

Huoranszki Ferenc

DÖNTÉSELMÉLET ÉS ERKÖLCSI NORMÁK II. (Válasz Ambrus-Lakatos Loránd kritikájára)

A döntéselméletet, annak társadalomtudományban és morálfilozófiában történő alkalmazhatóságát érintő viták sajnos gyakran a süketek párbeszédére emlékeztetnek. A sánc egyik oldalán az elméletet csupán felszínesen, vagy rosszabb esetben távolról sem ismerők zsigeri ellenszenv-nyilvánításaival találkozhatunk, míg a másik oldalon a merev ortodoxia utasít el minden módosítási javaslatot, a kritikai hangokat a szakmai kompetencia hiányára hivatkozva utasítva el. Ezért nagy öröm számomra, hogy írásomra olyan kritikus válaszolt, aki kiválóan ismeri a döntés- és játékelméletet, ugyanakkor jól látja az elmélet alkalmazhatóságával, esetleges korlátaival kapcsolatos problémákat is.

Mindazonáltal úgy gondolom, a bírálatra (vagy legalább is annak központi elemeire) van válaszom. Mindenekelőtt szeretném röviden megismételni azokat a főbb téziseket, amelyeket a tanulmányban védeni igyekeztem.

1. A döntéselmélet és a morális normák motiváló erejének kapcsolata problematikus.
2. A probléma nem csupán a döntéselméletet érinti, hanem egy tágabb filozófiai-módszertani hagyományt is.
3. A hagyomány paradigmatis eseteinek egyfelől a kantianus ihletésű weberi cselekvéselméletet, másfelől a Savage által kidolgozott döntéselméletet tekintem.
4. Azt állítom, hogy a motivációs erő kapcsán eltekinthetünk az értékelés tartalmától, tehát hogy a döntéselmélet szempontjából csak a motivációs struktúra érdekes.
5. Nem állítom, hogy ez azért van így, mert az értékelő kijelentések tartalma teljesen önkényes, épp ellenkezőleg, amellet érvelek, hogy az önkényesség a nem morális döntések esetében sem áll (a de gustibus elvet módszertani idealizációnak tekintem, nem pedig az értékelő kijelentések természetére vonatkozó tézisnek).
6. Azt állítom, hogy Jeffrey nyomán lehetséges úgy módosítani a klasszikus döntéselméletet, hogy a morális normák motiváló erejét ne helyezzük a döntéselmélet által vizsgált megfontolásokon kívül, és ezért ne feltételezzük, hogy szükség van egy olyan (irracionális) „metadöntésre”, amely meghatározza, hogy az egyik vagy a másik típusú megfontolás alapján cselekedjünk-e.

A következőkben megpróbálom sorra venni és megválaszolni azokat a kritikai megjegyzéseket, amelyeket Ambrus-Lakatos Loránd hoz fel az általam javasolt megoldással szemben, illetve általában e megközelítéssel kapcsolatban. Természetesen nem tárgyalom azokat a kérdéseket, amelyekben az az érzésem, egyetértünk.

Nem követem azt a sorrendet sem, amelyben ellenvetéseit tette. Előbb a könnyebben megválaszolhatókat igyekszem sorra venni, később áttérek a nehezebben megválaszolható kérdésekre, amelyek nem csak a tanulmány érveinek (talán világosabb formában történő) újrafogalmazását követelik meg, hanem új érveket is. A végére hagyom azokat a kérdéseket, amelyeket nem csupán nem tudok megválaszolni, de amelyekről azt is gondolom, hogy igazán meggyőző érveket nem lehet mellettük (de talán ellenük sem) felhozni. (Azt természetesen ennek ellenére feltételezem, hogy értelmesen lehet róluk beszélni.)

1

A kérdések első csoportjába azok tartoznak, amelyek a morális értékeléssel kapcsolatosak, illetve azzal, hogyan viszonyul egymáshoz döntésmélet és morális norma. Abban, úgy gondolom egyetértünk, hogy a morális normák 1. cselekvéseket illetnek (nem olyanok tehát, mint például egy minőségbizonyítvány); 2. hogy tartalmuk nem vezethető le pusztán az optimalizáló viselkedés fogalmából; 3. hogy motiváló erejük személyenként változhat (vannak bátrabb és gyávább, tisztességesebb és kevésbé tisztességes emberek).

Mindebből azonban *nem következik* két fontos állítás: 1. hogy a morális normák tartalma ne lenne objektív 2. hogy ne ítélnénk meg morális alapon embertársaink viselkedését.

1. Az értékelő kijelentések természetével kapcsolatos filozófiai viták szövevényesek, és azt hiszem, jelen pillanatban nem lennének hajlandó egyik álláspont mellett sem őszintén elkötelezni magam. Bár általában úgy gondolom, hogy az értékelő állítások egy részének igazságértéket lehet tulajdonítani, abban már nem vagyok biztos, hogy ezeket ugyanabban az értelemben kell-e objektíveknek tekintenünk, mint mondjuk a tudományos vagy akár mindennapi tényeket rögzítő kijelentéseket. Abban azonban biztos vagyok, hogy valamiféle interszubjektivitást el kell várnunk tőlük. Azt az embert, aki szerint sarat enni jó, bolondnak tartjuk. Azt, aki szerint a hazugság morális erény, *moral insanity*-nek.

2. Amikor azonban olyan cselekvést *magyarázunk*, ahol a moralitásnak szerepe van (s abban a morálfilozófusok többsége egyetért, hogy a moralitás fogalma elválaszthatatlan a cselekvés értékelésétől), a cselekvés szubjektív (vagy relatív) értékelésének van csak szerepe. Csakis az „internális indokok” segítségével rekonstruálhatjuk, ki mit miért tett – és a döntésméletnek véleményem szerint ebben, és csakis ebben lehet szerepe. Éppen azért, mert a döntésméletnek nem az a célja, hogy morálisan ítéletet alkosson a cselekvésről, hanem az, hogy hozzásegítsen annak megértéséhez, adott meggyőződések és vágyak miként vezetnek egy bizonyos cselekedetekhez. A racionalitást és az optimalizáló viselkedést sokszor szinonimaként kezeljük. Azt gondolom, abban egyetértünk, hogy ez nem helyes, általában véve. De ebből még nem következik, hogy ez az azonosítás specifikus célokra ne lenne alkalmazható. Erre a problémára később még visszatérek.

Mindezek alapján úgy gondolom, hogy Ambrus-Lakatos két ellenvetése is megválaszolható. Az egyik szerint az általam javasolt megoldásból (mely szerint a cse-

lekvéseknek önértéket kell tulajdonítanunk, és azt el kell választanunk a következők értékelésétől) az következne, hogy ha valaki immoralis lépések sorozatán keresztül jut el az általa vágyott következményekhez, s ezen következmények pozitív értéke számára meghaladja a morális normák áthágásából származó „negatív értéket”, akkor a cselekvő *„hitelesen védekezhet* avval, hogy az ő saját értékelése szerint az elérendő prudenciális haszon felülmúlja az erkölcsi költségeket” (126. old. – kiemelés tőlem). Ez azonban érzésem szerint a legkevésbé sem következik az általam védeni kívánt álláspontból. Csakis akkor következne, ha bevezetnénk a következő premisszát: „Tous comprendre, c’est tous pardonner”, vagyis megérteni annyi, mint megbocsátani. Ez a „premissza” azonban vaskos tévedés. Tökéletesen meg tudom érteni, hogy valaki gyávaságból cselekedett úgy, ahogy, de ez a legkevésbé sem teszi számomra megbocsáthatóvá (vagy erkölcsileg hitelessé) a cselekedetet. A racionális döntések elmélete viszont a cselekvés *megértés révén történő magyarázatának* alkotja részét, nem pedig a külső morális megítélésnek.

Ez utóbbi tényhez kapcsolódik válaszom Ambrus-Lakatos egy másik ellenvetésére is. Tanulmányomban amellet érvelek, hogy a humei–kanti–weberi tradíció egyik következménye, hogy a morális motivációk alapján végrehajtott cselekvés, vagy dedukálható kell legyen a cselekvés következményeinek optimalizálásából, vagy kívül kell kerülnön a racionalitás hatókörén: irracionális (illetve „transzracionális”) motivációk alapján végrehajtott cselekvés lesz. Ambrus-Lakatos szerint azonban a döntéseméleti megközelítés általam javasolt módosítása „mégis csak a fenti két lehetséges megoldás egyikét (gondolom az előbbit – H. F.) fogalmazza újra”. Őszintén szólva, ezt nem egészen értem. Hiszen világos, hogy ahol a moralitásnak az általam definiált értelemben szerepe van, ott nem csak a következmények alapján optimalizálunk. A döntésemélet alapján vett racionalitás: optimalizálás. És szerintem semmi baj nincs azzal, ha valaki optimalizál egy olyan szituációban, ahol morális szempontok is felmerülhetnek. Az érdekes kérdés az, *mit tekintünk relevánsnak* az optimalizáció során: csak a következményt, vagy a cselekvést magát is.

A „célracionális cselekvő”, mint az az elnevezésből kitűnik, csak az utóbbira koncentrál. (Kérdés persze, létezik-e ilyen személy vagy helyzet, de ezzel most nem kell foglalkoznunk.) Azzal, hogy az „értékracionális” cselekvést is „racionálisnak” neveztük el, még nem magyaráztuk meg, mit értünk itt racionalitás alatt. És legfőképp: nem magyaráztunk meg semmilyen cselekvést, amíg nem értettük meg, hogy a cselekvő hogyan integrálja cél-, illetve „értékracionális” megfontolásait. Mindebből azonban nem következik, hogy végső soron minden cselekedetünknek az önérdék az alapja.

Utóbbi csak akkor következne, ha feltételeznénk, hogy *az értékelés azonos az önérdékkel*. Ez azonban még az értékelés legtriviálisabb eseteiben sem igaz. Én a kínai konyhát szeretem, te a japánt. Mi értelme van ezt „önérdéknek” nevezni? Az „önérdék” fogalma nem a szubjektív értékeléshez, hanem egy cselekvés tartalmának objektív (külső) *megítéléséhez* tartozik. Hiszen gyakran mondjuk, hogy valaki önérdékből cselekedett, amikor csak cselekedeteinek várható (szubjektív) következményeit értékelte, a hozzájuk vezető cselekedeteket viszont nem (vagy a külső megfigyelő ítélete alapján nem megfelelő súllyal vagy módon) vette tekintetbe. Az „önérdék” fogalma tehát sok esetben *előfeltételezi*, hogy az eredményt és a cselekedetet külön-külön is értékeljük.

Ez talán választ ad arra is, miért gondolom úgy, hogy a morális cselekvést értékelnünk kell. Ambrus-Lakatos *en passant* megjegyzi, hogy ez a feltevés „motiválatlan”. Talán most sikerült megmutatnom, hogy nem az. Egyébként nem árt belegondolni, mi lehetne az alternatív megoldás. Én csak egy ilyenről tudok, mely szerint a morális motiváció: korlát. Azt gondolom azonban, hogy tanulmányomban részletesen érveltem e feltevés ellen, s bár elismerem, hogy talán újabb érveket is lehetne találni, erre most nem szeretnék kitérni.

2

Sokkal nehezebb megválaszolni azt az ellenvetést, amely a dolgozat központi tézisért érinti. A tézis szerint a Jeffrey-féle modellt alkalmazva a morális motivációk alapján történő cselekvés a döntéelmélet keretein belül tárgyalható. A tanulmányban a fogolydilemma típusú szituációt használok illusztrációként. Ha jól értem, Ambrus-Lakatos amellet érvel, hogy a Jeffrey-féle döntéelmélet e tekintetben végső soron nem különbözik a döntéelmélet savage-i változatától, s hogy „Jeffrey maga nem gondolta, hogy elméletével megoldotta volna a fogolydilemmát” (125. old.).

Azzal kezdeném, hogy természetesen én sem gondolom, hogy a problémát megoldotta volna, először is azért, mert Jeffrey-t a statisztika és a konfirmáció problémája foglalkoztatta, nem pedig a cselekvéelmélet, másodszor pedig azért, mert szerintem nincs is mit megoldani. A klasszikus fogolydilemma-szerű helyzet ugyan kellemetlen, de a szó logikai értelmében nem paradox vagy problematikus: valamennyi résztvevő számára létezik domináns stratégia, amelyet követni fog. A kollektív eredmény ugyan lesújtó, ami szomorúvá tehet bennünket, de ettől még sehol egy logikai dilemma. Ha van „dilemma”, az inkább társadalmi jellegű, arra vonatkozik, hogyan lehet olyan intézményeket tervezni, amelyek megakadályozzák, hogy ilyen helyzetbe kerüljünk. Ha a Jeffrey-féle elméletnek az lenne a következménye, hogy a fogolydilemma jellegű szituációkban a kooperatív stratégia választását ajánlja, az elmélet egyszerűen rossz lenne. De nem ez a helyzet. A probléma ugyanis nem az, hogy egy már klasszikus (egyszer játszott, egoista résztvevőket feltételező stb.) a fogolydilemmaként jellemzett szituációban mi lesz a játék megoldása, hanem sokkal inkább az, hogy *hogyan kell modelleznünk egy szituációt*. Azt gondolom, itt van köztünk az alapvető vitapont.

Ambrus-Lakatos szerint, ha a cselekedetekhez numerikus értékeket rendelhetünk, épp úgy, ahogyan azt a következmények esetében tesszük, akkor nem lesz értelme elválasztani egymástól a következmények és a cselekedetek értékét, hiszen ezek az optimalizáló megfontolás során úgyis összeadódnak. Ha ez így van, akkor visszakerültünk a klasszikus döntéelméleti modellhez: a cselekedetek értékelését pusztán következményeik és azok bekövetkezésének valószínűsége határozza meg. Szerintem azonban ez nincs így. Abból ugyanis, hogy a megfontolás során az értékek összeadódnak, még nem következik, hogy az a modell lenne helyes, amelyik nem választja szét őket. A várható érték kiszámítása során is összeadódnak a lehetséges következmények valószínűségekkal súlyozott összegei, de ez még nem ok arra, hogy a modellezés során ne próbáljuk szétválasztani őket! A következőkben szeretném

megmutatni, miért hasznos külön (és nem a következmény-függvény részeként) tekintetbe venni a cselekedetek értékét.

Mindenekelőtt egyáltalán nem mindegy, hogy egy cselekedetet célnak tekintek-e vagy sem. Amikor döntenem kell, hogy az olcsóbb, de unalmasabb futást vagy a drágább, de számomra élvezetesebb úszást választom-e szabadidő sportként, a cselekedet maga a cél. Amikor viszont azt kell eldöntenem, a fogyásomat az úszás vagy a futás segíti-e elő hatékonyabban, a cselekedet eszköz. Ugyanaz a típusú cselekedet tehát lehet cél is, eszköz is. De az a modell, amelyik a második esetben értékeket rendel magához a cselekvéshez, természetesen rossz lesz, hiszen feltevésünk szerint ebben az esetben én csak a fogyást értékelem! Hogy ez így van, az könnyen ellenőrizhető: ha felajánlanak nekem egy szobabiciklit, és tudom, hogy az hatékonyabban (és olcsóbban) fogyaszt, mint a futás vagy az úszás, rögvést azt fogom választani. Az úszás és futás (mint számomra kedvelt szabadidő sportok) közti választásomat viszont egyáltalán nem fogja befolyásolni. Röviden: a legtöbb esetben lehet olyan szituációt teremteni, amelyben el tudjuk dönteni, hogy egy cselekedetet célnak, vagy eszköznek tekint-e a cselekvő.

Természetesen az igazi nehézség akkor merül fel, amikor egy szituáció modellezése során azt kell feltételeznünk, hogy mind a cselekedeteket, mind a következményeket értékelnünk kell. Ilyenkor „az – mondja kritikusom – hogy hová számoljuk el a cselekedeteknek tulajdonított értékeket nem dönthető el az elmélet keretein belül” (126. old.). E kérdés kapcsán azonban először is hasznos lenne két problémát elkülöníteni. Az első, hogy szükséges-e és lehetséges-e olyan *modellt alkotni*, amelyben a kettőt el kell választani. Alább megpróbálom megmutatni, hogy erre feltétlenül szükség van, s hogy ez lehetséges is. A másik kérdés, hogy lehetséges-e a *modell érvényességét tesztelni*. Erre a kérdésre nem tudok határozott választ adni, de attól tartok, hogy a teszteléssel kapcsolatos probléma nem csak az általam vizsgált esetben, hanem valamennyi döntéseméleti modell kapcsán fennáll.

Képzeljük el a fogolydilemma klasszikus kerettörténetét. Azt javaslom, hasonlít-suk össze a következő eseteket (az egyszerűség kedvéért abból indulunk ki, hogy a foglyok személyiségjegyei egyformák):

1. a foglyok „egoisták” (a börtönben eltöltött évekhez arányosan negatív értékeket rendelnek);
2. a foglyok aszkéták (a börtönévek és az értékek függetlenek egymástól);
3. a foglyok mazochisták (a börtönévek értéke pozitív);
4. a foglyok bolondok (képtelenek a börtönben töltött éveknek értéket tulajdonítani);
5. a foglyok „forradalmárok” (az árulás stratégiájának alkalmazásához negatív értéket rendelnek).

Ha egoista foglyokról van szó, a klasszikus megoldás alkalmazható, a domináns stratégia az árulás lesz. Mi a helyzet a többi esetben? Világos, hogy az árulás nem lesz domináns stratégia. Ennek oka, hogy amennyiben „szituáció”-n az egyes stratégiakombinációk alkalmazása során az egyes résztvevőkre háruló hasznosságok szerkezetét értjük, a szituáció ezekben az esetekben nem lesz „fogolydilemma-szerű”. De az új szituáció, bárminek nevezzük is, nem lesz azonos a különböző típusú szereplők számára! Lehetséges, hogy az aszkéták, mazochisták és a bolondok is

a kooperatív stratégiát választják, de nem ugyanazon megfontolások alapján, mint azok, akik a cselekvést magát is értékelik! A cselekvés értékelését a következményektől elválasztani, ez az egyetlen mód arra, hogy megértsük, miért hajlandó valaki egy hasonló szituációban kooperálni, anélkül, hogy bolondnak vagy perverznek tekintenék. Ha a cselekedetek és a következmények értékelése közt nincs különbség, akkor azt a cselekvést, amelyet morális megfontolások is befolyásolnak, kénytelenek lennének a 2., 3., vagy 4. típusú szituáció valamelyikével azonosítani. De erre semmi szükség, hiszen egy egyszerű technikával értelmezni tudjuk a szituációk között (egyébként szerintem mindenki számára nyilvánvalóan fontos) különbséget.

El lehet persze vitatkozni azon, hogy a „forradalmárok” fogolydilemma-szerű szituációban vannak-e, mivel a következményekhez rendelt értékek számukra ugyanazok, mint az egoisták számára, vagy mivel cselekedeteik értékei a következményekhez hozzáadódva megváltoztatják preferenciáikat, ez már nem is a fogolydilemmája. Öszintén szólva ez engem a legkevésbé sem érdekel. Számomra ugyanis nem az a döntő kérdés, hogy minek nevezünk egy szituációt, hanem az, hogy helyesen modellezzük-e a résztvevők megfontolásait. Az a modell, amelyik nem tud, vagy nem akar különbséget tenni a morális normák alapján cselekvők és a perverzek vagy bolondok közt, számomra, finoman szólva, nem tűnik nagyon meggyőzőnek.

A vita persze nem is abban van köztünk, hogy meg kell-e tudnunk különböztetni e szituációkat, hanem abban, hogy megtehető-e ez a megkülönböztetés a döntésemélet eszköztára segítségével. Vajon lehetséges-e tesztelni azt a modellt, amelyik meg akarja különböztetni egymástól azokat a helyzeteket, amelyekben valaki a cselekedetek értéke alapján is dönt, azoktól, amelyekben csak a következményeket veszi figyelembe? Ez már fogósabb kérdés.

Nagyon egyszerű esetekben persze a válasz: igen. A technika, amelyet alkalmaznunk kell, a következő: ajánljuk fel egy személynek, hogy különböző cselekedetek választása révén, a bizonyosság körülményei között, ugyanazon nyereményeket érheti el. Ekkor nyilvánvaló módon a cselekedetek relatív, rendezési skálán mért értéke alapján fog választani. Elismerem, hogy ez távolról sem elégséges információ. Azt azonban bizonyítja, hogy semmi abszurd nincs abban a feltevésben, hogy cselekedeteket értékelünk: ha nem tennénk, egy ilyen szituációban teljesen közömbösnek kellene lennünk.

Azt is szokás persze mondani, hogy ilyenkor a cselekedetekhez „élvezeti értéket” rendelünk, és ezzel máris következményekké varázsoltuk őket. Ez azonban szerintem nem járható út. A döntéseméleti modell kiindulópontja ugyanis a következő: a cselekvések következményei (illetve bekövetkezésük valószínűsége az egyes cselekedetek végrehajtásának függvényében) a cselekvő számára *adottságok*, tehát nem befolyásolhatók. Amiről ő dönt, az a cselekvés. Ám ha a cselekvést is következménnyé alakítjuk, akkor vagy azt kell állítanunk, hogy a cselekvőnek nincs miről döntenie, vagy pedig azt, hogy a következményekről is dönthet, ami ellentmond eredeti feltevésünknek. (Éppen ez teszi nehezen elfogadhatóvá az Ambrus-Lakatos által is említett Harsányi-féle javaslatot, ahol is már csak a preferenciákat rendezgetjük, de valójában nincs miről döntenie.)

A kevésbé egyszerű esetekben (ahol eltérő következményekkel és szubjektív valószínűségekkel is számolnunk kell) persze nehéz lesz megmondani, mit értékel a cselekvő: csakis a következményeket, vagy a hozzájuk vezető cselekedeteket is.

Látunk kell azonban, hogy ez a probléma nem súlyosabb, mint az, amely a valószínűségek és a következmények értékelését illeti. Ez utóbbi esetben is abból indulunk ki, hogy a cselekvő különböző értékeket rendel a cselekedet következményeihez és különböző valószínűségeket rendel azok bekövetkezéséhez (ahol is a valószínűségek bizonyos esetekben, mint azt tanulmányomban részletesen tárgyalom, nem függetlenek a választott cselekedettől). Amikor azonban egy modell tesztelésére kerül a sor, és el kell tudnunk dönteni, egy adott szituációban és egy adott személy vonatkozásában melyek ezek az értékek, bizony bajban vagyunk.

Számos alternatív eljárást ismerünk (a legfontosabbak talán a Von-Neuman és Morgenstern, illetve a Ramsey által javasoltak), de mindegyiknek vannak hiányosságai. Az ezzel kapcsolatos technikai részleteket most nem elemezhetjük. Egy fontos problémát azonban szeretnék megemlíteni, mivel ez analóg azzal, ami a cselekedetek és következményeik értékének elkülönítését érinti. A klasszikus elméletek többségében azt feltételezzük, hogy a cselekvő *semleges* a kockázatot illetően: nem fél a kockázattól és nem is vágyik rá. Tudjuk azonban, hogy ez valójában nincs így. Más szavakkal: választásaink során nem csak a következményeket, hanem bekövetkezésük valószínűségét is értékeljük. El kell tudnunk tehát választani egymástól a következményeket értékelő függvényt és a kockázattalálási attitűdöt. Ezt bizonyos modellek elméletben képesek megtenni, az azonban már egy másik kérdés, van-e mód arra, hogy a modellt egyes esetekre alkalmazzuk, tehát képesek legyünk arra, hogy a megfigyelt cselekvés alapján eldöntsük, milyen mértékű valaki kockázattalálási attitűdje. Nyilvánvaló, hogy minél bonyolultabb a döntési helyzet, annál nehezebb ezt megtenni. De ebből még nem következik, hogy egy helyes modellnek ne kellene a kettőt elválasztani.

Következésképp azt gondolom, hogy a cselekedetek értékelése szükséges feltétele annak, hogy helyes modellt adjunk a cselekvést megelőző megfontolásokat logikai szerkezetéről. Ezzel többek között lehetővé válik, hogy elméletünk számot adjon a morális megfontolások motiváló erejéről. Fontosnak tartom azonban hangsúlyozni, hogy ezzel nem a „moralitás problémáját” akartam megoldani. Nem tudok választ adni arra a kérdésre: miért legyünk morálisak? Abban sem vagyok biztos, hogy van-e értelme ennek a kérdésnek. Nyilvánvaló módon abból, hogy a cselekvést is értékeljük racionális (optimalizáló) döntéseink során, önmagában még nem következik, hogy cselekedetünk morális szempontból releváns vagy éppenséggel helyes lesz. Hiszen cselekedeteket (vö. a dohányzás példáját tanulmányomban) nem csak morális okokból kifolyólag értékelhetünk. Valamint a cselekedetet morális szempontú értékelése nem vonja maga után, hogy a cselekedetet morálisan helyes döntés eredményének tartjuk. Az erős bosszúvágyból végrehajtott vagy kifejezetten gonosznak szánt cselekedet ugyan morális motivációk alapján történik, de nem helyes.

Az elmélet azonban (bár ismételten hangsúlyozni szeretném, hogy célja nem a cselekvés külső megítélése, hanem megértésének és ezzel magyarázatának elősegítése), hozzá tud járulni egy szituáció külső, morális értékeléséhez is. A következmények és a cselekedetek értékelésének elválasztásával ugyanis meg tudjuk különböztetni a morális döntési szituációk két igen fontos típusát. Az egyik esetben arról van szó, hogy valakinek mérlegelnie kell, hajlandó-e olyan jövőbeli kellemetlen következményeket vállalni, amelyek egy általa morálisan helyesnek vélt (tehát értékelt) cselekedetet végrehajtásából származnak. A következmények súlyosságától

(és az adott erkölcsi norma súlyától) függően különböző személyiségű egyének különbözőképp döntenek majd. Egy ilyen helyzet nyilvánvalóan nem azonos típusú azzal, amikor csak a következményeket vesszük tekintetbe.

Ámde különbözni fog attól a helyzettől is, amikor *csak* a cselekedetek értékelése dönt majd, mert az (adott személyre háramló) következmények cselekedetekhez viszonyított relatív súlya elenyésző, ámde az egyik értékelt cselekedet szükségképpen maga után vonja, hogy a másik értékelt cselekedetet nem hajthatjuk végre („Légy áruló, vagy minden tizedik embert kivégeztetek a faluban!” típusú szituációk.) Ezeket a helyzeteket nevezzük morális dilemmáknak. A morális dilemmák éppen abban különböznek a fentebb említett dilemmáktól, hogy míg utóbbiak esetében a cselekedetek és azok (szubjektív) következményei közt kell súlyoznunk, az utóbbi esetében maguknak a cselekedeteknek az értéke alapján kell döntenünk. Nyilvánvaló, hogy ez fontos különbség a motiváció szempontjából. Állításom szerint a cselekedetek megértés révén történő magyarázata során a döntésemélet segítségével képesek vagyunk érzékeltetni ezt a motivációban rejlő, de a cselekedet megítélése során is releváns különbséget.

3

Mint válaszom elején említettem, vannak nehéz kérdések, és vannak (legalábbis számomra) megválaszolhatatlan kérdések. Ez utóbbiak közé tartozik az a kérdés, miért kell egyáltalán komolyan vennünk a döntéseméletet, amikor a morális cselekedetek motivációs szerkezetének megértéséről van szó? Azért nem hiszem, hogy erre válaszolni tudnék, mert egy elméleti paradigma (nem szeretem ezt a rendkívül felvizezett tudományfilozófiai műszozt, de ebben az esetben kénytelen vagyok ezt használni) *egészéről* nagyon nehéz eldönteni, ragaszkodjunk-e hozzá vagy inkább vessük el mindenestől. A filozófiai vagy tudományos problémák valamely elméleti kereten belül nyernek értelmet. A filozófiai vagy tudományos elméletek ezért értelmüket vesztenék, ha minden egyes alkalommal, amikor egy elmélettel kapcsolatban nehézség merül fel (és természetesen minden elmélettel kapcsolatban állandóan nehézségek merülnek fel), kidobnánk az ablakon, ahelyett, hogy módosítani, kiegészíteni, pontosítani akarnánk. Sem a tudós, sem a filozófus nem regényíró. Nem lehet mindig előlről kezdeni egy történetet.

Mint említettem, nem szeretem a felvizezett kuhnianizmust, az „ez is csak egy paradigma” (vagy nyelvjáték, életforma stb.) jellegű érveket, de kétségtelen, hogy Kuhn felhívta egy nagyon fontos tényre a figyelmünket: a tudomány (és nem a tudósok!) konzervativizmusára. Egy elméleti keret teljes és radikális elhagyásának nagyon nagy az ára. Túl sok korábbi meggyőződésünkről derül ki, hogy fabatkát sem ér, és ezt senki nem veszi örömmel tudomásul. Ráadásul az új keret karrierje elején általában legalább annyi problémát vet fel, mint amennyit megold, ezért nagyon erős elkötelezettségre van szükség ahhoz, hogy szakítsunk a régivel.

A legfőbb ok, ami miatt úgy vélem, érdemes kitanunk a döntéseméleti megközelítés mellett, hogy a magam részéről nem látom az új és lenyűgöző alternatívát. Önmagában az ugyanis, hogy felhívjuk az elmélet korlátaira és gyengeségeire

a figyelmet, még nem ok elvetésére. Előbbi (a korlátok azonosítása) tulajdonképpen része kell legyen az elméletalkotásnak, utóbbi pedig csak egy feladat kijelölését, nem pedig egy elméleti keret értelmetlenségét kell hogy jelentse. Elképzelhető, a döntésemélet nem a legmegfelelőbb eszköz arra, hogy segítségével a cselekvést megelőző megfontolásokról modellt alkossunk azokban az esetekben, amikor feltételezzük, hogy a cselekvő a morális normákat is tekintetbe veszi. De mielőtt elvetnénk, világosan látnunk kellene, mik a számba jöhető alternatívák. Én azonban nem ismerek egyetlen más elméletet sem, amely fogalmilag tiszta képet tudna nyújtani arról, hogyan vezetnek az alternatív cselekvési lehetőségekről kialakított meggyőződések és a következmények értékelése a cselekvést meghatározó döntéshez. Márpedig azóta, hogy Arisztotelész a cselekedeteket a „gyakorlati következtetés” eredményeként értelmezte, ennek a kapcsolatnak a megértése minden cselekvésemélet egyik alapvető feladata.

Amikor azt mondom, a döntésemélet a jelenleg rendelkezésünkre álló legjobb fogalmi keret a cselekvést megelőző megfontolások megértéséhez, nem állítom, hogy ne merülne fel súlyos nehézségek vele kapcsolatban. Az egyik ilyen nehézség, melyet fentebb részletesen tárgyaltam, a tesztelhetőség problémája. Vajon empirikusan tesztelhető elmélet-e a döntésemélet? Vagy inkább a cselekvés (racionális értelmében vett) normája? Vagy tisztán matematikai elmélet? Egyik értelmezés sem állja meg teljesen a helyét. Annak tárgyalása, miért van ez így, túlságosan messzire vezetne. Azt azonban szeretném megemlíteni, mert témánk szempontjából fontos, hogy a tesztelhetőséggel kapcsolatos problémák nem vonják maguk után az elmélet használhatatlanságát. A tesztelés feltételezi az előrejelezhetőséget, márpedig számos olyan elmélet van, amelyik alig, vagy egyáltalán nem képes az előrejelzésre. Miután a cselekvés megértése *ex post actu* történik, a döntésemélet célja pedig az, hogy a cselekedetekhez vezető megfontolásokat rekonstruálja, kilátástalan az elméletet tesztelni, a szó klasszikus értelmében (amikor is a teszt a sikeres predikciót jelenti). Ez persze nem jelenti azt, hogy ne lehetne érvelni az elmélet alkalmazhatósága vagy alkalmazhatatlansága mellett.

Végül szeretném kiemelni, mit is jelent a döntésemélet azon igénye, hogy a gyakorlati következtetést egy meghatározott módon, a cselekvéshez vezető megfontolások logikai-matematikai modellje segítségével kívánja értelmezni. A „logikai-matematikai modell” emlegetése sokakban, különösen amikor a döntésemélet által vizsgált szituációkban morális motivációk is szerepet játszanak, ellenszenvet vált ki. A morális döntés pusztán számítás volt? Pontosan azt szerettem volna megmutatni, hogy ez nincs így. Abból, hogy megfontolásairól lehet logikai-matematikai modellt adni, nem következik, hogy az erkölcsi szereplő „számításból” cselekedne. Hiszen a modell nem zárja ki, épp ellenkezőleg, *feltételezi az erkölcsi személyiség fogalmát*. Az erkölcsi személyiség fogalma ugyanis éppen abban áll, hogy valaki képes felismerni egy cselekvés erkölcsi értékét, s annak következményeit is tekintetbe véve, ésszerű döntést képes hozni arról, mit tegyen. E megfontolások döntéseméleti rekonstrukciója természetesen nem az erkölcsi személyiségről alkotott elmélet. De nem látom be, miért ne lehetne része egy olyan elméletnek, amelynek célja az emberi cselekvés megértése.