

Kant hatása az utilitarizmus fejlődésére

DOI: <https://doi.org/10.70510/mfsz.19747>

I. BEVEZETÉS

A kanti morálfilozófia és az utilitarizmus az etikai gondolkodás két meghatározó jelentőségű hagyománya. E két normatív etikát rendszerint egymással szembeállítva, a különbségekre fókuszálva mutatják be. Kevesebb szó esik arról a hatásról, amelyet Kant fejtett ki az utilitarista etika fejlődésére John Stuart Milltől kezdődően, jóllehet ez a hatás rendkívül jelentősnek mondható.

Az utilitarizmus megalapítója, Jeremy Bentham nem ismerte Kantot, ahogy Kant sem ismerte Bentham munkásságát. Etikai elméletüket mindketten a 18. század végén dolgozták ki, Kant *Az erkölcsök metafizikájának alapvetése* című műve 1785-ben, Bentham átfogó igényű munkája, a *Bevezetés az erkölcsök és törvényhozás alapelveibe* 1789-ben jelent meg először.

A klasszikus utilitarizmus két másik nagy alakja, John Stuart Mill és Henry Sidgwick ugyanakkor ismerték a kanti filozófiát, s meglehetősen nagy tisztelettel írtak róla. Mill Kantról mint olyan figyelemre méltó emberről ír, akinek gondolatrendszere egyike a legjelentősebb mérföldkőnek a filozófiai spekuláció történetében (Mill 1879. 9). Sidgwicket állítása szerint lenyűgözte a kategorikus imperatívusz igazsága és fontossága, s főképp az univerzalizmusa – még ha minden csodálata ellenére el is veti, mert elégtelennek tűnt számára a kötelességek rendszerének felépítéséhez (Sidgwick 1962).

Mill az utilitarizmus elvét látta bele a kategorikus imperatívuszba, és úgy vélte, hogy nem értelmezhető másképp, mint a maxima univerzalizálásából fakadó következmények tesztjeként. Szerinte a kategorikus imperatívusz csak úgy értelmezhető, hogy a viselkedésünket olyan szabály szerint kell alakítanunk, amelyet minden értelmes lény elfogadhat a közös érdeküknek megfelelő módon (Mill 1879. 95). Míg Mill igyekezett Kantban az utilitaristát megtalálni, Sidgwick inkább vitázni próbált vele, igyekezett például kimutatni, hogy Kant a mások boldogsága előmozdításának kötelességét alátámasztó érvei helytelenek (Sidgwick 1962. 390–391), s vitába száll vele a szabad akaratot illetően is, amelynek egy külön mellékletet szentel a *The Method of Ethics* végén (Sidgwick 1962. 511). Sidgwick is (akárcsak Mill) a kategorikus imperatívuszt sajátosan, a maga módján értelmezte, mégpedig úgy, hogy ami számára helyes, annak helyesnek kell

lennie minden személy számára hasonló körülmények között (Sidgwick 1962. XVII) – s ezzel az értelmezésével megelőlegezte némileg Richard M. Hare univerzális preskriptivizmusát.¹

Sem Sidgwick, sem Mill, sem a későbbi, Kantból inspirációt nyerő utilitarista filozófusok nem fordítottak sok figyelmet arra, hogy mennyire alapvető különbséget jelent a kanti és az utilitarista etika között, hogy az utilitarizmus az empíriában gyökerezik, a kanti kategorikus imperatívusz elve viszont „a priori szintetikus tétel, mely nem alapozódik semmilyen, sem tiszta, sem empirikus szemléletre” (Kant 2004. 41). Az utilitarista filozófusok kategorikus imperatívusról szóló értelmezése így jelentősen különbözik attól, ahogyan azt Kant értette. Richard M. Hare, aki azon utilitarista filozófusok közé tartozik, akik különösképpen erősen hangsúlyozzák adósságukat Kant felé, a *Moral thinking: its levels, method, and point* című művének bevezetőjében elismeri, hogy Kant gondolatai bár hatással voltak rá, de nem feltétlenül úgy értelmezte őket, ahogyan azokat Kant szánta (Hare 1981). Ám hogy mennyire értették, vagy értelmezték félre a kanti gondolatokat az utilitarista filozófusok, mit sem számít abból a szempontból, hogy az utilitarista gondolkodást sikerrel megtermékenyítették-e – márpedig e megtermékenyítésre számos példát találhatunk.

Tanulmányomban Kant az utilitarista etika fejlődésére gyakorolt hatásának néhány fontos aspektusát veszem szemügyre: megvizsgálom az utilitarista szempontból különösen fontos elemeit a kanti rendszernek, úgymint Kant boldogságfogalmát, helyét Kant gondolkodásában és a mások boldogsága elősegítésének kötelezettségét; illetve megmutatom, hogy az utilitarizmus miképp újult meg a 20. század első felében a kanti gondolatok inspirációjának köszönhetően, s járult hozzá a szabályutilitarizmus létrejöttéhez. A gondolatok megfelelő és érthető bemutatásához természetesen szükséges az is, hogy a tanulmány elején röviden szóljak magáról az utilitarizmusról is.

Jelen tanulmány nem vállalkozhat arra, hogy a kanti hatástörténetet teljes gazdagságában bemutassa, így Richard M. Hare és a magyar származású John C. Harsanyi (Harsányi János) – bár e kiemelkedő utilitarista filozófusok nyíltan megvallották, hogy kifejezetten a kanti gondolatok által, vagy azok által is inspirálódva sikerült létrehozniuk az utilitarizmus szofisztikáltabb változatát – munkásságát éppen csak érintem, a szabályutilitarizmust megalapozó Roy F. Harrod munkáját azonban részletesebben elemzem.

¹ Az univerzális preskriptivizmus lényege röviden az, hogy számításba kell venni mindazon személyek preferenciáját egyenlő súllyal (egy adott erkölcsi ítéletnél), akik egy cselekvés kapcsán hasonló helyzetben lennének, időtől, tértől függetlenül.

II. AZ UTILITARIZMUS ÉS PROBLÉMÁI

Az utilitarizmus megalapítója Jeremy Bentham. Bár szoktak beszélni Bentham előtti utilitarista erkölcsi gondolkodókról is (mint a klasszikus utilitarizmus előfutáiról), de Bentham volt az első, aki tudatosan tette meg a hasznosság elvét erkölcsfilozófiája alapjául, s maga az utilitarizmus szó megalkotása is Bentham nevéhez fűződik (Crimmins 2014. 38).

Bentham szerint „a természet az emberi nemet két szuverén úr – a fájdalom és az élvezet – kormányzata alá helyezte”. Bentham leszögezi, hogy a hasznosság elve „elismeri ezen alávetettséget”, s erre az elvre kell alapozni az erkölcs teljes rendszerét: vagyis a cselekedetek megítélésének az válik alapjává, hogy az emberek boldogságát előmozdítja-e (Bentham 1977. 680). Nem másról van itt szó, mint hogy Bentham a hasznosság elvét tette meg a legfőbb jónak, minden jót, minden erkölcsi előírást, minden helyesnek vélt emberi célt e legfőbb jóból, a hasznosság elvéből vezetve le (a hasznosság elvéről Bentham az *A Fragment on Government*-ben azt írja, hogy azt is mondhatjuk helyette, mi több, jobb azt mondani, hogy a lehető legnagyobb boldogság elve – a lehető legtöbb ember lehető legnagyobb boldogságának elve a jó és a rossz mérőeszköze (Bentham 1891. 93). Bujalos István úgy fogalmaz, hogy az utilitarizmus alapkérdése „a legfőbb jó, s nem az igazságosság és nem is a *helyes* mibenléte”, s a válasz rá, hogy „[a] legfőbb jó a boldogság, mégpedig a lehető legtöbb ember boldogsága” (Bujalos 2007. 636–637). Hogy mi ennek a legfőbb jónak a pontos tartalma, arról az utilitarista hagyományban mindig is viták folytak,² amelyek – egyéb nézetkülönbségek mellett – hozzájárultak ahhoz, hogy az utilitarizmus számos változata jöjjön létre. Ezek taglalása szétfeszítené a tanulmány kereteit, ezért itt csak a hasonló vonásaik bemutatására szorítkozom.

² William H. Shaw összegzése szerint Jeremy Bentham és John Stuart Mill, a klasszikus utilitarizmus két nagy képviselője egyaránt egyenlőségjelet tett az öröm (*pleasure*) és boldogság, valamint a boldogtalanság és fájdalom (*pain*) közé (a megközelítésük abban viszont markánsan különbözött, hogy Bentham szerint mindegy, mi az öröm forrása, míg Mill szerint ez nagyon is számít). A kortárs utilitarista filozófusokra azonban nem jellemző, hogy elfogadják ezt az álláspontot, s rámutatnak arra, hogy az emberek nem egyszerűen csak olyasmiket akarnak, ami örömet okoz nekik, hanem számos más céljuk is lehet: például meglátogatni Barcelonát, megmászni a Kilimandzsárót, megtanulni franciául, befejezni egy vers írását stb. Ez a „vágy-kielégülés elmélet” (*desire-satisfaction theory*), vagy másképp a preferenciák (*preferences*) elmélete. E kortárs filozófusok szerint az ember jóllétéhez (*well-being*) e vágyak kielégítése is hozzátartozik (Shaw 1999. 37–53). Preferencia-utilitarista például az tanulmányban is tárgyalt R. M. Hare és John Harsanyi. Sajátos, és az etika fejlődésére rendkívül nagy hatást gyakorló elképzelést képvisel a jó fogalmáról az ideális utilitarista Moore, aki szerint a jó meghatározhatatlan, s éppúgy nem azonosítható a gyönyörrel, vagy bármi mással, ahogyan a sárga fogalma sem azonosítható a narancssal, vagy bármi egyéb sárgával – a jó ugyanakkor az erkölcs alapja (Moore 1903).

A benthami utilitarizmus fentebbi rövid leírásában is felismerhető azon öt jegy, amelyeket Geoffrey Scarre az utilitarizmushoz kötött, s amelyeket én is irányadónak tekintek, s az alábbiakban kivonatosan idézek Scarre *Utilitarianism* című művéből.

1. Jólét, jóllét. A „hasznosság” általában az emberek jóléte. Ez nem mondható kötelező ismérvnek. (Általában igaz, de nem mindig. Moore-nál és Rashdallnál, az úgynevezett idealista utilitaristáknál nem igaz).
2. Következményelvűség. Az utilitarizmus következményelvű (vagy teleologikus) doktrína abban az értelemben, hogy azt állítja, hogy az értékekre adott helyes válasz az, hogy előmozdítjuk azokat.
3. Aggregáció. Az utilitarizmus feltételezi, hogy lehetséges minden egyes ember hasznosságát összegezni (vagy átlagot képezni belőle), hogy egy átfogó hasznossági összértéket kapjunk. A hasznosságot mérhető jószágnak tekintik, amely különböző egyének között oszlik meg.
4. Maximalizálás. Az utilitarizmus maximalizáló eleme a gyakorlati ész látszólag kényszerítő erejű elvéből merít támogatást. Ha valami jó, akkor irracionálisnak tűnik, hogy kisebb mennyiségre törekedjünk, ha nagyobb mennyiség elérésére lehetőség van. Ha az emberi jólét értékes cél, akkor minél teljesebb mértékben előmozdítjuk, annál jobb.
5. Univerzalizmus. Az a gondolat, amelyet a benthami híres formula („*greatest happiness of the greatest number*”) második fele képvisel: „*of the greatest number*”. Az utilitaristák hagyományosan nem csupán a legnagyobb boldogság megteremtését szorgalmazták, hanem e boldogság minél szélesebb körben való elosztását is. (Scarre 1996.)

Az utilitarizmusnak számos támadható pontja van. Rendkívül kontraintuitív ítéletekhez vezethet bizonyos esetekben, ilyen például egy ártatlan ember megölése valami nagyobb jó érdekében. Híresen bírálta az utilitarizmust erről az oldaláról Anscombe, amikor azt mondta, hogy vannak bizonyos dolgok, amelyek tiltottak a következményektől függetlenül, mint például az ártatlanok megölésének választása bármilyen célból (Anscombe 1958). Az utilitarizmus – úgy tűnik – megengedi, hogy egyes emberek pusztán eszközként használhatók legyenek egy (igaz, magasztosnak tűnő) cél érdekében. Gauthier érve szerint – amelyre Parfit is hivatkozik a *Reasons and Persons*-ban (Parfit 1984) – az utilitarizmus úgy tesz, mintha az emberiség egyetlen „szuperszemély” lenne, s ezért csak a boldogság összege számíтана, ami azonban képtelenség (Gauthier 1963). Martha Nussbaumnak is – Gauthierhez hasonlóan – az egyik nagy problémája az utilitarizmussal, hogy az egyének egymástól való különállásának ténye nem kap különösebb jelentőséget az utilitarista elméletben. Ugyanakkor Nussbaum szerint ebben a világban „az én fájdalomam nem a tiéd, és a te fájdalomad nem az enyém”, így nem elfogadható, hogy az egyik ember elégedettségét egy másik ember szenvedése árán ériék el (Nussbaum 1992).

A fenti problémák a kanti etikával kapcsolatban nem jelentkeznek (más kifogások felvethetők, de a fentiek biztosan nem). Ami nyilvánvalóan nem azt jelenti, hogy Kant kész megoldásokkal szolgál az utilitarizmussal kapcsolatos problémákra, ugyanakkor talán legalábbis egyik oka lehet annak, miért nyűgözött le megannyi utilitarista filozófust Kant gondolkodása – még ha nem is értettek egyet vele –, s különösképpen Kant kategorikus imperatívusza, illetve szolgált az számukra inspirációul.

III. A BOLDOGSÁG HELYE A KANTI ETIKÁBAN

Az utilitarizmus különböző változatainak közös vonása (kivéve az úgynevezett idealista utilitaristáknál, Moore-nál és Rashdallnál) a welfarizmus, vagyis a „hasznosság” azonosítása általában az emberek jólétével – hedonista vagy valamiféle más értelemben, az utilitarizmus típusától függően. Kant gondolatai a boldogságról, illetve a mások jólétét előmozdító kötelezettségről ezért számos utilitarista filozófus figyelmét is felkeltették. Az elkövetkezőben Kant boldogságfogalmát és annak helyét vizsgálom meg az etikai rendszerében, illetve a mások jólétével való törődésnek a kötelezettségét.

1. Kant boldogságfogalma

Mi a boldogság Kant szerint? Mindenekelőtt fontos leszögezni: Kant gyakran hangsúlyozza a boldogság fogalmának bizonytalanságát.

Az *erkölcsök metafizikájának alapvetésében* Kant úgy fogalmaz: „a boldogság fogalma oly meghatározatlan fogalom, hogy noha minden ember vágyik az elérésére, mégsem tudja soha biztosan és egyértelmű megmondani, voltaképpen mit is kíván és akar” (Kant 1991. 48–49). A bizonytalanság oka, hogy „a boldogság fogalmához tartozó elemek mind empirikusak”, a boldogsághoz viszont valamiféle „abszolút teljességre, a jó közérzet maximumára” lenne szükség. Kívánhatunk magunknak ugyan bármit, amitől a boldogságot reméljük, s még az is lehet, hogy céljainkat elérjük – ám az már nem biztos, hogy csakugyan boldogok is leszünk tőle. Kant azt állítja, hogy nem tudjuk „teljes bizonyossággal meghatározni, hogy mi tesz valóban boldoggá bennünket, ehhez ugyanis mindentudónak kellene lennünk” (Kant 1991. 49). Kant *A gyakorlati ész kritikájában* is a boldogság tapasztalathoz kötöttségében látja a fogalom bizonytalanságának eredőjét, rámutatva arra, hogy „a boldogság megismerése csupa tapasztalati adaton nyugszik, mert az összes róla alkotott ítélet nagyon is függ az egyének saját véleményétől, mely ráadásul maga is igen könnyen változik” (Kant 2004. 47).

A fogalom bizonytalansága érvül szolgál Kant számára ahhoz, miért nem lehet szerinte (ellentétben az utilitaristákkal) a boldogság az alapja az erkölcsi előírás-

soknak – bizonytalan alapra nyilvánvalóan nem lehet építeni. A boldogság megismerésével kapcsolatban

megadhatók ugyan generális szabályok, vagyis olyanok, amelyek átlagban a leggyakrabban helyesek, de sohasem lehetségesek univerzális szabályok, melyeknek mindenkor és szükségszerűen érvényeseknek kell lenniük; ama megismerésre tehát nem alapozhatok gyakorlati törvények (Kant 2004.47).

Kant több meghatározással is szolgál műveiben a boldogságról. E definíciók mindegyike magában foglalja, hogy a boldogság minden pillanatban és minden embernél mást és mást jelent, így konzisztensek a meghatározások a fogalom Kant által hangsúlyozott bizonytalanságával – a szakirodalomban ugyanakkor vita van arról, ezek a definíciók mennyire konzisztensek egymással, erről a problémáról a későbbiekben ejtünk szót. Kant *Az erkölcsök metafizikájának alapvetésének* első szakaszában arról ír, hogy a boldogság voltaképp egy adott lény fennmaradása és jólléte, s a boldogság elérése sokkal inkább bízható az ösztönökre, mint az észre (Kant 1991. 22–23). *A tiszta ész kritikájában* a következőképp határozza meg a boldogságot:

A boldogság abban áll, hogy minden hajlamunk kielégül (mind extenzív tekintetben, sokféleségük vonatkozásában, mind intenzív tekintetben, fokmértékük vonatkozásában, mind pedig protenzív tekintetben, tartamuk vonatkozásában) (Kant 2009. [B 834]).

E definícióra utalva jegyzi meg Allen W. Wood, hogy Kant néha tisztán hedonista, sőt, feltűnően benthamista, amikor a boldogságról ír (Wood 2001). (Wood minden bizonnyal arra utal itt ezzel, hogy Bentham az élvezetek hosszának, intenzitásának, idejének mérésével igyekezett azokat az utilitarista kalkulus számára használhatóvá tenni.) *A gyakorlati ész kritikájában* a boldogság rövidebb meghatározását kapjuk Kanttól, amely a következőképp hangzik: „a boldogság a világbeli eszes lénynek az az állapota, amikor minden kívánsága és akarata szerint alakul létezésének egészében” (Kant 2004. 149). *Az erkölcsök metafizikájában* Kant a boldogságról úgy ír, mint az embernek az állapotával való elégedettségéről, „feltéve, hogy bizonyos ennek az állapotnak a fennmaradásában” (Kant 1991. 501). E meghatározások nem pontosan ugyanazt jelentik, ami vitákra ad lehetőséget Kant értelmezői között Kant boldogságfogalmának pontos tartalmát illetően. Alison Hills szerint Kant boldogság koncepciója a hajlamok, vágyak kielégítését, illetve az örömet, elégedettséget is magában foglalja, ami azért talányos, mert ezek nem ugyanazt jelentik (Hills 2006). Allan W. Wood szerint a boldogság különböző kanti definíciói összeegyeztethetők lennének, ha nem társulna a boldogság e leírásaihoz az a Kant által gyakran hangoztatott követelmény, hogy az ember legmélyebb hajlama a boldogságra törekvés, mert ez

a követelmény választásra kényszeríti az embert a leírások között. Vitatható, teremthető-e konzisztencia Kant különböző boldogságmeghatározásai között, s ha igen, hogyan; nehéz lenne vitatni ugyanakkor Allan W. Wood azon állítását – s ha az utilitarizmus tükrében vizsgáljuk a kanti boldogság fogalmát, akkor talán ez a legfontosabb vonása a kanti boldogság koncepciónak –, hogy Kant boldogságfogalma semmiképp nem lehet eudaimonista (Wood 2001). Kant ugyanis szakítva a klasszikus etika eudaimonista hagyományával, elválasztotta az erkölcsi jót a boldogságtól, s pontosan ezt tették az utilitaristák is.

Ha az utilitarista hagyomány tükrében megvizsgáljuk Kant boldogságfogalmát, akkor valójában nem eldönthető egyértelműen, hogy az utilitarizmus melyik típusának boldogságfogalmához áll legközelebb Kant boldogságkonceptiója. Vitatott a szakirodalomban, hogy a kanti boldogságfogalom hedonista jellegű-e.³ Kant egyes boldogságot érintő megfogalmazásai olykor benthamiaknak tűnnek (ahogy arra Wood is rámutatott), olykor viszont inkább a modern preferencia-utilitarizmusban használatos preferenciákra emlékeztetnek – Kant olykor a boldogságról, mint kívánságok (bármiféle kívánság) teljesüléséről ír (Kant 1991a. 603). E nehézség mögött vélhetően az is állhat, hogy Kant számára egyformán lényeges, hogy határozottan elhatárolja egymástól a boldogságot és az erkölcsi törvényt (ezen elhatárolás nélkül fogalmilag zűrzavarossá válna Kant jól felépített és koherens gondolati rendszere), ugyanakkor fenntartsa azt az álláspontját, hogy a boldogság pontos mibenlétét nem ismerjük. Hadakozik ezért azon álláspont ellen, hogy morális és fizikai boldogság megkülönböztethető, s megjegyzi, hogy a morális boldogság kifejezés ellentmondást tartalmaz (Kant 1991a. 501), illetve azon vélemények ellen is lándzsát tör, amelyek megkülönböztetnének magasabb, illetve alacsonyabb rendű gyönyöröket (Kant 2004. 29–31). Ugyanakkor elismeri az örömek és hajlamok sokféleségét (a táplálkozás és nemiség ösztöne mellett ír a tisztességre és ismereteink bővítésére irányuló ösztönről is (Kant 1991a. 308), s minthogy alapvetően szkeptikus a boldogság megismerhetőségét illetően, természetesen kész elismerni boldogságnak azt, amit mások (saját magukat illetően) annak tartanak (Kant 1991a. 502). Mindez azt eredményezi, hogy nehezen megragadhatóvá válik Kant boldogságkonceptiója, ami egyébként tökéletes összhangban áll Kant azon álláspontjával, hogy nehezen megragadható fogalomról van szó. S aligha lehet vitás, hogy a fogalom nehezen megragadható – az utilitarista hagyományban a boldogság fogalmának meghatározása a kezdetektől a viták kereszttüzében állt.

³ Olvasható egy rövid áttekintés erről például Allison Hills már idézett tanulmányában (Hills 2006. 257. 9. lábjegyzet).

2. Boldogság: cél helyett következmény

Kant rendszerében a boldogságnak kitüntetett helye van. Kant *A tiszta ész kritikájában* azt írja, hogy az ész valamennyi érdeke három kérdésben összpontosul:

- (1) Mit lehet tudnom?
- (2) Mit kell tennem?
- (3) Mit szabad remélnem? (Kant 2009. [B 833]).

A harmadik kérdés kapcsán Kant azt mondja, minden reménykedés tétje a boldogság (Kant 2009. [B 834]). A morális törvény azon az indítékon alapszik, hogy méltók legyünk a boldogságra (Kant 2009. [B 834]). Kant válasza az ész második és harmadik nagy kérdésére: „*cselekedd azt, ami méltóvá tesz rá, hogy boldog légy*” (Kant 2009. [B 837]). Ezen a ponton előtűnik az utilitarista és kanti boldogság felfogás egyik nagy különbsége: Kant szerint csak annak jár a boldogság, aki méltónak bizonyul rá – az utilitarizmusnál nincs ilyesfajta megfontolás, a cél a lehető legtöbb ember lehető legnagyobb boldogsága.

Kant szerint nem jó ötlet mindenkit boldoggá tenni. Bár a jó szívünk ugyan erre sarkallna bennünket, de az értelmünk ezt nem engedi meg. *Az erkölcsök metafizikájában* Kant az kérdezi: adnál-e párnát a lustának, bort a részegesnek, megnyerő modort a csalónak, erős öklöt az erőszakos embernek? „Hiszen ezek mind olyan eszközök, melyeket bárki kívánhatna, hogy a maga módján boldog legyen [...] A boldogságot mégsem adhatod oda meggondolatlanul bárkinek, aki érte nyúlna.” (Kant 1991. a 603.) – vonja le a következtetést (ez a szövegrész egyébként újfent felidézheti bennünk a boldogság preferenciákkal történő, a modern preferencia-utilitarizmusra jellemző módon történő megragadását, amelynél azonban az amorális preferenciák kizártak – akárcsak John C. Harsanyi elméletében).⁴

Az utilitarizmus és Kant boldogságkoncepciójában közös, hogy szorosan összefüggenek az erkölcsiséggel – csak éppen más módon függenek össze. Az utilitarizmusnál cél-eszköz viszonyról beszélhetünk, míg Kantnál „az erkölcs rendszere elválaszthatatlanul összefügg a boldogság rendszerével, ám csupán a tiszta ész ideájában” (Kant 2009. [B 836]). Vagyis abban a Kant által „morális világnak” nevezett ideában (Kant 2009. [B 836]), ahol minden az erkölcsi törvénnyel összhangban van (*Az erkölcsök metafizikájának alapvetésében* a koncepció később, mint a „célok birodalma” jelenik meg). Ez a morális világ nem az „érzéki világ” része, csak idea, de olyan „gyakorlati idea”, amelynek – Kant szerint – hatással

⁴ John C. Harsanyi a Bentham-féle boldogság helyett az egyéni preferenciákat helyezi bele az utilitarista „kalkulusba”. S nem mindegyik preferenciát, hanem csak a valódi preferenciákat („*true preferences*”), vagyis a tájékozott preferenciákat (amelyek nem tévedésen alapulnak), amelyekből kizártak az antiszociálisak – minthogy Harsanyi szerint az utilitarista elmélet szerint minden erkölcs és más emberek iránti elkötelezettségünk alapja az általános jóakarat és az emberi együttérzés.

kellene lennie arra az érzéki világra, amelyben élünk, közelítenünk kellene ehhez az ideához. Ebben a morális világban szükségszerű lenne Kant szerint az általános boldogság, maguk a morálisan cselekvő „eszes lények okoznák önmaguk – és egyszersmind mások – maradandó jólétét” (Kant 2009. [B 837]). Abban az érzéki világban viszont sajnos, amelyben élünk, mivel nem cselekszik mindenki morálisan, a moralitás nem részesülne a boldogság jutalmában – Isten létezése nélkül. Isten azonban összeköti a morális törvényt és a boldogságot:

hogya ha az a kérdés, mi volt *Isten végső célja* a világ megteremtésében, válaszul nem a világbeli eszes lények *boldogságát* kell megnevezni, hanem a *legfőbb jót*, mely e lények boldogságra irányuló kívánságához hozzácsatol egy feltételt, nevezetesen a boldogságra való méltóságot, azaz *erkölcsiségüket*; egyedül ennek mércéje szerint remélhetnek valamikor részesülni az eszes lények a boldogságban egy *bölcs* alkotó kezéből (Kant 2004. 156).

Kant szerint az erkölcsi törvény alapja nem lehet a boldogság, ugyanakkor azt mondja, hogy az erkölcsi törvény betartása a legnagyobb boldogsághoz vezet.

Az ész tiszta gyakorlati használata ugyanis a morális törvények előírásában áll. De ezek egyként a világban lehetséges *legfőbb jó* – amennyiben ez pusztán *szabadság* által lehetséges – eszméjéhez, az *erkölcsösséghez* vezetnek, másfelől viszont ahhoz is, ami nem csupán a szabadságon, hanem a *természet*en is múlik, nevezetesen a legnagyobb *boldogsághoz*, amennyiben ez az erkölcsösség arányában oszlik meg. (Kant 1974a. 117.)

Kant nem javasolja – eltérően az utilitaristáktól – hogy az általános boldogság legyen a törvénykezés alapelve, ugyanakkor azt vallja,

hogya az állam alkotmányának a *lehető legnagyobb emberi szabadságot* kell biztosítania, mégpedig olyan törvények alapján, melyek gondoskodnak róla, hogy *mindenki szabadsága összeférhessen mindenki másnak a szabadságával* (nem a lehető legnagyobb boldogságot kell célul kitűzni, mert ez már magától következni fog) (Kant 2009. [B 373]).

Nyilvánvaló, hogy az általános emberi boldogság Kant számára is kedves eszme – csak éppen helytelennek tartaná célul kitűzni. Úgy véli, hogy a lehető legnagyobb szabadságot helyesebb célul kitűzni, s az magával hozza majd az általános boldogságot is. Még ha a szabadság fontos érték is jellemzően az utilitarista filozófusoknál, – s különösen John Stuart Millnél, aki szintén erős összefüggést látott a szabadság és a boldogság között – ez lényeges és jelentős eltérés Kant és az utilitaristák között.⁵

⁵ Joggal felvethető ezen a ponton, hogy itt valójában áthidalhatatlan megközelítésbeli különbségről érdemes beszélni. Heller Ágnes arról ír egyik esszéjében, hogy Kant „arra törekedett, hogy rendszere egészében normatív módon alapozza meg a szabadság feltétlen

3. Kant és a mások boldogsága elősegítésének a kötelezettsége

Kant kétségtől kívül kötelességnek tartotta mások jólétének előmozdítását, ami több utilitarista filozófus figyelmét is felkeltette. Az *erkölcsök metafizikájának alapvetésében* Kant példákat hoz fel a kategorikus imperatívusz működésére. Kant negyedik példája arról szól, hogy nem lehet általános törvény az, hogy mindenki csak a maga boldogságával törődik, mert akkor, ha az ember bajba kerül, akkor senkinek a segítségére nem számíthat, amit természetesen nem akarhat, hiszen „egy ilyen, saját akaratunkból fakadó természettörvénnyel megfosztanánk magunkat a kívánt segítség minden reményétől” (Kant 1991b. 55). Richard Mervyn Hare utilitarista filozófus *Could Kant Have been a Utilitarian?* (Hare 1993). című nevezetes esszéjében Kant e példáját úgy értékeli, hogy mások céljait úgy kell kezelni a kategorikus imperatívusz értelmében, mint a saját céljainkat.

Hare ezt azzal is igyekszik igazolni, hogy idéz *Az erkölcsök metafizikájából* is: „a felebaráti szeretet kötelessége tehát így is kifejezhető: kötelességem mások *céljait* (feltéve, hogy nem erkölestelenek) a magamévá tennem” (Kant 1991a. 568). Hare szerint Kant végső soron azt mondja ezzel, hogy egyenlők a preferenciáink minden embertársunkkal szemben, így nem másról van szó, mint a benthami felszólításnak való engedelmeskedésről, hogy mindenki számítson egynek. Hare arról elfelejt szólni, hogy pár oldallal később Kant arról ír, hogy a jóakarat foka különböző lehet attól függően, hogy az illető személy mennyire áll közel hozzám, s ezzel szerinte nem sérül a maxima általánossága (Kant 1991a. 570), amire Hare esszéjére adott reflexiójában Timmerman is utal (Timmermann 2005). Mindemellet Kant szerint a kategorikus imperatívusz nem írja elő, hogy mások céljait úgy kezeljük, mint a saját céljainkat, amit teljes világossággal leír *A vallás a pusztá ész határain belül* című írásában.

prioritását a boldogsággal szemben”, majd jelzi, hogy Rawls ezt a kanti szabályt „»prioritási szabálynak« nevezte el” (Heller 1997. 79). Rawlsnál ugyanis a szabadságjogokra vonatkozó első elv megelőzi az egyenlőtlenségre vonatkozó második elvet, vagy ahogyan Rawls egy helyen fogalmaz: „a méltányosságként felfogott igazságosság szerint a helyes fogalma megelőzi a jót” (Rawls 1999. 53). Rawls fokozatosan jutott el addig, hogy az utilitarizmus nagy hatású kritikusaként váljon: 1960-ban még úgy írta le igazságosságelméletét, mint amely egyesíti az utilitarista és a kontraktualista elemeket, de az utilitarista hagyományból indul ki – 1962-ben viszont már az utilitarista hagyománnyal szemben fogalmazta meg az elméletét, mint amely a szerződéselméletekből származik (Gališanka 2019). Rawls *Az igazságosság elmélete* végleges szövegében azt is hozzáteszi, hogy elmélete „igencsak kanti jellegű” (Rawls 1999. 12). A szabadság elsődlegességének elfogadása lehetett valószínűleg az egyik olyan pont, amely nem tette lehetővé, hogy Rawls fenntartsa egy, az utilitarizmussal kompatibilis álláspontot – mindemellet azt is meg kell jegyezni, hogy Rawls utilitarizmuskritikája ennél természetesen többrétűbb.

Egy cselekedetet ugyanis mindenekelőtt önmagában kell megmérni a morális törvény mérlegén, még mielőtt mások boldogságára irányul. Ennek elősegítése tehát csak feltételes kötelesség, s nem szolgálhat a morális maximák legfőbb elvéül. (Kant 1974b. 135.)

IV. AZ UTILITARIZMUS KANT INSPIRÁLTA MEGÚJÍTÁSA: HARROD UTILITARIZMUSA

Roy F. Harrod a kanti kategorikus imperatívusz segítségével akarta megújítani az utilitarizmust a 20. század első felében – s ez sikerült is neki a *Utilitarianism Revised* című tanulmányával (Harrod 1936). Scarre szerint Harrod egy olyan időszakban tett javaslatot az utilitarizmus megreformálására, amikor az kikerült az etikai gondolkodás fő áramából, és úgy tűnt, hogy teljesen megszűnhet iránta az érdeklődés – de Harrod felélesztette azt, és meggyőzte az erkölcsfilozófusok legalábbis egy részét arról, hogy az utilitarizmus nem menthetetlen (Scarre 1996.)

Harrod úgy vélte, ha a „kanti elvet” követné mindenki, az végső soron több boldogsághoz vezetne, mint a benthami boldogságelv követése. Ezzel a megállapítással vélhetően Kant sem vitázna, hiszen egy morális világban, avagy a célok birodalmában (ahol mindenki a kategorikus imperatívuszt követi) szükségszerű lenne az általános boldogság. Ugyanakkor azt is rögtön hozzátenné, hogy nem számíthatunk arra, hogy mindenki követni fogja a kategorikus imperatívuszt, s természetesen tiltakozna az ellen, ha valaki úgy érvelne a kategorikus imperatívusz által előírtak teljesítése mellett, hogy az általános boldogsághoz vezet (a fentieket a korábbi fejezetben kifejtettem). Kantnak további vitája is lenne Harroddal: minden bizonnyal rámutatna, hogy amit Harrod „kanti elvként” emleget, az nem tőle, Kanttól származik, és alapvetően különbözik a kategorikus imperatívusztól (ugyanakkor azt nyilván nem vonná kétségbe, hogy Harrod valóban Kantot olvasva jutott a maga következtetéseihez). Harrod természetesen, ahogyan általában az utilitarista teoretikusok, a maga módján értelmezi a kategorikus imperatívuszt. Harrodot – Sidgwickhez hasonlóan – a kategorikus imperatívusz univerzalizmusa bűvöli el, s nem foglalkozik például azzal, hogy a kategorikus imperatívusz egy a priori szintetikus tétel. Harrod számára a kategorikus imperatívusz első megfogalmazása a fontos – olyan maxima által cselekedni, amely általános törvényként szolgálhat. Harrod szerint ez az elv nélkülözhetetlen az utilitarizmus számára. Harrod voltaképpen újítása azon gondolata, hogy nem a cselekedet konkrét következményeit, hanem a cselekedet általánosított következményeit (vagyis azon következményeket, hogy mi történne, ha hasonló körülmények közepette mindenki az adott módon cselekedne) kell mérlegre tennünk, amikor döntünk arról, hogy az adott cselekedet helyes vagy nem. Harrod a konkrét következmények mérlegre vételét nevezi utilitaristának, az

általánosított következmények értékeléséből fakadó előírásokat nevezi kötelességnek, illetve kanti elvnek (ismételten érdemes hangsúlyozni, a kantiánusok, illetve maga Kant ez utóbbit aligha tekintenék „kanti elvnek”, de Harrodot tagadhatatlanul Kant inspirálta).

Harrod tanulmányában a hazugságot hozza fel például. Harrod szerint lehetséges, hogy egymillió – meghatározott időbeli és térbeli korlátok között elhangzó – hazugság egyenként a haszon/veszteség mérleg alapján igazolható. Igazolható még akkor is, ha tekintetbe vesszük azt a kárt, amelyet a hazugságok a közbizalom csökkentése révén okoznak, s összeadjuk mind az egymillió eset által okozott kárt. Csakhogy – mint Harrod rámutat – az egymillió hazugság okozta általános bizalomvesztés nem írható le az egyes hazugságok okozta bizalomvesztések összegével. A bizalomvesztés valóságos mértéke valójában sokkal nagyobb lehet annál, mint ha csak összeadnánk az egyéni esetek okozta bizalomvesztéseket. (Hasonlít a jelenség a közgazdaságtan „összetétel csapdájá” fogalmához,⁶ amelyet Harrod közgazdász lévén jól ismert.) Nyilvánvaló, hogy aligha leszünk nagy bizalommal embertársaink felé, ha azt kell gondolnunk, hogy akkor mondanak csak igazat, amikor azt a hasznossági elv igazolja. Ezért Harrod szerint az utilitarizmus kifinomultabb formája az igazmondás általános kötelezettségét írja elő, mert a hazugság általánossá válása túl nagy károkat okozna még akkor is, ha egyes konkrét esetekben előnyös lenne.

Mint John C. Harsanyi rámutatott (Harsanyi 1977), Harrod az utilitarizmus olyan formáját javasolta, amelyet később szabályutilitarizmusnak nevezett el Richard Brandt (Brandt 1959), s e terminus azóta a vonatkozó szakirodalomban széles körben használatossá vált. (Ellene vethető annak, hogy Harrod alapította meg a szabályutilitarizmust, hogy valójában már Mill is részben szabályutilitarista volt. Valóban vita tárgyát képezi a szakirodalomban, hogy Mill mennyiben volt tett-, s mennyiben szabályutilitarista – a szisztematikus szabályutilitarizmus alapjait azonban Harrod rakta le.)

A szabályutilitarizmus számos olyan utilitarizmus elleni kifogásra választ képes adni, amelyre a benthami utilitarizmus nem. A hazudozás, az ígéretek be nem tartása, az ártatlanok megölése pusztán hasznossági okokból, vagyis mindazon kifogások egy része, amelyet az utilitarizmusnak fel szokás róni, a szabályutilitarizmusnak sokkal kevésbé róható fel, bár problematikus esetek azért a szabályutilitarizmusnál is akadnak (Scarre 1996). Összességében azonban az utilitarizmus könnyebben védhető formájának tekinthető, s létrejöttében Kant inspiráló hatása letagadhatatlan.

Harrod mellékesen megjegyzi, hogy a „kanti elve” nem követeli meg mindig az igazmondást. Harrod szerint akkor megengedett hazudni, ha a hazugság adott körülményei között (feltéve, ha mindenki hazudna hasonló körülmények

⁶ Ami a csoport egy tagját érinti csak, s nyereséget jelent számára, általános veszteséggé válhat, ha az mindenkit érint. A fogalomról bővebben: Nordhaus–Samuelson 2016.

között) a hazugsággal által megtakarított szenvedésből fakadó kár nagyobb a hazugságok okozta bizalomvesztésből eredő kárnál. Úgy tűnhet *prima facie*, hogy ez a legnagyobb és legárgulkozóbb különbség Harrod „kanti elve” és Kant tényleges elvei között. Scarre azt állítja, hogy az utilitarista és a kantiánus között alapvető szakadék marad, amit a hazugságról gondolnak (Scarre 1996). Scarre azért téved, mert bár Kant valóban így gondolkodott, de a kantiánus filozófusok már nem feltétlenül.

Mint Michael Cholbi tanulmányának, a *The Murderer at the Door: What Kant Should Have Said* elején megjegyzi: számos kortárs kantiánus filozófus tesz erőfeszítéseket azért, hogy kimutassa, a kanti etika igazolja a hazugságot bizonyos körülmények esetén, különösen akkor, ha egy gyilkosnak kell hazudni abból a célból, hogy megmentse tőle egy ártatlan személy életét. Christine Korsgaard neves, kortárs kantiánus filozófus a *The Right to Lie: Kant on Dealing with Evil* című tanulmányában amellet érvel, hogy kontextustól függ, mi tekinthető hazugságnak. Ha valótanságot állítunk, de pusztán azért, mert viccelünk, az nem hazugság, mert nem téveszt meg senkit. Ha a gyilkos a kapuban azt kérdezi: jöttem megölni a barátod, itt van a házban? – nincs senki, aki ebben a helyzetben azt mondaná, hogy itt van a házban. Ezt a gyilkos is tudja, így a nemleges válasz nem fogja megtéveszteni. Nem történik tehát megtévesztés, hazugság. A gyilkos viszont valószínűleg nem ilyen nyíltan fog érdeklődni, hanem megtévesztő módon, homályban tartva gyilkos szándékait. Korsgaard szerint a kategorikus imperatívusz első formulája értelmében ebben az esetben is letagadható a barát ottléte, mert a megtévesztő megtévesztése univerzalizálható, a világ működhet olyan módon, hogy a megtévesztőket megtévesztik (a tanulmánya későbbi részében azonban kimutatja, hogy a humanitásformula és a célok birodalma formula nem teszi lehetővé a megtévesztők megtévesztését sem). Korsgaard érvein túl is találhatunk indokokat arra, hogy a „minden körülmények között igazat kell mondani” maximája miért nem fér össze a kategorikus imperatívusszal – jómagam azt tartom perdöntő érveknek, hogy egy olyan világ működésképtelen lenne, ahol a gyilkos kezére kellene adnunk a barátunkat, s nem védhetnénk meg őt a gyilkos megtévesztésével, mert ebben a világban a barátság fogalma nem is lenne értelmezhető, sőt, talán még a felebaráti szeretet, a szolidaritás sem.

A szakadék azonban akkor is létezik az utilitarista és a kantiánus gondolkodás között, ha nem a hazugság kérdésénél van. Akad ugyan olyan szélsőséges nézet, mint az utilitarista J. J. C. Smarté, miszerint Kant valójában szabályutilitarista volt, s Harrod álláspontja valójában Kant nézeteinek kifinomultabb változata (Smart 1973) – de ezt a vélhetően sokakat megbotránkoztató állítást Smart pusztán azért írhatta le, mert a kategorikus imperatívuszt (mint oly sok más utilitarista filozófus) a maga módján értelmezte, amely olyan általános törvényről szól, amelyet egy emberséges és jószándékú ember kívánhat. Smart elismeri, hogy Kant tiltakozna az értelmezése ellen, mégis úgy véli, hogy nem lehetséges Kantot másképp értelmezni (Smart 1973). Smart figyelmen kívül hagyja, hogy a

kategorikus imperatívusz elve *a priori*, amelyet alkalmazunk azután a tapasztalati világra, Harrodnál azonban már maga az elv tartalmazza a tapasztalatiságot. Harrodnál végső soron a boldogság marad továbbra is a helyes cselekedet igazolása, Kantnál a helyes cselekedet igazolása soha nem lehet a nagyobb boldogság. Kant és Harrod filozófiai pozíciója között óriási a különbség. Kant gondolatai inspirálták a szabályutilitarizmust végső soron megalapító Harrodot, ám elméleteik lényegileg különböznek egymástól.

V. ÖSSZEFOGLALÁS

Azon hagyományos – és sok szempontból megalapozott – tankönyvi megközelítés, amely a kanti és utilitarista etikát jellemzően mint egymással szemben álló rendszereket igyekszik bemutatni, nem szabad, hogy elfedje előlünk azt a tényt, hogy a kanti gondolatok hatása nélkül egészen egyszerűen teljesen másképp alakult volna az utilitarizmus története. Nehezen találhatunk olyan nagyobb formátumú utilitarista gondolkodót John Stuart Milltől kezdődően, akire egyáltalán ne hatott volna Kant. Számos utilitarista filozófusnak felkeltette az érdeklődését, hogy Kant rendszerében fontos helye van a boldogságnak, illetve a mások boldogsága elősegítésének. Sokukat lenyűgözte a kategorikus imperatívusz, ahogy Kantot magát is.

Tagadhatatlan, hogy az utilitarista filozófusok hajlamosak voltak Kant gondolatait a maguk módján, Kant szándékaitól teljesen eltérően értelmezni – aminek az egyik példája Roy F. Harrod „kanti elve”, amely ugyan valójában nem kanti elv, de Kant inspirálta elv. Ám jól tudjuk, hogy a félreértések is bizonyulhatnak termékenynek. Kant gondolatai – ha olykor félreértve, félreértelmezve is – adtak lendületet a 20. századi utilitarizmusnak; nemcsak Harrod, s az általa végső soron megállapított szabályutilitarizmus, hanem a jelen tanulmányban terjedelmi okokból csak érintőlegesen tárgyalt John C. Harsanyi, és R. M. Hare munkái révén is.

IRODALOM

- Anscombe, Gertrude Elizabeth Margaret 1958. Modern Moral Philosophy. *Philosophy*. 33. évf. 124. szám. 1–19. <https://doi.org/10.1017/S0031819100037943>
- Bentham, Jeremy 1891. *Fragment on Government*. Oxford, The Clarendon Press.
- Bentham, Jeremy 1977. *Bevezetés az erkölcsök és törvényhozás alapelveibe*. Fordította: Fehér Ferenc. In Ludassy Mária (szerk.) *Brit moralisták a XVIII. században*. Budapest, Gondolat Kiadó.
- Brandt, Richard B. 1959. *Ethical Theory*. Englewood Cliffs/NJ, Prentice-Hall.
- Bujalos István 2007. Az angol utilitarizmus a 19. században. *Magyar Tudomány*. 168/5. 636–641.
- Cholbi, Michael 2009. The Murderer at the Door: What Kant Should Have Said. *Philosophy and Phenomenological Research*. 79/1. 17–46. <https://doi.org/10.1111/j.1933-1592.2009.00265.x>

- Crimmins, James E. 2014. Bentham and Utilitarianism in the Early Nineteenth Century. In Ben Eggleston – Dale E. Miller (szerk.) *The Cambridge Companion to Utilitarianism*. Cambridge, Cambridge University Press. 38–60. <http://dx.doi.org/10.1017/CCO9781139096737.003>
- Gališanka, Andrius 2019. *John Rawls: The Path to a Theory of Justice*. Cambridge/MA, Harvard University Press. <http://doi.org/10.4159/9780674239463>
- Gauthier, David P. 1963. *Practical reasoning*. Oxford, Clarendon Press.
- Hare, Richard M. 1981. Moral thinking: its levels, method, and point. Oxford, Oxford University Press. <https://doi.org/10.1093/0198246609.001.0001>
- Hare, Richard M. 1993. Could Kant Have been a Utilitarian? *Utilitas*. 5/1. 1–16. <https://doi.org/10.1017/S0953820800005501>
- Harrod, Roy F. 1936. Utilitarianism Revised. *Mind*. 45. évf. 178. sz. 137–156. <https://doi.org/10.1093/mind/XLV.178.137>
- Harsanyi, John 1977. Morality and the Theory of Rational Behavior. *Social Research: An International Quarterly*. 44/4. 623–656.
- Heller Ágnes 1997. *Életképes-e a modernitás?* Debrecen, Latin Betűk.
- Heydt, Colin 2014. Utilitarianism before Bentham. In Ben Eggleston – Dale E. Miller (szerk.) *The Cambridge Companion to Utilitarianism*. Cambridge, Cambridge University Press. 16–37. <https://doi.org/10.1017/CCO9781139096737.002>
- Hills, Andrew 2006. Kant on Happiness and Reason. *History of Philosophy Quarterly*. 23/3. 243–261.
- Kant, Immanuel 1974a. Tájékozódni a gondolatok között: mit is jelent ez? In Kant: *A vallás a puszta ész határain belül és más írások*. Ford. Vidrányi Katalin. Budapest, Gondolat. 109–127.
- Kant, Immanuel 1974b. A vallás a puszta ész határain belül. In Kant: *A vallás a puszta ész határain belül és más írások*. Ford. Vidrányi Katalin. Budapest, Gondolat. 134–363.
- Kant, Immanuel 1991a. Az erkölcsök metafizikája. In Immanuel Kant: *Az erkölcsök metafizikájának alapvetése. A gyakorlati ész kritikája. Az erkölcsök metafizikája*. Ford. Berényi Gábor, jegyzeteket kész. Láng Rózsa. Budapest, Gondolat. 295–463.
- Kant, Immanuel 1991b. Az erkölcsök metafizikájának alapvetése. (In: Immanuel Kant: *Az erkölcsök metafizikájának alapvetése. A gyakorlati ész kritikája. Az erkölcsök metafizikája*. Ford. Berényi Gábor, jegyzeteket kész. Láng Rózsa. Budapest, Gondolat. 13–117.
- Kant, Immanuel 2004. *A gyakorlati ész kritikája*. Fordította: Papp Zoltán. Budapest, Osiris kiadó.
- Kant, Immanuel 2009. *A tiszta ész kritikája*. Ford. Kis János – Papp Zoltán. Budapest, Atlantisz.
- Korsgaard, Christine M. 1986. The Right to Lie: Kant on Dealing with Evil. *Philosophy & Public Affairs*. 15/4. 325–349.
- Mill, John S. 1879. *Utilitarianism*. London, Longmans–Green and Co.
- Moore, George E. 1903. *Principia ethica*. Edited by Thomas Baldwin. Mineola/NY, Dover Publications.
- Nordhaus, William D. – Paul A. Samuelson 2016. *Közgazdaságtan*. Budapest, Akadémiai Kiadó.
- Nussbaum, Martha C. 1992. Human Functioning and Social Justice: In Defense of Aristotelian Essentialism. *Political Theory*. 20/2. 202–246. <https://doi.org/10.1177/0090591792020002002>
- Parfit, Derek 1984. *Reasons and Persons*. Oxford/GB, Oxford University Press.
- Rawls, John 1999. *Az igazságosság elmélete*. Ford. Krokavay Zsolt. Budapest, Osiris.
- Scarre, Geoffrey 1996. *Utilitarianism*. New York, Routledge.
- Shaw, William 1999. *Contemporary Ethics: Taking Account of Utilitarianism*. Malden_MA, Wiley-Blackwell.
- Sidgwick, Henry 1962. *The Methods of Ethics*. London, Palgrave Macmillan. <https://doi.org/10.1007/978-1-349-81786-3>

- Smart, John Jamieson Carswell – Bernard Williams 1973. *Utilitarianism: For and Against*. Cambridge, Cambridge University Press.
- Timmermann, Jens 2005. Why Kant could not have been an utilitarian. *Utilitas*. 17/3. 243–264. <https://doi.org/10.1017/S0953820805001639>
- Timmermann, Jens 2014. Kantian Ethics and Utilitarianism. In Ben Eggleston – Dale E. Miller (szerk.) *The Cambridge Companion to Utilitarianism*. Cambridge, Cambridge University Press. 239–57. <https://doi.org/10.1017/CCO9781139096737.013>
- Wood, Allen W. 2001. Kant versus Eudaimonism. In Predrag Cicovacki et al. (szerk.) *Kant's Legacy: Essays in Honor of Lewis White Beck*. Rochester/NY, Boydell–Brewer. 261–282. <https://doi.org/10.1017/9781580466011.013>