoriti motivaciju pod pritiskom razuma. Uloga razuma za hjumovce sastoji se u pružanju informacija o tome kako na najbolji način realizirati motivaciju koja je *već* prisutna i u korigiranju one motivacije koja je rezultat pogrešnog uvida u određeno stanje stvari. U okviru hjumovskog internalizma, to bi značilo da rasuđivanjem ne podvrgavamo našu motivaciju *autoritetu* razuma u svrhu utvrđivanja je li naša motivacija racionalna ili nije, već je samo nastojimo korigirati u svjetlu novih činjenica i saznanja o svijetu. Hjumovski internalist dakle ne može jamčiti da bi se, na primjer, nečije moralno dvojbene želje nužno

promijenile kada bi subjekt imao sva istinita vjerovanja i kada bi ispravno rasuđivao, već se takvoj promjeni može samo nadati. Kada bi nečije moralno neprihvatljive želje nekim slučajem "preživjele" proces rasuđivanja, hjumovski internalist bio bi obvezan tvrditi da bi ta osoba imala razlog na osnovi njih djelovati. Williamsov model nije hjumovski u pravom smislu riječi budući da se procesi rasuđivanja ne svode samo na instrumentalizam, već uključuju, kao što smo vidjeli, upotrebu imaginacije, razmišljanje o tome kako zadovoljiti različite elemente u S-u ukoliko između njih postoji konflikt itd. Međutim – i to je ključni razlog zašto se ovo stajalište može nazvati neo-hjumovskim – procesi rasuđivanja, u čemu god da se sastojali, moraju početi iz već postojeće motivacije jer sami po sebi nisu u stanju generirati motivaciju koja već nije prisutna.

Ovaj posljednji stav Christine Korsgaard nazvala je *motivacijskim skepticizmom* (Korsgaard, 1996a: 311–334). Središnja Korsgaardina teza jest da motivacijski skepticizam odnosno skepticizam u pogledu mogućnosti racionalne motivacije zapravo nema neovisnu snagu, već da je zasnovan na tzv. *sadržajnom skepticizmu*, tj. skepticizmu u pogledu pitanja imaju li formalna načela rasuđivanja ikakav sadržaj i mogu li biti vodičem u izboru i djelovanju. Njezina je namjera pokazati da internalist ne bi smio polaziti od motivacijskog skepticizma te tako unaprijed odrediti po kojim se načelima odvija praktično rasuđivanje. Upravo suprotno, on bi prvo trebao imati jasnu sliku o tome što bi se moglo smatrati načelom rasuđivanja, a tek onda izvlačiti zaključke o motivacijskoj snazi takvih načela. Jednostavnije rečeno, pogrešno je krenuti od pretpostavke da čisti razum nije u stanju motivirati te na osnovi te pretpostavke izvlačiti zaključke o prirodi načela praktičnog razuma, već je o motivacijskoj snazi razuma primjereno govoriti tek nakon što utvrdimo koja bismo načela uopće mogli smatrati načelima praktičnog razuma.

Zadržimo se još malo na Korsgaardinu prijedlogu. Pretpostavimo, slijedeći Korsgaard, da je Williams u pravu te da je jedno od načela praktičnog razuma instrumentalno načelo.

Prema tom načelu, podsjetimo se, subjekt ima razlog pribjeći onim sredstvima kojima može ostvariti prethodno određene ciljeve: ako S želi ostvariti cilj *x* i ima informaciju o tome da taj cilj može ostvariti ako učini *y*, onda S ima razlog učiniti *y*. Prihvaćajući ovu normu kao jednu od normi rasuđivanja, razumno je pretpostaviti da ljudi od te norme u nekim okolnostima mogu i odstupiti, a jedan očit način na koji bi se to moglo dogoditi jest da netko ne stvori motivaciju za poduzimanjem sredstva koje mu je potrebno da bi ostvario svoj cilj. Što reći o takvom slučaju? Što ako S nije motiviran učiniti *y* a zna da je to jedini način na koji može ostvariti *x*? Takav je slučaj paradigmatičan slučaj praktične *iracionalnosti*: razni "poremećaji" poput *akrasije*, depresije, bijesa i psihičke bolesti, smatra Korsgaard, mogu "blokirati" subjektovu motivaciju za realizacijom vlastitog cilja te njegovo djelovanje tako učiniti iracionalnim. Imajući ovo u vidu, ona nadalje primjećuje, Williamsov internalizam *ne* tvrdi da je nešto razlog samo ako nas može motivirati, već da je nešto razlog samo ako nas može motivirati ukoliko smo *racionalni*.[58] Naime, iz činjenice da osoba nije motivirana učiniti *y* još uvijek ne možemo zaključiti da ona nema razlog učiniti *y*, već samo da ona *ili* nema razlog učiniti *y ili* nije racionalna. No ako smo uvidjeli da netko može manifestirati iracionalnost u slučaju instrumentalnog rasuđivanja, onda se skepticizam u pogledu motivacijske snage ostalih oblika praktičnog rasuđivanja više i ne čini toliko osnovanim. Pogledajmo zašto.

Primjer kojim Korsgaard nastoji potkrijepiti svoju tvrdnju odnosi se na načelo prudencijalnosti. Kao što smo vidjeli kada je bilo riječi o Williamsovu internalizmu, Williams smatra da je želja za prudencijalnim ponašanjem samo jedan od mogućih elemenata koji se mogu nalaziti u subjektivnom motivacijskom skupu. Želja za vlastitom dobrobiti je kontingentna i ne postoji jamstvo da će ona biti dio svakog S-a. U slučaju da ta želja nedostaje u nečijem S-u, onda ne bismo mogli tvrditi da ta osoba ima *razlog* djelovati prudencijalno. To bi bio pokazatelj da očuvanje vlastite dobrobiti jednostavno nije nešto što ona želi ostvariti.

[58] Iako sâm Williams ne koristi izraz "racionalnost", takav opis njegove pozicije sigurno nije pogrešan. Stadij racionalnosti o kojem govori Korsgaard zapravo je stadij koji postižemo *nakon* što smo subjektivni motivacijski skup već korigirali procesom deliberacije.

Ili, poznatim Humeovim riječima: "…malo se protivi razumu da pretpostavim čak svoje priznato manje dobro svojemu većem dobru, i da vatrenije volim ono prvo negoli ovo drugo" (*R* 355). Korsgaardina razmatranja, međutim, sugeriraju da ovaj opis situacije nije jedini. Nedostatak motivacije za prudencijalnim djelovanjem, umjesto toga, može biti opisan kao manifestacija iracionalnosti. Najjednostavnije rečeno, moglo bi se tvrditi da bi netko – čak i kada ne bi imao takvu želju u svom S-u – *sigurno* bio motiviran djelovati prudencijalno kada bi bio racionalan. Činjenica da subjekt trenutno nema takvu motivaciju ne znači da u stadiju racionalnosti takvu motivaciju ne bi stvorio. Takva bi motivacija, ako je Korsgaard u pravu, tada predstavljala *zahtjev racionalnosti*.

Kojem ćemo se opisu gornjeg slučaja prikloniti ovisi o tome što smatramo da su načela praktičnog rasuđivanja. Mi ne možemo samo pretpostaviti da nas formalna načela ne mogu motivirati sama po sebi, već takvu tvrdnju treba i dokazati. Ako su Hume i Williams u pravu, onda bismo morali reći da onaj tko, umjesto većeg dobra, bira nešto što je manje dobro, ima za to najjaču želju, a samim time i najjači razlog. S druge strane, ako je Korsgaard u pravu, onda bi se takvo djelovanje ipak moglo okarakterizirati kao iracionalno. Pretpostavka je da bi motivacija za prudencijalnim djelovanjem bila stvorena u stadiju racionalnosti, što bi na koncu i pokazalo da smo sposobni stvoriti motivaciju pod pritiskom razuma odnosno načela po kojem se rasuđivanje odvija, a Williams u svojoj argumentaciji nije pokazao da ta pretpostavka nema osnovu. Sada možemo i bolje razumjeti sljedeće Korsgaardine riječi:

> Ako netko prihvati internalistički zahtjev, slijedi da će čisti praktični razum postojati ako i samo ako smo sposobni biti motivirani zaključcima operiranja čistog praktičnog razuma kao takvog. Nešto u nama nas mora učiniti sposobnima da budemo njima motivirani, a to nešto će biti dio subjektivnog motivacijskog skupa. Williams smatra da je to razlog za sumnju u postojanje čistog praktičnog razuma, dok ono što se čini da slijedi iz internalističkog zahtjeva jest ovo: ako možemo biti motivirani tvrdnjama koje proizlaze iz čistog praktičnog razuma,

> onda ta sposobnost pripada subjektivnom motivacijskom skupu svakog *racionalnog* bića. (Korsgaard 1996a: 328, moj kurziv)

Korsgaard dakle ne daje konkluzivni argument kojim pokazuje da je Hume nedvojbeno u krivu u tvrdnji da nas čisti praktični razum ne može motivirati. Njezina je namjera znatno skromnija: ona samo želi naznačiti da skepticizam u pogledu toga može li praktični razum imati utjecaja na ljudske postupke ne smije polaziti od te Humeove pretpostavke. Upravo suprotno, "...motivacijski skepticizam u pogledu praktičnog razuma *ovisi o* ... skepticizmu u pogledu mogućeg sadržaja racionalnih zahtjeva" (Korsgaard 1996a: 331–332, moj kurziv). Korsgaardina linija argumentacije otvara prostor za sljedeću mogućnost. Kada bi se pokazalo da postoje načela rasuđivanja koja bi mogla samostalno generirati motivaciju, internalistička bi pozicija i dalje mogla biti očuvana. Jedina, ali esencijalna razlika sastojala bi se u tome što bi se u tom slučaju izbjegle relativističke posljedice Williamsova internalizma te bi se ipak mogla zagovarati objektivnost razloga. Naime, motivacija više ne bi bila preduvjet za postojanje razloga, već bi nužno nastajala kao rezultat prepoznavanja kategoričkih razloga – sva bi racionalna bića, samim time što su racionalna, nužno bila motivirana (moralnim i prudencijalnim) razlozima.

Ne treba stoga isključiti mogućnost da pored instrumentalne norme (i njezinih, kako kaže Korsgaard "prirodnih ekstenzija") – za koju se obično smatra da je osnovna norma praktičnog zaključivanja – možda postoje i norme koje nam govore kakve bismo želje uopće *trebali* imati. U skladu s tim, ono što kantovski usmjereni oponenti hjumovskog internalizma moraju pokazati jest da instrumentalna norma racionalnosti nije jedina norma, već da postoje i druge norme racionalnosti iz kojih proizlazi postojanje kategoričnih razloga. Cullity i Gaut (1997: 23) ukazuju na nekoliko mogućih strategija koje kantovcima stoje na raspolaganju. Prvo, oni mogu tvrditi da instrumentalna norma samostalno nema nikakvu snagu te da je

njezino postojanje u potpunosti ovisno o postojanju norme koja se na nas odnosi bezuvjetno.[59] Drugo, oni mogu ukazati na neke formalne karakteristike praktičnog razuma, a potom iz tih formalnih karakteristika izvesti postojanje kategoričnih razloga. Treće, oni mogu tvrditi da kategorički razlozi proizlaze iz onoga što bi se moglo nazvati "konstitutivnim ciljem djelovanja".[60] Svrha ove podjele, koju svakako treba uzeti s oprezom, prvenstveno je pragmatična. Nije prijeporno da u pogledu odnosa etike i praktične racionalnosti danas postoji nepregledno mnoštvo argumenata i različitih pozicija i zato nam takva podjela može samo olakšati snalaženje u ovoj debati. Bilo kako bilo, smatram da glavna obilježja kantovskog internalizma najviše dolaze do izražaja u drugoj strategiji pa će upravo ona i biti predmetom mog glavnog interesa u ostatku ovog poglavlja. Dok sam Kant kategoričke razloge izvodi iz *univerzalnosti* praktičnog razuma, neki suvremeni neokantovski pokušaji fundamentalnu karakteristiku praktičnog razuma nalaze u zahtjevu za *koherencijom* naših mentalnih stanja. Krenimo od Kanta.

5.2. Kantov argument

Kada bi Kantovu filozofiju morala bilo potrebno svesti na jednu rečenicu, ne bi bilo netočno kazati da se radi o pothvatu čiji je cilj ustanoviti mogućnost postojanja kategoričkih moralnih

[59] Ovu strategiju brani Korsgaard (1997) koja argumentira da ne može postojati razlog za poduzimanjem sredstva ako pritom ne postoji razlog za sam cilj. Za Korsgaard to znači da mora postojati načelo koje nam govori koje ciljeve imamo *razloga željeti*. Ali, priznamo li postojanje takvog načela, ona nadalje ističe, tada je jasno da se odmičemo od instrumentalizma te da instrumentalno načelo ne može biti *jedino* načelo praktičnog razuma. (Instrumentalist smatra da su naši razlozi određeni onime što želimo a ne onime što imamo razloga željeti.) Ova kritika međutim polazi od potpuno drukčijeg shvaćanja praktičnih razloga – shvaćanja koje hjumovac uopće ne mora prihvatiti. On ne mora prihvatiti da razlozi moraju biti izvedeni iz nekih "osnovnijih" razloga koji nisu ovisni o subjektivnim, psihološkim čimbenicima. Kako kaže Donald Hubin: "Neohjumovci ne vjeruju – ili barem ne trebaju vjerovati – da je praktična racionalnost objektivna stvar. … Za neohjumovca djelatnikovi konačni ciljevi … sami po sebi nisu ni racionalno preporučljivi ni racionalno nepreporučljivi. … Djelatnikovi konačni ciljevi povjeravaju taj status politikama, planovima i radnjama ne zbog toga što ti ciljevi sami po sebi imaju neki normativan položaj. Oni to čine zato jer konkretno svojstvo 'biti racionalno preporučljiv' naprosto jest svojstvo 'biti valjano povezan s tim golim činjenicama'. … (Hubin, 2001: 466–467).

[60] Zastupnik ove strategije je Velleman (1996). Kao što je istina konstitutivni cilj vjerovanja, tako je, Velleman tvrdi, autonomija konstitutivni cilj djelovanja. Kao što se razlozi za vjerovanje odnose samo na onoga tko ima sklonost prema istini, tako se i razlozi za djelovanje odnose samo na onoga tko ima sklonost prema autonomiji. Međutim, ti su razlozi kategorički baš zato što tu sklonost *mora imati* svaki vjerovatelj i djelatnik. Imati sklonost prema istini odnosno autonomiji jest upravo ono što nas čini vjerovateljima odnosno djelatnicima.

zahtjeva – zahtjeva koji se na nas odnose bezuvjetno. Dokazivanje te teze središnji je zadatak njegova *Utemeljenja metafizike ćudoređa*, "najznačajnijeg djela moralne filozofije nakon Aristotela" (Williams, 1985: 55). Kako i sam Kant piše: "…ovo utemeljenje nije ništa više nego traženje i utvrđivanje *vrhovnoga principa* moralnosti" (*U* 64, kurziv u originalu). Završetak potrage za tim principom dobro je poznat: vrhovni princip moralnosti jest kategorički imperativ koji izvire iz same naše prirode kao racionalnih bića; samim time što smo racionalna bića nužno smo podložni univerzalnom moralnom zakonu. Očito je da Kant time dovodi u neodvojivu vezu kategorički imperativ i samu racionalnost, što u konačnici znači da nije riječ samo o vrhovnom moralnom zakonu, nego i o vrhovnom *racionalnom* zakonu. U *Utemeljenju metafizike ćudoređa* Kant nastoji ispuniti dva cilja: pokazati da je kategorički imperativ norma racionalnosti i pokazati da moralne norme imaju oblik kategoričkih imperativa. Pružiti vjernu rekonstrukciju Kantova dokaza nije jednostavno budući da je gotovo svaki korak u njegovoj argumentaciji predmetom brojnih diskusija i interpretacija. Za svrhu ove rasprave dovoljno je ukazati na neka ključna mjesta njegova etičkog sustava, a koja, u našem kontekstu, mogu poslužiti kao opravdanje zašto je Kanta uopće moguće smatrati internalistom u pogledu praktičnih razloga. Osim toga, razmatranje Kantova razumijevanja statusa moralnih razloga također može rasvijetliti i suvremena kantovska nastojanja utemeljenja etike u racionalnosti, a o kojima će kasnije biti više govora.

Tumačiti Kantov pothvat u svjetlu rasprave o unutarnjim i vanjskim razlozima već na prvi pogled može izgledati čudno. Sâm Kant svoju filozofiju morala nije razvijao imajući u vidu spomenutu distinkciju, a moglo bi se i prigovoriti da velike razlike u terminologiji također predstavljaju otežavajuću okolnost za vjerodostojnost takve interpretacije. Sklon sam vjerovati da se navedene poteškoće ipak lako mogu izbjeći. Da anticipiram nadolazeću raspravu, točno je da Kant nije razmišljao u kategorijama unutarnjih i vanjskih razloga, ali u njegovoj je etici ipak sadržan duh internalističke pozicije. To je vidljivo iz njegova uvjerenja

da moralni zahtjevi – unatoč tome što je njihova priroda kategorička – moraju na neki način biti povezani sa subjektovim motivacijskim sustavom. Konkretnije rečeno, da bi bilo moguće ustanoviti da je Kantova teorija kompatibilna s eksternalističkim intuicijama, ali da istodobno ostaje unutar internalističkih okvira, potrebno je pokazati da je u njoj (i) ostavljen prostor za objektivne, kategoričke razloge čije je postojanje neovisno o elementima nečijeg motivacijskog skupa, ali da (ii) između kategoričkih razloga i tog skupa mora postojati određena veza.

Budući da je cijelo Kantovo djelo posvećeno dokazivanju postojanja kategoričke norme racionalnosti odnosno norme koja generira objektivne, kategoričke razloge za djelovanje, ispravno je reći da Kant pokušava ispuniti prvi zahtjev. No ispunjenje prvog zahtjeva također je povezano s drugim zahtjevom. Mogućnost uspostavljenja odgovarajuće relacije između kategoričkog imperativa i našeg motivacijskog ustrojstva, neki smatraju, Kant vidi u tezi o racionalnoj prirodi kao svrsi po sebi odnosno u tezi o postojanju zajedničkog cilja svih racionalnih bića.[61] Središnja internalistička misao da normativnost mora imati odjek u motivaciji sadržana je, prema njima, u Kantovoj formuli svrhe po sebi. Kantova pozicija u potpunosti zadovoljava internalističko ograničenje budući da je univerzalno obvezujuća snaga moralnog zakona (kategoričkog imperativa) i zasnovana na činjenici da sva racionalna bića nužno *imaju zajednički cilj* – čovječnost ili našu umnu prirodu. Čovječnost ili umna priroda je nešto što bi svako savršeno racionalno biće, kako smatra Kant, nužno *htjelo* kao svoju svrhu, a bez te svrhe kategorički imperativ uopće ne bi bio ni moguć.[62] Meni se čini da glavnu poveznicu razloga i motivacije kod Kanta ipak treba tražiti u njegovu viđenju volje kao autonomne. Čovječnost kao svrha po sebi samo je jedna od karika koja Kanta na koncu

[61] Dva navedena cilja u Kantovoj su filozofiji isprepletena i teško se mogu promatrati odvojeno jedan od drugoga.

[62] Ovu tezu detaljno razrađuje i brani Julia Markovits u svojoj još neobjavljenoj doktorskoj disertaciji "Kantian internalism" (University of Oxford, srpanj 2006). Izuzev spomenute disertacije, za rekonstrukciju Kantova internalizma te njegova razumijevanja praktične racionalnosti također se oslanjam na Korsgaard (1996a), O'Neill (2004), Scruton (1982) te Brink (1997).

dovodi do stava da je volja "sebi sama zakon" (*U* 98). Jer što kazati o onome koji sam sebi postavlja zakon, a tim zakonom nije motiviran?

5.2.1. *Imperativi*

Kant vjeruje da se svi imperativi praktičnog razuma mogu podijeliti u one hipotetičke i one kategoričke. Prema Kantovu određenju, imperativi "kažu da bi bilo dobro nešto učiniti ili propustiti, ali oni to govore volji koja to ne čini uvijek samo zato što joj se predočuje da je to dobro činiti" (*U* 78). Budući da je naša volja ljudska i nesavršena, iluzorno je za očekivati da ćemo uvijek djelovati onako kako trebamo i zato su, između ostalog, i potrebni imperativi čija se funkcija sastoji u tome da usmjeravaju ljudsko ponašanje i da služe kao kriterij racionalnog djelovanja. Aplikacija imperativa na savršeno racionalnu osobu bila bi suvišna budući da bi ona, već samim time što je savršeno racionalna, imala volju koja je kompatibilna s *trebanjem,* tj. ne bi se nalazila pod "prinudom" činiti ono što joj imperativi nalažu.

Hipotetički su imperativi oni koji uvijek vrijede pod nekim uvjetom odnosno njihova je normativna snaga uvijek izvedena iz neke svrhe koju osoba želi postići. Jedan oblik hipotetičkog imperativa o kojemu govori Kant i nije ništa drugo nego već dobro poznat instrumentalni princip racionalnosti ili, kako ga on još naziva, "imperativ spretnosti". U tom je pogledu zanimljiva njegova napomena da svrhe mogu biti "sve moguće" jer "da li je svrha umna i dobra, to se ovdje i ne pita, nego samo to što se mora činiti da se postigne" (*U* 80).[63] Uz imperativ spretnosti Kant prepoznaje i imperativ razboritosti koji, prema njemu, također spada u grupu hipotetičkih imperativa. Riječ je ustvari o imperativu prudencijalnosti, tj. o imperativu koji nam nalaže da trebamo poduzeti ona sredstva putem kojih ćemo ostvariti

[63] Nije sasvim jasno kakav je inače status imperativa spretnosti u Kantovoj etici. Kant eksplicitno prepoznaje imperative spretnosti kao imperative praktičnog razuma, ali njegovo priznanje da moralni status svrha nije relevantan kada je o tim imperativima riječ, promatrana u kontekstu njegova cjelokupnog sustava, zvuči pomalo zbunjujuće. Vjerojatno David Brink to ima na umu kada piše: "Zašto bi netko imao razlog promovirati zadovoljenje svojih želja neovisno o sadržaju tih želja? U onoj mjeri u kojoj Kant razmatra hipotetičke imperative kao uvjetovane samo djelatnikovim htijenjima i željama, nije jasno da pretpostavlja da imanje relevantnih htijenja i želja automatski proizvodi razlog za djelovanje" (Brink, 1997: 284).

vlastitu sreću i promicati vlastiti interes.[64] Osnovna razlika između imperativa spretnosti i imperativa razboritosti jest u tome što u potonjem slučaju postoji "*jedna* svrha koja se kod svih umnih bića može pretpostaviti kao stvarna, … za koju se sigurno može pretpostaviti da je sva umna bića prema prirodnoj nužnosti i *imaju* (*U* 80). Nema dvojbe što Kant ovime želi poručiti. U slučaju imperativa spretnosti, nema ograničenja u pogledu ciljeva koje si netko može zadati, dok je u slučaju imperativa razboritosti krajnji cilj kod svih ljudi uvijek jednak. Ljudi "prema prirodnoj nužnosti" žele postići sreću i blaženstvo i zato se nalaze pod zahtjevom da biraju samo ona sredstva koja im mogu poslužiti da taj cilj i ostvare. S obzirom na činjenicu da nije uvijek jasno koja bi to sredstva bila, budući da sam pojam sreće nije dovoljno preciziran, Kant se ustručava princip razboritosti prozvati analitičkim principom. Drukčije formulirano, onaj tko želi neki cilj, smatra Kant, *nužno* želi i sredstvo za ostvarenje tog cilja: dovoljno je provesti analizu izraza "željeti neki cilj" da bi se došlo do zaključka da je u tome sadržano i to da se "već…pomišlja kauzalitet kao djelatni uzrok, tj. upotreba sredstava" (*U* 81). To, međutim, vrijedi samo u pogledu imperativa spretnosti gdje su individualni ciljevi jasno određeni. Imperativ razboritosti također bi se mogao nazvati analitičkim kada ne bi bilo dileme oko toga što sreća uključuje te koja su sredstva potrebna za njeno postizanje. Poteškoća je u tome, ističe Kant, što je "pojam blaženstva tako neodređen pojam da čovjek ipak nikada ne može reći određeno i suglasno sa samim sobom što zapravo želi i hoće iako svatko želi doći do blaženstva" (*U* 82).

Priznajući postojanje navedena dva imperativa, Kant se, barem na ovoj razini rasprave, nimalo ne udaljuje od hjumovskog razumijevanja razloga za djelovanje. Naglasak je na "barem na ovoj razini rasprave", jer za Kanta spomenuti imperativi ne iscrpljuju domenu praktičnog razuma. Ipak, ne postavljajući moralno ograničenje na svrhe koje bi netko želio

[64] Postoji opasnost da termin "vlastiti interes" navede na krivi put jer se, kada je riječ o prudencijalnosti, pod "vlastitim interesom" ne misli na trenutne, osobne interese koji se mogu razlikovati od pojedinca do pojedinca, već na nečiji interes sveukupno gledajući – na ono što je za pojedinca *dobro*. Norma prudencijalnosti nalaže da osoba na umu ima i svoje buduće interese te da svoje ciljeve pokuša uskladiti tako da ostvari ono što je za nju, cjelokupno gledajući, dobro.

ostvariti kada je riječ o imperativu spretnosti te nazivajući imperativ razboritosti (prudencijalnosti) hipotetičkim imperativom, Kant ne čini korake koje jednako tako ne bi mogao činiti hjumovski internalist. Prema hjumovskom internalistu, moralna procjena ciljeva nije nešto nad čim razum ima ingerenciju. Prihvaćanje suprotnog stava značilo bi odustajanje od instrumentalne koncepcije praktične racionalnosti. S druge strane, hjumovski internalist tvrdi da je nekome moguće pripisati razlog da čini ono što promiče njegovu dobrobit samo ako u njegovu motivacijskom skupu već postoji impuls za ostvarivanje vlastite dobrobiti odnosno promicanja vlastitog (trenutnog i budućeg) interesa. Ako je Kantova pretpostavka da je dobrobit svrha koju sva umna bića imaju prema "prirodnoj nužnosti" istinita, mislim da hjumovskom internalistu ta činjenica ne bi trebala zadavati značajnije poteškoće i da je on sasvim lako može uklopiti u svoje gledište. To samo znači da *svaka osoba* u svojem motivacijskom skupu ima element koji je povezan s ostvarenjem vlastite dobrobiti i da, sukladno tome, *svaka osoba* ima i *razlog* činiti ono što je u njezinoj dobrobiti. Jer praktično rasuđivanje, hjumovski internalist tvrdi, uvijek započinje iz motivacijskog skupa *što god da se u njemu nalazi*. U slučaju da se u njemu prema "prirodnoj nužnosti" nalazi motivacija za prudencijalnim ponašanjem, tim bolje.

Dokazati mogućnost postojanja hipotetičkih imperativa nije težak zadatak jer su oni, izuzmemo li problem nedovoljne određenosti pojma sreće, analitički istiniti. Problem za Kanta nastaje u trenutku kada treba dokazati postojanje druge vrste imperativa praktičnog razuma – postojanje kategoričkih imperativa. Zahtjevi koje nameću hipotetički imperativi prestaju vrijediti u trenutku napuštanja naših ciljeva[65], ali ta nam metoda nije dostupna kada je riječ o kategoričkim imperativima jer njihova obvezujuća priroda nije uvjetovana kontingentnim faktorima. Kategorički se imperativ "ne ograničava nikakvim uvjetom, pa se kao apsolutno-nuždan, premda praktično-nuždan, posve naročito može zvati zapovijed" (*U*

[65] Zanemarujem pitanje je li uopće moguće odreći se svrhe koju sva umna bića imaju prema *prirodnoj nužnosti*.

81). Moralno *trebanje* paradigmatičan je primjer kategoričkog trebanja i zato kategorički imperativ Kant još naziva i "imperativom ćudorednosti" (*U* 81). Moralni zakon, drugim riječima, treba biti izražen u obliku kategoričkog imperativa.

5.2.2. *Formule kategoričkog imperativa*

Ostavljajući formalno dokazivanje kategoričkog imperativa za kasniju raspravu i oslanjajući se samo na pretpostavku o njegovu postojanju, Kant sada pokušava konkretnije odrediti samu njegovu prirodu. Jedno od esencijalnih svojstava kategoričkog imperativa jest svojstvo univerzalnosti, a što Kanta i navodi da pruži njegovu opću formulu: "Djeluj samo prema onoj maksimi kojom ujedno možeš htjeti da postane opći zakon!" (*U* 84). Opća formula služi kao test provjere ispravnosti naših maksima odnosno subjektivnih principa na koje se oslanjamo u našem ponašanju. Što to konkretno znači nije teško razumjeti. Principi našeg djelovanja mogu biti bilo kakvi i zato je njihovu ispravnost potrebno provjeriti. Provjera maksima provodi se tako da se pitamo jesmo li spremni pristati da se i drugi oslanjaju na maksime koje mi prihvaćamo? Ili, mogu li naše maksime prihvatiti i drugi? Jednom riječju, kako znamo jesu li one ispravne? Ne postoji drugi način da to provjerimo nego da zamislimo svijet u kojem se svi ponašaju prema našoj maksimi te tako utvrdimo možemo li htjeti da naša maksima postane općim zakonom. Kant potom pokazuje da neke maksime ne samo da ne možemo htjeti, već da neke ne možemo ni misliti kao opći zakon jer bismo u oba slučaja upali u neku vrstu kontradikcije.

Prirodu maksima i pitanje njihove ispravnosti dodatno pojašnjava i Kantova rasprava o svrhama u kojoj on pruža još jednu formulu kategoričkog imperativa. Pogledajmo o čemu se radi. Normativna snaga hipotetičkih imperativa, vidjeli smo, zasnovana je na pojedinačnim i proizvoljno izabranim svrhama i snaga *trebanja* nestaje u trenutku njihova napuštanja: napuštanjem svrha nestaje i sam zahtjev da se poduzimaju sredstva za njihovo ostvarenje.

Moralnim zahtjevima, međutim, nije tako lako umaknuti. Proizvoljno odabrane svrhe ne mogu predstavljati temelj za kategorički imperativ jer iz proizvoljnosti i kontingentnosti nije moguće izvesti nužnost i objektivnost koje su esencijalna obilježja moralnog zakona. Kantova namjera je stoga pokazati da postoji nužna i objektivna svrha koja nije izvedena iz empirijske ljudske prirode (ljudskih nagnuća i želja) i na kojoj je moguće zasnovati kategorički imperativ. Ta svrha ne smije biti rezultat ljudskih inklinacija i nagnuća, već mora biti zajednička svim racionalnim bićima. Ali što je zajedničko svim racionalnim bićima? Izgleda da to ne može biti ništa drugo nego njihova racionalna priroda. Kant piše: "*Umna priroda egzistira kao svrha sama po sebi*" (*U* 89, kurziv u originalu). Princip prema kojemu racionalna priroda "egzistira kao svrha sama po sebi" nije samo subjektivni, već i objektivni princip djelovanja, a na osnovi tih razmatranja Kant ubrzo i izvodi formulu svrhe o sebi: "*Djeluj tako da čovječnost, kako u tvojoj osobi tako i u osobi svakoga drugoga, uvijek istodobno koristiš kao svrhu, nikada samo kao sredstvo!*" (*U* 90, kurziv u originalu). Čovječnost ili racionalna priroda, kao svrha po sebi, nije svrha koju je potrebno ostvarivati, već svrha protiv koje se ne smije djelovati: ona ima status negativne svrhe i zato predstavlja "vrhovni ograničavajući uvjet slobode djelovanja svakog čovjeka", ali i "vrhovni ograničavajući uvjet svih subjektivnih svrha" (*U* 91). Na taj način, uz filter univerzalizacije, formula čovječnosti također služi kao dodatan kriterij za provjeru valjanosti subjektivnih principa djelovanja.

Jedan od razloga zašto Kant upravo u racionalnoj prirodi nalazi svrhu po sebi po svemu sudeći leži u njegovoj teoriji vrijednosti. Kant vjeruje da racionalna priroda ima apsolutnu vrijednost i da su vrijednosti pojedinačnih i arbitrano odabranih svrha izvedene iz racionalne prirode kao ultimativnog izvora svih vrijednosti. Ono što Kanta dovodi do tog zaključka moglo bi se ukratko sažeti na sljedeći način. Naše je djelovanje u većini slučajeva usmjereno na realizaciju onih svrha koje nalazimo vrijednim. No uvođenje pojma vrijednosti

u diskusiju o svrhama otvara dilemu poput one u Platonovu dijalogu *Eutifron*: imaju li te svrhe vrijednost zato što ih mi nalazimo vrijednim ili ih mi nalazimo vrijednim zato što su te svrhe, same po sebi, vrijedne? Razmišljajući o toj dilemi, nekako nam izgleda da je teško opredijeliti se za bilo koju od opcija. S jedne strane imamo dojam da sfera vrijednosti nije potpuno nezavisna od našeg psihološkog ustrojstva i zato nam se opcija prema kojoj su vrijednosti ovisne "o nama" i čini privlačnom. Uostalom, vrlo često kažemo da je određena stvar dobra *za nas*; da neki predmet *za nas* ima određenu vrijednost, a time sugeriramo da su predmeti, sami po sebi, vrijednosno neutralni i da svoju vrijednost dobivaju tek kada postanu predmetima nagnuća i nečijeg interesa. Međutim, taj isključivo subjektivistički pristup ipak ne uspijeva zahvatiti vrlo jaku intuiciju da je o sferi vrijednosti moguće voditi racionalnu diskusiju i da vrijednosti kao takve nisu imune na racionalnu kritiku. Ljudi nerijetko za svoje svrhe određuju stvari koje nisu dobre i zato nismo uvijek skloni reći da je nešto dobro samo zato jer to netko želi ili prema tome ima neko nagnuće. Prije nam izgleda da su vrijednosti neovisne o kontingentnim ljudskim željama i afinitetima.[66]

Kantova solucija na postavljeni problem pokušava pomiriti obje intuicije. Izvor svih vrijednosti, Kant tvrdi, nalazi se u nama samima, ali te vrijednosti moraju biti u stanju dijeliti i drugi. Očito je da naša nagnuća ne mogu predstavljati izvor svih vrijednosti. Premda se nalaze u nama samima, nagnuća – uslijed svoje kontingentne prirode – ipak ne mogu zajamčiti objektivnost vrijednosti koju Kant želi dokazati. Jedini izvor koji se nalazi u nama samima, a koji pritom može zajamčiti takvu objektivnost, jest sama racionalna priroda. Ideja je da su svrhe, prema Kantu, dobre samo ako su izabrane racionalnim putem jer ih je samo na taj način i moguće opravdati drugima. Ili kako, interpretirajući Kantovu formulu čovječnosti, kaže Christine Korsgaard: "…ono što predmet vašeg racionalnog izbora čini dobrim jest to da on *jest* predmet racionalnog izbora" (Korsgaard, 1996a: 122).

[66] Rješenje ovog problema za koje sam se opredijelio u prošlom poglavlju, prisjetimo se, jest da su vrijednosti inherentne željama te da se o željama u pravom smislu riječi ne može ni govoriti ako one ne reflektiraju nečije vrijednosti.

Što zapravo Kant misli pod time da je racionalna priroda svrha po sebi sada bi već trebalo biti mnogo jasnije. Naše su pojedinačne svrhe i konkretni postupci opravdani samo u onoj mjeri u kojoj su rezultat racionalnog izbora, a sposobnost racionalnog izbora središnja je odlika nas kao racionalnih bića ili, u širem smislu, naše racionalne prirode. Racionalna priroda stoga služi kao izvor opravdanja, ali i kao izvor svih drugih vrijednosti. Kao što su naše vlastite svrhe dobre zato jer su rezultat našeg racionalnog izbora, na isti su način i svrhe drugih dobre zato jer su rezultat njihovih racionalnih izbora. Zato racionalna priroda može imati samo apsolutnu vrijednost i predstavljati objektivnu svrhu svih racionalnih bića – svrhu po sebi.

Obje spomenute formule (opća formula i formula svrhe o sebi) na neki način određuju prirodu kategoričkog imperativa. Dok je forma kategoričkog imperativa sadržana u općoj formuli, formula svrhe o sebi kategoričkom imperativu daje sadržaj. Kako kaže Korsgaard: "Usvajanje čovječnosti kao bezuvjetne svrhe dovodi do ponašanja koje formula univerzalnog zakona propisuje" (Korsgaard, 1996a: 18). No pored ove dvije formule, Kant izvodi i formulu autonomije, treću formulu kategoričkog imperativa: "[*Djeluj tako*] *da volja pomoću svoje maksime sebe samu ujedno može smatrati kao općenito zakonodavnu!*" (*U* 93, kurziv u originalu). Naša je volja samozakonodavna i iz toga izvire vjerojatno središnji pojam Kantove etike – pojam autonomije. Ideja autonomije usko je povezana s našom racionalnom prirodom kao svrhom po sebi. Činjenica da je racionalna priroda izvor svih vrijednosti i da su naši postupci opravdani samo u onoj mjeri u kojoj proizlaze iz naše racionalne prirode, sugerira da zakoni svoj izvor ne mogu imati nigdje drugdje nego u nama, a to ne znači ništa drugo nego da smo mi, kao racionalna bića, vlastiti zakonodavci i da smo podvrgnuti zakonima koje sami donosimo.

5.2.3. *Kategorički imperativ kao norma racionalnosti*

Sada preostaje odgovoriti na središnje pitanje: kako Kant opravdava svoju tezu da je kategorički imperativ zahtjev praktičnog razuma ili, nešto drukčije rečeno, koji je Kantov argument za tezu da je kategorički imperativ norma racionalnosti?[67] Osvrćući se na ovo pitanje, Cullity i Gaut pišu sljedeće: "Prema zrelom Kantu, norme razuma mogu biti utemeljene samo u formalnom načelu da razloge moraju dijeliti druga razumna bića i da ne smiju ovisiti o nametanju bilo kojeg autoriteta izvan onoga samog razuma" (Cullity i Gaut, 1997: 20). Formalno načelo o kojemu govore Cullity i Gaut Kantu služi kao osnova za izvođenje kategoričkog imperativa kao norme racionalnosti. Kategorički imperativ može imati status takve norme, Kant vjeruje, samo ako generira (i) one razloge koji su razlozi za svakoga i (ii) one razloge koji nisu nametnuti od strane nekog vanjskog autoriteta odnosno nekog autoriteta koji nije *sam razum*.[68] Iz dosadašnje je rasprave lako vidljivo da kategorički imperativ – ako postoji – uistinu zadovoljava spomenutim formalnim karakteristikama praktičnog razuma. Budući da sva racionalna bića dijele zajedničku svrhu, sva racionalna bića samim time dijele i iste razloge koji iz te svrhe proizlaze. Svako racionalno biće ima razlog činiti samo one stvari koje promiču čovječnost ili racionalnu prirodu. Zato razlozi koje generira kategorički imperativ zaista jesu razlozi koje dijele sva racionalna bića. S druge strane, očito je da ti razlozi nisu nametnuti od nekog vanjskog autoriteta, već da njihov izvor treba tražiti u našoj racionalnoj autonomiji. Porijeklo moralnog zakona nalazi se u nama samima, što povlači činjenicu da nije moguće racionalnim putem dovoditi u pitanje njegovu obvezujuću snagu. Oba su uvjeta sadržana u sljedećim Kantovim riječima: "Čovjeka su po njegovoj dužnosti gledali kao vezana na zakone, ali im nikada nije palo na pamet da je

[67] U nekoliko svojih članaka Korsgaard (1996) ističe da mnogi previđaju činjenicu da dedukciju kategoričkog imperativa Kant izvodi tek u trećem odsjeku *Utemeljenja metafizike ćudoređa* i da nam do tada Kant govori samo to kako bi taj imperativ trebao izgledati kada bi postojao. Slično napominje i Markovits (2006).

[68] Slično viđenje formalnih obilježja praktičnog razuma Kantu pripisuje i Onora O'Neill: "Fundamentalan potez o kojemu ovisi Kantova obrana razuma jest zahtjev da bude *prikladan za univerzalno korištenje*, radije nego prilagođen nekoj ograničenoj publici. … Ne nudim razloge za djelovanje svima ako predlažem princip djelovanja za koji znam da ga neki drugi ne mogu prihvatiti" (O'Neill, 2004: 101).

podložan *samo svojemu vlastitom*, ali ipak *općemu zakonodavstvu*…" (*U* 92, kurziv u originalu).

Ali čak i ako prihvatimo pretpostavku da mi jesmo autonomna bića u kantovskom smislu, zašto bi iz toga slijedilo da zakon koji sami sebi namećemo mora imati oblik kategoričkog imperativa? Zašto je, drugim riječima, racionalno pokoriti se upravo kategoričkom imperativu? Tvrdnja da smo kao autonomna bića podložni svojemu vlastitom zakonodavstvu je jedna stvar, ali tvrdnja da smo kao autonomna bića podložni kategoričkom imperativu jest nešto sasvim drugo. Jer ako je uistinu tako da se izvor zakona nalazi u nama samima, onda se čini da mi imamo izbor nametnuti bilo koji zakon. Sadržaj zakona nije unaprijed određen i do sada nije pružen nijedan argument koji bi pokazao da se nužno moramo opredijeliti za zakon koji ima oblik kategoričkog imperativa. Bilo kako bilo, ako Kant uspješno pokaže da je upravo kategorički imperativ ono čemu smo mi kao autonomna bića podložni, to će biti izravan dokaz njegova postojanja. Priroda kategoričkog imperativa jest takva da udovoljava uvjetima koje navode Cullity i Gaut, ali još uvijek nije dokazano da je postojanje tog imperativa zaista moguće i da se njegov izvor nalazi u samoj racionalnosti. Kako dakle Kant dokazuje njegovo postojanje i u kojoj mjeri se taj pokušaj izvođenja može smatrati uspješnim?

Rješenje je sadržano u Kantovu poimanju slobode. U posljednjem dijelu *Utemeljenja metafizike ćudoređa* Kant piše: "…budući da se ona [ćudorednost] mora izvesti samo iz svojstva slobode, mora se i sloboda dokazati kao svojstvo volje svih umnih bića" (*U* 104). Dovesti u vezu moralni zakon i racionalnost moguće je samo preko pojma slobode budući da "slobodna volja i volja pod ćudorednim zakonima jest jedno te isto" (*U* 103). Ukratko, pod pretpostavkom da je volja svih racionalnih bića slobodna volja te pod pretpostavkom da je slobodna volja podložna moralnom zakonu, tada neminovno slijedi da je volja svih racionalnih bića podložna moralnom zakonu. Struktura argumenta, kako vidimo, vrlo je

jednostavna, ali pitanje je uspijeva li Kant dokazati istinitost njegovih premisa: da je (i) sloboda zaista "svojstvo volje svih umnih bića" i da je zaista tako da (ii) "slobodna volja i volja pod ćudorednim zakonima jest jedno te isto".[69]

Prvenstveno treba imati na umu da Kantovo pozivanje na slobodu u ovom kontekstu nije nimalo proizvoljno. Biti podložan moralnom zakonu *pretpostavlja* mogućnost slobodnog i odgovornog ponašanja. Jer nema nikakvog smisla dokazivati postojanje zakona koji nam govori što trebamo činiti ako to ionako ne možemo činiti jer nismo slobodni. Kant je, naravno, svjestan da je za izvođenje dokaza slobode kao "svojstva volje svih umnih bića" neophodno imati barem neku naznaku o tome što sloboda uopće jest i zato na početku i pruža tzv. *negativno* određenje slobode. U najosnovnijem značenju, slobodu pripisujemo onim ljudima čije postupanje nije prisiljeno i koji se ponašaju na odgovarajući način zato jer su to sami odabrali. Ali već ova površna definicija otvara brojna pitanja. Što znači *biti prisiljen* nešto učiniti, a što znači da nešto odabiremo *sami*? Kant vjeruje da je pod prisilom svaka osoba koja nije oslobođena, da se poslužim njegovim terminom, "tuđih uzroka", a u koje on pritom ne svrstava samo vanjske uzroke, već i one čije se porijeklo nalazi u nama samima. Odmah postaje jasno o čemu je točno riječ sjetimo li se, na primjer, njegova stava da postoje takva nagnuća da "mora biti opća želja svakog umnog bića da bude potpuno slobodno od njih" (*U* 89). Kant ne daje primjer, ali njegova osnovna ideja nam ne bi trebala biti strana. Naša je sloboda, Kant ističe, sigurno dovedena u pitanje u onim situacijama u kojima djelujemo na osnovi nekih poriva ili nagnuća kojima ne želimo biti rukovođeni. Da bi se postupak koji je u potpunosti usmjeren takvim nagnućima mogao nazvati *mojim* postupkom, ja bih trebao odlučiti da se tom porivu zaista želim prepustiti. U suprotnom bi bilo netočno kazati da je moj postupak slobodan. To nas dovodi do središnjeg mjesta u Kantovu dokazivanju premise (i). Promotrimo sljedeće njegove riječi: "Um sam sebe mora smatrati

[69] Za Kantov argument također vidi Cullity i Gaut 1997: 21.

začetnikom svojih principa, neovisno o tuđim utjecajima. Prema tomu on se kao praktični um ili kao volja umnoga bića sam mora smatrati slobodnim, tj. njegova volja može samo pod idejom slobode biti vlastita volja…" (*U* 104). Kant, u najosnovnijim crtama, kaže da se moje rasuđivanje može okarakterizirati kao *moje* rasuđivanje samo pod uvjetom da rasuđujem pod *idejom* slobode. Radi se o tome da se u svom praktičnom rasuđivanju moram *smatrati* slobodnim. Kao što mnogi, interpretirajući ovaj dio Kantove argumentacije, naglašavaju (Darwall, 2006: 302; Korsgaard, 1996a: 162), ne radi se o tome da Kant tvrdi kako ja moram vjerovati da jesam slobodan, već samo to da svoje izbore moram temeljiti na ideji slobode odnosno da moram birati *kao* da jesam slobodan. Čak i ako na svoje odluke – kao što je ona da se prepustim nekom svojem porivu – zapravo nemam nikakav utjecaj jer je sve unaprijed determinirano, moje praktično rasuđivanje i dalje mora biti temeljeno na pretpostavci da ipak rasuđujem oslobođen "tuđih uzroka". Ukratko, sama činjenica da praktično rasuđujemo (da svakodnevno izabiremo i donosimo odluke o našim postupcima) navodi nas na prihvaćanje slobode. Sloboda se otkriva kroz naše praktično rasuđivanje i zato se, kako Kant smatra, također moramo vidjeti kao dio inteligibilnoga svijeta – svijeta u kojemu se naša volja nalazi izvan vanjskih utjecaja i u kojemu su naša djelovanja "potpuno u skladu s principom autonomije čiste volje" (*U* 108). Slobodu prakticiramo samo kao racionalna bića odnosno kao bića koja su dio inteligibilnoga svijeta, dok smo, kao osjetilna bića, uvijek podvrgnuti prirodnoj nužnosti. Kantovo razumijevanje slobode sažima Talanga: "Dakle ono što teorijski um nije uspio ni dokazati ni pobiti, a to je postojanje slobode, to dokazuje praktični um. … On nas prisiljava da vjerujemo u slobodu naše volje odnosno da je uzimamo kao stvarnu" (Talanga, 1999: 173).

Budući da smo prisiljeni slobodu naše volje uzimati kao stvarnu, tako smo prisiljeni i sebe promatrati kao bića koja su također dio razumskog svijeta, jer samo kao takva bića imamo volju koja je oslobođena svih vanjskih utjecaja. Time je, Kant vjeruje, ujedno i

dokazana premisa (i). No, kao što je već naznačeno, biti oslobođen vanjskih utjecaja znači biti slobodan *samo* u negativnom smislu. To samo znači da ništa "izvana" ne vrši pritisak na naše odluke i izbore, već da je izvor tih odluka i izbora u nama samima. Zato pored negativnog određenja slobode, Kant govori i o pozitivnoj slobodi, a u kojoj se krije ključ za dokazivanje premise (ii) njegova argumenta – premise prema kojoj se moralni zakon može izvesti samo iz svojstva slobode. Što uključuje sloboda u pozitivnom smislu? Premda je točno da je slobodna osoba izuzeta od vanjskih utjecaja i pritisaka, to ujedno ne znači da 'biti slobodan' znači biti izuzet od bilo kakvih zakona. Upravo suprotno, Kant slobodu definira kao kauzalnost, a kauzalnost, kako on naglašava, "donosi sa sobom pojam *zakona*" (*U* 103, kurziv u originalu). Iz toga slijedi da se o slobodnoj volji zapravo uopće ne može govoriti ukoliko ona nije podložna određenom zakonu. Naravno, "određeni" zakon o kojemu je ovdje riječ za Kanta je upravo moralni zakon odnosno kategorički imperativ. Argument kojim Kant pokazuje zašto je baš taj zakon onaj kojemu smo kao slobodna bića nužno podložni lijepo rekonstruira Korsgaard u sljedećem odlomku:

> Problem s kojim se suočava slobodna volja jest sljedeći: volja mora imati zakon, ali kako je volja slobodna, to mora biti njezin zakon. I ništa ne određuje što taj zakon mora biti. *Sve što mora biti jest to da je zakon*. Razmotrimo sada sadržaj kategoričkog imperativa kako je predstavljen u formuli univerzalnog zakona. Kategorički nam imperativ jedino govori da izaberemo zakon. Njegovo jedino ograničenje na naš izbor jest da ima oblik zakona. I ništa ne određuje što zakon mora biti. *Sve što mora biti jest to da je zakon*. Prema tome, kategorički imperativ je zakon slobodne volje. (Korsgaard, 1996b: 98)

Pod pretpostavkom da se racionalnim bićima zaista može pripisati slobodna volja i pod pretpostavkom da prihvatimo argument prema kojemu je kategorički imperativ zakon slobodne volje, dolazimo do zaključka da su sva racionalna bića podložna kategoričkom

imperativu (moralnom zakonu). Moralne norme nisu ništa drugo nego norme racionalnosti. Budući da su racionalna bića – *samim time što su racionalna* – podložna moralnom zakonu, to znači da moral nije nešto što racionalna bića mogu izbjeći. U kontekstu rasprave o internalizmu, ovaj zaključak ima dalekosežne implikacije. Već je rečeno da kantovski internalisti pokušavaju pronaći srednji put između hjumovskog internalizma i eksternalizma. Pronaći srednji put znači pokazati da postoje objektivni, kategorički razlozi za djelovanje, ali i to da bi takvim razlozima svaka racionalna osoba trebala biti motivirana. Pristupajući Kantovoj argumentaciji iz perspektive rasprave o internalizmu, stječe se dojam da su oba uvjeta zadovoljena. Kategorički imperativ, osim što predstavlja vrhovni moralni zakon, također predstavlja normu racionalnosti – normu koja generira kategoričke razloge za djelovanje. Razlozi koji proizlaze iz kategoričkog imperativa stoga nisu razlozi koji su ovisni o individualnim željama i preferencijama, već oni razlozi koji se odnose na sva racionalna bića. No kako Kant udovoljava drugom uvjetu odnosno kako on uspostavlja vezu između kategoričkih razloga i same motivacije? Odgovor je jednostavan. Norma kategoričkog imperativa nema eksternalan izvor, već je mi, kao autonomna bića, sami stvaramo. To ni u kojem slučaju ne znači da mi možemo stvoriti neki drugi zakon, a ne kategorički imperativ. Bilo bi sasvim pogrešno protumačiti Kanta na taj način. Kant upravo sugerira da je kategorički imperativ *jedini* zakon do kojeg dolazimo kao racionalna bića. Njegov je izvor u našoj racionalnosti i zato njegovu snagu nije moguće dovoditi u pitanje a da pritom ne dovedemo u pitanje samu našu racionalnost. Kako je prethodna rasprava sugerirala, jedan od razloga zašto je upravo kategorički imperativ ukorijenjen u racionalnosti mogao bi biti u činjenici da je racionalna priroda svrha po sebi. Sva racionalna bića nužno dijele zajedničku svrhu, a kategorički imperativ, kao univerzalno obvezujući zakon, regulira djelovanje u skladu s tom svrhom. U svakom slučaju, moralni zahtjevi imaju racionalni autoritet, a tako i *motivacija* da slijedimo moralne zahtjeve nastaje pod pritiskom razuma. Ne biti motiviran

zakonom koji nam nameće praktični razum znači biti iracionalan. Zato je kategoričkim razlozima nužno motivirana svaka *racionalna* osoba.

5.2.4. *Jedan problem za Kanta*

Moja prvenstvena namjera u gornjem prikazu bila je pokazati da Kantova filozofija morala sadrži sva relevantna obilježja internalističkog stajališta. Njegovu je poziciju moguće okarakterizirati kao internalističku baš zbog toga što kategoričnost moralnih razloga ne proizlazi iz vanjskog poretka stvari, već iz samog našeg ustrojstva. Izvor moralnih razloga nalazi se u nama samima i zato se ti razlozi s pravom mogu nazvati unutarnjima. Stoga Kantu – barem se tako čini na prvi pogled – i nije teško obrazložiti na koji nas način ti razlozi motiviraju: budući da su moralni razlozi produkt naše *racionalne* prirode, njima nećemo biti motivirani samo pod cijenu iracionalnosti. No pogledajmo sada u kojoj se mjeri Kantova argumentacija može smatrati uspješnom.

Nije prijeporno da Kantova redukcija morala na kategoričku racionalnost stoji na mnogim klimavim pretpostavkama, a od kojih u ovom kontekstu svakako treba izdvojiti njegovo shvaćanje slobode. Kako su mnogi s pravom isticali, govor o moralnom subjektu kao onom koji se nalazi u sferi nadosjetilnog i koji nije podložan kauzalnosti u najmanju ruku zvuči čudno i nerazumljivo i vjerojatno predstavlja najslabiju kariku u Kantovoj koncepciji morala. Jer, "[a]ko stvarno postoji subjekt koji na sebe uzima teret odlučivanja u moralno problematičnim situacijama, kakve je on prirode, kada nastaje, koje su njegove osobine i gdje se nalazi?" (Sesardić, 1984: 137). No to nije jedini problem. Kant, vidjeli smo, upozorava da je praktično rasuđivanje moguće samo ako rasuđujemo pod idejom slobode: mi jednostavno moramo sebe vidjeti kao dio nadosjetilnog svijeta u kojemu nismo podvrgnuti prirodnim zakonima odnosno u kojemu ne vlada determinizam. Bez te pretpostavke naše rasuđivanje zapravo ne bi imalo nikakvog smisla. Naime, zašto da uopće rasuđujemo o tome što učiniti

ako je već unaprijed određeno što ćemo učiniti? Naše je rasuđivanje dakle smisleno samo pod idejom slobode, a, kao moralna bića, zaista i *jesmo* slobodni. Sama sloboda, prema Kantovu mišljenju, izlazi izvan dosega naše spoznaje jer je dio noumenalnog svijeta, ali mi, kao moralni subjekti, tom svijetu ipak pripadamo. Teškoća s takvim gledištem jest u tome što relacija između noumenalnog i fenomenalnog svijeta ostaje nerazjašnjena i neprecizirana. Kako ističe Benn, nije jasno na koji način

> …rasuđivanje u noumenalnom području … može proizvesti bilo kakve učinke u fenomenalnom svijetu – što je u konačnici smisao naših rasuđivanja. Ako relacija nije kauzalna, kakva je onda? Kako razlog za djelovanje može proizvesti bilo kakav učinak osim ako ujedno nije i uzrok? (Benn, 1998: 100)

Ako smo praktičnim rasuđivanjem utvrdili što imamo razloga učiniti, onda bismo tim razlozima trebali biti i pokrenuti. U noumenalnom svijetu, međutim, nema prirodne kauzalnosti, što Kanta stavlja u prilično nezgodan položaj. Isključimo li postojanje uzročne veze, ostaje nedefinirano na koji nas način ti razlozi pokreću, kako se naše moralne odluke reflektiraju u području empirijske stvarnosti te na koji način moralni subjekt uopće može na išta utjecati. Stoga je očito da svaki simpatizer Kantova pothvata mora pronaći način kako prevladati poteškoće poput ovih (usp. Cullity i Gaut, 1997: 23; Brink, 1997: 267), a upravo ću u sljedećem dijelu razmotriti jedan takav pokušaj.

5.3. Smithov internalizam

U novije vrijeme jedno od središnjih mjesta u diskusiji o statusu normativnih razloga predstavljaju argumenti kojima se nastoji pokazati da središnja Williamsova ideja vrlo lako može izbjeći relativističke implikacije te da se postojanje *kategoričnih* razloga također može dokazivati unutar internalističkog okvira. Glavna značajka takve argumentacije jest – baš kao

što je to slučaj kod samog Kanta – izbrisati razliku između internalizma i eksternalizma odnosno potkopati pretpostavku da je *jedino* eksternalistima otvorena mogućnost za dokazivanje objektivnih, ne-relativnih razloga. Jedan od najpoznatijih zagovornika takvog pristupa zasigurno je Michael Smith. U svojoj poznatoj knjizi *The Moral Problem* (1994), kao i u drugim člancima na ovu temu (2004a, 2004b), Smith vrlo suptilnom argumentacijom pokušava pomiriti dvije tradicionalno suprotstavljene struje – hjumovsku i kantovsku. Praveći oštru distinkciju između motivirajućih i normativnih razloga, Smithova namjera jest pokazati da je internalistima, suprotno ranije navedenom Parfitovu mišljenju, *zaista* otvoren put za analizu pojma "jasna prosuđivanja" te da oni tom analizom ne moraju nužno odstupiti od svoje pozicije. Radi se o tome da ljudi koji jasno prosuđuju konvergiraju u svojim preferencijama i željama. Nitko ne spori da različiti ljudi imaju *različite* preferencije i želje unutar svojih motivacijskih skupova, ali iz toga još uvijek ne slijedi da oni ne bi imali *jednake* želje kada bi "jasno prosuđivali". Sve racionalne osobe, drugim riječima, imaju jednake elemente u svojemu motivacijskom skupu. Premda je moguće da netko zaista ima preferenciju za uništenjem cijelog svijeta, nije moguće da takva preferencija uopće prođe kroz filter procesa rasuđivanja pa stoga i nije moguće da netko s takvom preferencijom ima razlog za djelovanje. Dok Williamsov internalizam, kako smo vidjeli, ostavlja mogućnost da čak i motivacija za uništenjem cijelog svijeta "preživi" rasuđivanje, Smithov internalizam takvu mogućnost negira u samom početku.

Zalaganjem za mogućnost konvergencije internalistima se otvara put za zastupanje kantovskog shvaćanja normativnih razloga jer slijedi da su "tvrdnje o normativnim razlozima kategorički imperativi jer svi djelatnici koji se nađu u jednakim okolnostima imaju jednake razloge (Smith: 1994: 175)". Ali također uočavamo da to stajalište još uvijek ostaje vjerno internalističkoj ideji prema kojoj *svi* razlozi za djelovanje moraju u konačnici biti povezani s

motivacijom. U sljedećem dijelu detaljnije ću promotriti glavne aspekte Smithove pozicije, a potom ću pokušati pokazati da ona na koncu ipak ne uspijeva izbjeći određenim poteškoćama.

Smith predlaže sljedeću formulaciju teze internalizma razloga:

> (IR) A ima razlog učiniti Φ u okolnostima C akko bi A želio učiniti Φ u okolnostima C kada bi bio potpuno racionalan.

IR je, Smith ističe, podložan dvjema različitim interpretacijama. Prema prvoj interpretaciji, koju naziva *modelom primjera*, IR povlači da ono što imam razloga učiniti u okolnostima C ovisi o tome što bih ja, kada bih bio potpuno racionalan, želio učiniti kada bih se nalazio u okolnostima C. S druge strane, prema *modelu savjeta*, ono što imam razloga učiniti u okolnostima C ovisi o tome što bih ja, kada bih bio potpuno racionalan, želio da ja, *sada kada nisam potpuno racionalan*, učinim u okolnostima C. U svrhu boljeg objašnjenja ovih modela, Smith zamišlja dva moguća svijeta: svijet u kojem se nalazi subjektov potpuno racionalan"ja" (evaluirajući svijet) i svijet u kojem se nalazi subjektov manje racionalan "ja" (evaluiran svijet). Ideja se sastoji u tome da *ili* subjektov potpuno racionalan "ja" svojim ponašanjem u evaluirajućem svijetu daje *primjer* koji bi njegov manje racionalan "ja" trebao slijediti u evaluiranom svijetu *ili* subjektov potpuno racionalan "ja" svom manje racionalnom "ja" pruža *savjet* o tome što bi njegov manje racionalan "ja" trebao učiniti u evaluiranom svijetu. Smith argumentira da je model primjera neplauzibilan i da se IR može braniti samo ako je shvaćen kao model savjeta.

Pokušat ću Smithovu argumentaciju približiti koristeći primjer iz Homerove *Odiseje*. Odisej je, unatoč činjenici da bi zbog toga mogao izgubiti glavu, želio čuti opojan glas sirena. Stoga je naredio svojim ljudima da ga zavežu za jarbol i zanemare njegove naredbe da ga odvežu kada se začuje glas sirena. Interpretirajući ovu priču kroz prizmu modela savjeta, mogli bismo kazati sljedeće. Odisej, u stadiju u kojem jasno rasuđuje i u kojem je svjestan

svih mogućih opasnosti, *savjetuje* svom manje racionalnom "ja" da se, s obzirom na njegovu želju da čuje glas sirena, zaveže za jarbol. U okolnostima u kojima je Odisejeva moć jasnog rasuđivanja znatno oslabljena pjevanjem sirena, Odisej ima razlog učiniti ono što bi njegov potpuno racionalan "ja" želio da on učini u tim okolnostima – vezati se za jarbol. Model primjera, s druge strane, implicira da ono što Odisej ima razlog učiniti u okolnostima u kojima je njegova moć rasuđivanja oslabljena jest ono što bi njegov *potpuno racionalni "ja*" želio učiniti *u istim tim okolnostima*. Ali problem je u tome što mi jednostavno ne možemo zamisliti potpuno racionalnog Odiseja u *istim* okolnostima kao manje racionalnog Odiseja. Mi ne možemo zamisliti da je moć rasuđivanja *potpuno racionalnog* Odiseja oslabljena pjevanjem sirena jer bi to značilo da Odisej ipak nije potpuno racionalan. Želja da bude zavezan za jarbol kod potpuno racionalnog Odiseja ne bi bila formirana jer za formiranjem takve želje ne bi bilo nikakve potrebe. Ili, Smithovim riječima, "…ono što je subjektov potpuno racionalan 'ja" motiviran učiniti ovisit će o okolnostima u kojima se nađe i, po definiciji, te okolnosti nikada neće uključivati njegovu iracionalnost" (Smith, 2004a: 20).

Internalizam razloga, interpretiran isključivo kao model savjeta, prema Smithovu mišljenju, također može objasniti mehanizam rasuđivanja te uspješno pokazati kako motivacija nastaje čisto racionalnim putem. Kada rasuđujemo o tome što učiniti, mi zapravo nastojimo utvrditi razloge koje imamo za određenu vrstu postupanja. Ako smo rasuđivanjem formirali vjerovanje da imamo razlog učiniti Φ, onda je to vjerovanje, internalizam razloga nam govori, ekvivalentno vjerovanju da bismo željeli učiniti Φ u stadiju potpune racionalnosti. Ali ako vjerujem da bih želio učiniti Φ u okolnostima C kada bih bio potpuno racionalan, tada je *racionalno* da – ako vjerujem da se *sada* nalazim u okolnostima C – formiram želju (motivaciju) učiniti Φ. Psihologija koja, prema Smithovim riječima, sadrži vjerovanje o tome da bih želio učiniti Φ u C kada bih bio potpuno racionalan i želju da učinim Φ u C jest "koherentnija i ujedinjenija" od psihologije koja ne sadrži takvu kombinaciju

vjerovanja i želja. Slično objašnjenje nastanka motivacije racionalnim putem ne bi bilo moguće kada bi model primjera bio istinit. Osoba bi u tom slučaju mogla vjerovati da bi njezin potpuno racionalni "ja" želio učiniti Φ u okolnostima u kojima se njezin *potpuno racionalni "ja"* nalazi, ali takvo vjerovanje na nju ne bi moralo imati nikakvog utjecaja budući da takve okolnosti ne bi bile iste kao okolnosti u kojima se ona, kada nije potpuno racionalna, nalazi. No stvari se mijenjaju prihvatimo li model savjeta. U tom je slučaju subjekt pod određenim pritiskom razuma da stvori želju (motivaciju) djelovati u skladu sa savjetom svojeg potpuno racionalnog "ja" jer bi jedino na taj način njegova psihološka stanja postigla stadij koherencije. Smithov argument sigurno predstavlja jedno od centralnih mjesta njegove teorije budući da njime odstupa od glavnog obilježja Williamsova internalizma razloga – da procesi rasuđivanja, kakvi god da jesu, ne mogu sami po sebi generirati motivaciju. Takva motivacija, podsjetimo se, prema Williamsu mora biti prisutna prije nego što rasuđivanje uopće započne.

Koherencija želja, prema Smithu, jedan je od glavnih uvjeta koji moraju biti zadovoljeni da bismo za nekoga mogli reći da je potpuno racionalan. Smith kaže:

> Moj prijedlog jest da osoba, da bi bila potpuno racionalna, ne smije patiti od učinaka bilo kakvog fizičkog ili emocionalnog poremećaja, ne smije imati lažna vjerovanja, mora imati sva relevantna istinita vjerovanja i mora imati sistematski opravdljiv skup želja, naime skup želja koji je maksimalno koherentan i sjedinjen. (Smith, 2004a: 263)

Smith se ovdje velikim dijelom nadovezuje na uvjete koje je prethodno naznačio Williams, ali s dva ključna dodatka. Kao prvo, Smith, za razliku od Williamsa, smatra da emocionalni i fizički poremećaji moraju biti isključeni kada je riječ o potpuno racionalnoj osobi. Depresija, *akrasia*, bijes – sve su to fenomeni koji mogu znatno utjecati na proces rasuđivanja. Kao drugo, potpuno racionalna osoba mora imati želje koje međusobno nisu proturječne i koje

zajedno tvore jedan koherentan skup. Na taj način Smith predlaže metodu koja je po svojoj formi analogna tzv. metodi reflektivnog ekvilibrija i koja se najčešće povezuje s koherentističkom teorijom epistemičkog opravdanja.

Promatrajući tu metodu u kontekstu moralnog rasuđivanja, njezin glavni cilj jest postići stanje ekvilibrija između pojedinačnih moralnih vjerovanja i općih, teoretskih moralnih načela. To se stanje može postići jedino usklađivanjem, što znači da je na nekim mjestima potrebno revidirati teoriju, dok je ponekad potrebno odustati i od nekih pojedinačnih vjerovanja koja s njom ne mogu biti usklađena. Neko se moralno vjerovanje tako smatra opravdanim ako nije proturječno drugim moralnim vjerovanjima odnosno ako se zajedno s drugim vjerovanjima nalazi u stanju koherencije. Originalnost Smithova pristupa, međutim, prepoznaje se u njegovoj tvrdnji da se ova metoda može primijeniti i u pogledu opravdanja *želja*. Naše želje također mogu biti podvrgnute racionalnoj kritici pa je stoga, između dva skupa, uvijek *racionalniji* onaj skup u kojemu želje nisu proturječne odnosno onaj u kojemu se želje nalaze u međusobnoj ravnoteži. Tu ravnotežu, slično kao i kod vjerovanja, možemo postići *ili* napuštanjem nekih pojedinačnih želja koje imamo *ili* pak dodavanjem općenitijih želja koje tada mogu služiti kao opravdanje onih pojedinačnih. Oslanjanjem na metodu reflektivnog ekvilibrija, Smith tako dolazi i do svoje središnje tvrdnje – da proces opravdanja želja na koncu uvijek rezultira *samo jednim* maksimalno koherentnim skupom i da sve potpuno racionalne osobe, po definiciji, *konvergiraju* u svojim željama u pogledu toga što učiniti u određenim okolnostima. Time ujedno i dokazuje postojanje kategoričnih razloga jer sve osobe, u stadiju potpune racionalnosti, imaju i jednake razloge za djelovanje.

Od presudne je važnosti uvidjeti da je teza o konvergenciji isključivo *konceptualna* teza. Smith dakle ne smatra da će do takve konvergencije *de facto* i doći, već je njegovo glavno pitanje da li je "dio onoga što podrazumijevamo pod 'racionalan" da potpuno racionalne osobe konvergiraju u svojim željama, ili to nije dio onoga što podrazumijevamo

pod ‘racionalan’” (Smith 2004a: 27). Budući da smatra kako je u samom pojmu potpune racionalnosti sadržano da potpuno racionalne osobe imaju jednake želje, ta teza ne može biti falsificirana ukazivanjem na primjer u kojemu potpuno racionalne osobe nemaju jednake želje. Zanimljivo je da neki autori, raspravljajući o implikacijama internalističkog gledišta, to razlikovanje nemaju u vidu. Tako, primjerice, Jonathan Dancy vrlo otvoreno demonstrira svoj skepticizam u pogledu mogućnosti konvergencije želja u okviru internalističke teorije:

> Ne možemo kazati, na primjer, da postoji razlog za *svakoga* brinuti se o potrebama beskućnika, osim ako bismo mogli pokazati da bi sve osobe bez iznimke bile motivirane to učiniti u uvjetima C. Mogu li sve osobe, krećući od tamo gdje su sada, motivirati misli o beskućnicima kada bi bolje razmišljali i više znali? Nekako sumnjam. (Dancy 2000: 17)

Međutim, ako je Smith u pravu, onda Dancyev skepticizam nimalo ne ugrožava njegovo gledište. Dancy samo iskazuje sumnju u to da *svi* ljudi *de facto* mogu imati jednaku motivaciju u idealnim uvjetima, a to ni na koji način ne proturječi Smithovu stavu. Smithova se tvrdnja o konvergenciji ne može potkopati pružanjem primjera u kojem potpuno racionalne osobe divergiraju u svojim željama oko toga što učiniti u određenim okolnostima. Takav se primjer, ukoliko je Smith u pravu, ne može zamisliti: postojanje takvih osoba značilo bi da one jednostavno nisu u potpunosti racionalne.

Smithova teorija na vrlo uvjerljiv način pokazuje na koji bi se način mogle pomiriti dvije, naizgled potpuno suprotne, pretpostavke – pretpostavku da normativni razlozi moraju biti utemeljeni u prethodnoj motivaciji i pretpostavku da normativni razlozi moraju imati objektivno utemeljenje. No ipak, čini mi se da se ta teorija na nekim mjestima ipak suočava s nekim poteškoćama.

5.3.1. *Koherencija želja*

Smithu se može prigovoriti da opravdanje želja putem metode reflektivnog ekvilibrija možda i nema toliku snagu kao kada je riječ o primjeni te metode na sama vjerovanja. Kao prvo, zašto bismo morali smatrati da je nekoherentan skup želja *iracionalan*? Zašto bi nas činjenica da *imamo* želje koje su međusobno proturječne nužno činila iracionalnima? Ako A ima proturječne želje, to samo znači da na osnovi njih ne može istovremeno *djelovati* te da prije samog djelovanja mora utvrditi do koje mu je želje više stalo ili koja mu je želja najjača. No to još uvijek ne znači da sama činjenica da A *ima* proturječne želje nužno mora utjecati na njegovu racionalnost. Kao drugo, objašnjenje potrebe za koherentnošću naših vjerovanja može se pružiti zbog same prirode vjerovanja kao mentalnih stanja, dok nije očito u čemu bi se to objašnjenje moglo sastojati kada je riječ o željama. Smith, kao što smo vidjeli, tvrdi da ako A vjeruje da bi želio učiniti Φ u okolnostima C kada bi bio racionalan, onda je racionalno da on formira želju učiniti Φ ako je u okolnostima C. Kada A ne bi formirao takvu želju, to je slično kao da A vjeruje da bi vjerovao da *p* kada bi bio potpuno racionalan, ali unatoč tome *sada* odbija vjerovati da *p*. No zašto smatrati da su prvi i drugi slučaj identični odnosno da opis jedne situacije mora biti jednak opisu druge situacije? Ako je A iracionalan u slučaju kada njegova vjerovanja nisu koherentna, to još uvijek ne znači da je A iracionalan kada njegove želje nisu koherentne. Smith, čini se, prešutno pretpostavlja da postoji čvrsta analogija između teoretskog i praktičnog zaključivanja, ali to je pretpostavka koja se može dovesti u pitanje. Razlika između vjerovanja i želja jest u tome što vjerovanja mogu biti istinita ili lažna, dok želje ne mogu. Kada je riječ o vjerovanjima, mi smo pod određenim normativnim pritiskom da naša vjerovanja međusobno uskladimo. Ta činjenica može objasniti zašto postoji zahtjev za koherencijom kada je riječ o vjerovanjima, ali porijeklo tog zahtjeva

kada je riječ o željama ostaje misteriozno. Drugim riječima, kako objasniti porijeklo normativnog pritiska za usklađivanjem želja?[70]

5.3.2. *Protiv konvergencije*

Čak i pod pretpostavkom da Smith može uspješno odgovoriti na gornji prigovor, još uvijek ostaje razmotriti njegove argumente za konvergenciju želja kod potpuno racionalnih osoba. Smith navodi dva takva argumenta. Prvim se nastoji pokazati da se o normativnosti razloga uopće ne bi moglo govoriti kada bi razlozi bili relativni, dok se drugim pokazuje da fenomen neslaganja govori u prilog ne-relativnosti razloga odnosno da predstavlja indirektan dokaz za konvergenciju. Evo i Smithove formulacije prvog argumenta:

> Tvrdnja da bi potpuno racionalna bića konvergirala u svojim željama jaka je tvrdnja i branim je na sljedeći način. Kažem da misliti drukčije […] znači pretpostaviti […] da nešto sasvim arbitrarno […] može na neki način konstituirati normativnu činjenicu: činjenicu o racionalnom opravdanju. Ali to je nekoherentno jer jedinu odlučnu tvrdnju koju možemo izreći o normativnosti jest da arbitrarnost, kao takva, uvijek potkopava normativnost. […] Internalni razlozi, prema relativističkom poimanju, stoga *nemaju normativni značaj*. A ako je to točno, tada slijedi da relativni internalni razlozi uopće nisu razlozi. (Smith, 2004a: 265, moj kurziv)

Hjumovski internalizam, Smith tvrdi, ne može objasniti normativnost razloga jer bi u tom slučaju normativnost bila zasnovana na željama koje su potpuno arbitrarne. Govor o normativnosti naime ima smisla jedino ako *trebanje* nije ovisno o subjektivnim, arbitrarnim psihološkim karakteristikama i stoga se unutar hjumovske koncepcije o normativnosti ne može ni govoriti. Napomenimo još jednom da, prema hjumovskom internalizmu, elementi

[70] Možda bi trebalo naglasiti da moja namjera nije dovesti u pitanje činjenicu da naše želje mogu biti nekoherentne, već samo ukazati na problematičnost teze da je *iracionalano* imati nekoherentne želje. Moj prigovor Smithu djelomično je inspiriran prigovorom koji je Smithu uputio Sayre-McCord (vidi u Smith: 2004).

koje A ima u svom motivacijskom skupu nakon rasuđivanja uvijek zavise o elementima koje A ima u tom skupu prije rasuđivanja. Ali elementi nečijeg motivacijskog skupa prije rasuđivanja, hjumovski internalisti tvrde, potpuno su arbitrarni: A može, ali i ne mora imati želju učiniti Φ, što znači da to *ima li razlog* A učiniti Φ na koncu konca ovisi samo o tome ima li A tu želju u svom motivacijskom skupu ili nema. Ako A nema tu želju, ne postoji jamstvo da će je stvoriti nakon rasuđivanja, a ako A ima tu želju, ne postoji jamstvo da će ta želja "preživjeti" proces rasuđivanja. Hoće li A stoga imati normativni razlog učiniti Φ zapravo je relativna stvar, jer to ovisi samo o A-ovoj arbitrarnoj motivaciji. No tako shvaćena normativnost, Smith nastoji pokazati gornjim argumentom, uopće nije normativnost: pojam *relativnog* normativnog razloga jednostavno je *contradictio in adjecto*.

Smithov internalizam, s druge strane, vrlo lako izbjegava relativističke posljedice. Koje će elemente A imati u svom motivacijskom skupu nakon rasuđivanja ni na koji način ne ovisi o tome koje je elemente A imao prije nego što ih je podvrgnuo sistematskom opravdanju jer "sam razum određuje sadržaj naših potpuno racionalnih želja" (Smith, 2004a: 33). Budući da *sve* potpuno racionalne osobe konvergiraju u svojim željama u pogledu toga što učiniti u određenim okolnostima, više ne postoji opasnost od relativizma, a razlozi ne gube svoj normativni značaj. Jer odgovor na pitanje ima li A *normativni* razlog učiniti Φ u C ne ovisi više o tome bi li A želio učiniti Φ u C kada bi bio racionalan, već o tome bi li *sva* potpuno racionalna bića željela učiniti Φ u C.

Smithov je argument, čini mi se, dijelom posljedica njegova vrlo jasnog razlikovanja normativnih i motivirajućih razloga. Jer kako uopće možemo govoriti o *normativnosti* razloga ako, kao što to hjumovski internalisti tvrde, normativni razlozi zapravo ne postoje?[71] No vjerujem da, čak i ako distinkcija između normativnih i motivirajućih razloga nije tako oštra, to još uvijek ne znači da za normativnost nema mjesta unutar hjumovske koncepcije. Na

[71] Williams na jednom mjestu napominje da "mora biti pogreška jednostavno razdvojiti eksplanatorne i normativne razloge" (Williams, 1995: 39). O tome također vidi Stampe (1987).

primjer, vidjeli smo da je jedan od učinaka rasuđivanja to da osoba može utvrditi da je neka njezina motivacija rezultat pogrešnog shvaćanja stanja stvari. Ako je A stvorio motivaciju da učini Φ zato jer nije razmotrio sve relevantne činjenice ili možda zato jer je njegovo vjerovanje o tome što će postići tim postupkom lažno, onda ima smisla reći da on nema razlog učiniti Φ. On nema razlog to učiniti budući da on sam to ne bi želio učiniti kada bi znao relevantne činjenice. Osim toga, A možda može imati neki element u svom S-u koji ga može potaknuti na neko postupanje, ali A, uslijed brojnih čimbenika, tog elementa ne mora biti svjestan u danom trenutku ili u danim okolnostima. Tada izgleda primjereno kazati da A ima razlog učiniti Φ premda u svom S-u ne može pronaći ništa što bi ga potaknulo na činjenje Φ-a. Iz toga proizlazi da ponekad možemo izricati normativne tvrdnje o tome što A ima ili nema razlog učiniti, ali takve su tvrdnje primjerene jedino ako su izrečene *u odnosu na* A-ov već postojeći motivacijski skup. Takve tvrdnje, naravno, imaju smisla samo ako se promatraju u relaciji s nečijom motivacijom, ali one su ipak normativne. Očito je da se ovako shvaćena normativnost razlikuje od Smithova shvaćanja normativnosti budući da su razlozi i dalje zavisni od arbitrarnih elemenata koji čine nečiji motivacijski skup. Ali nije jasno zašto bismo u prvom redu i prihvatili Smithovo shvaćanje normativnosti? Ako je točno da unutar hjumovske koncepcije ima smisla govoriti o tome što bi netko *trebao ili ne bi trebao* učiniti, tada smatram da je to dovoljno za obaranje Smithove tvrdnje prema kojoj razlozi unutar hjumovske koncepcije "nemaju normativni značaj".

Osvrnut ću se sada na drugi Smithov argument u prilog konvergencije odnosno protiv relativističkog shvaćanja razloga. Smith smatra da svakodnevna (moralna) praksa nedvojbeno pokazuje da su razlozi ne-relativni. Njegovim riječima:

> Mi se potencijalno slažemo ili ne slažemo jedni s drugima oko toga što konstituira razlog, a što ne. Zato u slučaju neslaganja – kao na primjer u slučaju neslaganja oko toga postoji li ili ne postoji razlog otići na godišnji odmor u određenim okolnostima – uvijek imamo opciju

> upustiti se u raspravu da bismo pokušali ustanoviti tko je u pravu a tko u krivu. Mišljenja drugih ljudi o tome koji su razlozi za djelovanje konstituira potencijalne izazove mojim vlastitim mišljenjima. … Ovo je sve nekonzistentno s tvrdnjom da je naš pojam razloga za djelovanje relativan u odnosu na pojedinca; da tipično znači $razlog_{ja}$ iz mojih usta, $razlog_{ti}$ iz tvojih, $razlog_{njezin}$ iz njezinih, i tako dalje. Prije sugerira da je naš pojam razloga nepobitno ne-relativan (Smith, 2004a: 32–33).

Mislim da Smithov argument u prilog ne-relativnosti razloga počiva na jednoj pretpostavci koja se vrlo lako može dovesti u pitanje i oko koje se u metaetičkim diskusijama još uvijek vode žustre rasprave. Da bih lakše istaknuo što imam u vidu, osvrnimo se ukratko na sam argument te pokušajmo Smithovo ne-relativističko shvaćanje razloga suprotstaviti onom relativističkom kakvo zastupa Williams. Što bi to značilo da su razlozi relativni? Pretpostavimo da se A, koji preferira filmove, i B, koji preferira kazališne predstave, dogovaraju oko večernjeg izlaska. A-u njegova preferencija pruža razlog da ide u kino, dok B-u njegova preferencija pruža razlog da ide u kazalište. Da li je ovo relativizam kakav Williams ima u vidu? Smith tvrdi da nije. Različitost njihovih preferencija jest ujedno i ključan čimbenik zašto se oni nalaze u različitim situacijama i zašto imaju različite razloge, ali to još uvijek ne znači da ovaj primjer podržava postojanje *relativnih* razloga. Jer kada bi se A, primjerice, nalazio u B-ovoj situaciji, onda bi on isto imao razlog ići u kazalište a ne u kino. Dakle, da bi se utvrdio relativizam razloga, nije dovoljno pokazati da ljudi imaju različite razloge kada se nalaze u različitim okolnostima, već da ljudi imaju različite razloge kada se nalaze u *relevantno sličnim* okolnostima. No takvu je situaciju, Smith argumentira, zapravo i teško zamisliti. Jer, kao što se uostalom i nastoji pokazati gornjim argumentom, kada bi dvoje ljudi – koji se nalaze u istoj situaciji – tvrdili da imaju različite razloge za djelovanje, to bi, Smith ističe, samo impliciralo *da je netko od njih u krivu*. Ali zašto bi netko od njih morao biti u krivu? Zašto ne pretpostaviti da oboje mogu biti u pravu? Zato što je ta pretpostavka – što

Smith u gornjem odlomku prešutno pretpostavlja – konstitutivan dio naše moralne prakse. Ne možemo zaključiti da su i A i B u pravu jer je očito da se oni *ne slažu* oko toga kako postupiti. Ali ako se radi o pravom *neslaganju*, onda je evidentno da *samo jedan od njih* može biti u pravu.[72] Sama moralna praksa favorizira ne-relativnost razloga.

Snaga gornjeg argumenta posebno dolazi do izražaja ukoliko taj argument sagledamo iz šire perspektive. U metaetičkim je raspravama, pogotovo kod realistički usmjerenih autora, često prisutno pozivanje na moralnu fenomenologiju. Moralni realisti, pokušavajući obraniti svoju poziciju, nerijetko ukazuju na same karakteristike moralnog diskursa te smatraju kako one svjedoče u prilog tome da je etika objektivna odnosno da postoji moralna stvarnost koju moralni sudovi opisuju. Nama se čini da moralni sudovi referiraju na objektivne moralne činjenice za koje pretpostavljamo da postoje nezavisno od naših vjerovanja. Također nam se čini da postoje neslaganja oko moralnih pitanja, vjerujemo da postoji opravdanje za mnoga naša moralna uvjerenja te smatramo da su neki ljudi "bliže" moralnoj istini od nekih drugih. Sagledavajući ove karakteristike u cjelini, možemo kazati da nam se moral *pojavljuje* na način da o njemu mislimo kao o nečemu objektivnom, kao o području koje možemo spoznati i koje nije ovisno o našim subjektivnim uvjerenjima. Kada i sami ne bismo smatrali da ne postoji moralna istina, tada se ne bismo niti upuštali u neslaganja o moralnim pitanjima, ne bismo pokušavali opravdati naše moralne sudove i ne bismo smatrali da nam neki ljudi mogu biti od pomoći kada je riječ o moralnim dilemama. Ali budući da ovo jesu glavna obilježja moralne prakse, zašto ne zaključiti da moral zaista jest takav kakvim ga mi doživljavamo? Autori koji negiraju postojanje objektivne moralne stvarnosti moraju pružiti barem nekakvo objašnjenje činjenice da moralni diskurs sadrži implicitno pozivanje na takvu stvarnost. Intuitivna

[72] Na jednom drugom mjestu Smith kaže sljedeće: "…trivijalno je da ako A tvrdi da je Φ vrijedno truda, a B tvrdi drukčije, onda je, kada je vrijednost u pitanju moralna vrijednost, samo jedan u pravu. … Suprotstavimo to slučaju u kojem se to neslaganje ne tiče nečega što ima moralnu vrijednost, već nečega što je samo stvar ukusa: recimo, da li je dobro jesti sladoled. … Jesti sladoled može biti dobro iz A-ova gledišta, ali ne iz B-ova gledišta. Oboje mogu biti u pravu. Međutim, tražiti takve skrivene relativnosti i učiniti da oboje budu u pravu čini se da nikada nije primjeren način rješavanja očevidnog konflikta kada je vrijednost koja je u pitanju moralna vrijednost. Kao što se to ponekad kaže, moralno neslaganje jest *neslaganje*, ne puka *razlika*. (Smith, 2004a: 252)

plauzibilnost realističkog rješenja leži u njegovoj jednostavnosti, dok antirealisti moraju ponuditi rješenje koje bi objasnilo realističku prirodu moralnog iskustva, ali koje ujedno ne bi impliciralo moralni realizam. Jer, ukoliko u tome ne uspiju, zašto tada ne bismo zaključili da etika uistinu jest o nečemu objektivnom? Ne ulazeći dublje u problem moralnog realizma, sada postaje sasvim očito gdje se skriva osnovna poteškoća. Uzimajući u obzir fenomenologiju moralnog iskustva, moramo priznati da nam se u mnogim situacijama (poglavito moralnim) čini da razlozi za djelovanje ne ovise o našim subjektivnim preferencijama. To poborniku objektivnih razloga već na samom početku omogućuje bolju "startnu poziciju" budući da samo treba ukazati na činjenicu da se ljudi vrlo često ponašaju *kao* da razlozi za djelovanje nisu relativni. Smith u svom argumentu u prilog konvergenciji koristi upravo tu strategiju i na taj način s lakoćom prebacuje teret dokazivanja na svog protivnika. Jer zašto bi se na koncu A i B upuštali u raspravu oko toga što imaju razloga učiniti kada sami ne bi smatrali da postoji *samo jedan* istiniti odgovor na to pitanje?

Unatoč tome što se ova linija argumentacije može činiti plauzibilnom, vrlo lako možemo uočiti jednu njezinu slabost ako barem na trenutak obratimo pozornost na pretpostavke na kojima je zasnovana. Možemo li naime iz činjenice da nam se *čini* da je neko stanje stvari takvo i takvo zaključiti da ono uistinu *jest* takvo i takvo? Ili, ako se u svakodnevnoj životnoj praksi zaista ponašamo *kao* da su razlozi ne-relativni, je li opravdano otići korak dalje i zaključiti da razlozi *zaista* jesu ne-relativni? Ali opet, Smith bi vjerojatno replicirao, čak i ako je neopravdano izvesti takav skok u zaključivanju, još je uvijek nesporno da taj zaključak predstavlja zaključak *na najbolje objašnjenje*. Najkraće rečeno, ako se svima *čini* da je određeno stanje stvari takvo i takvo, onda tu činjenicu najbolje možemo objasniti pretpostavkom da ono zaista i jest takvo i takvo.

Smithu se međutim može prigovoriti da oslanjanje na moralnu fenomenologiju ne može poslužiti kao konkluzivni argument za izgradnju određene metaetičke pozicije. Premda

teško možemo osporiti da karakteristike moralnog diskursa favoriziraju kognitivizam i objektivnost u etici, brojni su moralni filozofi ta svojstva nastojali objasniti ujedno zadržavajući nekognitivni pristup. Emotivistički pristup problemu moralnog neslaganja, na primjer, zasnovan je na pretpostavci da pravo moralno neslaganje ustvari ne postoji; da je moralno neslaganje zapravo neslaganje oko činjenica te da se svrha moralne argumentacije sastoji u tome da sugovornika natjeramo na promjenu moralnog stava. U okviru ove teorije nema nikakve osnove za tvrdnju da samo jedan može biti u pravu: ako su strane u sporu postigle konsenzus oko svih činjenica, ali prema tim činjenicama i dalje zauzimaju oprečne moralne stavove tada je svako daljnje neslaganje bespredmetno. "U tom slučaju", kako kaže Ayer, "nema smisla pitati koje je od konfliktnih gledišta istinito. Budući da izražavanje vrijednosnog suda nije propozicija pitanje istine i laži ovdje se ne javlja" (Ayer, 1952 [1936]: 22). Ovdje ne želim sugerirati da je emotivizam (ili neka verzija te teorije) istinita moralna teorija koja uspješno pristupa problemu moralne fenomenologije, već samo nastojim istaknuti da postoje i druga objašnjenja moralne prakse koja ujedno ne povlače realizam ili objektivizam u etici.

Netko bi na ovom mjestu mogao razmišljati na sljedeći način. "Kada ne bi postojala istina u etici, tada bi svatko mogao činiti što god želi i u tom slučaju ne bi postojala osnova za moralnu procjenu i kritiku; kada bi naši moralni sudovi imali subjektivan izvor, tada bi naše moralne dužnosti bile neutemeljene; kada moralne dužnosti ne bi bile eksternalne moralnim djelatnicima, tada one ne bi imale nikakvu snagu. Imamo sklonost vjerovati da moralne dužnosti svoj autoritet dobivaju upravo na osnovi toga što predstavljaju zahtjeve koji su o nama potpuno neovisni. Rezultat prihvaćanja emotivizma bi stoga mogao dovesti do neželjenih posljedica. Ljudi naprosto imaju *osjećaj* da su dužnosti izvan nas, da one predstavljaju zahtjeve kojih se trebamo držati neovisno o tome koje su naše želje i preferencije. Gubitak ovog osjećaja, koji može nastupiti uslijed objašnjenja da moralne

dužnosti predstavljaju puke projekcije naših emocija, može dovesti do stava da izvršavanje dužnosti zapravo uopće nema smisla".

Meni se čini da to ipak nije tako. Samo objašnjenje prave prirode određene prakse ne povlači da će se ljudi, jednom kada su upoznati s takvim objašnjenjem, početi i drukčije ponašati. Čak i ako su moralne vrijednosti samo naša projekcija, to nipošto ne znači da ljudi *ne mogu* imati stvarne dileme o tome što treba činiti, o tome što je moralno ispravno te koje su njihove moralne dužnosti. Blackburnovim riječima:

> Nije toliko iznenađujuće da dobrim stvarima u životu možemo pridavati vrijednost premda znamo da je vrijednost samo izražavanje naših subjektivnih čuvstava. To ne mora biti ništa manje čudno nego kada nalazimo stvari smiješnima, bolnima, vrijednima truda ili lijepima, premda je Bog mrtav, ili premda prihvaćamo subjektivni izvor ovih reakcija (Blackburn, 1993: 156).

Pod pretpostavkom da je *istinito* da moralne vrijednosti ne postoje, zašto bi to nužno značilo da se ljudi ne mogu ponašati *kao* da takve vrijednosti postoje? Čak i ako bismo pružili objašnjenje prirode morala koje ne bi bilo u skladu s našim svakodnevnim i intuitivnim razumijevanjem, to ne znači da prihvaćanje takvog objašnjenja mora imati značajan utjecaj na način na koji ljudi o moralu razmišljaju u svakodnevnom životu. Zašto bi onda Smithovo objašnjenje bilo najprirodnije? Zašto bismo na osnovi činjenice da dvoje ljudi imaju različite prosudbe oko toga što imaju razloga učiniti morali zaključiti da samo jedan od njih može biti u pravu? Smithov argument jednostavno pretpostavlja kognitivističku interpretaciju fenomenologije moralnog iskustva, ali ne nudi dodatno objašnjenje zašto bismo tu interpretaciju trebali prihvatiti. Kao što sam već napomenuo, mi još uvijek ne možemo kazati kakvu ulogu moralna fenomenologija ima u metaetičkim diskusijama i puko oslanjanje na fenomenologiju moralnog iskustva nije dovoljno da bismo se opredijelili za bilo koju

metaetičku teoriju. Jer čak i ako je tako da mi u svakodnevnim moralnim situacijama pretpostavljamo postojanje objektivne moralne stvarnosti, nije isključeno da bi *istinito* objašnjenje naše moralne prakse takvu pretpostavku smatralo krajnje neopravdanom. Jedno je pitanje kakva je *naizgled* priroda morala, a sasvim drugo kakva je *uistinu* priroda morala.

6. Zaključak

Objektivnost je, kako je to napisao Thomas Nagel, "središnji problem etike" (Nagel, 1986: 138). Moralni su filozofi – barem oni realističke orijentacije – oduvijek nastojali pokazati da moralnost nije samo stvar naših subjektivnih uvjerenja i stavova, da postoje moralni standardi koji su univerzalno važeći te da je na pitanja o ispravnom i pogrešnom, dobrom i lošem u konačnici moguće dati istinite odgovore. Problem objektivnosti, međutim, nije problem koji samo zadire u područje moralne ontologije, već i onaj koji za sobom povlači određene epistemološke konzekvence – zagovarati postojanje objektivne moralne stvarnosti nema previše smisla ako pritom ne mislimo da je ta stvarnost u načelu spoznatljiva. Za objektivistički nastrojene autore etika ima *kognitivnu* osnovu: s obzirom da moralne prosudbe mogu biti istinite ili lažne, one su, kako oni tvrde, nalik *vjerovanjima*. Možda najveću prepreku ovakvom pokušaju zasnivanja etike, kako je već sugerirano u samom početku ovoga rada, predstavlja motivacijska priroda morala. Ponovimo samo ukratko. Moralne prosudbe, barem nam se tako čini, ne koristimo samo zato (ako i uopće) da bismo opisali objektivno postojeće moralno stanje stvari, nego njihovom upotrebom također nastojimo izvršiti i određeni praktični utjecaj. Mi smatramo da su moralne prosudbe nešto što nas pokreće ili barem nešto što bi nas trebalo pokretati: motivacijski aspekt morala očituje se u našoj spremnosti da se ponašamo u skladu s prosudbama koje prihvaćamo. Međutim, ako moralne prosudbe motiviraju te ako se izvor motivacije, kao što se obično smatra, nalazi u *konativnom* dijelu ljudske prirode (*željama*), onda praktičnost morala nije jednostavno integrirati u objektivistički usmjerenu moralnu teoriju. Objektivisti smatraju da moralne prosudbe imaju kognitivni sadržaj, a time ostavljaju nerazjašnjenim pitanje njihova motivacijskog utjecaja.

Kantovski program, kao što se može iščitati iz diskusije u prethodnim poglavljima, predstavlja pokušaj objedinjavanja ovih dvaju elemenata. Shvatimo li moralne činjenice,

kantovci smatraju, kao činjenice o tome što imamo razloga činiti, tada je lako uvidjeti da objektivistički shvaćena moralnost ne mora biti u konfliktu s njezinim motivacijskim obilježjem. Dok neki tvrde da je za uspjeh takvog programa presudno napustiti duboko ukorijenjenu pretpostavku o željama kao pravom izvoru motivacije, neki pokazuju da čak ni ta pretpostavka nije u sukobu s kantovskim intuicijama. No, u čemu god da se sastojale njihove međusobne razlike, mislim da se kantovski pokušaj u glavnim crtama može okarakterizirati na sljedeći način. Moralnost pruža objektivne razloge za djelovanje, a govor o razlozima smislen je samo ako nas ti razlozi mogu motivirati. Razlozi ne motiviraju bilo koga, već samo onoga tko je racionalan, što znači da su samo racionalne osobe motivirane ponašati se moralno. Svi *imaju* dobar razlog činiti ono što moral od njih zahtijeva, a samo će oni potpuno racionalni biti motivirani to i činiti. Iz ovog je također vidljivo zašto postoji bliska veza između *moralnih prosudbi* i motivacije. S obzirom da su moralne prosudbe prosudbe o tome što imamo razloga činiti, onda će, lako je zaključiti, samo racionalne osobe biti motivirane djelovati u skladu sa svojim moralnim prosudbama.

Moj glavni cilj bio je pokazati da je ovakvo razumijevanje moralnosti pogrešno te obraniti alternativno objašnjenje moralne motivacije i moralnih razloga – objašnjenje čije su glavne karakteristike inspirirane filozofijom Davida Humea. Hjumovci, za razliku od svojih kantovskih oponenata, ne misle da je moguće pružiti takvo shvaćanje moralnosti koje bi istovremeno sadržavalo elemente objektivnosti i sačuvalo motivacijsku snagu moralnih sudova. Stoga oni, da se poslužim riječima Živana Lazovića, pružaju "…dosta uvjerljivo objašnjenje motivacijskog porijekla moralnog ponašanja, ali pod cijenu odricanja od moralnog realizma (Lazović, 1988: 176).

Ipak, tvrdnju da je hjumovsko objašnjenje moralne motivacije "dosta uvjerljivo" tek treba dokazati, a upravo sam to i nastojao učiniti u prvoj cjelini ovog rada. Sažeto rečeno, tamo je moja intencija bila obraniti argument čiju bismo strukturu, slijedeći Humea, mogli

prikazati ovako: (1) Izvor motivacije jest u željama; (2) Moralne prosudbe imaju motivacijsku snagu; (3) Dakle, moralne prosudbe moraju sadržavati želje. Obrani prve premise, koja prezentira središnju ideju hjumovske teorije motivacije, bilo je posvećeno drugo poglavlje. U tom sam poglavlju tvrdio da neki od najčešćih prigovora s kojima se obično suočavaju zastupnici hjumovskog modela ne uspijevaju taj model ozbiljnije uzdrmati te da upravo kognitivistički shvaćena motivacija u tom smislu nailazi na nesavladive prepreke. Ako želja nema funkciju središnjeg motivatora, onda je teško razumjeti zašto je njezina prisutnost u motivacijskom mehanizmu uopće potrebna, a ako prisutnost želje nije potrebna, onda se teško oteti dojmu da motivacija u nekom trenutku mora poprimiti oblike misterioznog procesa. Pitanje odnosa između morala i motivacije bilo je glavni predmet trećeg poglavlja. Za razliku od onih koji argumentiraju da između moralnih sudova i motivacije zapravo i ne postoji tako bliska veza kao što se obično misli te da je priroda tog odnosa samo kontingentna, tvrdio sam da moralni sudovi nužno motiviraju i da za generiranje motivacije – izuzev samog moralnog suda – nije potrebno nikakvo dodatno psihološko stanje. Pristajući dakle uz stajalište da su upravo želje ona mentalna stanja koja daju pravi motivacijski impuls te pokušavajući obraniti tvrdnju (2) odnosno internalističko shvaćanje moralne motivacije, očito je da sam se opredijelio za nekognitivistički pristup moralnosti, pristup koji "ne ostavlja mjesta za moralnu istinu" (McNaughton, [2010] 1988: 38).

Ako nekognitivizam ne ostavlja mjesta za moralnu istinu, onda ne ostavlja mjesta ni za objektivne moralne razloge. Zbog toga sam u drugoj cjelini ovog rada pokušao, između ostalog, dovesti u sumnju fundamentalnu pretpostavku cjelokupnog kantovskog pothvata – da moralne činjenice nisu ništa drugo nego činjenice o tome što imamo razloga činiti. Moralni zahtjevi, prema ovoj pretpostavci, imaju racionalni autoritet i zato je svako njihovo kršenje ekvivalentno kršenju zahtjeva racionalnosti. Ovakvom zasnivanju etike suprotstavio sam stajalište koje se s pravom može nazvati hjumovskim i prema kojemu razlozi za djelovanje

imaju isključivo instrumentalni karakter. Zagovaranje instrumentalne prirode praktičnih razloga u korijenu blokira svako kantovsko nastojanje postizanja objektivnosti i zato nije iznenađujuće da se najveći spor vodi oko pitanja normi po kojima se odvija praktično rasuđivanje. Budući da upravo procesom praktičnog rasuđivanja dolazimo do zaključaka o tome što imamo razloga činiti, jasno je da ti naši zaključci velikim dijelom ovise o tome kakva je priroda tog procesa. Dok hjumovci vjeruju da taj proces u osnovi mora biti instrumentalan zato jer je samo na taj način moguće objasniti motivacijsku snagu razloga za djelovanje, kantovci argumentiraju da postoje kategorične norme rasuđivanja te da motivacija nastaje kao rezultat prepoznavanja kategoričnih razloga koje te norme generiraju. Tako je središnje obilježje hjumovskog internalizma – pozicije u prilog koje sam argumentirao u četvrtom poglavlju – stav da razlozi za djelovanje ne mogu postojati bez prisutnosti adekvatne motivacije (*želje*) kod subjekta. Moralni razlozi u tom smislu ne predstavljaju nikakvu iznimku: to da je neki postupak "moralno izvrstan … uopće nije razlog za osobu da ga učini – osim ako njegova izvrsnost nije nešto do čega joj je stalo ili osim ako će joj taj postupak, bez obzira koliko bio izvrstan, donijeti nešto drugo što ona želi" (Shafer-Landau, R. i Cuneo, T., 2007: 281). Osigurati poseban status moralnim razlozima svakako nije moguće bez dodatne argumentacije, a upravo sam neke pokušaje pružanja takve argumentacije – one kantovskih internalista – kritizirao u petom poglavlju. Za razliku od hjumovskih internalista, koji vjeruju da subjekt ima razlog samo ako je motiviran u određenim idealnim uvjetima, kantovski internalisti tvrde da je subjekt nužno motiviran u idealnim uvjetima zato jer ima razlog. Motivacija, prema ovome gledištu, nije preduvjet prisutnosti razloga, već nastaje pod pritiskom razuma – kao posljedica prepoznavanja razloga koji nije ovisan o našoj kontingentnoj psihološkoj konstituciji. U kategoriju tih razloga, naravno, spadaju i oni moralni. U poglavlju sam pokušao pokazati zašto bi se samog Kanta moglo tumačiti kao

zagovornika takvog stajališta te sam također kritički razmotrio jedan suvremeni pokušaj argumentiranja u tom pravcu.

Sasvim je očito da naše razumijevanje praktičnih razloga baca određeno svjetlo i na pitanje opravdanja morala: netko tko traži odgovor na pitanje "zašto biti moralan?" ne traži ništa drugo nego razloge koji bi ga mogli motivirati na moralno postupanje. No kakve bismo razloge takvoj osobi uopće mogli pružiti? U svom poznatom članku "Does Moral Philosophy Rest on a Mistake?" (1968 [1928]) Prichard u osnovi sugerira da problem opravdanja počiva na potpuno pogrešnim pretpostavkama odnosno da je sam *zahtjev* za pružanjem razloga koji bi opravdali moralno djelovanje jednostavno "nelegitiman". Prichardova je ideja sljedeća: ako bismo moral pokušali opravdati *moralnim* razlozima, onda bismo se u takvom opravdanju zapravo pozivali na ono što jest predmet spora, tj. na ono što tek treba dokazati, a ako bismo moral pokušali opravdati *izvanmoralnim* razlozima (pozivanjem na nešto što osoba želi), onda se to, najjednostavnije rečeno, ne bi moglo nazvati opravdanjem. Jer, ideja je pokazati da trebamo postupati moralno, a ne da želimo postupati moralno. Za razliku od samog Pricharda – koji je ovu dilemu nastojao izbjeći priklanjajući se etičkom intuicionizmu – kantovci, vidjeli smo, vjeruju da razloge za moralno djelovanje ima svako racionalno biće odnosno da moralne zahtjeve ne treba promatrati kao nešto odvojeno i različito od samih zahtjeva racionalnosti. Zato se u okviru kantovskog programa pitanje "zašto biti moralan?" zapravo svodi na pitanje "je li racionalno biti moralan?", a, kako su mnogi utvrdili, tražiti razloge za racionalno ponašanje nije koherentno i taj bi zahtjev mogao postaviti samo netko tko ne razumije prirodu razloga kao takvih. Pitanje "zašto biti racionalan?" već sadrži pozivanje na racionalnost jer se njime traže razlozi, a razlog za neko djelovanje možemo imati samo ako je to djelovanje racionalno (usp. Dreier, 1997: 84). Međutim, ako racionalnost ovdje nije shvaćena hipotetički – na način da uključuje referiranje na subjektove želje, izbore i ciljeve – u kojem je onda smislu moralno ponašanje *racionalno*? William K. Frankena ovako odgovara na upravo

postavljeno pitanje: "Pretpostavimo da se slažemo da "Racionalno je biti moralan?" znači "Osoba bi izabrala biti moralna kada bi bila potpuno razborita, logična te na jasan način potpuno informirana o sebi i svemu što je relevantno" (Frankena, 1980: 90). Moralno je djelovanje, Frankena tvrdi, ono djelovanje koje bismo – uzevši sve u obzir – izabrali pod određenim idealnim uvjetima. Ima li smisla stoga kazati da bismo *svi* izabrali biti moralni kada bismo bili "potpuno razboriti, logični te potpuno informirani"? Frankena zaključuje da to ne možemo sa sigurnošću tvrditi te da je još uvijek moguće da u stadiju potpune razboritosti i informiranosti neki ljudi *ne* izaberu moralni način življenja. Međutim, on nastavlja, čini se razumnim "…*pretpostaviti* da su svi po prirodi tako konstituirani … da bi izabrali živjeti moralni život kada bi naznačeni uvjeti bili ispunjeni. To ne može biti dokazano te mora ostati pretpostavkom, ali također ne može biti opovrgnuto…" (Frankena, 1980: 94, kurziv u originalu). Uočimo da je u Frankeninoj posljednjoj opservaciji sadržano glavno razlikovanje između hjumovskih i kantovskih internalista razloga. Dok kantovski internalisti Frankeninu pretpostavku nastoje dokazati, hjumovski internalisti je pokušavaju opovrgnuti. Dok kantovski internalisti svoje stajalište grade na uvjerenju da bi u stadiju potpune razboritosti, logičnosti te informiranosti svaka osoba nužno izabrala [željela] postupati moralno, hjumovski internalisti argumentiraju da nema nikakvog jamstva da bi netko – čak i u tom stadiju – učinio takav izbor. Ono što sam, između ostalog, u ovom radu pokušao pokazati jest da hjumovac u ovom sporu odnosi pobjedu, da moral ne može biti opravdan, barem ne u onom smislu u kojem bi to opravdanje bilo opravdanje *za sve* te da je to, kako kaže James Dreier, "razočaranje s kojim ćemo morati naučiti živjeti" (Dreier, 1997: 88).

Bibliografija

Anscombe, G. E. M. 1963. [1957] *Intention* (Cambridge Mass.: Harvard University Press).

Ayer, A. J. 1952. [1936] *Language, Truth and Logic* (New York: Dover Publications, Inc.).

Benn, P. 1998. *Ethics* (London: UCL Press).

Berčić, B. 2009. "Deontološka etika" u M. Hudoletnjak Grgić (ur.), *Vrijednosti, djelovanje i znanje, Godišnjak za filozofiju* (Zagreb: Institut za filozofiju), 1–44.

Björnsson, G. 2002. "How Emotivism Survives Immoralists, Irrationality, and Depression", *The Southern Journal of Philosophy*, Vol. LX., 327–344.

Blackburn, S. 1993. "Errors and the Phenomenology of Value" u Blackburn, *Essays in Quasi-Realism* (Oxford: Oxford University Press), 149–165.

Blackburn, S. 2000. *Ruling Passions* (Oxford: Oxford University Press).

Brink, D. O. 1997. "Kantian Rationalism: Inescapability, Authority, and Supremacy" u G. Cullity, i B. Gaut, (ur.), *Ethics and Practical Reason* (Oxford: Oxford University Press), 255–291.

Brink, D. 1996. *Moral Realism and The Foundations of Ethics* (New York: Cambridge University Press).

Brink, D. 1997. "Moral Motivation", *Ethics* 108: 4–32.

Broome, J. 1999. "Normative Requirements", *Ratio* 12: 398–419.

Broome, J. 2004. "Reasons" u R. J. Wallace, P. Pettit, S. Scheffler i M. Smith (ur.), *Reason and Value*: *Themes from the Moral Philosophy of Joseph Raz* (Oxford: Oxford University Press), 28–55.

Broome, J. 2005. "Have We Reason To Do as Rationality Requires? A Comment on Raz", *Journal of Ethics and Social Philosophy* 1.

Chang, R. 2004. "Can Desires Provide Reasons for Action?" u R. J. Wallace, P. Pettit, S. Scheffler i M. Smith (ur.), *Reason and Value*: *Themes from the Moral Philosophy of Joseph Raz* (Oxford: Oxford University Press), 56–90.

Cohon, R. 1986. "Are External Reasons Impossible?", *Ethics* 96: 545–556.

Cowley, C. 2005. "A New Defence of Williams's Reasons-Internalism", *Philosophical Investigations* 28/4: 346–368.

Cullity, G. i Gaut, B. 1997. "Introduction" u Cullity i Gaut (ur.), *Ethics and Practical Reason* (Oxford: Oxford University Press), 1–27.

Dancy, J. 1993. *Moral Reasons* (Oxford: Blackwell).

Dancy, J. 2000. *Practical Reality* (Oxford: Oxford University Press).

Darwall, S. 1997. "Reasons, Motives, and the Demands of Morality: An Introduction" u Darwall, A. Gibbard i P. Railton, P. (ur.), *Moral Discourse & Practice: Some Philosophical Approaches* (Oxford: Oxford University Press), 305–312.

Darwall, S. 2006. "Morality and Practical Reason: A Kantian Approach" u D. Copp (ur.), *The Oxford Handbook of Ethical Theory* (Oxford: Oxford University Press), 282–320.

Davidson, D. 1980a. "Actions, Reasons, and Causes" u Davidson, *Essays on Actions & Events* (Oxford: Clarendon Press), 3–19.

Davidson, D. 1980b. "How is Weakness of the Will Possible?" u Davidson, *Essays on Actions & Events* (Oxford: Clarendon Press), 21–42.

Dreier, J. 1997. "Humean Doubts about the Practical Justification of Morality" u G. Cullity i B. Gaut (ur.), *Ethics and Practical Reason* (Oxford: Oxford University Press), 81–99.

Falk, W. D. 1952. [1947–48] "'Ought' and Motivation" u W. Sellars i J. Hospers (ur.), *Readings in Ethical Theory* (New York: Appleton Century Crofts), 492–510.

Finlay, S. i Schroeder, M. 2008. "Reasons for Action: Internal vs. External", *The Stanford Encyclopedia of Philosophy* (Fall 2008 Edition), Edward N. Zalta (ur.), URL = http://plato.stanford.edu/archives/fall2008/entries/reasons-internal-external/.

Flew, A. 1969. "On not deriving 'ought' from 'is'" u W. D. Hudson (ur.), *The Is/Ought Question* (London: The Macmillan Press Ltd.), 135–143.

Foot, P. 2002. [1972] "Morality as a System of Hypothetical Imperatives" u Foot, *Virtues and Vices* (Oxford: Clarendon Press), 157–173.

Frankena, W. K. 1997. [1958] "Obligation and Motivation in Recent Moral Philosophy" u T. L. Carson i P. K. Moser (ur.), *Morality and the Good Life* (Oxford: Oxford University Press), 399–419.

Frankena, W. K. 1980. *Thinking about Morality* (Michigan: The University of Michigan Press).

Frankfurt, H. G. 1988. "Freedom of the Will and the Concept of a Person" u Frankfurt, *The Importance of What We Care About: Philosophical Essays* (Cambridge: Cambridge University Press), 11–25.

Garrard, E. i McNaughton, D. 1998. "Mapping Moral Motivation", *Ethical Theory and Moral Practice* 1: 45–59.

Gert, B. 2004. *Brute Rationality: Normativity and Human Action* (Cambridge: Cambridge University Press).

Hare, R. M. 1998 [1952] *Jezik morala* (Zagreb: Hrvatski studiji. Preveo Filip Grgić).

Hare, R. M. 1969. "The Promising game" u W. D. Hudson (ur.), *The Is/Ought Question* (London: The Macmillan Press Ltd.), 144–156.

Hare, R. M. 2000. [1963] *Freedom and Reason* (Oxford: Oxford University Press).

Hooker, B. 2001. "Williams' Argument against External Reasons" u E. Millgram (ur.), *Varieties of Practical Reasoning* (Cambridge, Mass: MIT Press), 99–101.

Hooker, B. i Streumer, B. 2004. "Procedural and Substantive Practical Rationality" u A. R. Mele i P. Rawling (ur.), *The Oxford Handbook of Rationality* (Oxford: Oxford University Press), 57–74.

Hubin, D. C. 1999. "What's Special about Humeanism", *Nous* 33: 30–45.

Hubin, D. C. 2001. "The Groundless Normativity of Instrumental Rationality", *The Journal of Philosophy* 98/9: 445–468.

Hubin, D. C. 2003. "Desires, Whims and Values", *The Journal of Ethics* 7: 315–335.

Humberstone, I. L. 1992. "Direction of Fit", *Mind* 101/401: 59–83.

Hume, D. 1983. [1739–1740] *Rasprava o ljudskoj prirodi* (Sarajevo: Veselin Masleša. Preveo Borivoj Nedić).

Johnson, R. N. 2001. "Comments on Svavarsdóttir" (Brown Electronic Article Review Service).

Kant, I. 2001. [1785] *Utemeljenje metafizike ćudoređa* u J. Talanga (ur.), *Klasični tekstovi iz etike* (Zagreb: Hrvatski studiji).

Kolodny, N. 2005. "Why Be Rational?", *Mind* 114: 509–563.

Korsgaard, C. 1996a. *Creating the Kingdom of Ends* (Cambridge: Cambridge University Press).

Korsgaard, C. 1996b. *The Sources of Normativity* (Cambridge: Cambridge University Press).

Korsgaard, C. 1997. "The Normativity of Instrumental Reason" u G. Cullity i B. Gaut (ur.), *Ethics and Practical Reason* (Oxford: Oxford University Press), 215–254.

Lazović, Ž. 1988. *Razlozi, uzroci i motivi (Hjum i Kant)* (Beograd: FDS).

Lewis, D. 2000. "Dispositional theories of value" u Lewis, *Papers in Ethics and Social Philosophy* (Cambridge: Cambridge University Press), 68–94.

Mackie, J. L. 1977. *Ethics: Inventing Right and Wrong* (Harmondsworth: Penguin).
Markovits, J. 2006. *Kantian Internalism* (neobjavljena doktorska disertacija, University of Oxford).

McClellan, J. E. i Komisar, B. P. 1969. "On deriving 'ought' from 'is'" u W. D. Hudson (ur.), *The Is/Ought Question* (London: The Macmillan Press Ltd.), 157–162.

McDowell, J. 1995. "Might There Be External Reasons?" u J. E. J. Altham i R. Harrison (ur.), *World, Mind and Ethics: Essays on the ethical philosophy of Bernard Williams* (Cambridge: Cambridge University Press), 68–85.

McDowell, J. 1998. "Are Moral Requirements Hypothetical Imperatives?" u McDowell, *Mind, Value, and Reality* (Cambridge Mass.: Harvard University Press), 77–94.

McDowell, J. 2006. "Response to Philip Pettit and Michael Smith" u C. Macdonald i G. Macdonald (ur.), *McDowell and His Critics* (Oxford: Blackwell Publishing Ltd.), 170–179.

McNaughton, D. 2010. [1988] *Moralni pogled:uvod u etiku* (Zagreb: Hrvatski studiji. Preveo Tomislav Bracanović).

Mele, A. R. 2002. "John Searle: Rationality in Action" (Book review), *Mind* 111: 905–909.

Mele, A. R. 2003. *Motivation and Agency* (Oxford: Oxford University Press).

Mill, J. S. 2003. [1861] *Utilitarianism* u R. Crisp (ur.), *Utilitarianism/J.S.Mill* (Oxford: Oxford University Press).

Miller, A. 1996. "An objection to Smith's argument for internalism", *Analysis* 56/3: 169–174.

Miller, A. 2003. *An Introduction to Contemporary Metaethics* (Cambridge: Polity Press).

Millgram, E. 1996. "Williams' Argument Against External Reasons", *Nous* 30:2, 197–220.

Millgram, E. 1997. *Practical Induction* (Cambridge, Harvard University Press).

Millgram, E. 2001. "Practical Reasoning: The Current State of Play" u Millgram (ur.), *Varieties of Practical Reasoning* (Cambridge, Mass: MIT Press), 1–26.

Millgram, E. 2005. "Was Hume a Humean" u Millgram, *Ethics Done Right: Practical Reasoning as a Foundation for Moral Theory* (Cambridge: Cambridge University Press), 198–217.

Nagel, T. 1970. *The Possibility of Altruism* (Princeton: Princeton University Press).

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

O'Neill, O. 2004. "Kant: Rationality as Practical Reason" u A. R. Mele i P. Rawling (ur.), *The Oxford Handbook of Rationality* (Oxford: Oxford University Press), 93–109.

Parfit, D. 1997. "Reasons and Motivation", *Proceedings of the Aristotelian Society* 71: 99–130.

Pettit, P. i Smith, M. 2006. "External Reasons" u C. Macdonald i G. Macdonald (ur.), *McDowell and His Critics* (Oxford: Blackwell Publishing Ltd.), 142–170.

Platon 1997. *Država* (Zagreb: Naklada Jurčić. Preveo M. Kuzmić).

Price, H. 1989. "Defending Desire-as-Belief", *Mind* Vol. 98, No. 389: 119–127.

Prichard, H. A. 1968 [1928]. "Does Moral Philosophy Rest on a Mistake" u Prichard, *Moral Obligation* (Oxford: Oxford University Press), 1–17.

Prichard, H. A. 1968 [1928]. "Duty and Interest" u Prichard, *Moral Obligation* (Oxford: Oxford University Press), 201–238.

Quinn, W. 2002. "Putting Rationality in its Place" u R. Hursthouse, G. Lawrence i Quinn (ur.), *Virtues and Reasons: Philippa Foot and Moral Theory* (Oxford: Oxford University Press), 181–208.

Railton, P. 2006. "Humean Theory of Practical Rationality" u D. Copp (ur.), *The Oxford Handbook of Ethical Theory* (Oxford: Oxford University Press), 265–281.

Raz, J. 1997. "Incommensurability and Agency" u R. Chang (ur.), *Incommensurability, Incomparability, and Practical Reason* (Cambridge Mass.: Harvard University Press), 110–128.

Raz, J. 2005. "The Myth of Instrumental Rationality", *Journal of Ethics and Social Philosophy*, 1.

Ross, D. 2002. [1930] *The Right and the Good*, tiskano s uvodom Phillip Stratton-Lakea (Oxford: Clarendon Press).

Sadler, B. J. 2003. "The Possibility of Amoralism: A Defence Against Internalism", *Philosophy* 78: 63–78.

Scanlon, T. 1998. *What We Owe to Each Other* (Cambridge: Harvard University Press).

Schueler, G. F. 1995. *Desire: Its Role in Practical Reason and the Explanation of Action* (Cambridge, Mass.: MIT).

Scruton, R. 1982. *Kant* (Oxford: Oxford University Press).

Sesardić, N. 1984. *Fizikalizam* (Beograd: Istraživačko izdavački centar SSO Srbije).

Shafer-Landau, R. 1998. "Moral Judgement and Moral Motivation", *The Philosophical Quarterly* 48/192: 353–358.

Shafer-Landau, R. 2003. *Moral Realism: A Defence* (Oxford: Oxford University Press).

Shafer-Landau, R. 2004. *Whatever Happened to Good and Evil?* (Oxford: Oxford University Press).

Shafer-Landau, R. i Cuneo, T. 2007. "Introduction" u Shafer-Landau i Cuneo (ur.), *Foundations of Ethics: An Anthology* (Oxford: Blackwell Publishing Ltd.), 281–285.

Shaw, D. 1989. "Hume's Theory of Motivation", *Hume Studies* Vol. XV/1: 163–183.

Sinhababu, N. 2009. "The Humean Theory of Motivation Reformulated and Defended", *Philosophical Review,* Vol. 118/4: 465–500.

Skorupski, J. 2007. "Internal Reasons and the Scope of Blame" u A. Thomas (ur.), *Bernard Williams* (Cambridge: Cambridge University Press), 73–103.

Searle, J. 1969. "How to Derive 'Ought' from 'Is'"? u W. D. Hudson (ur.), *The Is/Ought Question* (London: The Macmillan Press Ltd.), 120–134.

Searle, J. 2001. *Rationality in Action* (Cambridge: MIT).

Smith, M. 1994. *The Moral Problem* (Oxford: Basil Blackwell).

Smith, M. 1996. "The argument for internalism: reply to Miller", *Analysis* 56/3: 175–184.

Smith, M. 2004a. *Ethics and the A Priori* (Cambridge: Cambridge University Press).

Smith, M. 2004b. "Humean Rationality" u A. R. Mele i P. Rawling (ur.), *The Oxford Handbook of Rationality* (Oxford: Oxford University Press), 75–92.

Sobel, D. i Copp, D. 2001. "Against direction of fit accounts of belief and desire", *Analysis* 61: 44–53.

Stampe, D. W. 1987. "The Authority of Desire", *The Philosophical Review* 96: 335–381.

Stevenson, C. L. 1952. [1937] "The Emotive Meaning of Ethical Terms" u W. Sellars i J. Hospers (ur.), *Readings in Ethical Theory* (New York: Appleton Century Crofts), 415–429.

Svavarsdóttir, S. 1999. "Moral Cognitivism and Motivation", *The Philosophical Review* 108: 2, 161–219.

Svavarsdóttir, S. 2001. "Replies to Critics" (Brown Electronic Article Review Service).

Svavarsdóttir, S. 2006. "How Do Moral Judgments Motivate?" u J. Dreier (ur.), *Contemporary Debates in Moral Theory* (Oxford: Blackwell Publishing Ltd.), 163–181.

Talanga, J. 1999. *Uvod u etiku* (Zagreb: Hrvatski studiji).

Thomson, J. i Thomson, J. 1969. "How not to derive 'ought' from 'is' u W. D. Hudson (ur.), *The Is/Ought Question* (London: The Macmillan Press Ltd.), 163–167.

Velleman, J. D. 1996. "The Possibility of Practical Reason", *Ethics* 106: 694–726.

Wallace, R. J. 1990. "How to Argue about Practical Reason", *Mind* 99/395: 355–385.

Wallace, R. J. 2006. “Moral Motivation” u J. Dreier (ur.), *Contemporary Debates in Moral Theory* (Oxford: Blackwell Publishing Ltd.), 182–196.

Williams, B. 1981. “Internal and External Reasons” u Williams, *Moral Luck* (London: Cambridge University Press), 101–113.

Williams, B. 1985. *Ethics and the Limits of Philosophy* (London: Fontana).

Williams, B. 1995a. “Internal reasons and the obscurity of blame” u Williams, *Making Sense of Humanity* (Cambridge: Cambridge University Press).

Williams, B. 1995b. “Replies” u J. E. J. Altham i R. Harrison, R. (ur.), *World, Mind and Ethics: Essays on the ethical philosophy of Bernard Williams* (Cambridge: Cambridge University Press), 185–224.

Williams, B. 2001. “Postscript: Some Further Notes on Internal and External Reasons” u E. Millgram (ur.), *Varieties of Practical Reasoning* (Cambridge, Mass: MIT Press), 91–97.

Williams, B. 2006. “Values, Reasons, and the Theory of Persuasion” u A. W. Moore (ur.), *Philosophy as a Humanistic Discipline* (Princeton, NJ: Princeton University Press), 109–118.

Zangwill, N. 2007. “Externalist Moral Motivation” u R. Shafer-Landau i T. Cuneo (ur.), *Foundations of Ethics: An Anthology* (Oxford: Blackwell Publishing Ltd.), 243–251.

Zangwill, N. 2009. “Non-cognitivism and Motivation” u C. Sandis (ur.), *New Essays on the Explanation of Action* (Basingstoke: Palgrave Macmillan), 416–424.

Životopis

Matej Sušnik rođen je 1978. u Zagrebu gdje završava osnovnu školu. Nakon završene srednje škole u Kairu (Cairo American College), 1998. upisuje studij filozofije i povijesti na Hrvatskim studijima Sveučilišta u Zagrebu, a diplomu stječe 2004. Iste godine upisuje poslijediplomski studij filozofije na Hrvatskim studijima, a od 2005. zaposlen je na Filozofskom fakultetu Sveučilišta u Rijeci kao znanstveni novak u suradničkom zvanju asistenta gdje je angažiran na znanstvenim projektima "Racionalnost" i "Normativnost u teoriji spoznaje i etici". Sudjelovao je u izvođenju većeg broja kolegija na Filozofskom fakultetu u Rijeci, izlagao je na više domaćih i međunarodnih filozofskih konferencija u Zagrebu, Rijeci i Bledu te je objavio znanstvene radova u časopisima *Croatian Journal of Philosophy, Prolegomena* te *Filozofska istraživanja*. Član je Udruge za promicanje filozofije te Hrvatskog društva za analitičku filozofiju.

Sažetak

Cilj ovog rada jest obraniti tvrdnju da se o motivima i razlozima ne može govoriti bez referiranja na subjektove želje te ujedno razmotriti kakve implikacije ta teza ima specifično za naše razumijevanje prirode *moralne* motivacije te prirode *moralnih* razloga. Dok se s jedne strane opredjeljujem za stajalište prema kojemu nas moralnost nužno motivira, s druge se strane zalažem za relativizam u pogledu razloga te tvrdim da ne postoje objektivni razlozi za moralno djelovanje. Budući da je moje gledište hjumovsko, najvećim se dijelom suprotstavljam kantovski usmjerenim autorima koji – utemeljujući moralnost u praktičnom razumu – pružaju sasvim drukčije objašnjenje motivacijske prirode morala i statusa moralnih razloga.

U uvodnom poglavlju pružam glavne smjernice u okviru kojih će se kretati moja rasprava. U drugom poglavlju branim hjumovsku teoriju motivacije – teoriju u čijem se središtu nalazi tvrdnja da nas motiviraju želje. Upotrebljavajući neke ključne distinkcije koje nalazimo kod samoga Humea te ih smještajući u suvremeni kontekst, pokušavam pokazati kako one mogu poslužiti zastupniku ovog modela u obrani od nekih standardnih prigovora. U trećem poglavlju razmatram kakvo razumijevanje *moralne* motivacije slijedi iz prihvaćanja teze da želje motiviraju. Pokazujući na koji je način hjumovska teorija kompatibilna s *internalizmom* i s *eksternalizmom motivacije*, opredjeljujem se za prvu poziciju. Internalizam motivacije jest stajalište koje tvrdi da moralne prosudbe nužno motiviraju, dok eksternalisti relaciju između moralnih prosudbi i motivacije smatraju raskidivom i kontingentnom. U poglavlju tvrdim da neki od najutjecajnijih eksternalističkih prigovora predstavljaju opasnost samo za, kako ću to ovdje nazvati, kantovski interpretiran internalizam motivacije, ali ne i za hjumovsku verziju tog gledišta.

U drugoj cjelini rada (četvrto i peto poglavlje), analiziram prirodu razloga za djelovanje. U četvrtom poglavlju nastojim obraniti shvaćanje razloga koje se u širem smislu može nazvati hjumovskim. Preciznije rečeno, branim gledište poznato pod nazivom *internalizam razloga*, a prema kojemu je razlog za djelovanje ispravno pripisati samo onome tko ima odgovarajuću želju, tj. onome tko je već motiviran na odgovarajući način. *Eksternalisti razloga*, s druge strane, nastoje pokazati da postoje razlozi za djelovanje koji su neovisni o tome je li subjekt motiviran djelovati u skladu s njima. U ovom poglavlju nastojim eliminirati četiri eksternalistička prigovora. Najveći izazov hjumovskom internalizmu razloga predstavlja kantovska interpretacija tog stajališta. Kantovski orijentirani autori nastoje osporiti relativističke implikacije hjumovskog internalizma te žele pokazati da je unutar internalističkog okvira također ostavljen prostor za dokazivanje postojanja kategoričnih razloga. Napadu na kantovski pristup posvećeno je peto, posljednje poglavlje ovog rada.

Ključne riječi: želje, praktični razum, motivacija, razlozi, internalizam motivacije, internalizam razloga, Hume, Kant

Abstract

This thesis aims to argue that we can only consider motives and reasons in view of the desires of the subject, and examine what are the implications of this claim for the understanding of the nature of *moral* motivation and the nature of *moral* reasons. While on one side, I agree with the view that morality necessarily motivates us, on the other side I argue for the relativism in relation to reasons and I claim that there are no objective reasons for moral behaviour. My view is Humean, which means that I mostly oppose Kantian authors who – by placing the foundation of morality in practical reason – offer an entirely different explanation of the motivational nature of morality and status of moral reasons.

In my introductory chapter I offer main directions indicating the trajectory of my argument. In the second chapter I defend the Humean theory of motivation – a theory centered around the notion that we are motivated by our desires. By relying on some key distinctions that we encounter in Hume's work and by placing them in a contemporary context, I attempt to prove that they can serve the one advocating this model to resist some standard objections. In the third chapter, I analyse the understanding of the moral motivation that follows from the acceptance of the claim that desires are motivating. Showing the way in which the Humean theory is compatible with *internalism* and *externalism about motivation*, I argue for the former. Internalism about motivation is the position that argues that moral judgments are necessarily motivating, while externalists regard the relation between moral judgement and motivation contingent. In the second chapter I argue that some of the most influential externalist objections pose a danger to, what I refer to here as Kantian internalism about motivation, but not to the Humean version of that position.

In the second part of the work (fourth and fifth chapter), I analyse the nature of reasons for action. In the fourth chapter I attempt to defend the Humean understanding of reasons. More precisely, I defend the viewpoint known as *internalism about reasons*. According to this position it is appropriate to attribute a reason for action solely to the one who has a corresponding desire, that is, to the one who is already motivated accordingly. *Externalists about reasons*, on the other hand, attempt to show that there are reasons for action that are independent of the question whether the subject is motivated to act accordingly or not. In this chapter I am trying to eliminate four externalist objections. The biggest challenge to the Humean internalism about reasons is Kantian interpretation of this view. Kantians are trying to refute the relativistic implications of Humean Internalism and want to show that within internalistic framework there is also a space for advocating the existence of categorical reasons. The fifth, final chapter of this thesis is dedicated to the rebuttal of this Kantian approach.

Key words: desires, practical reason, motivation, reasons, internalism about motivation, internalism about reasons, Hume, Kant.