

A SZABADSÁG LEHETŐSÉGE

RÉZ ANNA

Bernáth László:

Morális felelősség, érdem és kontroll

A morális felelősség metafizikája

L'Harmattan, Budapest, 2019., 206 oldal, 2500 Ft

Bernáth László vállalkozása egyszerre régimódi és heroikus. Régimódi, mert miközben a szabad akarat kortárs szakirodalma egyre inkább technikai részletkérdések megoldásával és a kívülálló számára értelmezhetetlen „meta-vitákkal” van elfoglalva, Bernáth egy teljes és egységes felelősségelméletet dolgoz ki, annak hagyományos kérdésfelvetéseivel és eszköztárával. És heroikus, mert a szabadakarat-vita – Peter Strawson kifejezésével élve – annyira „sűrűn lakott” területe a filozófiának,¹ hogy a szerzőnek minden egyes új állítás kimondásához komplex és szövevényes vitákon kell keresztülgáznia.

Ezt pedig lelkiismeretesen és magas színvonalon meg is teszi. Ám mielőtt belevágnánk a kötet méltatásába, érdemes megjegyezni: a könyv szakmai színvonala éles (már-már ordító) ellentétben áll annak kiadói kivitelezésével. A szöveg úgy fest, mint amit utoljára egy lelkiismeretes, ám megfáradt szerző olvasott át alaposan: elütések, vesszőhibák, az újraírások során elvétett egyeztetések garmada csökkenti a szöveg hitelét (arra már ki sem térek, hogy egy alapos szerkesztő – a semmilyen helyett – az angolszász szakterminusok magyarítására is értelmes javaslatokat tehetett volna, ezzel komoly szolgálatot téve a szakmának). De a L'Harmattan Kiadó már-már rutinszerű hanyagságát mi sem demonstrálja jobban, mint hogy a fülszöveg helyére a legteljesebb nyugalommal illesztettek be egy olyan szöveget, ami konferencia-absztraktnak is túlon túl körmönfont lenne. A reménybeli olvasó – aki a mostani szak-kiadványi tendenciák mellett már alig több, mint zavaró fikció – így a következő mondattal találkozik először a hátlapon: „A könyv amellel érvel, hogy a morális felelősség nem lehet összeegyeztethető a természeti törvények determinisztikusságával, mivel a morális felelősség feltétele egy olyan erős kontroll, ami csak akkor teljesülhet, ha a cselekvők olyan szubsztanciák, melyek választásai nem előzetesen determináltak.” Igazi kedvcsináló.

És hogy mit jelent? Bernáth az ágens-okságon alapuló *libertarianizmusnak* egy *restriktivista* elméletét dolgozza ki. Menjünk végig sorban! A szabadakara-

rat-vita központi kérdése a következő: vajon létezhet-e szabadság és erkölcsi felelősség egy olyan világban, amelyben az emberi cselekvések – csakúgy, mint az általunk ismert természeti világ eseményei – oksági-
lag determináltak? Az oksági vagy fizikai determinizmus tézise azt mondja ki, hogy a világ egy adott időpontbeli állapotából és a természeti törvényekből együttesen szükségszerűen következik a világ állapota az összes többi időpillanatban. Libertáriánusnak azokat az elméleteket nevezzük, amelyek szerint

1. a szabad akarat *nem* egyeztethető össze a fizikai determinizmus tézisével (ezt hívjuk tágabban inkompatibilizmusnak) és

2. az emberi cselekedetek (vagy legalábbis egy részük) fizikailag nem determináltak, ezért létezik erkölcsi felelősség és szabad akarat. Mindez azonban a Bernáth által kidolgozott elméletre csak feltételesen igaz: Bernáth állítása az, hogy *ha* létezik szabad akarat, akkor ilyen, libertáriánus természetű szabad akaratnak *kell* lennie – amellel azonban nem hoz fel érveket, hogy ez a libertáriánus szabad akarat ténylegesen létezik vagy hogy mai természettudományos tudásunk szerint létezhet-e.

Az ágens-okságon alapuló elméletek a libertarianizmus egyik válfaját alkotják. A másik nagy elméletcsoport az esemény-oksági libertáriánusoké,² akik a következőképp képzelik el a szabad cselekvést: vannak olyan döntési helyzetek, amelyekben egyformán erős érvek, indokok szólnak két döntési lehetőség mellett: vágyaink, céljaink, értékeink nem teszik egyértelművé (és így nem determinálják), mit kellene tennünk.³ Ilyenkor nyílik tér a szabad cselekvésre, amelyet egy indeterminált neurális folyamat alapoz meg: az egymásnak ütköző indokok, vágyak, meggyőződések keveredéséből végül mind az egyik, mind a másik döntés megszülethet.

1 ■ Peter F. Strawson: Szabadság és neheztelés. Ford. Veres Máté. In: Réz Anna (szerk.): *Vétkek és választások: A felelősség elméletei*. Gondolat, Bp., 2013. 239–264. old.

2 ■ Lásd például: Robert Kane: *The Significance of Free Will*. Oxford University Press, 2004.

3 ■ Egy tipikus példa: tegyük fel, épp egy fontos és a karrierem hosszú távon meghatározó megbeszélésre sietek, amikor észreveszek egy vérző, sérült embert az utcán, akinek kizárólag én tudok segíteni – ám ha a segítségére sietek, akkor egész biztosan el fogok késni a megbeszélésről. Hogyan fogok dönteni?

4 ■ Fellow of the Royal Society.

5 ■ Peter van Inwagen: *When is the will free? Philosophical Perspectives*, 3 (1989), 411. old.

Az esemény-oksági elméletek egyik legkomolyabb hátránya, hogy nagyon védtelenek a libertariánus elméletekkel szemben legtöbbször felmerülő szerencse-érvek különböző változataival szemben. A szerencse-érvek így vagy úgy, de mind arra épülnek, hogy amennyiben nem maga a cselekvő, hanem valamilyen, az elméjében lejátszódó indeterminált neurális esemény „hozza meg” végső soron a döntést, akkor azt a cselekvő ugyanúgy nem kontrollálja, mintha determinált oksági eseménysorról lenne szó, csak ebben az esetben nem kényszerítő természeti törvények, hanem a szerencse dönt arról, hogy a személy végül melyik cselekedet mellett teszi le a garast.

Részben a szerencse-érveket kivédendő, az ágens-okság (tegyük hozzá, meghökkentően csekély számban előforduló) védelmezői azt feltételezik, hogy a szabad döntéseket nem a mentális állapotaink okozzák (indeterminált módon), hanem szó szerint *mi*. Ehhez feltételezniük kell a személynek valamifajta egyedi *szubsztanciáját*. A szubsztancia kettős funkciót tölt be az elméletben: egyrészt, összhangban a fogalom arisztotelészi használatával, ez a szubsztancia „hordozza” a tulajdonságainkat (bármit jelentsen is ez), miközben ezen tulajdonságoktól független (bármit jelentsen is ez). Másrészt az ágens-oksági elméletek szerint ez a szubsztancia nem pusztán hordozó, de sajátos hatóképességgel is rendelkezik: ez a hatóképesség pedig nem más, mint hogy a személy (szubsztanciája) képes tulajdonságaiból levezethetetlenül dönteni a felmerülő, metafizikailag nyitott cselekvési lehetőségek között. Így történhet meg, hogy nem vágyaink, indokaink, meggyőződéseink hozzák meg „helyettünk” a döntéseinket, hanem *mi* magunk.

Bár a szubsztancia fogalma nem véletlenül tűnik spekulatív, sőt misztikus metafizikai konstrukciónak (erre a későbbiekben még visszatérek), vitathatatlanul vannak teoretikus előnyei. Ezek közül a legkiemelkedőbb, hogy messze az ágens-oksági elméletek tükrözik a legpontosabban a szabad cselekvés fenomenológiáját: amikor tudatos döntést hozunk, nem pusztán azt gondoljuk, hogy metafizikailag nyitott lehetőségek közül választunk, de azt is, hogy döntésünkkel egy új oksági láncot hozunk létre – és hogy *mi magunk* (és nem a vágyaink, hiteink és indokaink) vagyunk azok, akik ezt az új oksági láncot elindítják.

Mi több: Bernáth elmélete a szabadakarat-vita egy központi és mostanában sokszor elhanyagolt nagy témájához tér vissza: az *akarat* szabadságához. Ahhoz, érvel Bernáth, hogy valóban megérdemljük tetteinkért a hibáztatást, nem elegendő, ha maguk a cselekedeteink szabadok, mondjuk abban az értelemben, hogy megfelelő tudás birtokában, külső kényszer nélkül és saját céljainkat és vágyainkat követve hajtottuk végre őket – ahhoz, hogy a cselekedeteink rajtunk múljanak, szükségünk van valamifajta, az emberi tudatosságot igénylő „meta-kontrollra”, amelynek segítségével szabadon eldönthetjük, hogy *melyik* metafizikailag lehetséges cselekedetet választ-

szuk. A Bernáth által feltételezett szubsztancia (az a tulajdonságainktól, mentális állapotainktól és minden egyébtől független dolog, ami *mi* vagyunk) és annak sajátos hatóképessége pontosan ezt a munkát végzi el: nemcsak dönteni képes metafizikailag nyitott lehetőségek között, de azt is kontrollálni tudja, *hogyan* döntsön.

Már amikor Bernáth elmélete ugyanis *restriktivista*: azt állítja, hogy közel sem minden döntésünk szabad ebben az értelemben. Számos olyan döntési helyzettel nézünk szembe nap mint nap, amikor hiteink, vágyaink, indokaink, az általunk vallott értékek és célok valóban nem adnak teret a szabad döntésnek. Álljon itt Peter van Inwagen Bernáth által idézett példája:

„Nightingale a Royal Society tagja szeretne lenni – az 1980-as évek kedvelt kifejezésével élve, felvételét annyira akarja, hogy már szinte érzi a siker ízét a szájában. [...] Most tegyük fel, hogy szegény Nightingale a választás napján a körmeit rágva ül a szobájában, arról álmodozva, hogy milyen lesz, ha saját magát az »F. R. S.«⁴ címmel illetheti. Ekkor megcsörren a telefon. Felkapja a telefont, beleszól: »Nightingale« – megsüketítve hívóját.”⁵

Ebben az esetben különösebb aggályok nélkül gondolhatjuk úgy, hogy szükségszerű volt, hogy Nightingale felvegye a telefont (bár, talán, méltóságát megőrzendő, várhatott volna még egy-két csörgést) – egyszerűen semmi okunk azt feltételezni, hogy a döntés nem Nightingale vágyaiból és hiteiből, hanem valamifajta magasan kontrollált „meta-döntésből” született.

A restriktivizmus melletti – amúgy észszerűnek tűnő – elköteleződés több szempontból is gyengíti Bernáth elméletét. Egyfelől az elmélet ezen a ponton mégiscsak lemond arról, hogy híven tükrözze a cselekvés fenomenológiáját: talán a legkontrollálatlanabb indulatkitöréseinket leszámítva, minden tudatosan választott cselekedetünket (és így feltehetőleg a fiktív Nightingale is a sajátját) úgy szoktuk megélni, mint ami mellett *mi* döntöttünk, egy új oksági láncot indítva el általa. Bernáth elméletének fényében ebben *időnként* igazunk is van – máskor azonban nem.

Másfelől a restriktivizmus nyilvánvaló gyakorlati és episztemikus problémákat vet fel: mégis honnan tudnánk megmondani, kinek melyik döntése volt metafizikailag szabad, amikor ehhez a tényhez a cselekvőnek magának sincs hozzáférése? És ha nem tudjuk megmondani, akkor mégis hogyan tudnánk a gyakorlatban megkülönböztetni a felelős és a nemfelelős cselekvést? Bernáth szerint szinte sehogy: bele kell törődnünk, hogy mindennapi felelősségtulajdonítási gyakorlatunkban többnyire – mondjuk, amikor letolunk egy démonikus rosszindulatú ügyintézőt vagy artikulálatlanul üvöltünk a sofőr után, aki kis híján elütött a zebra közepén – nincs megfelelő

mennyiségű és minőségű információ a birtokunkban ahhoz, hogy jó okkal feltételezhessük: az illető szabad akaratából cselekedett (vagy cselekedete megfelelő következménye⁶ volt egy korábbi szabad döntésének). Ehhez sokkal többet kellene tudnunk egyfelől az illető céljairól, vágyairól, indokairól (hogy megítélhesük, azok determinálták-e a cselekedetét), másfelől múltbeli szabad cselekedeteiről, hogy kiderülhessen, mostani viselkedése belőlük fakad-e (ebben az esetben ugyanis felelős lenne érte).

De Bernáth ezzel meglegszik: úgy tűnik, még az sem aggasztaná különösebben, ha az elmélete ahhoz a következményhez vezetne, hogy lényegében soha senkit nem tudunk jogosan hibáztatni. Mindez nem azért érdekes, mert alapvető revíziót igényel a morális gyakorlatunkban – az inkompatibilista elméletek többsége ezt várja tőlünk, még hozzá nagyon akceptálható és humánus indokokra hivatkozva. Azért érdekes, mert elárul valamit a projekt alapvető motivációiról: ha egy komoly gonddal és nagy precizitással kidolgozott felelősségelmélet azt állítja, hogy létezik ugyan szabad akarat, de ennek a szabad akaratnak a létezése legfeljebb hébe-hóba jogosít fel minket a felelősségre vonásra, akkor okkal gyanakodhatunk arra, hogy az elmélet tétje egyáltalán nem a felelősségre vonással kapcsolatos. Úgy hiszem, a jelen kötet végső tétje sokkal inkább az ember metafizikai státusa és ebből fakadó méltósága: szabad akaratra nem azért van szükségünk, hogy megalapozhassuk hol magasatos, hol végtelenül kicsinyes indokainkat a másik hibáztatására, hanem azért, hogy olyan típusú létezők lehessünk, akikkel kapcsolatban egyáltalán felvethető, hogy vannak erkölcsi kötelességeik, és akikkel szemben másoknak erkölcsi kötelességeik lehetnek – olyan létezők, akiknek kapcsán van értelme moralitásról beszélni.

Am Bernáth elméletének – mint az ágens-oksági elméleteknek általában – egy ennél súlyosabb eldöntetlenséggel is szembe kell néznie – és feltehetőleg ennek köszönhető, hogy minden vonzereje ellenére az ágens-okság nem örvend nagy népszerűségnek a szabadakarat-vitában.⁷ Természettudományos szempontból már az esemény-oksági libertarianizmus is szintiszta spekuláció: nem azért, mert teljes tudományos képtelenség lenne, hogy indeterminisztikus neurális folyamatok vezessenek a szabad döntéseinkhez – ennek lehetőségét éppenséggel a kvantummechanika teremtette meg –, hanem azért, mert az égvilágon semmilyen tudományos megfigyelés nem utal arra, hogy *ténylegesen* ez lenne a helyzet. Az esemény-oksági elméletek tehát mindössze annyit tudnak felmutatni, hogy mai természettudományos ismereteink szerint nem elképzelhetetlen az általuk tételezett szabad cselekvés, és egy napon akár még az is kiderülhet, hogy ilyenek valóban léteznek. Az ágens-okság elmélete egy lépéssel továbbmegy: az a hatóképesség, amelyet az emberek egyedi szubsztanciáinak tulajdonít, *elvileg is* konfirmálhatatlan. Mivel a szubsztancia és annak tevékenysége defini-

ció szerint levezethetetlen a személy megfigyelhető tulajdonságaiból, így az sem derülhet ki, egyáltalán rendelkezik-e vele.

De a konkrét felelősségelméletig, amelyet Bernáth az utolsó terjedelmes zárófejezetben fejt ki, hosszú út vezet. Bernáth gondolatmenete a következő fő lépéseken megy végig:

1. Erkölcsi értelemben felelősnek lenni valamiért azt jelenti, hogy elviekben⁸ megérdemeljük a hibáztatást miatta. (1. fejezet)

2. Csak olyan dologért érdemelhetjük meg a hibáztatást, amiről tehetünk. (2. fejezet)

3. Csak azokról a dolgokról tehetünk, amelyek kontrollált aktivitásunk eredményei. (2. fejezet)

4. Ennek a fajta kontrollnak magasabb szintűnek kell lennie, mint amelyet gépek képesek kifejteni. (3. fejezet)

5. Ez a „fokozott” vagy *végző közvetlen kontroll* csak akkor jöhet létre, ha a személy

(i) képes olyan változást létrehozni, amely egy másik lehetséges világban, amelyben ugyanaz a kiinduló állapot és a természeti törvények, nem jön létre (*libertariánus kritérium*) és

(ii) képes dönteni arról, hogy a lehetséges hatások közül melyiket fejtse ki (*metakontroll-kritérium*). (3. fejezet)⁹

Az első állítás – amely mellett Bernáth a könyv első harmadában érvel – arra a kérdésre ad választ, hogy egyáltalán mit értünk erkölcsi felelősségen. Bár eléggé evidensnek tűnik, hogy egy felelősségelmé-

6 ■ Bernáth elég fifikás elméletét dolgozza ki a következményekért viselt felelősségnek. Némileg leegyszerűsítve: a cselekvő akkor hibáztatható szabad döntése következményéért (legyen az a következmény egy esemény, egy mentális állapot vagy egy későbbi cselekvés), ha ez a következmény oly módon reprezentálódott a cselekvő számára a szabad döntés pillanatában, mint ami erkölcsileg rossz – és csak *abban a mértékben* hibáztatható, *amennyire* rosszként az a cselekvő számára reprezentálódott.

7 ■ Köszönettel tartozom a recenzeált könyv szerzőjének, amiért felhívta a figyelmemet az itt következő összefüggésre.

8 ■ Az „elviekben” kitétel azért lényeges, mert úgy szoktuk gondolni, hogy sok dicséretes, illetve morálisan irreleváns cselekedetünkért is felelősek vagyunk. Az „elviekben” jelen összefüggésben annyit jelent, hogy *amennyiben* az általunk elfogadott normatív etikai elmélet fényében az adott cselekedet erkölcsileg helytelen, *akkor* hibáztathatók vagyunk érte.

9 ■ A szerző ennél természetesen pontosabban és nagyobb technikai apparátussal fogalmazza ezt meg (148. old.).

10 ■ Az erkölcsi felelősség jelentésének egy nagyon hasonló felosztása mellett érveltem itt: Anna Réz: *Responsibility as Attributability: Control, Blame, Fairness*. Dissertation, CEU, Bp., 2013.

11 ■ Mivel Bernáth több lábjegyzetben is a szememre hányta, hogy nem világos, mit értek doktori disszertáciomban „normatívn”, e helyütt megpróbálom világosabbá tenni az állítástomat. A kérdéses normativitást nagyon gyenge értelemben használtam: amellett érveltem, hogy bármely felelősségelméletnek számot kell tudni adnia arról, hogy mitől és milyen alapon „megfelelő” egy adott válaszreakció (mondjuk a nehezítés vagy a hibáztatás) egy adott cselekedet (vagy mulasztás, mentális állapot stb.) fényében. Ebben az értelemben még azok az újgenerációs naturalista és strawsoniánus elméletek (pl. David Shoemaker: *Response-Dependent Responsibility; or A Funny*

letnek illene meghatározni saját tárgyát, Bernáth kifejtésének mind terjedelme, mind kidolgozottsága kivételes a szakirodalomban. A felelősségelméletek túlnyomórészt azzal a kérdéssel foglalkoznak, hogy milyen körülmények fennállása szükséges és elégséges feltétele az erkölcsi felelősségnek – ám hogy pontosan mit értenek felelősségen, az legtöbbször homályban marad.

Bernáth ezzel szemben nagyon világosan felteszi a felelősség fogalmával összefüggésben felmerülő legfontosabb kérdéseket:¹⁰

1. Milyen értelemben helyes/igazságos/megérdemelt/megfelelő reakció a felelősségre vonás (hibáztatás, büntetés stb.) a felelősség alapját képező dologért (legyen az cselekedet, mulasztás, következmény vagy mentális állapot)?¹¹

2. Kik lehetnek elvben felelősek? Csak individuumok vagy csoportok is?

3. Milyen típusú dolgokért lehetünk elvileg morálisan felelősek? Csak a szándékos cselekedeteinkért, vagy (olykor szándékolatlan) mulasztásainkért, esetleg meggyőződéseinkért és érzelmeinkért is felelősséggel tartozhatunk?

4. Mit értünk felelősségre vonáson? A magunkban elmormolt sommás erkölcsi ítéletektől („Na ez aztán egy undok fráter!”) a hibáztatás érzelmes és kevésbé érzelmes módozatain át az intézményes büntetéssel sok mindennel azonosíthatjuk az erkölcsi felelősséggel összefüggő válaszreakciót.

5. Mi a felelősségre vonás végső alapja, vagy más-
ként fogalmazva: mi az a tulajdonsága az adott cse-

lekedetnek (mulasztásnak, mentális állapotnak stb.), amely feljogosít minket, hogy valakit miatta hibáztassunk? (Bernáth a maga elméletében ezt „érdemalapnak” nevezi.)

Ez utóbbi válik a második fejezet alapkérdésévé, amelyben Bernáth egy hosszú kitérőt tesz, hogy bírálat tárgyává tegye a kompatibilista elméletek egyik kortárs válfaját. A Bernáth által tárgyalt elméletek, melyeket – Thomas Scanlon és Angela Smith szóhasználatát követve¹² – együttesen attribucionizmusként (vagy attribucionalizmusként, a tisztességes magyar fordítás sajnos még várat magára) azonosítanak, lényegében a libertarianizmustól legtávolabb eső elméletcsoport.¹³ Az attribucionisták kompatibilisták (tehát a szabad akaratot összeegyeztethetőnek tartják a determinizmus tézisével); elvetik, hogy a felelősséghez szükség lenne alternatív lehetőségekre;¹⁴ mi több: azt is, hogy a felelősséghez szükség lenne valamilyen magas szintű kontrollra.

Ehelyett úgy gondolják, hogy az erkölcsi felelősség (Bernáth megfogalmazásában) „végső érdemalapja” a személy normatív identitása: azok az értékítéletek (Scanlonnál cselekvési indokokra vonatkozó ítéletek), amelyeket cselekedetei során figyelembe vesz vagy éppenséggel nem vesz figyelembe – és amelyek kifejeződnek nemcsak a cselekedeteiben, hanem szándékolatlan mulasztásaiban, jellemvonásaiban, meggyőződéseiben, érzelmeiben és attitűdjeiben is. Ennek köszönhetően az attribucionista elméletek szerint a felelősség köre meglehetősen tág (ez amúgy összhangban van mindennapi tapasztalatainkkal): nekünk tulajdonítható, és ezáltal felelősségre vonás tárgya minden olyan cselekvésünk és mentális állapotunk, amelyben kifejeződnek értékítéleteink – függetlenül attól, hogy mit tettünk vagy tehetünk volna ezen értékítéletek megváltoztatása érdekében. Ily módon megmagyarázható, miért hibáztatunk másokat jellemgyengeségeikért, rasszista meggyőződéseikért és hanyag figyelmetlenségükért – mindezek ugyanis kifejezik azt, ahogyan ők más emberekről és a hozzájuk fűződő kapcsolatukról ítélnék.

Bernáth számára azért kulcsfontosságú az attribucionista elméletek vizsgálata, mert úgy azonosítja őket, mint amelyek szerint a felelősség végső érdemalapja nem valamifajta aktivitás, hanem egy passzív tulajdonság – márpedig Bernáth éppen ezt tagadja. Azonban nem egyértelmű, hogy az attribucionisták ezt állítanák: hiszen az elmélet szerint *végső soron* nem a cselekvő jellemvonásait vagy érzelmeit (meggyőződéseit, mulasztásait) hibáztatjuk, hanem azokat az értékítéleteket, amelyeket ezek a jellemvonások, érzelmek stb. tükröznek vagy kifejeznek. Márpedig az attribucionista szakirodalomban – erre Bernáth helyesen ki is tér – egyáltalán nem egyértelmű, hogy az értékítéletek passzív adottságok lennének, vagy pedig valamilyen aktív ítélő aktus eredményei. Ám ha az attribucionisták elköteleznék magukat amellett, hogy

Thing Happened on the Way to Blame. *Philosophical Review*, 126 [2017], 4. szám, 481–527. old.) is „normatívak”, amelyek szerint erkölcsileg felelősnek lenni pusztán annyit jelent, hogy a cselekedetre megfelelő (*fitting*) válaszreakció a harag vagy a felháborodás – ebben az esetben magukba a felelősséggel összefüggő erkölcsi érzelmekbe van „beépítve” ez a megfeleltetési viszony. Ha például azt gondoljuk, hogy a harag formális tárgya a másik személy szándékos és normasértő károkozása, akkor a harag csak akkor megfelelő reakció, ha a másik cselekedete ténylegesen megfelel ennek a leírásnak. Semmiképp nem szerettem volna tehát amellet hitet tenni, hogy azok a felelősségelméletek, amelyek nem valamilyen *acti generis* erkölcsi normával igazolják a felelősnek tartást, ne lennének adekvát vagy teljes elméletek. És ebből következően normatív elméletnek tekintem Bernáth érdemelméletét is, akkor is, ha az az érdem egy robusztus metafizikai értelmezésére támaszkodik. (Ez ügyben Bernáth végül nem foglal egyértelműen állást, de érezhetően ezzel az értelmezéssel szimpatizál.)

12 ■ Thomas Scanlon: *What We Owe to Each Other*. Belknap Press of Harvard University Press, Cambridge, 1998. Angela M. Smith: Responsibility for Attitudes: Activity and Passivity in Mental Life. *Ethics*, 115 (2005), 2. szám, 236–271. old., Uő: Attributability, Answerability, and Accountability: In Defense of a Unified Account. *Ethics*, 122 (2012), 3. szám, 575–589. old.

13 ■ Az attribucionista elméletek elődei és közeli rokonai az „igazi én”-elméletek és a „hierarchikus” elméletek, vö. Harry Frankfurt: Az akarat szabadsága és a személy fogalma; Gary Watson: Szabadon cselekedni. Ford. Balogh Zsuzsanna. In: Réz (szerk.): *Vétke*, 23–39. old. és 40–57. old.; Susan R. Wolf: *Freedom within Reason*. Oxford University Press, New York, 1990.

14 ■ Ez nem minden kompatibilista elméletre igaz, lásd Huoranszki Ferenc: *Freedom of the Will: A Conditional Analysis*. Routledge, New York, 2011.

ezek az értékítéletek aktív döntés eredményei, azzal mégiscsak megnyitnák az utat egy erősebb kontrollfogalom felé, és elindulnának az inkompatibilizmus felé vezető úton – márpedig ez nyilvánvalóan nem áll szándékukban. Következésképpen amikor kenyértörésre kerül sor, az attributionisták elismerik, hogy állításuk szerint az emberek akkor is hibázthatók azért, mert önzők, kegyetlenek vagy megbízhatatlanok, ha *egyáltalán nem tehetnek arról*, hogy önzővé, kegyetlenné, megbízhatatlanná váltak – ettől még ugyanis ezek a jellemvonások tükrözik értékítéleteiket.

Bernáth szerint ez viszont nem lehet igaz: már-már analitikus igazságnak tűnik, hogy valaki csak olyan dolgokért érdemli meg a hibáztatást, amelyekről tehet. Ez az elvi, morális meggyőződésünk annyira alapvető,¹⁵ hogy a legtöbb szerző meg sem próbál érvelni mellette (nemhogy ellene). Természetesen erre az ellenvetésre az attributionista szerzők is fel vannak készülve: pontosan ennek az ellenvetésnek a kivédése érdekében úgy szokták jellemezni a hibáztatást, mint ami kiérdemelhető pusztán a tulajdonságaink alapján is. Ennek legjobb példája Thomas Scanlon hibáztatáselmélete,¹⁶ amely a hibáztatást elsődlegesen a két ember közötti kapcsolat változásaként írja le, amelyre az egyik viselkedése vagy tulajdonságai adnak okot. Ha például egy barátom rendre elárulja a bizalmamat, akkor jogos reakció megláztatni a köztünk lévő szoros és intim kapcsolatot; ha egy kollégám soha nem tartja be a szavát, akkor vitathatatlanul megérdemli, hogy ne higgyek az ígéreteinek. Ezek az állítások pontosan olyan típusú érdemviszonyokat fejeznek ki, amilyenekkel Bernáth is operál, az igazságukhoz pedig, úgy tűnik, egyáltalán nincs szükség arra, hogy fontolóra vegyük, miként vált az illető illojális baráttá vagy megbízhatatlan kollégává.

Mivel azzal Bernáth is egyetért, hogy az általa használt érdemfogalom kiterjeszthető a tulajdonságaink alapján megérdemelt reakciókra is, ezért egyetlen választása marad: azt kell tagadnia, hogy mindezen reakciók egyesével vagy együttesen kimerítenék a *hibáztatás* fogalmát. Itt a diszkuszió már-már körkörösé (valójában csak redundánsá) válik: bár Bernáth nem adja teljes meghatározását az általa preferált hibáztatásfogalomnak, annyit leszögez, hogy a hibáztatásnak szükséges eleme az a *gondolat*, hogy az illető tehet arról, amiért hibáztatják. Hogy miért? Azért, mert – ahogy ez az előző bekezdésben is elhangzott – alapvető meggyőződésünk, hogy valaki csak olyasvalamiért érdemelheti meg a hibáztatást, amiről tehet.¹⁷ Ha azonban ez a végső ok, amiért az attributionista elméletek semmiképp nem lehetnek helytállóak, akkor feleslegesnek tűnik az a nagyon alapos és tiszteletre méltó munka, amelyet előzetesen a szerző az attributionista elméletek elemzésére ölt.

De vannak más, járulékos problémák is ezzel a megoldással. Az első a meg nem érdemelt büntetés-sel kapcsolatos: ugyan elképzelhető, hogy a bizalom megvonása, a barátság megszakítása, a munkakapcsolat felmondása önmagukban nem számítanak

hibáztatásnak, de ettől még kétségtelenül hátrányos következményekkel járnak. Ezeknek a reakcióknak pontosan ugyanolyan szankciójellegük van, mint a hibáztatásnak – márpedig Bernáth gondolatmenetét követve nehéz belátni, hogyan lehetne megérdemelt *bármifajta* szankció olyasvalakivel szemben, aki nem tehetett arról, ami miatt szankcionálják. Végső soron úgy tűnik, nem tudjuk elkerülni annak elfogadását, hogy emberek „ártatlanul” (értsd: olyan tulajdonságaik miatt, amelyekről nem tehetnek) szenvedjenek hátrányt.

Mindez egy nagyon fontos különbségre mutat rá Bernáth (és tágabban a libertáriánus elméletek) és az attributionisták (és egyéb, „laza” feltételekkel dolgozó kompatibilisták) között. Az attributionista elméletekben nem pusztán az vonzó, hogy képesek egységesen kezelni és igazolni az erkölcsi felelősség „nem sztenderd” eseteit (amilyen a meggyőződésekért, szándékolatlan mulasztásokért, érzelmekért, jellemvonásokért viselt felelősség) – ami azt illeti, a legtöbbünk szívesen meglenne ezeken a területeken a mostaninál kevesebb morális hibáztatással. Az attributionizmus legfőbb erénye az, hogy a felelősségtulajdonítás és azon belül a hibáztatás gyakorlatát képes szervesen beágyazni mindennapi erkölcsi praxisunk egészébe. Mindennapi gyakorlatainkat elemezve nem tűnik úgy, mintha az erkölcsi hibáztatás valami *minőségileg* különböző dolog lenne, mint a jóindulatú dorgálástól a kapcsolat megszakításáig, a magunkban eldűnnyögött korholó szavaktól az intézményes büntetésig terjedő skálákon elhelyezkedő reakciók bármelyike. Mint ahogy a gyakorlatban az a krisztusi előírás is kilátástalanul bonyolultnak

15 ■ George Sher: Kantian Fairness. *Philosophical Issues*, 15 (2005), 1. szám, 179–192. old.

16 ■ Thomas Scanlon: *Moral Dimensions: Permissibility, Meaning, Blame*. Belknap Press of Harvard University Press, Cambridge, 2008.

17 ■ Ezen a ponton egy kicsit – de tényleg csak egy kicsit – igazságtalan vagyok a szerzővel, aki további gondolatkísérleteket és párhuzamokat is felhoz állítása mellett. Ezek azonban messze nem annyira erősek, mint az alapvető erkölcsi meggyőződésünkre vonatkozó érv.

18 ■ Kivételt jelentenek ez alól a naturalista felelősségelméletek, lásd Peter F. Strawson: Szabadság és neheztelés. In: Réz (szerk.): *Vétkek*, 239–264. old. Am a naturalista elméletek vagy *nem-teljes* felelősségelméletek (amennyiben egyáltalán nem foglalkoznak azzal, hogy mitől erkölcsileg helyes, elfogadható, megfelelő valakit felelősségre vonni), vagy – ahogy Bernáth interpretálja őket – lényegileg *relativista* felelősségelméletek, amelyek azt állítják, hogy a felelősnek tartás feltételei kizárólag az adott közösség aktuális erkölcsi normáitól függenek. Egyik sem túl vonzó opció.

19 ■ Bernard Williams: Morális szerencse. In: Réz (szerk.): *Vétkek*, 335–357. old.

20 ■ Amellett, hogy a felelősséghez szükséges kontroll meghatározásához először a kontroll felelősségtől független értelmét kell feltérképeznünk, én is érveltem, lásd Réz Anna: Befolyásoljuk a meggyőződéseinket? *Magyar Filozófiai Szemle*, 2012. 3. szám, 53–64. old.

21 ■ Nem ez az egyetlen olyan gondolatkísérlet Bernáth könyvében, amely mesterkéltége vagy komplikáltsága folytán nem bírja el azt az érvelési terhet, amelyet a szerző rápakol. Az első fejezetben a felelősség igazságossági elméleteit „cáfolja” Bernáth egy könnyebben elgondolható, de végtelenül mesterkéltd gondolatkísérlettel (lásd. a fogadásos példát, 39–40. old.)

tűnik, hogy csak a vétket gyűlöljük, ne a vétkest – az erkölcsi számonkérés legtöbbször kifejez valamit arról, hogy a másik *milyen*, nem csak arról, hogy *mit választott*. A legtöbb kompatibilista elmélet pontosan az ilyesfajta belátásokat igyekszik megmagyarázni és rendszerbe foglalni.

Természetesen semmi akadályja annak, hogy úgy gondoljuk: mindezen gyakorlatok erkölcsi értelemben alapvetően elhibázottak – ami azt illeti, a legtöbb felelősségelmélet előír *valamifajta* revíziót a mindennapi erkölcsi gyakorlatunkban.¹⁸ Bernáth azonban nem is nagyon *próbál* „kiegyezni” a felelősségtulajdonítás mindennapi tapasztalataival (miközben, némileg paradox módon, azokat az absztrakt pre-filozófiai elveinket, amelyek valamilyen erkölcsi vagy metafizikai elköteleződést vonnak maguk után, nagyon is komolyan veszi). Mind az erkölcsi felelősség etikáját, mind az erkölcsi felelősség metafizikáját tekintve mélyen kantianus projektet visz végig: megpróbálja kihasítani az általunk ismert erkölcsi elvek és metafizikai tézisek közül (mögül, alól?) azt a gondosan lehatárolt területet, azt a – Bernard Williams szavaival élve – „végső menedéket”¹⁹, amelyben az emberi szabadság, és következtetésképpen az erkölcsi kötelesség és az erkölcsi felelősség egyáltalán létrejöhet.

A harmadik nagy fejezetben bontakozik ki a már elemzett libertariánus vízió, amelynek kiindulópontja tehát az, hogy csak azokért a dolgokért érdemlünk hibáztatást, amelyek valamilyen aktivitás eredményei – ez az aktivitás pedig kontrollálja azt, amit teszünk. Az ezt követő eszmefuttatás egyetlen pofonegyszerű, mégis nagyon újszerű gondolatnak, az eszköz-érvnek a kibontása. Az eszköz-érv valójában nem több, mint a következő kérdés: vajon mi lehet az a különbség az emberek és az – akár nagyon bonyolult és finom kontrollt gyakorló – gépek között, ami miatt csak az előbbiekre lehet igaz az, hogy „tehetnek arról”, amit létrehoztak?

Bernáth hiánypótló munkát végez, amikor alapos kifejtését adja annak, mit szoktunk általában kontrollon érteni²⁰ – és ezzel egy időben megalapozza saját hatóképesség-elméletét is. Ráadásul az eszköz-érv annyira magától értetődő, és annyira pontosan adja vissza a determinizmus tézisével kapcsolatos központi aggodalmat (hogy ti. egy determinált világban az emberek sem mások, mint igen bonyolult gépek), hogy az ember azon csodálkozik, hogyhogya ez senkinek nem jutott még eszébe.

Bernáth gondolatmenete végül a kontroll fentebb már tárgyalt libertariánus és metakontroll-feltételeihez lyukad ki: akkor gyakorol valaki a gépekhez képest fokozottabb, a felelősségtulajdonításhoz elégséges kontrollt, ha

(i) képes olyan változást létrehozni, amely egy másik lehetséges világban, amelyben ugyanaz a kiinduló állapot és ugyanazok a természeti törvények, nem jön létre, és

(ii) képes dönteni arról, hogy a lehetséges hatások közül melyiket fejtse ki.

Am megkérdőjelezhető, hogy a fenti két kritérium szükségszerűen vagy akárcsak különösebben organikusán következne a kontrollfogalom ezt megelőző tárgyalásából. Különösen igaz ez a libertariánus kritériumra: amennyiben már amúgy is feltételezzük egy olyan szubsztancia létezését, amely képes a személy mentális állapotaitól és tulajdonságaitól függetlenül eldönteni, hogy melyik alternatívát válassza (ez a metakontroll-kritérium), miért kell még azt is feltételeznünk, hogy ezek metafizikailag (és nem, mondjuk, csak episztemikusan) nyitott alternatívák.

Bernáth érintőlegesen foglalkozik ezzel a felvetéssel, amikor azt a kérdést teszi fel, nem lehetséges-e, hogy az embereknek nem több, csak más típusú kontrolljuk van a cselekedeteik felett, mint a gépeknek, és ez a típuskülönbség az emberi tudatosságból és intencionalitásból ered. Erre egy gondolatkísérlettel válaszol (*Mentális mechanika*, 140. old.):

„Tegyük fel, hogy valamikor a közeljövőben legyártanak egy igen különös klímaberendezést. Ez a klímaberendezés kicsit másképpen működik, mint a maiak. Amikor megnyomjuk a gombjait, nem pusztán az áramköreikben okozunk változást. Ha valaki beüti, hogy 24 Celsius-fokosra szeretné lehűteni a szobát, akkor az áramkörök stimulálásán keresztül vágyakat és mentális célokat alakítunk ki, amelyek a közvetlen okai a klímaberendezés későbbi aktivitásának, amelynek eredménye a szoba 24 fokos lehűtése. A klímaberendezés nem képes saját magának vágyakat vagy célokat kialakítani. Céljai mindig összhangban kell hogy legyenek azon utasításokkal, melyeket a gombokon keresztül kap.”

Bernáth ezzel a gondolatkísérlettel mindössze azt szándékozik megmutatni, hogy a tudatos, mentális, intencionális állapotok jelenléte önmagában sem több kontrollt, sem több szabadságot nem biztosít a cselekvőnek – és ebben nyilvánvalóan igaza is van. Am lássuk be, ez a gondolatkísérlet azért eléggé próbára teszi a fantáziánkat: kérdés, hogy egyáltalán *el tudunk-e gondolni* gépeket, amelyeknek vágyaik és mentális céljaik vannak – na, pláne egy ilyen *klímaberendezést*.²¹ Ezzel összefüggésben pedig az is világos, hogy a *Mentális mechanika* klímaberendezésének működése nem analóg azoknak az embereknek a helyzetével, akik egy fizikailag determinált világban cselekszenek. A klímaberendezés esetében nem tudunk másra gyanakodni, mint hogy tervezői meglehetősen pazarlóan bántak az erőforrásaikkal: ugyan a berendezés vágyai és céljai okozzák a szoba lehűlését, de ezekre a vágyakra és célokra funkcionális szempontból az égvilágon semmi szükség nincs. Ugyanez az emberekre nem igaz: egyfelől nem igaz, hogy egy determinált világban ne lennének képesek

saját magunknak „vágyakat és célokat kialakítani” (bár természetesen ezek a vágyak és célok következni fognak korábbi vágyainkból, céljainkból, meggyőződéseinkből), másfelől az sem igaz, hogy vágyaink és céljaink működése egyszerű redundancia lenne egy determinált rendszerben – bár determinált módon, de tényleges változásokat idéznek elő. Az embereknek ugyanis – szemben a klímaberendezésekkel – senki nem szab célt „kívülről”, épp ellenkezőleg: egy okságilag determinált világot a céltalan, „vak” szükségyszerűség jellemez.

A kérdés kompatibilista oldalról tehát a következőképp hangzik: miben tud többet egy olyan létező, amely determinált módon, de sajátos hatóképességénél fogva, tulajdonságaitól függetlenül dönteni képes a felmerülő episztemikus alternatívák közül, mint egy olyan létező, amelyik ugyanezt indeterminált módon teszi? Noha a kompatibilista ágens-okság meglehetősen motiválatlan elképzelésnek hathat, de teoretikusan nem elképzelhetetlen. Ha pedig így van, akkor az inkompatibilistákon van a bizonyítás terhe a tekintetben, hogy megmondják, mit tesz hozzá a libertariánus feltétel a metakontroll-feltételhez.

Az inkompatibilista erre természetesen válaszolhatná azt, hogy ez az elmélet teljes mértékben megmagyarázatlanul hagyja, miért dönt egy ilyesfajta létező mindig azon episztemikus lehetőség mellett, amelynek megtételére amúgy is determinálva van. Ez teljességgel jogos aggály – csak hogy a libertariánus ágens-oksági elméletekkel kapcsolatban éppígy felvethető. Amennyiben testünk része a fizikai világnak (márpedig ez nehezen tagadható), akkor ágens-okság ide vagy oda, testünk alá van vetve determinisztikus-szükségszerű vagy indeterminisztikus-probabilisztikus természeti törvényeknek. Márpedig, ahogy Derk Pereboom fogalmaz:

„Semmi furcsa vagy hihetetlen nincs abban, hogy *egy konkrét esetben* az ágens-ok szabad választása pontosan arra a cselekvésre esik, amelynek fizikai komponensei szükségszerűen következnek az esemény oksági előzményeiből. De azt már nehezen fogadnánk el, hogy az emberi létezés bármely hosszabb-rövidebb időszakában minden szabad döntés úgy zajlik le, hogy az ágens-okok révén pontosan azt a lehetséges cselekvést választjuk, amelynek fizikai komponensei szükségszerűen következnek az esemény oksági előzményeiből, és soha nem valamelyik másikat. Hacsak nem vagyunk idealisták, és nem gondoljuk úgy, hogy a fizikai világot maguk az ágensek hozzák létre, akkor nem hihetünk ennyi véletlen egybeesésben.”²²

Mindezen érvek és ellenérvek sem most, sem máskor nem fogják eldönteni a kompatibilisták és inkompatibilisták közti vitát. De talán nem is ez a lényeg: hanem az, ahogy egy komplex és szerteágazó elmélet hátterében lassan kirajzolódik az emberi

szabadság és moralitás egy sajátos víziója. Bernáth László új könyvének megjelenéséért nemcsak azért lehetünk hálásak, mert a szabadakarat-vita gazdagabb lett számos eredeti meglátással, korábban elhanyagolt elemzéssel és újszerű érveléssel, hanem azért is, mert a vállalkozás, amelyet végigvisz, a legmélyebb értelemben emberi és filozófiai: annak a fogalmi és metafizikai térnek a feltárása, amelyben az ember szert tehet arra a sajátos erkölcsiségre és méltóságra, amely – úgy hisszük – elidegeníthetetlenül az övé. □

22 ■ Derk Pereboom: Determinizmus al dente. Ford. Bárány Tibor. In: Réz (szerk.): *Vétkek*, 94. old. Habár itt még Pereboom csak a fizikai determinizmusra alkalmazza az érvét, de a későbbiekben megmutatja, hogy a „véletlen egybeesések” problémája akkor is fennáll, ha hiszünk a makroszintű indeterminisztikus folyamatok létezésében.