

Gyakorlati racionalitás és akaratgyengesség

I. A RACIONÁLIS ÁLLAT

Egy mindenki által jól ismert – bár, mint látni fogjuk, sok szempontból vitatható – meghatározás szerint az ember racionális állat. A hagyományos meghatározás tartalma azonban többféleképpen is értelmezhető. A *genus* utalhat arra a tényre, hogy az ember biológiai organizmus. Miután azonban nem minden organizmus állat, a *genus proximumot* értelmezhetjük úgy is, hogy az ember olyan élőlény, amely rendelkezik viselkedés jellegű attribútumokkal, a *differentia specificát* pedig úgy, hogy viselkedése olyan mentális állapotok és folyamatok következménye, amelyeket racionálisként jellemezhetünk. Csakhogy az embert nem úgy jellemzi a racionalitás, ahogyan mondjuk az, hogy testét nem fedi toll, vagy pedig az, hogy pontosan két lába van.

Először is a racionalitás, akárcsak az, hogy az ember természeténél fogva két lábon közlekedik, egy képességre, nem pedig egy folytonosan manifesztálódó tulajdonságra utal. De túl ezen, van egy másik fontos eltérés is a racionalitás és a korábban említett tulajdonságok közt. A racionalitás, szemben a két lábon járással, *normatív* fogalom, éppen ebből adódik a hagyományos meghatározás nehézsége. A racionalitás képességének valamely elemi mértéke elengedhetetlen ahhoz, hogy valaki viselkedését mint specifikusan emberi viselkedést értelmezzük, de a racionalitás teljes hiánya sem zárja ki, hogy valaki ember legyen. Akár helyes azonban a meghatározás, akár nem, kétségtelen, hogy az ember racionális lény annyiban, amennyiben a biológiai organizmusok közül – egy különös értelemben – csak az emberre alkalmazzuk a racionalitás normáját. Érdekes módon azt, hogy milyen értelemben, nem feltétlenül a racionalitás teljes vagy részleges hiánya magyarázza. Sok esetben ugyanis más organizmusokról is állítjuk, hogy „okosan viselkedtek”, ami többé-kevésbé szinonim azzal, hogy racionálisak. És amikor megállapítjuk, hogy a kutyák vagy majmok, „bármilyen okosak is”, nem tudnak szorozni, akkor racionalitásuk korlátaira hívjuk fel a figyelmet.

A kutyákról, majmokról vagy más organizmusokról azonban bizonyosan nem állítanánk, hogy *irracionalisan* viselkedtek. Az irracionalitás ugyanis nem egyszerűen egy képesség hiánya vagy korlátja, hanem a racionalitás képességének sajátos diszfunkciója, azaz természetes céljától eltérő használata. Ezért annak

vizsgálata, hogy milyen körülmények között állíthatjuk valakiről, hogy irracionális, bizonyos értelemben jobban megvilágítja a specifikusan emberi értelemben vett racionalitás természetét, mint az, hogy milyen körülmények között állíthatjuk valakiről (valamiről), hogy racionális. Racionalitást tulajdonítani viszonylag könnyű – véleményem szerint túlságosan is könnyű, amint azt az „intencionális rendszerek” alkalmazásának minden határt túlszárnyaló kiterjesztése bizonyítja –, az irracionalitást megérteni nehéz. És éppen azért olyan nehéz, mert az irracionális viselkedés megértése segít leginkább a specifikusan emberi racionalitás természetének vizsgálatában.

A „racionalitás” kifejezést elsődleges értelmében hitekre, vélekedésekre vagy meggyőződésekre alkalmazzuk. Egy vélekedés akkor racionális, ha konzisztens a vélekedést valló egyén egyéb hiteivel, illetve a rendelkezésére álló evidenciákkal. Az ismeretelmélet egy fontos kérdése, hogy az evidenciák maguk hitek-e, vagy pedig olyan tények, amelyek a bizonyíték jellegű meggyőzések tartalmát adják. A racionalitás szempontjából azonban az evidenciák természeténél fontosabb az evidenciák és az evidenciákon alapuló meggyőzések viszonyának kérdése. Egy meggyőződés tartalma konzisztens lehet számos olyan vélekedéssel, amely se nem támasztja alá, se nem cáfolja az egyén meggyőződését. Az evidenciáknak azonban sajátos viszonyban kell állniuk vélekedéseinkkel. Tartalmuk szerint, és ha racionálisak vagyunk, megerősíthetik (konfirmálhatják), vagy épp ellenkezőleg, gyengíthetik (diszkonfirmálhatják) meggyőződéseinket. A sajátosan emberi racionalitás egyik fontos filozófiai problémája ezért éppen az, hogy miképpen használja valaki a számára rendelkezésre álló evidenciákat.¹

Fontos megjegyezni, hogy az ily módon értett racionalitás *nem* egyfajta tendencia arra, hogy igaz meggyőzésekre tegyünk szert. Akik a racionalitásra pusztán mint az események magyarázatának vagy előrejelzésének eszközére hivatkoznak, nemcsak azt feltételezik, hogy az általuk magyarázni kívánt viselkedés racionális, hanem azt is, hogy a viselkedés szempontjából releváns meggyőzések egyben igazak is (Dennett 1998. 265–296). Az emberi racionalitás *képességének* gyakorlása azonban ilyesmit nem követel meg. Miután az, hogy milyen evidenciák állnak rendelkezésünkre, esetleges, az is esetleges, hogy racionális meggyőzések igazak-e. Vannak persze filozófusok, akik úgy vélik, hogy csak olyasvalami lehet evidencia, ami igaz meggyőzést eredményezhet.² Véleményem szerint azonban ez lehetetlenné tenné annak megértését, hogyan lehetnek valaki episztemikus értelemben racionális meggyőzései hamisak.

Márpedig ilyen meggyőzések nyilvánvalóan lehetségesek, amint azt például a manipuláció jelensége bizonyítja. A manipuláció nem egyszerű tévedés, inkább az érzékcsalódáshoz hasonlítható, amennyiben manipulálni csak racionális egyént lehet, mégpedig *éppen azon az alapon*, hogy racionális. A manipu-

¹ Erről lásd Davidson 2004, 13. fejezet.

² Mi több, az evidencia maga mindig tudás. Vö. például Williamson 2000, 9. fejezet.

látó nem egyszerűen becsapja a másikat azzal, hogy valótlan-ságot állít. A manipuláció az adott esetben igaz evidenciák olyan szelektív elrendezése, amely a potenciálisan megerősítő és cáfoló bizonyítékok közül csak az egyik csoportba tartozó evidenciák elérhetőségét teszi lehetővé a manipulált egyén számára. A manipuláció azonban feltételezi, hogy a manipulált egyén racionális: képes a rendelkezésre álló evidenciák alapján konzisztens vélekedéseket kialakítani. Éppen ezért a kognitív szempontból irracionális személynek rendelkeznie kell a racionalitás képességével. Az úgynevezett „önbecsapás” jelenségének leghatékonyabb módja éppen az, hogy valaki „szándékosan” manipulálja az önmaga számára rendelkezésre álló evidenciákat. De csak akkor „csaphatja be” ilyen módon önmagát, ha egyben arra is képes, hogy a rendelkezésre álló evidenciák alapján racionális meggyőződéseket alakítson ki.

Hogy a fentebb említett pszichológiai folyamatot helyes-e „önbecsapásnak” nevezni, és ha igen, milyen konkrét pszichológiai mechanizmusok révén valósulhat meg, az nem a jelen tanulmány tárgya. Mindössze annyit állítok, hogy az ilyen típusú pszichológiai folyamatok nyilvánvalóan irracionálisak, és hogy ennek szükséges, de nem elégséges feltétele, hogy az egyén inkonzisztens meggyőződésekkel rendelkezzen, ugyanakkor képes legyen arra, hogy a rendelkezésre álló evidenciák alapján racionális meggyőződéseket alakítson ki. Bizonyosan mindnyájan számos inkonzisztens vélekedéssel rendelkezünk, ez azonban nem jelenti azt, hogy kognitív szempontból irracionálisak is vagyunk, mivel az irracionálisnak az is lényegi feltétele, hogy az inkonzisztens meggyőződések egy olyan mechanizmussal történő „visszaélés” termékei legyenek, amely általában racionális meggyőződésekhez vezet. Éppen ezért segíthet bennünket a *sajátosan emberi racionalitás* megértésében az irracionális tanulmányozása.

II. GYAKORLATI RACIONALITÁS ÉS KONZISZTENCIA

A következőkben tehát azt próbálom megérteni, mit jelent pontosan a csakis az emberre jellemző *gyakorlati* racionalitás. Amellett fogok érvelni, hogy ezt a gyakorlati irracionális vizsgálatán keresztül érthetjük meg legjobban. A gyakorlati irracionális az irracionális azon formája, amely bizonyos körülmények között *a cselekvésben*, vagy épp ellenkezőleg, a cselekvéstől való tartózkodásban nyilvánul meg. Állításom az, hogy a cselekedetek irracionálisának megértésén keresztül vezet az út a racionális cselekvés feltételeinek megértéséhez, hasonló módon ahhoz, ahogyan a vélekedések irracionálisának megértése segíthet hozzá bennünket a kognitív racionalitás jobb megértéséhez. Az is nyilvánvalónak tűnik, hogy a cselekedetek irracionálisága is kapcsolódik valamilyen módon a konzisztencia fogalmához. Ám konzisztencia és racionalitás kapcsolata bonyolultabb és nehezebben értelmezhető a cselekedetek, mint a meggyőződések esetében.

Először is, mint fentebb említettem, a vélekedések pusztán inkonzisztenciája nem elégséges ahhoz, hogy kognitív irracionálisról beszéljünk. Ez utóbbihoz arra is szükség van, hogy azonosítsunk egy olyan mechanizmust, amely *általában* racionális meggyőződésekhez vezet, és amellyel az irracionálisan gondolkodó egyén valamilyen módon „visszaél”. Az azonban világos, mit jelent a konzisztencia a vélekedésekre vonatkoztatva: két vélekedés akkor inkonzisztens, ha tartalmuk egymással logikailag összeegyeztethetetlen. Akkor rendelkezik tehát valaki inkonzisztens vélekedésekkel, ha arról is meg van győződve, hogy egy proposíció hamis, és arról is, hogy igaz.³ Az ellenben sokkal bonyolultabb kérdés, hogy mit jelent az inkonzisztencia a viselkedés tekintetében. Ugyanis nyilvánvaló módon képtelenség, hogy valaki végrehajtsa egy cselekedetet, amit ugyanakkor nem is hajt végre.

A gyakorlati irracionális megértését tovább nehezíti, hogy a kognitív irracionális leggyakoribb formáinak is vannak viselkedései megnyilvánulásai. Előfordulhat, hogy valaki úgy ítéli, épp itt az ideje, hogy megszabaduljon sok gondot okozó járművétől. Mindezek ellenére szemmel láthatóan őszintén – nem pedig a félrevezetés egyértelmű szándékával – azt hajtogatja, hogy az autó a lehető legjobb állapotban van, és egyébként is vérzik a szíve, hogy meg kell válnia tőle. Az ilyen viselkedés jó magyarázatának tűnik, hogy az illető önmagát is félrevezeti saját meggyőződését illetően. Tehát – szemben a tudatos hazudozóval – bizonyos értelemben irracionálisan viselkedik. Viselkedését mégsem a *gyakorlati* irracionális megnyilvánulásának tekintenénk. Pedig az irracionális gyakorlati formája sem állhat pusztán valamely cselekedet végrehajtásában vagy valamely cselekedettől való tartózkodásban. Bármely cselekedet csak a hozzá kapcsolódó mentális folyamatok kontextusában lehet irracionális. Ám a cselekvéshez vezető mentális folyamat irracionális önmagában még nem gyakorlati irracionális. Ez a belátás alapvető jelentőségű annak megértésében, hogy mit is jelent valójában racionálisan cselekedni.

A racionális cselekvés fogalma közvetlenül kapcsolódik a *racionális kontroll* kérdéséhez, vagyis ahhoz, hogy miképpen kell értenünk annak a folyamatnak a természetét, *magának a cselekvőnek a szempontjából*, amely egy meghatározott viselkedéshez vezet. Mint látni fogjuk, a viselkedésnek nem csupán a cselekvő mentális állapotai és folyamatai kölcsönöznek racionalitást. A cselekvés racionalitásának megértése szempontjából azonban elengedhetetlen a mentális állapotok és folyamatok szerkezetének helyes megértése. Ezért első megközelítésben talán a cselekvés esetében is érdemes az irracionális az inkonzisztenciával

³ Fontos szem előtt tartani, hogy ebből természetesen nem következik az a meggyőződés, hogy ugyanaz a proposíció igaz is és hamis is. A kognitív irracionális természetére vonatkozó egyik alapkérdés éppen az, miért nem képes valaki meggyőződéseit megfelelő módon integrálni. Vö. Davidson 2004, 13. fejezet.

kapcsolatba állítanunk. Az első megközelítésnek azonban két nehézséggel is szembe kell néznie.

Először is, mint láttuk, a konzisztenciának csak akkor van szerepe a kognitív racionalitás megértésében, ha az valamilyen módon az evidenciákhoz való viszonyban nyilvánul meg. A gyakorlati racionalitás elsődleges szerepe azonban nem az, hogy a racionális képesség használata révén az egyén bizonyos mentális tartalmakra tegyen szert, hanem az, hogy e képessége használatával *kontrollálja*, hogy mit tesz. Ám a meggyőződés és a viselkedés közti kapcsolat sokkal komplexebb, mint az evidenciák és az evidenciákon alapuló vélekedések közti viszony. Ennek oka az, hogy a kognitív inkonzisztencia esetében a vélekedések *tartalma* egyértelműen meghatározza, konzisztens meggyőződésekkel rendelkezik-e valaki vagy sem. Fenti példánkat alapul véve világos, hogy az eladni kívánt jármű értékéről vallott meggyőződés és a releváns viselkedést meghatározó vélekedés tartalma nem konzisztens. Miután azonban adott helyzetben az egyén csak egy bizonyos módon viselkedhet, ennek alapján nem állíthatjuk, hogy maga a cselekedete „inkonzisztens”.

Másodszor, a cselekedeteket meghatározó mentális állapotok esetében sem egyszerű kérdés, hogy mit tekinthetünk inkonzisztenciának, mivel a cselekvés magyarázata általában nem csupán egy típusú mentális állapotot, vélekedést vagy meggyőződést feltételez. A racionális meggyőződés nem feltétlenül motivált, a cselekvés viszont az kell hogy legyen. Másképp fogalmazva, részben a cselekvő céljai vagy preferenciái magyarázzák, amit tesz. Ámde a preferenciák nem meggyőződések, és ezért racionalitásfeltételeik sem szükségképpen esnek egybe a meggyőződésekével. Ha valaki úgy véli, hogy egy adott pillanatban gépjárműve már sok gondot okoz neki, és ugyanakkor azt is gondolja, hogy rendkívül jó állapotban van, akkor valószínűleg inkonzisztens meggyőződései vannak (még ha ezek a meggyőződések nem is feltétlenül tudatosak). Abban azonban semmi inkonzisztens nincs, ha valaki szeretne is elmenni az Északi-sarkra, meg nem is. Tisztában lehet azzal, hogy mindkét kívánsága nem teljesülhet: vágyai *tartalmuk szerint* inkonzisztensek. De a preferenciák olyan kontextust teremthetnek, amelyben az inkonzisztens tartalmak nem vonnak maguk után inkonzisztens mentális állapotokat. Általánosabban, ha *A* és *B* inkonzisztens propozíciók, akkor az, aki felismeri ezt, és mégis azt hiszi, hogy *A*, és azt is hiszi, hogy *B*, inkonzisztens mentális állapotban van. De aki arra vágyik, hogy *A*, és arra is vágyik, hogy *B*, az se nem inkonzisztens, se nem irracionális.

Ennek ellenére van egy fontos analógia a gyakorlati és a kognitív konzisztencia között. Mint láttuk, a racionalitás szempontjából különös szerepe van az evidenciák és a hozzájuk kapcsolódó meggyőződések közti viszonynak. Ennek oka az, hogy az evidenciák és a releváns meggyőződések közt valamilyen tág értelemben vett – tehát nem feltétlenül deduktív – következtetési viszony áll fenn, a szónak abban az értelmében, hogy az *evidenciák érvként (reason) szolgálnak* más meggyőződések kialakulásában és igazolásában. A kognitív irracionális va-

lőszínűleg ott kezdődik, ahol az evidenciák nem az elméleti érvelésben nekik szánt szerepet töltik be. A gyakorlati racionalitás, illetve irracionalitás szintén a gyakorlati érveléshez kötődik. A cselekedetek racionalitásának feltételeit nem érthetjük meg anélkül, hogy a *gyakorlati érvelés* (*practical reasoning*) természetét ne vizsgálánánk (Harman 1976, Bratman 1987, Velleman 2000, Audi 2006). Természetesen „az érvelés” kifejezés egyik esetben sem feltétlenül tudatos tevékenységre utal. A racionalitás nem követeli meg, hogy a meggyőződése vagy a cselekedetei mellett szóló érveit valaki explicit módon kifejtse, vagy akár tudatosan végiggondolja. Csupán azt kell feltételeznünk, hogy az evidenciák olyan folyamatokkal kapcsolódjanak a meggyőzésekhez, amelyek explicit érvek formájában is kifejezhetők; illetve hogy bizonyos cselekedetekhez vezető mentális folyamatok explicit gyakorlati érvekkel is rekonstruálhatók (Audi 1993, 6. fejezet).

III. GYAKORLATI SZILLOGIZMUS

A gyakorlati racionalitásra, illetve irracionalitásra vonatkozó alapvető kérdés ezért az, hogy mi a gyakorlati érvelés célja és természete. Az egyik, Arisztoteléstől származó hagyomány szerint a gyakorlati érvelés nem más, mint helyes következtetés. Ezért a gyakorlati érvelés célja végső soron éppúgy megfelelő premisszákból helyes következtetések levonása, ahogyan az elméleti érvelésé helyes elméleti következtetések levonása. Gyakorlati racionalitással kapcsolatos fejtegetéseinket érdemes e hagyomány elemzésével kezdeni, természetesen anélkül, hogy elkötelezném magam Arisztotelész saját, sokat vitatott elméletének valamely meghatározott értelmezése mellett. A következőkben inkább az eredeti szövegek által inspirált kérdésfeltevéseket tekintem kiindulópontnak.⁴

Mivel Arisztotelész szerint a következtetés formája szillogisztikus, a gyakorlati érvek premisszái, éppúgy, ahogyan az elméleti érveké, egy univerzális és egy szinguláris kijelentésből kell, hogy álljanak. A gyakorlati következtetés konklúziója azonban nem egy kijelentés igazságára vonatkozó meggyőződés, hanem egy a konkrét cselekvés értékét kifejező szándék.⁵ Mivel Arisztotelész szerint a szándék abban az esetben, ha a cselekvő képes a cselekedet végrehajtására, és semmi sem akadályozza annak végrehajtásában, magában a cselekedetben manifesztálódik, ezért a gyakorlati szillogizmus következményeként magát a

⁴ Ahogyan azt Anscombe klasszikus művének megjelenése óta a legtöbb a témát érintő filozófus teszi. Vö. Anscombe 1957; Davidson 1963, 1970; McDowell 1978, 1979; Rorty 1980; Wiggins 1998.

⁵ Ez az értelmezés a gyakorlati szillogizmus *Nikomakhoszi etikában* található megfogalmazásából indul ki. Kevésbé világos, hogy alkalmazható-e a gyakorlati szillogizmus Arisztotelész más műveiben, például *Az állatok mozgásában* (701a7–25) fellelhető alkalmazásaira. Köszönmő Betegh Gábornak, hogy erre felhívta a figyelmemet.

cselekedetet is tekinthetjük. Például: minden édes dolog fogyasztása jó; a tábla csokoládé előttem édes; jó lesz, ha megeszem, aminek *ceteris paribus* rögvest neki is látok (NE. 1147a).

Az így értett gyakorlati következtetés rendkívül flexibilis keretet nyújt a cselekedetek mentális eredetének vizsgálatára, mivel csak azt a formális feltételt köti ki, hogy a racionális cselekedetek magyarázata megköveteli egy az általános cselekvéstípus értékére vonatkozó és egy az adott szituáció észlelésére vonatkozó meggyőződés azonosítását, és azt feltételezi, hogy ezek következményét maga a cselekedet végrehajtása manifesztálja. Az univerzális ítélet általánossági foka azonban meghatározatlan. Lehet olyasmi, mint „minden esetben, amikor éhes vagyok, nem fáj a fogam, és előző nap választások voltak, édeset fogyasztani jó”. De lehet olyasmi is, hogy „minden, ami az élet fenntartásához kell, az jó”. A fontos csupán az, hogy a második premissza tartalmazza azt a tényt, hogy a cselekvő felismeri: fennállnak azok a partikuláris körülmények, amelyek között az adott cselekedet a cselekvő értékítéletének manifesztációja. A gyakorlati következtetés kiindulópontja pedig egy értékítélet.

Az arisztotelészi modellben tehát a cselekvés racionalitása részben attól függ, hogy ezen értékítélet helytálló-e, másrészt pedig attól, hogy a cselekvő képes-e helyesen értelmezni azt a szituációt, amely lehetővé teszi az ítélet cselekvésben történő megnyilvánulását. Ennek megfelelően a cselekedet racionalitása csorbulhat, ha a két ítélet közül bármelyik téves, *annak ellenére*, hogy a cselekvő képes lett volna helyes ítéletet alkotni (Rorty 1980). További kérdés azonban az, mennyiben képes ez a modell a sajátosan a cselekvésre vonatkozó gyakorlati irracionálitást megragadni.

A probléma a következő. Kétségtelen, hogy a partikuláris helyzet cselekvés szempontjából releváns észlelését eltorzíthatják a cselekvő érzelmei vagy motívációi. Elismerheti valaki, hogy bajba jutott embertársán segítenie kell, ha erre alkalma nyílik, és nem jár különösebb kockázattal. De lehetséges, hogy ha az illető szerelmi vetélytársa jut bajba, hajlamos ezt észre sem venni. Ez lehet emberi gyengeség és így bizonyos értelemben irracionális viselkedés, de nem sajátosan a cselekvésben megnyilvánuló irracionáltság. Túlságosan is hasonlít ugyanis arra az esetre, amikor az evidenciákat hagyjuk figyelmen kívül azért, hogy az általunk vágyott meggyőződésekre tegyünk szert.

Mi a helyzet azonban akkor, amikor valaki a „felső premissza” – megközelítőleg: valamely többé vagy kevésbé általános cselekvéstípusra vonatkozó értékítélet – tekintetében téved? Ez a kérdés az arisztotelészi felfogás – sőt általában a gyakorlati irracionáltság – egyik alapvető problémájára mutat rá. Ahhoz, hogy a felső premisszában szereplő értékítélet lehessen a gyakorlati irracionáltság forrása, a cselekvőnek inkonzisztens értékítéletekkel kell rendelkeznie. Nem világos azonban, hogy a cselekvés szempontjából releváns értékítéletek esetében ez miként lehetséges, éppen azért, mert ezek az ítéletek a cselekvő preferenciáit fejezik ki. A diétára szoruló egyén gondolhatja azt, hogy csokoládét enni

jó, és gondolhatja azt is, hogy csokoládét enni rossz, ettől még preferenciái a legkevésbé sem inkonzisztensek. Ez világossá válik, ha belegondolunk, hogy az ilyen cselekvéstípusokra vonatkozó preferenciákat csaknem kivétel nélkül feltételes állítások fejezik ki (Davidson 1970). Például gondolhatja valaki úgy, hogy csokoládét enni jó, amennyiben kellemes az íze, de csokoládét enni rossz, amennyiben túl sok kalóriát és csekély hasznos tápanyagot tartalmaz. E két állítás egyáltalán nem mond ellent egymásnak. Ezért bármelyik is szolgáljon a gyakorlati következtetés felső premisszájaként, nem állíthatjuk, hogy az adott következtetés alapján cselekvő személy irracionális.

Ezt a következményt akár a modell előnyének is tekinthetjük, amennyiben egyszerűen tagadja annak lehetőségét, hogy valaki irracionálisan cselekedjék. Eszerint a szándékos cselekedet mindig racionális. Ez nem azt jelenti, hogy a cselekvő minden esetben helyesen cselekedne. Sőt azt sem jelenti, hogy a cselekvő maga ne lehetne bizonyos értelemben irracionális. Mint fentebb láttuk, lehetséges, hogy a cselekvő helytelen módon észleli a cselekvés szempontjából releváns szituációt. És az is lehetséges, hogy értékítéletei irracionálisak. Bár az értékítéleteknek, amennyiben a cselekedetek magyarázatának alapjául szolgálnak, preferenciákat kell kifejezniük, semmi sem követeli meg, hogy a preferenciákat – a közgazdasági modellt követve – egyszerű magyarázhatatlan adottságoknak tekintsük. Az arisztotelészi modellben tehát nem lehetetlen magyarázatot találni arra, miért tekinthetünk *valakit* irracionálisnak. A sajátosan *a szándékos cselekvésben megnyilvánuló* irracionális lehetőségét azonban ez a modell nem magyarázza.⁶

Ez a meglátás nem új keletű. Maga Arisztotelész is tisztában volt vele, és gondot is okozott a számára. Szókratésszel ellentétben úgy vélte ugyanis, hogy a szándékos, ám irracionális cselekedet nem lehetetlen, miután vannak aktuális megnyilvánulásai, az aktualitásból pedig következik a lehetőség (NE. 1145b). Ezért aztán magyarázatot kell találnunk az efféle viselkedés lehetőségére. Arisztotelész megoldása azonban megmarad az intellektualista hagyomány keretein belül, amennyiben a gyakorlati irracionális végső forrásaként egy *kognitív hibát* jelöl meg. A sokat vitatott elképzelés egyik eleme, hogy az irracionális cselekvés megköveteli, hogy a cselekvő rangsorolja a cselekvéshez vezető szillogizmus felső premisszájában szereplő értékítéleteket: ítéletet kell alkotnia arról, hogy melyik (Davidson kifejezésével) *prima facie* preferenciát kifejező premissza alapján helyes cselekedni: ez az úgynevezett „legjobb ítélet” (*best judgment*).⁷ Amikor azonban a cselekvés pillanata elérkezik, ezen ítéletét „nem megfelelően használja”, és egy olyan gyakorlati szillogizmus alapján cselekszik, amelynek felső premisszája nem racionális értékítéletét, hanem „appetítív” pre-

⁶ Ahogyan azt Wiggins 1998 részletesen megmutatja.

⁷ E problémáról és annak jelentőségéről lásd Jackson 1984; Grice–Baker 1985; Bishop 1989; Gosling 1990.

ferenciáit fejezi ki. Ilyen lehet a csokoládé kellemes ízén alapuló preferencia, szemben az általános egészségügyi megfontolásokat mozgósító, racionális belátáson alapuló preferenciákkal.

IV. A „LEGJOBB ÍTÉLET”

Arisztotelész javasolt megoldásával kapcsolatban több kérdés is felvethető. Először is úgy tűnik, hogy az arisztotelészi koncepció alapján a gyakorlati irracionális ismét csak gyanúsán kognitív színezetet ölt. Igaz ugyan, hogy – szemben a szókratikus felfogással – az irracionális cselekvő nem egyszerűen ostoba, tehát nem arról van szó, hogy képtelen lenne valamely cselekvés értékéről helyes meggyőződést kialakítani, és ezért cselekszik helytelenül. Hiszen Arisztotelész elismeri, hogy a cselekvésében irracionális – Arisztotelész terminológiájában *akratikus* – személy nemcsak képes felismerni egy cselekedettípus értékét, de annak alapján is *ítél*. Azonban annak, hogy mégsem ezen ítélet alapján cselekszik, az az oka, hogy a rendelkezésre álló evidenciákat helytelenül használja: az adott szituációban eltekint azoktól, amelyek a cselekedet végrehajtása ellen szólnak, és csak a mellette szólókat veszi figyelembe (vagy fordítva). A gyakorlati irracionális ily módon mégiscsak a kognitív irracionális egy formája lesz. Ez nemcsak azért tűnik problematikusnak, mert összeköti a két eltérő típusú irracionálisitást. Egyáltalán nem képtelen feltételezés ugyanis, hogy a kettőnek közös gyökere van. A probléma az, hogy a közös gyökér inkább a gyakorlati irracionálisból ered, nem a kognitív. Hiszen épp az különbözteti meg a kognitív irracionálisitást a racionalitás pusztá korlátaitól, hogy az inkonzisztenciát nem a képességek hiánya, hanem a képességek motivációkkal, következtetésképp az egyén preferenciáival magyarázható helytelen használata okozza.⁸

Arisztotelész megoldásával kapcsolatban azonban felmerülnek olyan nehézségek is, amelyek nem a gyakorlati és elméleti irracionális megkülönböztetésének talán túlságosan is szubtilis és ezért könnyen vitatható hiányából adódnak, hanem abból, hogy Arisztotelész az irracionális viselkedésnek csak bizonyos eseteit tárgyalja, és nem világos, hogy ebből milyen a gyakorlati irracionálisra vonatkozó általános következtetéseket vonhatunk le. Arisztotelész gyakorlati irracionálisról adott magyarázata ugyanis feltételezi, hogy világos különbséget tudunk tenni „appetitív” és racionális alapú értékítéletek és megfontolások között. Továbbá hogy amennyiben a különböző értékítéletek az adott szituációban egymással összeegyeztethetetlen viselkedést követelnek meg, akkor a „legjobb ítéletet” mindig az appetitívvel szembenálló ítélettel kell azonosítanunk.

⁸ Az irracionális típusainak eltérő osztályozásáról és annak magyarázatáról lásd Elster 1983 és Mele 1987.

Azonban még ha el is ismerjük – ami, mint látni fogjuk, nem vitán felül álló bizonyosság –, hogy a „legjobb ítéletnek” megfelelő cselekedetben testesül meg a gyakorlati racionalitás, mitől lesz egy ítélet jobb, mint az összes többi? Pusztán az a tény, hogy milyen megfontolás alapján hajtunk végre egy adott szituációban valamely cselekedetet, aligha tünteti ki a rá vonatkozó ítéletet. Ellenkező esetben bármilyen ítélet, amely ténylegesen egy szándékos cselekedethez vezetett, egyben „jobb” is lenne, mint az összes többi. A korábbi példát alapul véve, ha a csokoládé élvezeti értékére vonatkozó ítélet alapján cselekszik valaki, akkor az lesz a „legjobb ítélet”. Egy másik értelmezés szerint az arisztotelészi elemzést voltaképp a szókratikus felfogás módosításaként kell értenünk. Eszerint azért cselekszik valaki irracionálisan, mert irracionálisan ítél. Azt pedig, hogy mely ítélet irracionális, az ítélet alapjául szolgáló gyakorlati érvelés felső premisszája határozza meg.

A kérdés azonban az, hogy miként. Mi tünteti ki a csokoládéfogyasztás egészséges életre gyakorolt hatását a csokoládéfogyasztás okozta élvezeti értékkel szemben az *egyén gyakorlati racionalitása* szempontjából? Természetesen lehet vitatkozni azon, hogy az egyén hosszú távú érdeke szempontjából mi fontosabb: az egészsége vagy a pillanatnyi élvezetek. De akárhogy is vélekedjünk a tőlünk eltérő nézeteket vallókról, ez a véleménykülönbség nem fog magyarázatot adni a konkrét cselekedetben megnyilvánuló irracionalitás természetéről. Tegyük föl ugyanis, hogy valaki úgy véli: a pillanatnyi élvezetek sokkal fontosabbak az ember életében, mint az, hogy húsz év múlva egészséges lesz-e. Neveltetésének okán azonban fél az élvezetektől; szörnyű, bár önmaga által is nevetségesnek tartott szorongás fogja el, ha lehetősége nyílik, mondjuk, egy tábla csokoládé elfogyasztására. Ezért aztán mikor megkínálják vele, nem képes igent mondani, bár ezt ítéli helyesnek. Viselkedése – saját meggyőződése szerint – irracionális. De irracionalitása nem abban gyökerezik, hogy viselkedését az „appetitív” megfontolások és ítéletek határozták meg.

A „legjobb ítéletet” a gyakorlati érvelésben szerepet játszó „felső premisszák”, tehát az egyes cselekvéstípusokra vonatkozó értékítéletek tartalma alapján azonosító elmélettel kapcsolatban azonban nem ez a legnagyobb probléma. Tegyük föl ugyanis, hogy képesek vagyunk a cselekvéstípusok értékére vonatkozó ítéleteket attól függetlenül osztályozni, hogy ténylegesen melyik vezetett egy adott cselekvéshez. És tegyük fel azt is, hogy minden egyes szituációban, amikor a két különböző típusba tartozó értékítélet egymással összeegyeztethetetlen viselkedést követel meg, a „legjobb ítélet” nem lehet az appetitív értéket kifejező premisszából következő ítélet. Ez azonban kizárni látszik annak lehetőségét, hogy lehetséges lenne az *egyes ítélettípusokon belüli* inkonzisztencia, és hogy ez magyarázza a „legjobb ítélettel” összeegyeztethetetlen viselkedést. De miért is lenne ez lehetetlen?

A példa kedvéért fogadjuk el, hogy a morális alapú ítéletek nem az „appetitív” alapú ítéletek sorába tartoznak. Azonban léteznek, sőt mindennapiak az olyan

morális konfliktusok, amelyekben különböző *prima facie* helyesnek tekintett premisszákból kiinduló gyakorlati következtetések egymással összeegyeztethetetlen viselkedést írnak elő. Megígértem a kislányomnak, hogy délután állatkertbe megyünk. Ugyanakkor másnapra ki kellene javítanom néhány dolgozatot. Tudom, hogy ha délután nem csinálom meg, a dolgozatjavítással másnapra nem leszek kész. Úgy ítélem, hogy a dolgozatjavítás halaszthatatlan kötelesség, az állatkerti látogatás nem az. Viszont amikor a lányom szemébe nézek, saját „legjobb ítéletem” ellenére, az állatkertet választom. Nem azért, mert az állatotkat kicsit is jobban kívánom látni, mint a dolgozatokat, és az appetitív ítéletem elfelejteti velem a racionálist. Az állatkertbe is morális okoknál fogva megyek: mivel megígértem, és az ígéretek megtartása nem állítható szembe a racionális típusú cselekedetekkel. Az adott helyzetben cselekedetem mégis irracionális abban az értelemben, hogy a „legjobb ítéletemmel” összeegyeztethetetlen módon viselkedem.

V. OPTIMALIZÁLÁS

Ha tehát az irracionális cselekvésben rejlő inkonzisztenciát úgy akarjuk értelmezni, hogy az az adott körülmények közt „legjobb ítélettel” összeegyeztethetetlen cselekedet, akkor további vizsgálódás tárgyává kell tenni, mit is jelent pontosan a „legjobb ítélet”. Az arisztotelészi gyakorlati szillogizmus, úgy tűnik, ehhez nem lesz elegendő, mivel szinte bármely viselkedés leírható úgy, mint egy ehhez hasonló szillogizmus következménye. Arisztotelész azonban a gyakorlati érvelés egy másik, nem szillogisztikus formáját is megemlíti, amely talán segítségünkre lehet a „legjobb ítélet” természetének megértésében. A gyakorlati érvelés tipikus formája a cél-eszköz jellegű érvelés: adott célok tekintetében igyekszünk kiválasztani azt a cselekedetet, amely céljaink beteljesülését a leghatékonyabb módon segíti elő (NE. 1139a). Az efféle érvelésben a „felső premissza” szerepét az egyén meghatározott céljait kifejező proposíciók töltik be. A „felső premissza” továbbra is a cselekvő preferenciájára vonatkozik. A preferencia azonban nem valamely cselekvéstípus értékét fejezi ki, hanem egy elérni vágyott célt. A másik premissza pedig a cselekvő azon meggyőződésére vonatkozik, hogy egy adott helyzetben milyen a cél elérése szempontjából releváns cselekedettípusok állnak az egyén rendelkezésére. A gyakorlati érvelést pedig az teszi racionálissá, hogy a cselekvő a rendelkezésre álló cselekedettípusokból a leghatékonyabbat hajtja végre.⁹

⁹ Érdemes megjegyezni, hogy Arisztotelész ebben a kontextusban a gyakorlati érvelést egyfajta *kalkulációnak* tekinti. Ez is azt támasztja alá, hogy nem indokolatlan elméletét a modern döntéseméleti megközelítésekkel rokonítani, ahogy azt a következőkben teszem. Davidson klasszikus értelmezését a gyakorlati érvek szerkezetéről nyilvánvalóan ez a hason-

Első hallásra ez a gondolat, bár plauzibilis elképzelés arról, milyen típusú gyakorlati érvek hatására döntünk bizonyos cselekedetek végrehajtása mellett, aligha segít abban, hogy a „legjobb ítélet” természetét megértsük. Az érvelés ugyanis nem mond semmit magukról a célokról. Márpedig, mint fentebb láttuk, bizonyos preferenciák tartalmának logikai összeegyeztethetlenségéből nem következik sem a preferenciák inkonzisztens volta, sem pedig az irracionalitás. Valakinek lehet az a célja, hogy édes dolgokat egyen, mert ezek élvezeti értéke számára magas. És lehet valakinek az is a célja, hogy egészségét megőrizze. Ha abban a helyzetben találja magát, hogy csokoládéval kínálják, az előbbi célokat figyelembe véve a csokoládé elfogyasztása mellett dönt, az utóbbiak figyelembe vételével pedig tartózkodik attól. Bármit tegyen is, nem tűnik úgy, hogy cselekedete irracionális lenne, vagy hogy gyakorlati érvelési képességével visszaélt volna. Mindössze annyi történt, hogy adott helyzetben egyes célok által megkívánt cselekedetet a más célok által meghatározott cselekedetnél magasabbra értékelte.

Bizonyos feltételek kielégülése mellett azonban a gyakorlati érvelésnek ez a formája talán mégis segítséget nyújthat a gyakorlati irracionalitás megértéséhez. A cél-eszköz érvelés keretében ugyanis felvethető az a kérdés is, hogy maguk a célok racionális rendszerbe állíthatók-e. Az alapgondolat az, hogy az egymással ellentétben álló cselekvéstípusokat megkövetelő „alacsonyabb rendű” célokat „magasabb rendű” rendű célok szempontjából értékelhetjük. A hagyományos – görög eredetű, de még a középkorban is csaknem egyetemesen elfogadott – értelmezés szerint a „magasabb rendű” kifejezés a célok *tartalmára* utal. Minden racionális lény számára egyetemes cél lehet például az – általában nem a pusztán fizikai élvezettel azonosított – gyönyör vagy a boldogság elérése.¹⁰ Amikor tehát egyes konkrét céljaink megvalósítása egy adott helyzetben egymással összeegyeztethetetlen cselekedetek végrehajtását követeli meg, annak alapján kell döntenünk, hogy melyik járul hozzá nagyobb mértékben, azaz hatékonyabban a – hosszú távú – gyönyör vagy boldogság megvalósításához. Ennek alapján kell kialakítanunk „legjobb ítéletünket” arról, melyik cselekedetet hajtsuk végre. Az a személy pedig, aki ennek az ítéletnek ellenére tesz, irracionálisan viselkedik.

E megközelítés – történetileg is bizonyítható – nehézsége, hogy további vitákat generál a tekintetben, mit követel meg a gyönyör, illetve miben is áll a boldogság. E viták erkölcsfilozófiai vonatkozását meg kell különböztetnünk attól a kérdéstől, hogy közelebb visznek-e a gyakorlati racionalitás, illetve irracionalitás megértéséhez.¹¹ Az ugyanis, hogy ki mit gondol a gyönyör vagy boldogság lé-

lőség inspirálta (Davidson 1963). E modell általánosításáról és a mellette szóló érvekről vö. Hurley 1989, 8. fejezet és Smith 1994, 4. fejezet.

¹⁰ Az álláspont klasszikus megfogalmazása Platón *Prótagoras* című dialógusában található (351–358).

¹¹ Előbbiről lásd Huoranszki 1999a, II. fejezet.

nyegéről, egyetlen cselekedetet sem fog sajátosan gyakorlati szempontból irracionálissá tenni. Ha valaki egy adott szituációban másképp cselekszik, mint azt a gyönyör vagy boldogság elérése megköveteli, abból pusztán az következik, hogy vagy a gyönyörre, vagy a boldogságra vonatkozó meggyőződése hamis, vagy adott helyzetben helytelenül választotta meg az eléréséhez vezető eszközt. Visszajutottunk tehát a szókratikus, a gyakorlati irracionálisitást a kognitív irracionálisra redukáló állásponthoz.

Hasonló problémával kell szembenéznie a kortárs döntéselméleten alapuló megközelítésnek is. E keretben nem kell feltételeznünk, hogy léteznek olyan „magasabb rendű célok”, amelyek alapján az egyéb célok által megkövetelt és egymással ellentétes tartalmú ítéleteket rangsorolhatjuk. Pusztán egy olyan formális fogalomra van szükségünk, amely a különböző célok által megkövetelt cselekedetek értékelését „közös nevezőre hozza” azáltal, hogy *mérhetővé* teszi őket. E formális fogalom a preferenciák mértékegységeként értelmezett szubjektív hasznosság. Fontos hangsúlyozni, hogy a „hasznosság” itt nem valamilyen pszichológiai tartalomra vagy állapotra utal. A „szubjektív hasznosság” pusztán *mértékegység*, amely éppoly kevésbé határoz meg egy pszichológiai tartalmat, ahogyan a „méter” sem határozza meg, hogy minek a hosszát vagy távolságát mérjük vele. E mértékegység használata azonban elvileg lehetővé teszi, hogy a különböző célok által meghatározott cselekedetek értékét egy közös értékmérő alapján vessük össze.¹² A „legjobb ítélet” pedig azt határozza meg, hogy a cselekvő minden releváns preferenciáját és az adott körülményeket tekintetbe véve melyik cselekedet végrehajtása kecségtet a legnagyobb várható haszonnal. Az, aki az ily módon meghatározott ítéletével szemben cselekszik, gyakorlati szempontból irracionális.

Bár ez a megközelítés „technikai szempontból” világossá teszi, mit jelent a „legjobb ítélet”, a gyakorlati irracionális megértéséhez mégsem visz közelebb. A probléma abból ered, hogy a gyakorlati irracionális csak akkor érthető meg e felfogás segítségével, ha azt, hogy mi az egyén számára egy cselekedet várható értéke, attól függetlenül tudjuk azonosítani, hogy mit tett. Nem világos azonban, hogy a döntéselmélet keretei között ez miként lehetséges. Egy cselekvőnek ugyanis éppoly kevésbé kell tisztában lennie választott cselekedete várható hasznosságával ahhoz, hogy az meghatározza a cselekedetét, mint amennyire tisztában kell lennie azzal, hogy hány kiló, ahhoz, hogy súlya legyen. És éppoly, ahogyan önmaga és mások is csak a mérleg állásából következtethetnek a súlyára, mások is és önmaga is csak a bizonyos körülmények között végrehajtott cselekedeteiből következtethet arra, mit gondol egy cselekedet hasznosságok segítségével mért várható értékéről. Ezért aztán ha a várható érték fejezi ki a „legjobb ítéletet”, a cselekvés gyakorlati szempontból ismét csak nem lehet irracionális. Ez persze nem zárja ki, hogy a cselekvő önmaga vagy mások cseleke-

¹² Erről bővebben lásd Huoranszki 1999b.

detét ne tarthassa *bizonyos szempontból* irracionálisnak. Lehetséges például, hogy utólag úgy látja, hogy „rosszul ítélt”, és más cselekedetet kellett volna végrehajtania. De ismét csak: ez önmagában nem *gyakorlati*, magában a cselekvésben megnyilvánuló irracionális. Ha a cselekvő úgy érzi, „rosszul kalkulált”, az inkább elméleti, mintsem gyakorlati képességeinek korlátaira mutat rá. Cselekedete csak ennek kifejeződése, és nem a sajátosan a cselekvésben megnyilvánuló irracionális fejezi ki.

VI. FORDÍTOTT AKRASZIA

A gyakorlati irracionális a „legjobb ítélet” fogalma segítségével elemző valamennyi elméletnek szembe kell néznie azonban egy további közös problémával: minden ilyen elmélet azt feltételezi, hogy valamiképp egy *a cselekvés végrehajtásának értékére vonatkozó ítélet* kölcsönöz racionalitást a cselekedetnek. Az eddig tárgyalt problémák elsősorban azt érintették, hogy miképp azonosíthatjuk a releváns ítéletet, és abból indultak ki, hogy a viselkedés racionalitása megköveteli, hogy legyen egy olyan ítélet, amelynek alapján adott szituációban a racionális cselekvőnek cselekednie kell. Ámde, először is, vajon lehetetlen-e irracionális módon cselekedni abban az esetben, ha a cselekvő nem a „legjobb”, csak relatíve jónak ítélt egy cselekedetet? És másodszor, vajon minden további nélkül elfogadhatjuk-e, hogy egy cselekedetnek az kölcsönöz racionalitást, hogy az éppen egy ilyen ítélet alapján történik? Másképp fogalmazva, mi tünteti ki az ítéletet, és ezen belül mi tünteti ki – a korábban vizsgált értelmezések valamegyikében – a „legjobb ítéletet” a cselekvés racionalitása tekintetében?

Ami az általánosabb kérdést illeti, fontos hangsúlyozni – ami talán már eddig is világossá vált –, hogy a „legjobb ítélet” fogalma nem valamiféle „objektív legjobb ítéletre” vonatkozik, bármit is jelentsen ez. Mint láttuk, a racionalitás kérdése – szemben például az igazság kérdésével – mindig a *cselekvő perspektívájából* merül fel. A legjobb ítélet tehát az, amit a cselekvő saját mérlegelése alapján alakít ki. Ez az a mentális karakter, amely az eddig vizsgált elméletek szerint racionalitást kölcsönöz a cselekedetnek azáltal, hogy a cselekedetben egy efféle ítélet tartalma nyilvánul meg. A hagyományos elmélet szerint ezért ha a cselekedet nem tükrözi az ítélet tartalmát, hanem ellentmond annak, akkor egyben irracionális is. Néhány filozófus azonban úgy véli, semmi sem indokolja, hogy a „legjobb ítéletnek” ilyen kiemelt szerepet tulajdonítsunk a cselekedetek racionalitásának vizsgálatokor.¹³ Véleményük szerint az ítélet csak akkor kölcsönözhet racionalitást a cselekedetnek, ha a releváns gyakorlati érvek alapján maga az ítélet helyes. Ugyan miért kellene egy helytelen ítélet alapján végrehajtott cselekedetnek racionálisnak tekintenünk csak azért, mert abban olyan ítélet tartalma

¹³ Lásd például McIntyre 1990, Audi 1993, 12. fejezet és Arpaly 2000.

nyilvánul meg, amelyet a cselekvő – tévesen – a „legjobbnak” tart? A kérdés költőinek tűnik, de egyáltalán nem az, miután a válasz korántsem egyértelmű.

Hasonló eseteket már Arisztotelész is tárgyal, mégpedig egy olyan érv kapcsán, amit *szofisztikusnak* nevez (NE. 1146a). Az érv szerint egy olyan személy, aki rosszul ítél, majd ezek után ítéletével ellentétesen cselekszik, érényesnek tekinthető, hiszen végső soron a helyes dolgot teszi. Ám ez a válasz nyilván „paradox” (a szónak hétköznapi, nem speciálisan logikai értelmében). Miért tekintenénk valakinek a viselkedését helyesnek csak azért, mert előbb helytelenül ítél, majd képtelen ennek megfelelően cselekedni, tehát kétszeresen is hibázik? Lehetséges, hogy az ilyen cselekedet eredményét, önmagában tekintve, pozitívan értékeljük. De maga a cselekvő nyilván híján van valaminek, mégpedig éppen a racionalitás tekintetében. Először rosszul ítél: ez a racionalitás valamiféle korlátjára utal. Aztán pedig nem az ítéletének megfelelően cselekszik: ez ismét csak valamilyen hiányosságra mutat rá. Hogyan tehetné az utóbbi hiányosság racionálissá a cselekedetet?

A következők mellett fogok érvelni. Egyfelől elismerem, hogy azoknak a kortárs filozófusoknak, akik azt állítják, hogy a „legjobb ítélet” ellenében történő viselkedés nem minden esetben irracionális, igazuk van. Másfelől azonban ebből nem következik, hogy a cselekedeteknek ne az kölcsönözne racionalitást, hogy a hozzájuk vezető gyakorlati érvek konklúziójának tartalmát fejezik ki. Különbség van ugyanis aközött, hogy valaki milyen gyakorlati érveket használ, és aközött, hogy ezekből milyen konklúziót von le. Lehetséges, hogy valakinek a gyakorlati érvei helyesek, de az érvek alapján mégis helytelen ítéletet alkot. Az ilyen esetekben bizonyos értelemben nem állítható, hogy irracionális lenne annak alapján cselekedni, ami a helyes következmény és ítélet *lenne*. Ebből azonban nem következik, hogy az ekképpen cselekvő személy megfelelő módon használta racionális képességeit.

Jonathan Bennett egy klasszikus példáján jól érzékeltethető, miről is van szó (Bennett 1974). Huckleberry Finn döntésre kényszerül, földjára-e szökésben lévő rabszolga barátját vagy sem. A törvény szava szerint ezt kell tennie. De a „szíve” mást diktál. Huckleberry úgy ítéli, hogy helyesebb a törvénynek engedelmeskedni. Ámde a döntő pillanatban mégsem az ítéletének megfelelően cselekszik, és a legtöbbször ezt aligha helytelenítenénk. De ennél több is állítható, bár Bennett maga ezt a következményt nem vonja le. (Bennettet elsősorban a morális normák és az emóciók viszonya érdekelte, nem a racionalitás.) Ha Huckleberry értékrendjében szerepet játszott a barátság és a szolidaritás, akkor valamilyen értelemben cselekedete nem tekinthető irracionálisnak még akkor sem, ha az nem a „legjobb ítéletét” fejezi ki. A rendelkezésére álló gyakorlati érvek alapján ugyanis gyakorlati ítélete volt téves, amennyiben az összes érvet figyelembe véve a konkrét cselekedet vonatkozásában „helytelen” következtetést vont le. Vessük össze ezt azzal az esettel, amikor hasonló helyzetben a barátság és szolidaritás semmilyen szerepet nem játszik az illető megfontolásaiban, vi-

szont, mondjuk, lusta riasztani a szökésben lévő rabszolgát keresőket. Az utóbbi esetben nincsenek olyan releváns gyakorlati érvek (tegyük fel, hogy lustaságát a cselekvő maga sem tekintené annak), amelyek racionalitást kölcsönözhetnének viselkedésének. Ezzel szemben Huckleberry esetében vannak ilyen érvek, és viselkedése (a példa feltevése szerint) éppen azt manifesztálja, hogy ezeknek az érveknek komoly súlyuk is van: csak éppen nem vezetnek a megfelelő ítélethez.

Véleményem szerint azonban az ilyen és ehhez hasonló esetek, amelyeket „fordított” vagy „inverz” akrasziónak is szokás nevezni, egyáltalán nem bizonyítják, hogy *általában* a cselekedetek racionalitását ne abban kellene keresnünk, hogy azok a cselekvő gyakorlati ítéletét manifesztálják. Abból ugyanis, hogy Huckleberry viselkedése nem volt irracionális, még nem következik, hogy racionálisnak tekinthetnénk. A „legjobb ítéleten” alapuló megközelítés szerint a cselekvés racionalitását az adja, hogy egy – nem feltétlenül tudatos – gyakorlati érvelés konklúzióját fejezi ki. Ezért tipikusan azokat a cselekedeteket tekintjük irracionálisnak, amelyek az ily módon értelmezett gyakorlati konklúzióval ellentétes tartalmúak, hiszen ezek mutatják, hogy a cselekvő *helytelen módon használta releváns képességét*: levonta ugyan a következtetést, ítéletet formált, de ezt nem cselekedete irányítására használta. Huckleberry viselkedése esetében valami hasonló történik: nem a cselekedet értékéről alkotott ítélete alapján tette, amit tett, ezért cselekedetében nem nyilvánul meg racionális kontroll.

A gyakorlati érvelés funkciója az, hogy a *konklúziójaként kialakuló ítélet révén* a cselekvő *racionális módon kontrollálja*, hogy mit tesz. Huckleberry ugyan a rendelkezésére álló legjobb érvek alapján cselekedett, mivel azonban ezt nem a helyes konklúzió levonása által tette, cselekedetében mégsem nyilvánul meg racionális kontroll. Ezért azokban az esetekben, amikor saját gyakorlati érveit használva – és ezek viszonylatában – a cselekvő helytelen konklúzióra jut, de a vélhetően helyes konklúzió alapján cselekszik, cselekedete nem tekinthető racionálisnak sem.

Ideiglenes konklúzióként tehát a következőt állapíthatjuk meg. A cselekedeteket sok esetben az teszi irracionálissá, hogy nem a rendelkezésre álló gyakorlati érvek alapján levont következtetés tartalmát fejezi ki. Ezekben az esetekben a cselekvő a racionális kontroll számára rendelkezésre álló eszközt használja ugyan, de nem arra, amire természetes szerepénél fogva az hivatott lenne. Ahelyett, hogy kontrollálná, mit tesz, gyakorlati ítéletével, akár egy külső szemlélő, *pusztán értékeli* saját cselekedetét. Az értékelésnek és a cselekedetnek ez a kettéválása alkotja a gyakorlati irracionális lényegét. Egyes esetekben azonban a cselekedet nem a tényleges levont következtetést, hanem a – feltételezhető módon – helyesen *levonandó* következtetést fejezi ki. A „kívülről szemlélve” ostoba vagy értelmetlen módon gonosz cselekedet a cselekvő racionalitásának korlátaira mutathat rá, de ettől még nem gyakorlati irracionális. A gyakorlati racionalitás akkor csorbul, amikor a gyakorlati érvek konklúziói nem töltik be

kontrollfunkciójukat, és pusztán önmagunk viselkedéséről szóló értékítéletekké válnak, annak ellenére, hogy eredeti szerepük nem ez.

Az értékítélet és a kontrollfunkció kettéválása magyarázza, miért tekintjük a gyakorlati irracionalitás okán elkövetett vétséget kevésbé rossznak, mint a megfontoltan rosszindulatú vagy gonosz cselekedetet. Gyakorlati érveinken alapuló ítéleteink éppannyira kifejezik karakterünket, mint cselekedeteink. Aki önmaga is helytelennek tartja azt, amit tesz, sokszor kevésbé hibáztatható, mint akinek cselekedeteit visszataszító gyakorlati érvek alapján hozott ítéletei vezérlik. Ettől függetlenül azonban csak az a viselkedés tekinthető gyakorlati szempontból racionálisnak, amelyet az egyén képes legjobb ítélete segítségével kontrollálni. A kontrollfunkció hiánya még akkor is a gyakorlati racionalitás hiányát bizonyítja, ha a cselekvő egyébként mindent összevetve saját gyakorlati érvei alapján helyesen cselekedett.

VII. AKARATGYENGESÉG

Miután a gyakorlati irracionalitás fogalmát a racionális *kontroll*, nem pedig a gyakorlati érvelés diszfunkciójaként értelmezem, ennek megfelelően az *akaratgyengeség* egy formájának is tekintem. Hogy a gyakorlati irracionalitás alapvető formája az akaratgyengeség, az sokáig szinte magától értetődő bizonyosság volt. Manapság azonban nem az. Sokan gondolják úgy, hogy az, amiről eddig beszéltem, az *akraszia* arisztotelészi fogalmához kapcsolódik, és valóban az irracionalitás egyik formája. Az akraszia azonban nem azonos az akaratgyengeséggel.

A rivális felfogás szerint Huckleberry Finn esete az olyan jellegű akaratgyenge cselekvés példája, amely *nem* akratikus, mivel nem irracionális. Ezzel szemben viszont léteznek az akratikus viselkedésnek olyan formái, amelyeknek nincs közük az akaratgyengeséghez. Például az az egyén, aki meg van győződve róla, hogy az állati eredetű táplálék fogyasztása etikailag helytelen, de soha még csak nem is gondol arra, hogy vegetáriánus legyen, akratikus ugyan, de nem akaratgyenge (Holton 1999. 253). Az akaratgyenge cselekedet lényege az *előzetes szándék és az attól való eltérés*, nem pedig az, hogy a viselkedés nem tükrözi a gyakorlati (a konkrét cselekedet értékére vonatkozó) ítélet tartalmát. Ez utóbbi ugyanis önmagában nem tekinthető a cselekvésben megnyilvánuló kontroll hiányának, az akaratgyengeség viszont szükségképpen az. Az akraszia és az irracionalitás viszont nem a cselekvésben megnyilvánuló kontroll defektusai.

A hagyományos felfogás szerint az akratikus cselekvő irracionalitása abban áll, hogy az adott cselekedet értékére vonatkozó ítélete nem nyilvánul meg intencionális viselkedésében: ebből származik az efféle irracionalitást szükségképpen jellemző belső konfliktus. Az alternatív megközelítés szerint azonban a gyakorlati kontroll hiánya nem feltételez értékítéletet, illetve az értékítélettel szem-

benálló viselkedést. Az akaratgyengeség inkább az *elköteleződés bizonytalanságára* és a *kitartás hiányára* utal.

Abban a kérdésben, hogy ez az értelmezés vajon az akaratgyengeség hétköznapi fogalmának megfelelő leírását adja-e, ahogyan azt védelmezői állítják, néhez állást foglalni. Egy ilyen állásfoglalás ugyanis azt feltételezi, hogy létezik az akaratgyengeség *egy meghatározott* „hétköznapi” fogalma, a filozófiai pszichológiában kidolgozott elméletek tesztje pedig az, hogy mennyiben felelnek meg a fogalom e hétköznapi használatának. Véleményem szerint azonban a legtöbb hétköznapi nyelvben használt fogalomnak nincs egy teoretikus keretet megelőző, ámde meghatározott és szigorú (akár implicit) szabályok szerint működő használata. Aligha vitatható, hogy az akaratgyengeséget használjuk a „kitartás hiánya” értelmében is. Ámde használjuk sok olyan kontextusban is, amelyben az akaratgyengeséget valamely előzetes elköteleződéssel szembeni viselkedésként értelmezni meglehetősen erőltetett, mivel az akaratgyenge viselkedés csak az adott cselekvés pillanatában nyilvánul meg, míg – az itt releváns értelmében – a szándék valamilyen hosszabb távú elköteleződést feltételez.¹⁴

Milyen elméleti érvek szólhatnak amellett, hogy az akaratgyengeséget az elköteleződés és kitartás fogalmi segítségével értelmezzük? Korábban már láttuk, hogy az akaratgyengeséget akrasziaiként értelmező elmélet ellen vethető, hogy a „legjobb ítélet” nem feltétlenül fejezi ki jól, hogy mi következik a cselekvő gyakorlati érveiből, mivel megfontolásainak tényleges következménye és tudatos ítélete eltérhetnek egymástól. Az akaratgyengeség és irracionális fogalmát, úgy tűnik, ebben az esetben el kell választanunk egymástól, hiszen a vizsgált esetekben a cselekvő *annak ellenére* sem cselekszik irracionálisan, hogy akaratgyenge a hagyományos értelmezés szerint: cselekedete ellentmond a cselekedetről alkotott ítéletének. Azonban, bár az akaratgyengeség és az akraszia jelensége ezekben az esetekben is elválik egymástól, ez önmagában nem bizonyítja, hogy az akaratgyengeséget megérthetnénk pusztán a kitartás és az elköteleződés fogalmi segítségével. Az utóbbi felfogásnak ugyanis az a célja, hogy az akaratgyengeség jelenségét *függetlenítse* a racionalitás problémájától, a „fordított akraszia” lehetősége pedig csupán azt bizonyítja, hogy a „gyakorlati ítélet” nem biztosítja minden esetben a gyakorlati racionalitást. A „kitartás-elmélet” melletti érveket ezért máshol kell keresnünk.

Az alapgondolat talán a következőképpen ragadható meg. Az akaratgyengeség az akaraterő ellentéte (Holton 2003. 42). Az akaraterő elsődleges fogalma azonban semmiképp sem kötődik a racionalitáshoz, sokkal inkább a kitartáshoz. Először is, egy tökéletesen irracionális cselekedet mellett is ki lehet tartani: valaki bámulatra méltó eltökéltséggel merheti vödörrel a vizet a Dunából, hogy

¹⁴ A pillanatnyi, tehát hosszú távú intenciókat, viselkedési stratégiákat nem feltételező akaratgyengeség problémájáról lásd Bigelow–Dodds–Pargetter 1990.

ezzel segítse az árvíz elleni védekezést. Másodszor, és ez a fontosabb, a célok melletti kitartás kérdése teljesen független lehet a célok racionalitásától. Ha valaki eldönti, hogy délután kétezer métert fog úszni, de ezeröttszáz méter után feladja, az kevésbé kitartó, mint aki teljesíti kitűzött célját. De önmagában annak, hogy kétezer vagy ezeröttszáz métert úszik, semmi jelentősége nincs; nem versenyre készül, nem lesz tőle se szebb, se egészségesebb. Kitartásának nincs racionálisan indokolható oka, de ettől még nem állítható, hogy az illető, aki teljesíteni kívánja a maga elé tűzött célt, irracionális lenne. Ennek megfelelően az akaratérőt az előzetes szándékok melletti kitartás, míg az akaratgyengeséget az előzetes szándékoktól történő eltérés jellemzi.

Az, hogy ez az elmélet mennyire plauzibilis értelmezést ad az akaratgyengeség jelenségéről, legalább két kérdés megválaszolásától függ. Egyrészt kérdés, hogy a pusztán kifejezésbeli analógián túl, igazolja-e bármi is, hogy az akaratérőt és az akaratgyengeséget – a fizikai erő, illetve a fizikai gyengeség mintájára – egymás ellentétéként értelmezzük. Másodszor, pontosabban meg kell értenünk annak feltételeit, hogy milyen körülmények között értelmes *módosítani* egy korábbi szándékot. Függetlenül ugyanis attól, hogy maga a szándék racionális gyakorlati megfontolás eredménye-e vagy sem, sok esetben egyáltalán nem állítanánk, hogy a szándék módosítása akaratgyengeség következménye.

E két kérdés nem független egymástól. Használjuk ismét a hobbiúszó példáját. Ha az illető ezernégyszáz méter tájékan érzi, hogy nemsokára görcsbe áll a lába, nevenséges lenne – ahogyan a rossz edző teszi – a kétezer méteres terv feladását *akaratgyengeségnek* tekinteni. De ugyanez áll a pusztán szellemi teljesítményekre is: ha valaki fejben össze akar szorozni két négyjegyű számot, és folyton elakad, majd feladja a próbálkozást, az nem feltétlenül akaratgyengeség. A szellemi és fizikai teljesítőképesség korlátja önmagában nem az „akarat” gyengesége. És még ha el is fogadjuk, hogy az előzetes szándéktól való eltérést valamilyen értelemben az akarat fogalmához köthetjük, helyesebb azt állítani, hogy ez az „akarat” (mint egyfajta szellemi fakultás) gyengeségére utal (*weakness in will*), nem pedig arra a jelenségre, amelyet hagyományosan akaratgyengeségnek (*weakness of will*) nevezünk (Audi 2006. 69).

Természetesen azt, hogy az előzetes szándéktól való eltérés önmagában még nem feltétlenül irracionális, senki sem tagadja. A kérdés azonban az, hogy van-e értelme akaratgyengeségről beszélni olyankor, amikor az eredeti szándéktól történő eltérést racionális érvek támasztják alá. A viselkedés irracionálisitását és az akaratgyengeséget egymástól független jelenségeként kezelő elmélet szerint az akaratgyengeséget azért nem érthetjük meg a „legjobb ítélet” fogalma segítségével, mert a cselekvés tartalmát sok esetben nem a rá vonatkozó ítélet, hanem csak egy olyan előzetes szándék határozza meg, amely nem értelmezhető a cselekvésre vonatkozó értékítéletként. Ilyen előzetes szándékok alapján cselekedni, mint fenti példánk is bizonyítja, valóban le-

hetséges.¹⁵ Ahhoz azonban, hogy az akaratgyengeség fogalmát elválaszthassuk az irracionalitás kérdésétől, többet kell bizonyítanunk. Amint azt a fordított akraszia esetei mutatják, egy cselekedet lehet szándékolt és racionális még abban az esetben is, amikor nem tükrözi a cselekvő „legjobb ítéletét”. Következésképp szándék és gyakorlati ítélet szembekerülhetnek egymással. Ha így van, akkor azt is állíthatjuk, hogy a szándékos cselekedetek bizonyos esetekben függetlenek lehetnek a gyakorlati ítéletektől. Ebből azonban még nem következik, hogy az akaratgyenge viselkedés megértése független lehetne a gyakorlati irracionalitástól.

A kérdés, mint fentebb láttuk, ugyanis az, hogy mely esetekben módosítható egy előzetes szándék *anélkül*, hogy a cselekvő akaratgyenge lenne. A válasz pedig eléggé egyértelműnek tűnik: azokban az esetekben, amikor a *módosítás elmaradása* irracionális lenne. Az, ha valaki olyankor sem képes megváltoztatni előzetes szándékait, amikor emellett erős gyakorlati érvek szólnak, nem akaratérőről tanúskodik, hanem ostoba csökönységéről, vagy rosszabb esetben kényszeres viselkedésről.¹⁶ Következésképp a szándékok melletti kitartás hiányaként értelmezett akaratgyengeség fogalma éppúgy feltételezi a gyakorlati racionalitás képességét, mint a hagyományos elméletek.

VIII. KONKLÚZIÓ: A GYAKORLATI RACIONALITÁS

A gyakorlati racionalitás lényege a cselekvések feletti *racionális* kontroll. A racionális kontroll nem azonos az intencionális kontrollal: egy cselekedet lehet az utóbbi értelemben kontrollált abban az esetben is, ha a racionális kontroll hiányzik.¹⁷ A gyakorlati racionalitás feltételezi, hogy a cselekvő képes adott helyzetben gyakorlati érvei alapján ítéletet alkotni arról, milyen cselekedetet lenne helyes végrehajtani. A racionális kontroll azonban nem azt követeli meg, hogy minden szándékos cselekedet egy ilyen ítéletet fejezzon ki, mivel a legtöbb a legkevésbé sem irracionális cselekedet sem fejez ki szükségképpen ilyen ítéletet. Amit a gyakorlati racionalitás képessége megkövetel, az sokkal inkább egy kondicionális állítás segítségével ragadható meg: *ha* adott körülmények között egy cselekedet végrehajtása (vagy annak elmulasztása) *ellentétes* a cselekvő gyakorlati megfontolásokból következő ítéletével, *akkor* az adott cselekedet elmulasztása (vagy végrehajtása) irracionális. Pontosan ez magyarázza, hogy – amint

¹⁵ A szándékos, illetve a cselekvő értékítéletét kifejező és racionális kontroll alapján végrehajtott cselekedetek elválasztása melletti további érvekről lásd Stocker 1971, Hursthouse 1991 és Velleman 2000.

¹⁶ A kényszeres viselkedés és az akaratgyengeség különbségének eltérő értelmezéseiről lásd Watson 1977, Smith 2003 és Huoranszki 2011b.

¹⁷ Ennek jelentőségéről bővebben lásd Huoranszki 2011a, 7. fejezet.

azt a tanulmány elején említettem – miért csak a gyakorlati irracionalitás vizsgálata segítségével érthetjük meg helyesen a gyakorlati racionalitás lényegét.

Az a hagyományos elképzelés tehát, amely szerint a gyakorlati racionalitás abban áll, hogy a cselekvő tette egy adott helyzetben a cselekedet értékéről alkotott ítéletét fejezi ki, a gyakorlati racionalitás megértésének helyes *kiindulópontja*. Eredeti formájukban azonban az erre a belátásra építő elméletek az embert egyoldalúan racionális lényként mutatják be. Ezt az bizonyítja, hogy viselkedésével kapcsolatban olyan elvárásokat fogalmaznak meg, amelyek megsértését még nem tekintenénk irracionálisnak. Önmagában nem irracionális, ha egy cselekedetet olyan szándék alapján hajtunk végre, amely nem valamely gyakorlati érvelés következménye. Sok esetben nincs szükség gyakorlati következtetésre, más esetekben különböző gyakorlati érvek is rendelkezésünkre állnak, amelyek egymással összeegyeztethetetlen cselekedetek végrehajtását követelik meg, és amelyek közül további érvek alapján nem választhatunk. Amit a gyakorlati racionalitás megkövetel, az az, hogy *ne* hajtsunk végre olyan cselekedetet, amely *ellentmond* a cselekedet értékéről alkotott ítéletünknek, *ha* van ilyen. Ha rendelkezésre álló érveink ellenére téves ítéletet alkotunk, lehetséges, hogy bizonyos értelemben nem helytelen az ítélet ellenében cselekednünk. De miután a gyakorlati racionalitás lényege az, hogy a cselekedeteink értékére vonatkozó ítéletekkel képesek legyünk szándékos cselekedeteinket kontrollálni, e képességünk nem manifesztálódhat azokban az esetekben, amelyekben gyakorlati érveink ítéletünk ellenében, mintegy „kerülő úton” határozzák meg cselekedeteinket.

IRODALOM

- Anscombe, G. E. M. 1957. *Intention*. Oxford, Basil Blackwell.
- Audi, R. 1993. *Action, Intention, and Reason*. Ithaca, Cornell University Press.
- Audi, R. 2006. *Practical Reasoning and Ethical Decision*. New York, Routledge.
- Arisztotelész 1971. = NE. *Nikomakhoszi etika*. Ford. Szabó Miklós. Budapest, Magyar Helikon.
- Aristoteles 2010. *On the Motion of Animals*. Ford. A. S. L. Farquharson. eBooks@Adelaide.
- Arpaly, N. 2000. On Acting Rationally against One's Best Judgement. *Ethics*. 110. 488–513.
- Bennett, J. 1974. The Conscience of Huckleberry Finn. *Philosophy*. 49. 123–134.
- Bigelow, J. – S. Dodds – R. Pargetter 1990. Temptation and the Will. *American Philosophical Quarterly*. 27. 39–49.
- Bishop, J. 1989. *Natural Agency*. Cambridge, Cambridge University Press.
- Bratman, M. E. 1987. *Intentions, Plans, and Practical Reason*. Cambridge/MA, Harvard University Press.
- Davidson, D. 1963. Actions, Reasons, and Causes. In uő *Essays on Actions and Events*. 3–20.
- Davidson, D. 1970. How is Weakness of the Will Possible? In uő *Essays on Actions and Events*. 21–42.
- Davidson, D. 1980. *Essays on Actions and Events*. Oxford, Clarendon Press.
- Davidson, D. 2004. *Problems of Rationality*. Oxford, Clarendon Press.

- Dennett, D. C. 1998. *Az intencionalitás filozófiája*. Ford. Pap Mária – Pléh Csaba – Thuma Orsolya. Szerk. Pléh Csaba. Budapest, Osiris.
- Elster, J. 1983. *Sour Grapes*. Cambridge, Cambridge University Press.
- Gosling, J. 1990. *Weakness of the Will*. London, Routledge.
- Grice, P. – Baker, J. 1985. Davidson on „Weakness of the Will”. In B. Vermazen – M. B. Hintikka (szerk.) *Essays on Davidson: Actions and Events*. Oxford, Clarendon Press. 27–49.
- Harman, G. 1976. Practical Reasoning. *Review of Metaphysics*. 29. 431–463.
- Holton, R. 1999. Intention and Weakness of the Will. *Journal of Philosophy*. 96. 241–262.
- Holton, R. 2003. How Is the Strength of Will Possible? In B. Stroud – C. Tappolet (szerk.) 2003. *Weakness of the Will and Practical Irrationality*. Oxford, Oxford University Press. 39–67.
- Hurley, S. 1989. *Natural Reasons*. Oxford, Oxford University Press.
- Huoranszki F. 1999a. *Filozófia és utópia*. Budapest, Osiris.
- Huoranszki F. 1999b. Döntésmélet és erkölcsi normák. *Szociológiai Szemle*. 9/1. 3–30.
- Huoranszki F. 2011a. *Freedom of the Will. A Conditional Analysis*. New York, Routledge.
- Huoranszki F. 2011b. Weakness and Compulsion: The Essential Difference. *Philosophical Explorations*. 14. 51–67.
- Hursthouse, R. 1991. Arational Action. *Journal of Philosophy*. 88. 57–68.
- Jackson, F. 1984. Weakness of Will. *Mind*. 93. 1–18.
- McDowell, J. 1978. Are Moral Requirements Hypothetical Imperatives? *Proceedings of the Aristotelian Society. Supplementary Volume*. 52. 13–29.
- McDowell, J. 1979. Virtue and Reason. In R. Crisp – M. Slote (szerk.) *Virtue Ethics*. Oxford, Oxford University Press. 141–162.
- McIntyre, A. 1990. Is Akratic Action Always Irrational? In O. Flanagan – A. Rorty (szerk.) *Identity, Character, and Mind: Essays in Moral Psychology*. Cambridge/MA, MIT Press. 379–400.
- Mele, A. R. 1987. *Irrationality*. New York, Oxford University Press.
- Platón 2007. *Prótagoras*. Ford. Bárány István. Budapest, Atlantisz.
- Rorty, A. 1980. Where Does the Akratic Break Take Place? *Australasian Journal of Philosophy*. 58. 333–346.
- Smith, M. 1994. *The Moral Problem*. Oxford, Blackwell.
- Smith, M. 2003. Rational Capacities, or: How to Distinguish Recklessness, Weakness, and Compulsion. In B. Stroud – C. Tappolet (szerk.) 2003. *Weakness of the Will and Practical Irrationality*. Oxford, Oxford University Press. 39–67.
- Stocker, M. 1979. Desiring the Bad: An Essay in Moral Psychology. *Journal of Philosophy*. 76. 738–753.
- Stroud, B. – C. Tappolet (szerk.) 2003. *Weakness of the Will and Practical Irrationality*. Oxford, Oxford University Press.
- Velleman, J. D. 2000. *The Possibility of Practical Reason*. Oxford, Clarendon Press.
- Watson, G. 1977. Skepticism about Weakness of Will. *Philosophical Review*. 86. 316–339.
- Watson, G. 2004. *Agency and Answerability*. Oxford, Oxford University Press.
- Williamson, T. 2000. *Knowledge and its Limits*. Oxford, Oxford University Press.
- Wiggins, D. 1998³. Weakness of Will, Commensurability and the Objects of Deliberation and Desire. In uő *Needs, Values, Truth. Essays in the Philosophy of Value*. Oxford, Oxford University Press. 239–267.