ANαliZA

časopis za kritično misel

Letnik 8, številka 4, december 2004 · ISSN 1408-2969

SVOBODA IN VOLJA

Harry G. Frankfurt

Svobodna volja in pojem osebe

prevedla Milan Franc in Danilo Šuster

Citiranje: Harry G. Frankfurt, »Svobodna volja in pojem osebe«, *Analiza: časopis za kritično misel* 8, št. 4 (2004): 33–45.

Prevod. Licenca Creative Commons, ki velja za arhiv revije, za prevode ne velja; pravice izvirnika pripadajo njegovim imetnikom.

Izdajatelj: Društvo za analitično filozofijo in filozofijo znanosti, Ljubljana.

Elektronska različica je bila leta 2026 rekonstruirana iz izvirnih stavnih datotek in preverjena s tiskano izdajo. Številke strani ustrezajo tiskani izdaji; prelomi strani znotraj prispevka se lahko nekoliko razlikujejo.

Harry G. Frankfurt
Prevedla Milan Franc in Danilo Šuster

SVOBODNA VOLJA IN POJEM OSEBE*

Kar filozofi zadnje čase sprejemajo kot analizo pojma osebe, pravzaprav sploh ni analiza *tega* pojma. Strawson, katerega raba je sodobni standard, pojem osebe opredeli kot "pojem tipa bitnosti, ki je takšna, da lahko tako predikate, ki pripisujejo stanja zavesti, kot predikate, ki pripisujejo telesne značilnosti, enakovredno pripišemo individuumu posameznega tipa".[1] Vendar pa obstaja razen oseb še mnogo drugih bitnosti, ki imajo tako mentalne kot psihične lastnosti. Zato se dogaja, čeprav se zdi nekoliko nenavadno, da za tip bitnosti, ki jo ima v mislih Strawson, v angleškem jeziku ne najdemo ustreznega izraza. Gre za tip, ki ne vključuje samo ljudi, ampak tudi različne vrste živali. Kljub temu pa to komaj opraviči neustrezno rabo dragocenega filozofskega termina.

Tega, ali so pripadniki določene živalske vrste osebe, vsekakor ne moremo dognati zgolj tako, da določimo, ali je pravilno, da na njih apliciramo predikate, ki pripisujejo stanja zavesti in ne le tiste, ki pripisujejo telesne značilnosti. Aplikacija termina 'oseba' na številna bitja, ki imajo hkrati psihološke in materialne lastnosti, vendar se ne manifestirajo kot osebe v nobenem normalnem pomenu te besede, je posiljevanje jezika. Takšna jezikovna zloraba nedvomno ne zagreši nobene teoretične napake. Čeprav je spotika zgolj verbalna, povzroči pomembno škodo, saj neupravičeno osiromaši filozofski besednjak in poveča možnost prezrtja pomembnega področja raziskav, ki ga kar najbolj samoumevno povezujemo s pojmom 'oseba'. Lahko bi pričakovali, da za filozofe ni bolj osrednjega in trajnega problema od poskusa razumevanja tega, kaj v bistvu smo. A ravno to vprašanje je tako vsesplošno zanemarjeno, da smo dopustili celo krajo njegovega poimenovanja, ne da bi nas pri tem mučil kak hud občutek izgube.

Beseda 'oseba' lahko pomeni samo edninsko obliko pojma 'ljudje', pri tem pa oba termina ne označujeta nič več kot pripadnost določeni biološki vrsti. V tistih pomenih besede, ki so predmet filozofskega zanimanja, pa merila za določanje bistva osebe primarno niso namenjeni razločevanju pripadnikov naše vrste od pripadnikov drugih vrst. Merila za določanje bistva osebe naj bi zajela tiste značilnosti, ki so predmet našega najglobljega zanimanja za nas same in izvor tega, kar štejemo za najpomembnejše in najbolj problematično v našem življenju. Tisto, kar nas najbolj zanima v naši človeški situaciji, nas ne bi nič manj zanimalo, tudi če bi šlo za značilnost drugih bitij.

---

* Prevedeno po H. G. Frankfurt, "Freedom of the Will and the Concept of a Person", *Journal of Philosophy*, zv. 68, št. 1 (14. jan. 1971), str. 5-20. Oštevilčene opombe so izvirne.

[1] P. F. Strawson, *Individuals* (London, Methuen, 1959), 101-102. Ayerjeva raba termina 'oseba' je zelo podobna: 'za osebe je v tem smislu značilno, da so poleg raznih fizičnih lastnosti ... obdarjene tudi z različnimi oblikami zavesti' (A. J. Ayer, *The Concept of Person* (New York: St. Martin's, 1963), 82). Kar zanima Strawsona in Ayerja, je v prvi vrsti problem povezanosti duha in telesa, za razliko od precej drugačnega problema, kaj pomeni biti bitje, ki nima samo duha in telesa, marveč je tudi oseba.

Našega lastnega pojmovanja nas samih kot oseb zatorej ne gre razumeti kot pojma, sestavljenega iz atributov, ki so nujno specifični za našo vrsto. Konceptualno gledano je namreč možno, da bi bili kot osebe pojmovani tudi člani novih ali celo znanih nečloveških vrst, pa tudi, da nekateri člani človeške vrste niso osebe. Dejansko sicer zares privzemamo, da noben pripadnik drugih vrst ni oseba. V skladu s tem velja predpostavka, da vse tisto, kar je bistveno za osebo, sodi med zbir lastnosti, za katere na splošno domnevamo – pa najsi bo po pravici ali po krivici —, da so edinstveno človeške.

Po mojem mnenju lahko eno od bistvenih razlik med osebami in drugimi bitji najdemo v zgradbi volje osebe. Človeška bitja niso edina, ki imajo želje in motive ali sposobnost izbiranja. Te lastnosti delijo s pripadniki določenih drugih vrst in za nekatere izmed njih se celo zdi, da presojajo ter sprejemajo odločitve na osnovi predhodnega razmisleka. Sposobnost oblikovanja tega, kar bom imenoval ‘želje drugega reda’, pa se vendar zdi edinstvena lastnost ljudi.

Poleg tega, da hočejo, izbirajo in so motivirani, *da storijo* to ali ono, ljudje prav tako lahko hočejo imeti (ali ne imeti) določene želje in motive. Zmožni so hoteti biti drugačni, kot so v svojih preferencah in namenih. Za mnoge živali se namreč zdi, da so sposobne imeti to, kar bom imenoval ‘želje prvega reda’, ki so preprosto želje storiti ali ne storiti to ali ono. Kljub vsemu pa se dozdeva, da nobena žival razen človeka nima zmožnosti reflektivnega samovrednotenja, ki se izpričuje v oblikovanju želja drugega reda.[2]

I.

Pojem, ki ga označujemo z glagolom ‘hoteti’ [*want*], je nenavadno težko opredeljiv. Trditev oblike *“A* hoče *X”* nam sama po sebi in vzeta iz konteksta, s katerim razložimo ali specificiramo njen pomen, posreduje izjemno malo informacij. Takšna trditev je, na primer, združljiva z vsako od naslednjih: (a) pričakovanje, da bo storil *X*, ne izzove nobenega občutka ali introspektivnega čustvenega odziva v *A*; (b) *A* se ne zaveda, da hoče *X; (c) A* je prepričan, da noče *X*; (d) *A* se želi vzdržati početja *X*; (e) *A* hoče y in je prepričan, da je zanj nemogoče, da bi hkrati storil *Y* in *X*; (f) *A* “v resnici” noče storiti *X*; (g) *A bi raje umrl*, kot pa storil *X*, in tako naprej. Formuliranje razlike med željami prvega in drugega reda na način, kot ga jaz predlagam in po katerem ima nekdo željo prvega reda, ko želi ali ne želi storiti nečesa, in željo drugega reda, ko želi ali ne želi imeti želje prvega reda, je zato komaj zadostno.

Pri mojem razumevanju trditve oblike *“A* hoče *X”* pokrivajo kar precejšen obseg

[2] Zavoljo preglednosti se bom ukvarjal samo s tem, kar kdo hoče ali želi, ter pri tem zanemaril sorodne fenomene, kot so izbire in odločitve. Predlagam, da uporabljamo glagola ‘hoteti’ [*to want*] in ‘želeti’ kot zamenljiva, čeprav nikakor nista popolnoma sinonima. Moj vzgib za opustitev uveljavljenih pomenskih odtenkov izhaja iz dejstva, da glagol ‘hoteti’, ki sicer pomensko bolje ustreza mojemu namenu, ni tako pripraven za tvorbo samostalnikov kot glagol ‘želeti’. Morebiti bi bilo sprejemljivejše, čeprav grdo, govoriti v množini o “hotnostih” [*wants*] določene osebe. Toda govoriti o hotnosti določene osebe v ednini bi bilo gnusno. [V prevodu je ‘want’ dosledno preveden kot ‘hoteti’ ali ‘hotenje’, ‘desire’ pa kot ‘želeti’ ali ‘želja’. – op. prev.]

možnosti.[3] Lahko so resnične tudi tedaj, ko so resnične trditve od (a) do (g): ko se *A* ne zaveda nobenih občutkov, ki zadevajo *X*; ko se *A* ne zaveda, da hoče *X*; ko se slepi o tem, kar hoče, in je zmotno prepričan, da noče *X*; ko ima še druge želje, ki so v nasprotju z *X*; in ko je *A* ambivalenten. Omenjene želje so lahko zavestne ali nezavedne, lahko da niso nedvoumne in *A* se lahko o njih tudi moti. Vendar obstaja dodaten vir negotovosti, povezan s trditvami, ki identificirajo želje dane osebe in tu moramo biti strožji.

Najprej razmislimo o tistih trditvah oblike "*A* hoče *X*", ki opredelijo želje prvega reda in v katerih se *X* nanaša na dejanje. Trditev te vrste sama po sebi ne nakazuje relativne moči A-jeve želje po *X*. Iz nje ni razvidno, ali ta želja sploh igra odločilno vlogo pri tem, kar *A* v resnici stori ali poskuša storiti. Lahko bi bilo povsem pravilno reči, da *A* hoče *X* celo v primeru, ko je njegova želja po *X* samo ena od njegovih želja, in četudi je daleč od tega, da bi bila najodločilnejša med njimi. Res pa je lahko tudi to, da *A* hoče *X* kljub dejstvu, da ga takrat, ko stori, kar stori, ne motivira želja po *X*. Po drugi strani pa lahko nekdo, ki trdi, da *A* hoče *X*, s tem meni, da je ravno ta želja (po *X)* tista, ki dejavnika *A* motivira ali vodi do tega, da stori, kar dejansko stori, ali da ga bo k ravnanju navedla prav ta želja (razen *če* si ne premisli).

Samo v tem drugem smislu gre za trditev, ki določa A-jevo voljo, glede na posebno rabo pojma "volja", kot jo jaz predlagam. Določiti voljo dejavnika pomeni določiti željo (ali želje), ki ga motivirajo za dano dejanje, ki ga izvede, ali določiti želje, ki ga bodo ali bi ga motivirale v primeru, ko dejanje izvede ali bi ga izvedel. Volja dejavnika je potemtakem istovetna z eno ali več njegovih želja prvega reda. Vendar pojem volje, kot ga tu uporabljam, nima enakega obsega kot pojem želje prvega reda. Ne gre samo za pojem nečesa, kar zgolj do neke mere nagiba dejavnika, da ravna na določen način. Gre za pojem *učinkovite* želje, tiste, ki spodbudi (ali bo oz. bi spodbudila) osebo k dejanju. Zato pojem volje nima enakega obsega s pojmom tega, kar dejavnik namerava storiti. Četudi ima kdo trden namen storiti *X*, lahko še vseeno stori nekaj drugega, ker je njegova želja storiti *X* – kljub njegovi nameri – šibkejša ali manj učinkovita od neke nasprotne želje.

Sedaj pa si oglejmo še trditve oblike '*A* hoče *X*', ki določajo želje drugega reda, se pravi trditve, v katerih se termin *X* nanaša na željo prvega reda. Tudi tukaj obstajata dve vrsti situacij, v katerih je lahko res, da *A* hoče hoteti *X*. *V* prvem primeru bi za *A* lahko veljajo, da hoče imeti željo po *X*, kljub dejstvu, da ima enoznačno željo, ki je nedvoumna in ji nič ne nasprotuje, da bi se vzdržal tega, da bi storil *X*. Z drugimi besedami, nekdo bi lahko hotel imeti določeno željo, vendar bi hkrati nedvoumno hotel, da ta želja ostane nezadovoljena.

Vzemimo, da je zdravnik, ki se ukvarja s psihoterapijo narkomanov, prepričan, da bi svojo sposobnost pomoči zasvojencem izboljšal, če bi bolje razumel, kaj za

[3] To, o čemer govorim v tem odstavku, se ne nanaša samo na primere, v katerih se 'da bi *X* nanaša na možno dejanje ali nedejavnost, temveč tudi na primere, v katerih se 'da bi *X* nanaša na željo prvega reda in v katerih je trditev 'X hoče *X* skrajšana verzija za '*A* hoče hoteti *X*, ki označuje željo drugega reda.

njih pomeni želeti si mamilo, s katerim so zasvojeni. Denimo, da ga to napelje k hotenju, da bi imel željo po mamilu. Če je želja, ki jo hoče imeti, pristna, potem ne le da hoče občutiti stanje, značilno za počutje zasvojencev, ko jih grabi želja po mamilu, ampak (v kolikor hoče imeti to željo) hoče biti do določene mere nagnjen ali spodbujen k zaužitju mamila.

Povsem mogoče pa je, da kljub hotenju, da bi imel željo, ki bi ga spodbudila k jemanju mamila, noče, da bi ta želja bila učinkovita. Mogoče ne želi, da ga vodi prav do končnega dejanja. Ni nujno, da ga zanima, kaj pomeni vzeti mamilo. In v kolikor zgolj hoče, da bi ga *hotel* vzeti, ne pa tudi, da bi ga *resnično* vzel, v njegovem trenutnem hotenju ni ničesar takega, kar bi bilo zadovoljeno z mamilom samim. Pravzaprav lahko ima v tem trenutku enoznačno željo, da mamila *ne* vzame; lahko torej tako preudarno uredi okoliščine, da si onemogoči zadovoljitev želje, ki bi jo imel, če bi bila njegova želja po hotenju mamila pravočasno zadovoljena.

Iz dejstva, da si zdravnik trenutno hoče želeti, da bi vzel mamilo, bi bilo zatorej napačno sklepati, da si v resnici že želi, da bi ga vzel. Iz njegove želja drugega reda, da bi bil motiviran za vzetje mamila, še ne sledi, da ima željo prvega reda po mamilu. Če bi mu dali mamilo, to morda ne bi zadovoljilo nobene želje, ki je implicirana v njegovi želji po hotenju, da bi odmerek vzel. Kajti ko lahko hoče, da bi hotel vzeti mamilo, morda nima *nobene* želje po tem, da bi ga zaužil. Morda je *vse*, kar si želi, okusiti željo po mamilu. To pa pomeni, da njegova želja, da bi imel določeno željo, ki je nima, ni nujno hkrati tudi želja, da bi bila njegova volja kakorkoli drugačna, kot je.

Kdor zgolj na tak okrnjen način hoče *X*, je na meji izumetničenosti in dejstvo, da hoče hoteti *X*, ni primerno za določitev njegove volje. Vendar pa obstaja še drugačna vrsta situacij, ki jih lahko opišemo kot *"A* hoče *X"*, in kadar trditev uporabimo za opis situacije te vrste, se le-ta nanaša na tisto, za kar *A* hoče, da je njegova volja. V teh primerih trditev pomeni, da *A* hoče, da je želja po *X* tista želja, ki ga spodbudi (vodi) k učinkovitemu dejanju. Ne gre za to, da hoče imeti željo po *X* med željami, ki ga v večji ali manjši meri motivirajo ali spodbujajo k dejanju. *A* hoče, da je ta želja tudi učinkovita, da priskrbi motiv za to, kar *A* dejansko stori. Ko na ta način uporabimo trditev, da *A* hoče hoteti *X*, iz tega sledi, da *A že* ima željo po *X*. Ne more biti hkrati resnično, da *A* hoče, da ga želja po *X* spodbudi k dejanju in da hkrati noče *X*. Samo v primeru, ko hoče *X*, lahko koherentno hoče željo po *X* ne zgolj kot eno izmed svojih želja, pač pa – kar je bolj odločujoče – kot svojo voljo.[4]

Vzemimo, da nekdo hoče v svojem ravnanju biti motiviran z željo po

[4] Ni povsem jasno, ali tukaj opisana relacija sledenja drži za tiste vrste primerov, za katere sicer sam menim, da so neobičajni in v katerih je bistvena razlika med običajnimi in neobičajnimi primeri v vrsti opisa, s katerim je označena dana želja prvega reda. Torej predpostavimo, da *A* tako klečeplazno občuduje *B*, da četudi ne ve, kaj *B* želi početi, *A* hoče biti spodbujen k dejanju od katerekoli želje, ki k dejanju učinkovito navede *B*. Z drugimi besedami, ne da bi vedel, kakšna je B-jeva volja, *A* hoče, da je njegova enaka B-jevi. Vsekakor pa iz tega ne sledi, da *A* že ima med svojimi željami prav takšno, ki tvori B-jevo voljo. Na tem mestu se ne bom ukvarjal z vprašanjem, ali za trditev, vsebovano v besedilu, obstajajo pravi protiprimeri oziroma kako bi jo bilo treba spremeniti, če bi obstajali.

osredotočenju na svoje delo. Če je ta predpostavka pravilna, potem je nujno res, da se že hoče osredotočiti na svoje delo. Ta želja se sedaj nahaja med njegovimi željami. Toda vprašanje, ali je njegova želja drugega reda izpolnjena ali ne, ni odvisno samo od tega, ali je želja, ki jo hoče, ena od njegovih želja, temveč od tega, ali je ta želja, kot to hoče sam, tudi njegova učinkovita želja ali volja. Če je v trenutku, ko so vse stave položene, res njegova želja po osredotočenju na delo tista, ki ga spodbudi k temu, kar naredi, potem je to, kar ob tistem času hoče, zares (v relevantnem smislu) to, kar hoče hoteti. Če pa ga k dejanju spodbudi kakšna druga želja, potem to, kar hoče v tistem trenutku (v relevantnem smislu), ni tisto, kar hoče hoteti, in to kljub dejstvu, da želja po osredotočenju na delo še naprej ostaja med njegovimi željami.

II.

Željo drugega reda imamo lahko bodisi tedaj, ko preprosto hočemo imeti določeno željo, bodisi v primeru, ko hočemo, da določena želja postane naša volja. V drugem primeru bom želje drugega reda imenoval drugoredna hotenja ali hotenja drugega reda [*volitions of the second order*]. Posedovanje hotenj drugega reda in ne le posedovanje želja drugega reda je tisto, kar je zame bistveno za osebo. Četudi skrajno neverjetno, je vseeno logično možno, da obstaja dejavnik z željami drugega reda, ki pa nima hotenj drugega reda. Po mojem mnenju takšno bitje ne bi bilo oseba. Za dejavnike, ki imajo želje prvega reda, vendar pa niso osebe, ker nimajo hotenj drugega reda, ne glede na to, ali imajo ali nimajo želje drugega reda, bom uporabljal termin 'hotnež' [*wanton*].[5*]

Bistvena značilnost hotneža je, da mu ni mar za njegovo voljo. Njegove želje ga spodbujajo k ravnanju, ne da bi to hkrati pomenilo, da hoče biti vzpodbujen s temi željami ali da bi bil raje motiviran z drugimi željami. Razred hotnežev vključuje vse nehumanoidne živali in vse zelo majhne otroke. Morda vključuje tudi nekaj odraslih ljudi. V vsakem primeru so lahko tudi odrasli bolj ali manj hotneži. Bolj ali manj pogosto lahko ravnajo kot hotneži v svojih odzivih na tiste želje prvega reda, glede katerih nimajo hotenj drugega reda.

Dejstvo, da hotnež nima hotenj drugega reda, še ne pomeni, da je vsaka izmed njegovih želja prvega reda brez pomisleka in takoj prevedena v dejanje. Lahko se namreč zgodi, da nima priložnosti, da bi ravnal v skladu z nekaterimi od želja. Še več, prevod njegovih želja v dejanja je lahko preložen ali onemogočen zaradi nasprotujočih si želja prvega reda ali zaradi kakega razmisleka. Kajti tudi hotnež lahko poseduje in uporablja racionalne sposobnosti na visokem nivoju. Nič v pojmovanju hotneža le-temu ne odreka razuma ali presoje o tem, kako storiti to,

[5] Bitja z željami drugega reda, vendar brez hotenj drugega reda, se pomembno razlikujejo od živali brez razuma in bi jih za nekatere namene bilo smiselno obravnavati kot osebe. Moja raba, ki jim odreka oznako biti oseba, je zato do neke mere arbitrarna. V glavnem jo privzemam zato, ker poenostavi formulacijo nekaterih dognanj, ki jih želim poudariti. Torej, kadarkoli obravnavam trditve oblike "*A* hoče hoteti *X*", imam v mislih trditve, ki določajo hotenja drugega reda, ne pa trditve, ki določajo želje drugega reda, ki niso hotenja drugega reda.

[*] 'Hotnež', ki nastopa v prevodu, konotira pohotnost, ki v angleški besedi 'wanton' ni primarna. Vendar tak prevod ohranja koren glagola 'hoteti' in vsaj tako ustreza angleški dvojici 'want' in 'wanton'. – op. prev.

kar hoče. Racionalni hotnež se od drugih racionalnih dejavnikov loči po tem, da ga ne skrbi zaželenost njegovih želja samih. Vprašanje, kaj oz. kakšna naj bo njegova volja, preprosto ignorira. Ne le, da sledi tistemu toku dejanj, h kateremu je pač najmočneje nagnjen, temveč mu sploh ni mar, katero od njegovih nagnjenj je najmočnejše.

Racionalno bitje, ki reflektivno ocenjuje primernost svojih želja glede na ta ali oni tok dejanj, je kljub temu lahko hotnež. Ko trdim, da bistvo osebe ni v razumu, ampak v volji, sem daleč od tega, da bi trdil, da je bitje brez razuma lahko oseba. Kajti zgolj v okviru racionalnih zmožnosti se oseba lahko kritično zave svoje lastne volje in oblikovanja hotenj drugega reda. Skladno s tem šele struktura volje dane osebe predpostavlja, da gre za racionalno bitje.

Razloček med osebo in hotnežem lahko ponazorimo z razliko med dvema zasvojencema z mamili. Privzemimo, da so fiziološki pogoji, ki pojasnjujejo zasvojenost, pri obeh enaki in da se oba neogibno vdajata ponavljajoči se želji po mamilu, s katerim sta zasvojena. Eden od zasvojencev svojo zasvojenost sovraži in se vselej obupno bori proti navalu želje po drogi, čeprav brez uspeha. Poskusi vse, kar bi mu kakorkoli pomagalo premagati željo po drogi. Vendar pa je želja kljub temu premočna, da bi se ji lahko uprl, zato se nazadnje vda. To je primer upirajočega se zasvojenca, ki je nemočen plen svojih želja.

Upirajoči se zasvojenec ima nasprotujoči si želji prvega reda: mamilo hoče vzeti, hkrati pa se hoče droge vzdržati. Poleg teh dveh želja prvega reda pa ima tudi hotenje drugega reda, kajti v sporu med željo po mamilu in željo po vzdržanju ni nevtralen. Druga želja in ne prva je tista, za katero hoče, da bi tvorila njegovo voljo in za katero hoče, da bi bila učinkovita ter mu priskrbela namen, ki ga bo skušal udejaniti v svojem ravnanjem.

Drugi zasvojenec je hotnež. Njegova dejanja odražajo ekonomijo želja prvega reda, ne da bi ga pri tem skrbelo, ali so želje, ki ga spodbudijo k dejanjem, hkrati tudi želje, za katere hoče, da ga motivirajo v delovanju. Če naleti na težave pri preskrbi z mamilom ali pri jemanju samem, lahko njegovi odzivi na njegove potrebe vključujejo razumno presojo. Vendar pa se nikoli ne vpraša, ali hoče, da odnos med njegovimi željami vodi do tega, da ima voljo, ki jo ima. Hotni odvisnik je lahko tudi žival, in tako nezmožen biti zaskrbljen za svojo voljo. V vsakem primeru se, kar zadeva njegovo hotno pomanjkanje skrbi za svojo voljo, ne razlikuje od živali.

Drugi zasvojenec lahko trpi zaradi spora med željami prvega reda, ki je podoben tistemu sporu, ki ga preživlja upirajoči se zasvojenec. Ne glede na to, ali gre za človeka ali ne, lahko tudi hotnež (morda zaradi pogojevanja) hoče vzeti mamilo in hkrati hoče, da bi se vzdržal. Za razliko od upirajočega se zasvojenca pa hotni zasvojenec noče, da bi katera od želja prevladala nad drugimi. Nima preference po tem, da bi bila katera od nasprotujočih si želja prvega reda tvorila njegovo voljo. Zavajajoče bi bilo, če bi dejali, da je glede konflikta med svojimi željami nevtralen, ker bi to lahko pomenilo, da so zanj vse enako sprejemljive. Ker sam

nima nobene identitete onstran želja prvega reda, zanj ne drži niti to, da ima nekatere izmed želja raje kot druge, niti to, da je glede tega raje nepristran.

Upirajočemu se zasvojencu, ki je oseba, je pomembno, katera od njegovih nasprotujočih si želja prvega reda zmaga. Gotovo je, da sta obe želji njegovi; in četudi navsezadnje vzame mamilo ali pa se ga vzdrži, ravna tako, da bi zadovoljil tisto, kar je v dobesednem pomenu njegova lastna želja. V obeh primerih stori nekaj, kar hoče storiti in tega ne naredi zavoljo kakšnega zunanjega vpliva, katerega cilj je po naključju enak njegovemu, marveč zaradi lastne želje, da bi to storil. Upirajoči se zasvojenec se v oblikovanju hotenja drugega reda poistoveti z eno od nasprotujočih se želja prvega reda. Ena od želja postane bolj resnično njegova lastna, s tem pa se odmakne od druge. Zaradi tega poistovetenja in odmika, do katerega pride v oblikovanju hotenja drugega reda, lahko upirajoči se zasvojenec povsem pomenljivo zatrdi sicer analitično težko razumljive trditve, da sila, ki ga vodi k jemanju mamila, ni njegova lastna in da tega ne dela po svoji lastni volji, ampak proti svoji volji.

Hotnega odvisnika ne more skrbeti ali ga ne skrbi, katera od njegovih nasprotujočih si želja bo prevladala. Vendar razlog za njegovo pomanjkanje zaskrbljenosti ni v njegovi nezmožnosti, da bi našel prepričljivo osnovo za to, čemu dati prednost. Gre za pomanjkanje sposobnosti za refleksijo ali pa za lahkomiselno ravnodušnost do tega, da bi ovrednotil lastne želje in motive.[6] Boj med njegovimi željami prvega reda se lahko razreši samo tako, da je ena ali druga želja močnejša. Ker ga spodbujata obe želji, ne bo čisto zadovoljen s svojim ravnanjem, ne glede na to, katera od obeh je učinkovita. Toda *zanj* ni pomembno, ali bo prevladalo poželenje ali vzdržnost. V spor med njima ni vložil ničesar, zato, v nasprotju z upirajočim se zasvojencem, v boju, v katerega je vpleten, ne more niti zmagati niti izgubiti. Kadar deluje *oseba*, je želja, ki dejanje motivira, bodisi volja, ki jo oseba hoče, ali pa volja, ki je noče. Ko pa deluje *hotnež*, ni ne eno ne drugo.

III.

Obstaja tesna zveza med zmožnostjo oblikovanja hotenj drugega reda in še eno zmožnostjo, ki je za osebe bistvena in ki se velikokrat pojmuje kot tisto, kar je posebna razlikovalna značilnost človekovega položaja v svetu. Samo zato, ker ima hotenja drugega reda, oseba lahko uživa svobodo volje ali pa je lahko takšna, da svobode volje nima. Pojem osebe torej ni samo pojem tipa bitnosti, ki ima tako želje prvega reda kot hotenja drugega reda. Lahko ga razložimo tudi kot pojem tipa bitnosti, za katero je svoboda njene volje lahko problem. Tak pojem izključuje vse hotneže, tako pripadnike nižjih vrst kot ljudi, saj ne zadostijo

[6] Ko govorim o vrednotenju lastnih motivov in želja kot značilnosti oseb, ne želim reči, da hotenja drugega reda dane osebe nujno manifestirajo njeno *moralno* držo do želja prvega reda. Ni nujno, da oseba vrednoti svoje želje prvega reda z vidika moralnosti. Še več, oseba je lahko pri oblikovanju svojih hotenj drugega reda muhasta in neodgovorna, lahko da njihove pomembnosti sploh ne jemlje resno. Hotenja drugega reda izražajo vrednotenja samo zato, ker izražajo preference. Ni pa omejitev glede načina in osnove, če sploh kakšna je, na kateri so utemeljena.

bistvenega pogoja za uživanje svobode volje. Izključuje pa tudi vsa tista nadčloveška bitja, če taka so, katerih volje so nujno svobodne.

Torej, kakšne vrste svoboda je svoboda volje? To vprašanje zahteva prepoznanje posebnega področja človeškega izkustva, za katerega je pojem svobodne volje, za razliko od pojmov drugih vrst svobode, še zlasti značilen. Pri ukvarjanju s tem pojmom bo moj cilj primarno določitev problema, s katerim se oseba najbolj neposredno ukvarja, ko se ukvarja s svobodo svoje volje.

V skladu z znano filozofsko tradicijo smo v osnovi svobodni takrat, ko ravnamo tako, kot hočemo. No, pojem dejavnika, ki stori to, kar hoče, je vse prej kot jasen: tako ravnanje kot hotenje pa tudi ustrezna zveza med njima zahtevata natančnejšo pojasnitev. Toda čeprav je treba fokus tega pojmovanja še izostriti in njegovo formulacijo izboljšati, sem vendarle prepričan, da to pojmovanje res zajame vsaj del tistega, kar je implicirano v ideji dejavnika, ki *ravna* svobodno. Vendar pa ta pojem popolnoma zgreši posebno vsebino precej drugačne ideje dejavnika, katerega *volja* je svobodna.

Za živali ne predpostavljamo, da uživajo svobodo volje, čeprav priznavamo, da lahko žival svobodno steče, kamorkoli hoče. Zatorej imeti svobodo, po kateri storimo to, kar hočemo, ni zadosten pogoj za posedovanje svobodne volje. Ni pa niti nujni pogoj. Če koga prikrajšamo za svobodo dejanj, to še ne pomeni, da smo izničili svobodo njegove volje. Če se dejavnik zaveda, da so nekatera dejanja takšna, da v njih ni svoboden, to nedvomno vpliva na njegove želje in omejuje obseg izbir, za katere se lahko odloči. Toda privzemimo, da je nekdo resnično izgubil svobodo dejanj ali bil prikrajšan zanjo, ne da bi se tega zavedal. Čeprav ne more več svobodno storiti tega, kar hoče, lahko njegova volja ostane enako svobodna, kot je bila prej. Navzlic dejstvu, da ne more svobodno prevesti svojih želja v dejanja ali pa ravnati v skladu z odločitvami svoje volje, lahko še vedno oblikuje te želje in sprejema odločitve tako svobodno, kot da njegova svoboda ravnanja ne bi bila okrnjena.

Ko se namreč vprašamo, ali je volja neke osebe svobodna, se ne sprašujemo, ali ima možnost pretvoriti svoje želje prvega reda v dejanja. To je vprašanje, ali lahko ravna, tako kot se mu zljubi. Vprašanje volje dane osebe ne zadeva odnosa med tem, kar stori in kar hoče storiti. Gre za želje same. Toda za kakšno vprašanje o željah tu gre?

Naravno in obenem koristno se mi zdi obravnavati vprašanje, ali je volja dane osebe svobodna, po analogiji z vprašanjem, ali dejavnik uživa svobodo dejanj. Svoboda dejanj je (vsaj v grobem) svoboda, da storiš to, kar hočeš. Torej po analogiji trditev, da oseba uživa svobodo volje, pomeni (prav tako v grobem), da oseba svobodno hoče to, kar hoče hoteti. Natančneje, to pomeni, da je svoboden hoteti to, kar hoče hoteti, ali da ima voljo, ki jo hoče. Tako kot je vprašanje svobode dejanja vprašanje o tem, ali gre za dejanje, ki ga dejavnik hoče izvesti, tako je vprašanje svobode njegove volje vprašanje o tem, ali gre za voljo, ki jo hoče imeti.

V tem, da uskladi svojo voljo s svojimi hotenji drugega reda, oseba izpriča svobodo volje. V neskladnosti med voljo in hotenji drugega reda, ali pa v zavedanju, da njun sklad ni plod njene lastne volje, ampak zgolj srečno naključje, oseba, ki svobode volje nima, občuti njeno pomanjkanje. Volja upirajočega se zasvojenca ni svobodna. To smo pokazali z dejstvom, da ne gre za voljo, ki jo hoče. Prav tako je res, čeprav na drugačen način, da volja hotnega zasvojenca ni svobodna. Odvisnik-hotnež nima niti volje, ki jo hoče, niti volje, ki se razlikuje od volje, ki jo hoče. Ker sploh nima hotenj drugega reda, svoboda volje zanj ne more biti problem. Že v samem izhodišču ni deležen svobode volje.

Ljudje so v splošnem veliko bolj zapleteni, kot bi lahko sklepali iz mojega približnega orisa zgradbe volje oseb. Glede želja drugega reda obstaja prav toliko možnosti za ambivalentnost, konflikt in samoprevare kot glede želja prvega reda. Kadar pri komu obstaja nerazrešen konflikt med željami drugega reda, potem obstaja nevarnost, da bo ostal brez hotenj drugega reda. Če konflikta ne razreši, tudi nima nobenih preferenc glede tega, katera od njegovih želja prvega reda naj bo njegova volja. Če je konflikt tako resen, da mu preprečuje, da bi se v zadostni meri poistovetil s *katerokoli* od nasprotujočih si želja prvega reda, ga bo to uničilo kot osebo. Konflikt bo paraliziral njegovo voljo in mu preprečil kakršnokoli delovanje ali pa ga bo ločil od njegove volje, tako da bo volja delovala brez njegove udeležbe. V obeh primerih postane, podobno kot upirajoči se zasvojenec, čeprav na drugačen način, zgolj nemočen opazovalec sil, ki ga ženejo.

Dodatne zaplete predstavlja dejstvo, da ima lahko oseba – še posebej če so njene želje drugega reda v konfliktu – želje in hotenja, ki so višjega reda od drugega. Teoretske meje za dolžino zaporedij želja vedno višjih redov ni; nič drugega kot zdravi razumi ter morda odrešujoče utrujenosti ne preprečuje posamezniku, da ne bi obsedeno zavračal poistovetenja s katerokoli od lastnih želja, dokler ne oblikuje želje naslednjega višjega reda.

Težnja po tvorbi takšnega zaporedja aktov oblikovanja želja, kar bi pomenilo "podivjanje" humanizacije, prav tako vodi do izničenja osebe.

Takšno zaporedje aktov pa je mogoče prekiniti, ne da bi ga samovoljno presekali. Ko se oseba *odločilno* poistoveti z eno od želja prvega reda, ta zavezanost "odzvanja" skozi potencialno neskončno vrsto višjih redov. Zamislimo si osebo, ki brez zadržkov ali konfliktov hoče, da jo motivira želja po osredotočenju na svoje delo. Dejstvo, da je njeno hotenje drugega reda, tj. hotenje biti spodbujen s strani želje po osredotočenju na delo, odločilno, pomeni, da v tem primeru ni prostora za vprašanja, ki zadevajo primernost želja ali hotenj višjih redov. Denimo, da osebo vprašamo, ali hoče hoteti, da je osredotočena na svoje delo. Povsem umestno lahko vztraja, da to vprašanje, povezano z željo tretjega reda, sploh ne nastopi. Napak bi bilo trditi, da je ta oseba zaradi tega, ker ni pomislila, ali hoče imeti hotenje drugega reda, ki ga je oblikovala, ravnodušna do vprašanja, ali hoče, da je njena volja skladna s tem ali kakšnim drugim hotenjem drugega reda. Odločilnost njene zavezanosti pomeni, da se je odločila, da ni več nobenih

vprašanj o njenih hotenjih drugega reda, ki bi jih postavljala s kakšnega višjega reda. Pri tem je relativno nepomembno, ali to pojasnimo tako, da rečemo, da takšna zavezanost bodisi generira neskončno zaporedje potrjujočih želja višjega reda, ali pa rečemo, da je ta zavezanost nekaj takega kot izničenje smiselnosti vseh nadaljnjih vprašanj, ki zadevajo višje rede želja.

Primeri, kot ga ponazarja upirajoči se zasvojenec, nas lahko navedejo na misel, da morajo biti hotenja drugega ali višjih redov oblikovana namerno in da je zanje značilen trud oseb, da bi jih zadovoljile. Vendar pa je skladnost volje dane osebe z njenimi hotenji višjega reda lahko veliko bolj samodejna in spontana. Nekatere ljudi že po naravi motivira prijaznost, kadar želijo biti prijazni, in neprijaznost, kadar želijo biti neprijazni, brez kakršnegakoli vnaprejšnjega razmisleka in brez potrebe po energičnem samonadzoru. Druge motivira neprijaznost, kadar hočejo biti prijazni, in prijaznost, kadar nameravajo biti neprijazni, prav tako brez vnaprejšnjega razmisleka ali aktivnega odpora proti tem kršitvam njihovih želja višjega reda. Nekaterim je uživanje svobode dano z lahkoto, drugi pa se morajo naprezati, da bi jo dosegli.

IV.

Moja teorija o svobodi volje enostavno pojasni, zakaj nismo pripravljeni dopustiti, da bi bila takšna svoboda dostopna pripadnikom katerekoli nižje vrste od naše. Zadosti pa tudi še enemu pogoju, ki ga mora izpolniti vsaka podobna teorija, saj je po njej takoj razvidno, zakaj je svoboda volje tako zaželena. Uživati svobodno voljo pomeni zadovoljitev določenih želja – želja drugega ali višjih redov, nesvoboda volje pa pomeni frustracijo teh želja. Zadovoljitve želja, za katere gre, so tiste, ki sodijo k osebi, za katero lahko rečemo, da je njena volja njena lastna volja. V primeru neizpolnitve teh želja pa gre za frustracije osebe, za katero lahko rečemo, da je sama sebi odtujena ali pa je popolnoma pasivna in nemočna opazovalka sil, ki jo ženejo.

Za osebo, ki svobodno stori, kar stori, še ni nujno res, da ima voljo, ki jo hoče. Denimo pa, da uživa tako svobodo dejanj kot svobodo volje. Tedaj ni samo svobodna storiti to, kar hoče, ampak tudi svobodno hoče, kar hoče hoteti. Zdi se mi, da ima v tem primeru vso svobodo, ki si jo je mogoče želeti ali zamisliti. Morda obstajajo druge dobre stvari v življenju, ki jih takšna oseba nima. Toda glede svobode ji ne manjka prav ničesar.

Vse prej kot jasno je, ali tudi druge teorije svobode izpolnjujejo te temeljne, vendar bistvene pogoje: da je razumljivo, zakaj si to svobodo želimo in zakaj je nočemo pripisati živalim. Oglejmo si, denimo, Chisholmovo nenavadno verzijo doktrine, po kateri iz človeške svobode izhaja odsotnost vzročne določenosti.[7] Kadarkoli oseba izvede svobodno dejanje, je to za Chisholma čudež. Kadar oseba premakne roko, je sam gib zgolj rezultat zaporedja fizičnih vzrokov; toda eden

[7] 'Freedom and Action', v *Freedom and Determinism*, Keith Lehrer ur., (New York: Random House, 1966), 11-44.

izmed dogodkov v tem zaporedju, "in domnevno eden tistih, ki so se zgodili v možganih, je bil povzročen s strani dejavnika, ne pa z drugimi dogodki" (18). Svobodni dejavnik ima zatorej "posebne ugodnosti, ki jih nekateri pripisujejo le Bogu: kadar ravnamo, je vsak od nas v vlogi prvega gibalca, ki se sam ne giblje" (23).

Po tem pojasnilu pa ni razlogov za dvom v to, da so tudi nečloveške živali deležne svobode, ki jo takšna definicija opredeljuje. Chisholm ne trdi nič takega, zaradi česar bi se zdelo manj verjetno, da tudi zajec napravi čudež, ko premakne nogo, podobno, kot je to pri človekovem premiku roke. Toda, navsezadnje, zakaj bi kogarkoli *skrbelo*, ali lahko prekine naravni red vzrokov na način, kot ga Chisholm opisuje? Chisholm ne navede nobenega razloga za prepričanje, da obstaja kakršnakoli opazna razlika med izkustvom človeka, ki s premikom svoje roke čudežno sproži zaporedje vzrokov, in človekom, ki svojo roko premakne brez vsakršne prekinitve normalnega vzročnega zaporedja. Zdi se, da ni nobene konkretne osnove, zaradi katere bi bilo bolj zaželeno biti vpleten v eno od obeh stanj stvari in ne v drugo.[8]

V splošnem se predpostavlja, da, poleg tega, da izpolni dva pogoja, ki sem ju navedel, zadovoljiva teorija svobodne volje nujno oskrbi tudi analizo enega od pogojev moralne odgovornosti. Najobičajnejši novejši pristop k problemu razumevanja svobodne volje je bil zares ta, da raziščemo posledice predpostavke, da je oseba moralno odgovorna za svoja dejanja. Toda po mojem mnenju so odnos med moralno odgovornostjo in svobodo volje v glavnem razumeli napačno. Ni res, da je oseba moralno odgovorna za tisto, kar je storila, samo, če je bila njena volja svobodna, ko je to storila. Oseba je lahko moralno odgovorna za to, kar je storila, čeprav njena volja sploh ni bila svobodna.

Volja osebe je svobodna samo, če je svobodna imeti voljo, ki jo hoče. To pomeni, da je glede na katerokoli njeno željo prvega reda oseba svobodna, da naredi to željo za svojo voljo ali pa naredi kako drugo željo prvega reda za svojo voljo. Kakršnakoli že je njena volja, volja osebe, ki ima svobodno voljo, bi lahko bila drugačna. Lahko bi ravnala drugače, kot pa da je osnovala svojo voljo tako, kot jo je. Kako natanko razumeti stavek "lahko bi storila drugače" v teh kontekstih, je sporno vprašanje. A četudi je to vprašanje pomembno za teorijo svobode, pa nima vpliva na teorijo moralne odgovornosti. Kajti iz predpostavke, da je oseba moralno odgovorna za to, kar stori, ne sledi, da je oseba bila v položaju, v katerem je imela voljo, kakršno je hotela imeti.

Iz te predpostavke (o moralni odgovornosti) *zares* sledi, da je oseba svobodno storila, kar je storila, ali da je ravnala po svoji svobodni volji. Napak pa bi bilo verjeti, da nekdo ravna svobodno samo, če lahko svobodno stori, karkoli hoče, ali da ravna po svobodni volji samo, če je njegova volja svobodna. Denimo, da je oseba storila, kar je hotela, da je to storila zato, ker je tako hotela in da je bila

[8] Ne želim navajati na misel, da je dozdevna razlika med obema stanjema stvari nepreverljiva. Nasprotno, fiziologi bi lahko pokazali, da Chisholmovi pogoji za svobodno dejanje niso izpolnjeni, če bi dokazali, da ni ustreznega dogodka v možganih, za katerega ne bi mogli najti zadostnega fizičnega vzroka.

volja, ki jo je motivirala, njena volja, saj je bila to volja, ki jo je oseba hotela. Tedaj je oseba to dejanja izvedla svobodno in po svoji lastni svobodni volji. Tudi ob domnevi, da bi lahko ravnala drugače, ne bi ravnala drugače. In celo ob predpostavki, da bi lahko bila imela drugačno voljo, ne bi hotela, da bi se njena volja razlikovala od volje, kakršno je imela. Še več, ker je bila volja, ki jo je spodbudila k dejanju, njena volja, saj je tako tudi hotela, ne more trditi, da ji je bila ta volja vsiljena ali da je bila ta oseba pri njenem oblikovanju zgolj pasivna opazovalka. Pod temi pogoji je za oceno njene moralne odgovornosti precej nepomembno, ali so ji bile alternative, proti katerim se je odločala, dejansko na voljo.[9]

V ponazoritev si oglejmo še tretjo vrsto odvisnika. Denimo, da ima njegova odvisnost enako fiziološko osnovo in neubranljivo moč kot odvisnosti upirajočega se in hotnega odvisnika, vendar je ta odvisnik na splošno izjemno zadovoljen s svojim stanjem. Gre za prostovoljnega zasvojenca, ki ne bi hotel, da bi bile stvari kakorkoli drugačne. Če bi začel primež njegove odvisnosti kakorkoli popuščati, bi storil vse, da bi obnovil njegovo moč. Če bi njegova želja po mamilu začela popuščati, bi naredil vse, da obnovi njeno intenzivnost.

Volja prostovoljnega zasvojenca ni svobodna, saj bi njegova želja po mamilu učinkovala ne glede na to, ali on hoče, da ta želja tvori njegovo voljo, ali ne. Toda kadar vzame mamilo, to stori svobodno in po svoji lastni svobodni volji. Zato se nagibljem k temu, da bi to situacijo razumel kot naddoločenost njegove želje prvega reda po jemanju mamila. Ta želja deluje zaradi njegove fiziološke odvisnosti. Učinkovita pa je tudi zato, ker on tako hoče. Njegova volja ni pod njegovim nadzorom, toda zaradi njegovega hotenja drugega reda, da naj želja po mamilu deluje, je to njegova lastna volja. Če torej izhajamo iz dejstva, da njegova želja po mamilu ni učinkovita samo zaradi njegove odvisnosti, je lahko tak zasvojenec moralno odgovoren za jemanje mamila.

Za moje pojmovanje svobode volje se zdi, da je nevtralno glede problema determinizma.

Lahko si zamislimo, da bi bilo vzročno določeno, da je oseba svobodna, da hoče, kar hoče hoteti. Če si to lahko zamislimo, potem je lahko vzročno determinirano, da oseba uživa svobodo volje. V tem, da je vzročno določeno, neizbežno in s silami, nad katerimi nimajo nadzora, da nekateri ljudje imajo svobodno voljo, drugi pa ne, gre samo za neškodljivi videz paradoksa. V tezi, da je za dejstvo, da dana oseba poseduje ali pa ne poseduje svobode volje, odgovoren (*celo moralno* odgovoren) kak drug dejavnik in ne ta oseba sama, ni nobene nekoherence. Mogoče je tudi, da je kakšna oseba moralno odgovorna za tisto, kar stori po svoji svobodni volji, vendar pa je tudi neka druga oseba moralno odgovorna za to, da je prva oseba to storila.[10]

[9] Za dodatno razpravo o pomislekih, ki izražajo dvom v načelo, da je oseba za svoja dejanja moralno odgovorna le tedaj, če bi lahko v dani situaciji ravnala drugače, glej moj članek 'Alternate Possibilities and Moral Responsibility', *Journal of Philosophy*, 1969, 829-839.

[10] Obstaja razlika med biti *popolnoma* in biti *izključno* odgovoren. Recimo, da je prostovoljni zasvojenec postal odvisnik zaradi preračunljivega ravnanja druge osebe. Tedaj lahko drži, da sta oba, odvisnik in ta druga oseba

Prav tako pa je pojmljivo, da je po naključju prišlo do tega, da je oseba svobodna v tem, da ima voljo, ki jo hoče. In če je to mogoče, potem bi lahko bila stvar naključja tudi to, da nekateri ljudje uživajo svobodo volje, drugi pa ne. Morda je prav tako pojmljivo, kot je prepričano nekaj filozofov, da stanja stvari nastanejo na kak drug način in ne po naključju ali po nekem zaporedju naravnih vzrokov. Če je zares pojmljivo, da ustrezna stanja stvari nastanejo na kak tretji način, potem je prav tako možno, da oseba na tak tretji način pride do tega, da uživa svobodo volje.

popolnoma odgovorna za odvisnikovo jemanje mamila, čeprav nobeden od njiju ni izključno odgovoren. Razlika med popolno moralno odgovornostjo in izključno moralno odgovornostjo je razvidna v naslednjem primeru. Neka luč se lahko prižge s pritiskom na katerokoli od dveh stikal, ki ju istočasno pritisneta dve osebi, vendar se druga druge ne zavedata. Nobena od njiju ni izključno odgovorna za prižig luči, niti si odgovornosti ne delita v smislu, da je vsaka od njiju delno odgovorna, temveč je vsaka od njiju popolnoma odgovorna.