

## **A tudományos közösség (in)toleranciája a fenomenológiai eliminativizmus, a normatív hibaelmélet, valamint a mesterséges intelligencia morális felelősségének tükrében \***

Ebben a tanulmányban azt a kérdést szeretném feltenni, hogy a tudományos közösség miért nem egyformán reagál olyan elméletekre, amelyek majdnem biztos, hogy hamisak. A majdnem biztosan hamis elméletek némelyikére a tudományos közösség érzékenyen, sőt, sokszor ellenségesen reagál, máskor viszont teljesen közömbös velük szemben, vagy egyenesen támogatja őket. Megvizsgállok néhány kézenfekvő hipotézist, és amellet érvelek, hogy ezek nem működnek. Végül felvetek egy olyan hipotézist, amely ugyan működőképes, de úgy vélem, mindenképpen találni kellene helyette egy másik magyarázatot. Ugyanis e hipotézis szerint a tudományos közösség azért intoleráns olyan – a társadalom tagjainak méltóságát egyébként nem veszélyeztető – elméletekkel szemben, mint a fiatal és a lapos Föld-elmélet, valamint az érvényben lévő jogrendszerek alternatív (és egyben szélsőséges) értelmezését propagáló megközelítések, mert ezek az elméletek alapvetően kétségbe vonják a tudományos közösség szakértelmét.

### I. BEVEZETÉS

Vannak nézetek, amelyek mellett ha valaki nyilvánosan kiáll, akkor kiírja magát a tudományos közösségből. Csak három példát sorolok fel. Ha valaki egy tanulmányban amellet érvel, hogy a Föld csupán hatezer éves, akkor bármilyen magára valamit is adó lap anélkül elutasítja a tanulmányt, hogy tüzetesebben megvizsgálná azt, hogy a szerző hogyan is érvel (nemcsak természettudományos, de

\* A szerző a tanulmány megírásakor a Magyar Tudományos Akadémia Bolyai János Kutatási Ösztöndíjában (BO/00432/18/2), valamint a Nemzeti Kutatási, Fejlesztési és Innovációs Hivatal – NKFIH Posztdoktori OTKA Posztdoktori Kiválósági Program (PD131998) támogatásában részesült. A tanulmányhoz vezető kutatás során a szerző az OTKA K132911 és K123839 projektjeinek, illetve az ELTE BTK Felsőoktatási Intézményi Kiválósági Program „Önvezető autók, automatizáció, normativitás: Logikai és etikai kutatások” kutatócsoportjának támogatását élvezte.

filozófiai lapok is így tennének – már amennyiben „valamit is adnak magukra”). Ehhez hasonlóan, ha valaki amellett érvelne bármilyen tanulmányban, hogy a Föld valójában lapos, csupán a tudósok és a politikusok világméretű összeesküvése miatt hiszi azt mindenki, hogy nem így van, akkor anélkül utasítanák el a tanulmányt, hogy az absztraktot tovább olvassák (és nemcsak a természettudományos, de a filozófiai lapok is). És végül: nem lehet a tudományos közösség által elfogadott lapban közölni egy olyan tanulmányt, amely szerint *a hatályos jogszabályok alapján jogi nonszensz az, hogy az Amerikai Egyesült Államokban adót szednek az állampolgároktól* (sem jogtudományi, sem politikai filozófiai folyóiratban nem lenne esélye egy ilyen tanulmánynak eljutnia a bírálókhoz).

A félreértések elkerülése végett: osztom azt az általánosan elterjedt véleményt, hogy helyes az, hogy a tudományos közösség és az általa működtetett folyóiratok nem adnak helyet ilyen nézeteknek. Azzal a gyakorlattal is egyetértek, hogy bárki, aki nyilvánosan hirdet ilyen nézeteket, kiírja magát a tudományos közösségből. Korábban azt gondoltam, hogy tudom, miért nem jelenhetnek meg ilyen tanulmányok még filozófiai folyóiratokban sem (amelyek, azt hiszem, a folyóiratok közül a legtoleránsabbak abban a tekintetben, hogy milyen pozíciók mellett lehet bennük érvelni). Úgy véltem, hogy azért, mert ezek a nézetek nyilvánvalóan ellentmondanak a bizonyítékoknak, aláássák a tudományos közösségben bevett episztémikus eljárásokat, és nagy a valószínűsége, hogy ha elterjednének ezek a nézetek, annak összességében negatív társadalmi hatása lenne. Ám ahogy jobban megismertem a filozófiai szakirodalmat, fel kellett ismernem, hogy ezeknek a magyarázatoknak egyike sem működik. Tekintélyes filozófiai folyóiratokban *megjelennek* olyan elméletek, amelyek legalább annyira nyilvánvalóan ellentmondanak a bizonyítékoknak, legalább annyira szembenállnak a tudományos gyakorlatban elfogadott episztémikus elvekkkel, és elterjedésük (valószínűleg) legalább annyira káros lenne, mint azoknak a nézeteknek az esetében, amelyeket a tudományos közösség kiátkozott.

A tanulmány első felében ismertetem azokat a filozófiai elméleteket, amelyek ugyanolyan nagy valószínűséggel hamisak, ugyanúgy feszültségben állnak a tudományos kutatásban alkalmazott és elfogadott episztémikus elvekkkel, sőt, elterjedésük (valószínűleg) ugyanúgy negatív társadalmi következményekkel járna, mint a fiatal Föld-elmélet, a lapos Föld-elmélet vagy a hatályos jogszabályok egészen abszurd értelmezése. Mindegyik filozófiai elmélet ismertetése után rámutatok, hogy miért szolgáltat ellenpéldát egy-egy olyan népszerű hipotézissel szemben, amely annak magyarázatára tesz kísérletet, hogy a tudományos közösség miért intoleráns a már említett, általában nem filozófiai természetűnek tekintett elméletekkel szemben. A tanulmány második, rövidebb részében ismertetek egy olyan magyarázatot, amelyet nem cáfolnak a felsorolt ellenpéldák, ám amely helyett szerencsés lenne, ha rendelkeznenk egy másik magyarázattal.

## II. A FIATAL FÖLD ELMÉLET, A FENOMENOLÓGIAI ELIMINATIVIZMUS ÉS A BIZONYÍTÉKOK

Fiatal Föld-elméletnek azt az elméletet szokás nevezni, amely szerint a Föld életkora nem több 12 000 évnél. Az elmélet legfőbb támaszát a Biblia ún. „szó szerinti” értelmezése adja, ugyanis a Föld életkorát a Bibliában található leszármazási táblázatokból és a Teremtés könyvében található teremtéstörténet(ek)-ből vezetik le. Ezt alapul véve az elmélet képviselői igyekeznek olyan alternatív értelmezéseit felvázolni azoknak a bizonyítékoknak, amelyeket geológusok és más tudományterületek kutatói tettek le az asztalra, és amely bizonyítékok nagy erővel mutatnak abba az irányba, hogy a Föld legidősebb kőzeteinek életkora nagyjából 4 milliárd éves. Mivel ez az elmélet azt is állítja, hogy a Földet Isten teremtette néhány ezer évvel ezelőtt, ezt a megközelítést kreacionizmusnak is nevezik (annyiban némileg félrevezetően, hogy a szó szoros értelmében azok a filozófiai és teológiai elméletek is kreacionisták, amelyek szerint mindent Isten teremtett, jóllehet nem mindent közvetlenül – és például az univerzumot több mint 13 milliárd éve).

Ebben a tanulmányban nem az a célom, hogy világosan elmagyarázzam, miért teljesen elhanyagolható az esélye annak, hogy a fiatal Föld-elmélet igaz. Csupán annyit szeretnék röviden megjegyezni, hogy azért szokás a fiatal Föld-elméletet tudományosan reménytelennek tekinteni, mert azok a kormeghatározási eljárások, amelyek megbízhatósága elméletileg jól megalapozott és tesztelhető is, mind abba az irányba mutatnak, hogy a Föld több milliárd éves. Röviden: úgy tűnik, a fiatal Föld-elmélettel a legfőbb probléma az, hogy a bizonyítékok határozott ellentmondásban állnak az elmélettel. Ez alapján fel lehet állítani egy általános érvényű feltevést arra vonatkozóan, hogy a tudományos közösség miért nem tolerál bizonyos elméleteket. Azért, mert a bizonyítékok határozottan cáfolják ezeket az elméleteket.

Bár helyesnek tartom, hogy a tudományos közösség nem tekinti komoly alternatívának a fiatal Föld-elméletet, és azzal a gyakorlattal is egyetértek, hogy ez az elmélet nem jelenhet meg nagy tudományos presztízzsel rendelkező fórumokon, úgy vélem, ezt a gyakorlatot nem magyarázhatjuk azzal, hogy a bizonyítékok határozottan ellentmondanak az elméletnek. Azért nem, mert a filozófiai szakirodalomban találkozhatunk olyan elméletekkel, amelyeknek határozottan ellentmondanak a bizonyítékok, mégis nagy presztízzsel rendelkező lapok és kiadók jelentetik meg őket.

Vegyük a fenomenológiai eliminativizmus példáját. E szerint a filozófiai elmélet szerint *nem léteznek* fenomenális karakterrel rendelkező események, sőt, fenomenális tulajdonságok sem léteznek (Humphrey 2011; Pereboom 2011; Frankish 2016 – bár érdemes megjegyezni, hogy Pereboom 2011 nem áll ki teljességgel az elmélet mellett). Mivel talán az olvasóban felmerül a gyanú, hogy rosszul foglaltam össze az elméletet, álljon itt egy idézet az elmélet egyik képviselőjétől:

A tudatosság elméletei általában komolyan veszik a kemény problémát. Elfogadják, hogy a fenomenológiai tudatosság valós és kísérletet tesznek arra, hogy megmagyarázzák, létezése hogyan lehetséges. Ám van egy másik megközelítés is, amely szerint a fenomenológiai tudatosság illuzórikus, és arra tesz kísérletet, hogy megmagyarázza, miért tűnik úgy mégis, hogy a fenomenológiai tudatosság létezik. Ezt a megközelítést hívhatjuk a fenomenális tudatossággal kapcsolatos eliminativizmusnak. Viszont ez a kifejezés nem ideális, mivel azt sugalja, hogy a fenomenális tudatosságba vetett hit egy egyszerű teoretikus hiba, és hogy a fenomenológiai realizmus tagadása csupán egy része a népi pszichológia általában vett tagadásának, valamint hogy a fenomenális tulajdonságokról való beszédnek nincs semmilyen szerepe sem – ezek azonban olyan állítások, amelyek egyáltalán nem lényegi elemei a fenomenológiai tudatossággal kapcsolatos eliminativizmusnak. Egy másik címke az „irrealizmus”, de ennek is vannak szerencsétlen konnotációi; az illúziók maguk nagyon is valóságosak és nem elhanyagolható befolyással rendelkezhetnek. (Frankish 2016, 11.)

Összefoglalva: a fenomenológiai eliminativizmus szerint úgy tűnik, hogy a tapasztalatoknak van fenomenális jellegük, de valójában nem rendelkeznek ezzel. Jóllehet *annak illúziója*, hogy tudatos tapasztalataink rendelkeznek fenomenális karakterrel, nagyon is valós (ahogy valós az az illúzió is, hogy a vízbe mártott evező megtörik a vízben), tapasztalataink fenomenális jellege egyáltalán nem valóságos (mint ahogy az sem valóságos, hogy az evező megtörik, amikor a vízbe bemártjuk).

Érdemes elhatárolni az ebben az értelemben vett fenomenológiai eliminativizmust attól a nézettől, amely szerint a fenomenális karakterrel rendelkező eseményeket meglévő fenomenális fogalmainkkal rosszul ragadjuk meg. Daniel Dennett például amellet érvel, hogy a *kválé* fogalmával nem ragadhatóak meg tudatos tapasztalataink, mivel a tudatos tapasztalatainkra alkalmazva a *kválé* fogalmát azt az eredményt kapjuk, hogy tudatos tapasztalatainkhoz esszenciálisan hozzátartozik azok fenomenális karaktere, valamint hogy tévedhetetlenek vagyunk tudatos tapasztalataink típusának megállapításával összefüggésben (Dennett 1988). Még ha valaki egyet is ért Dennett-tel abban, hogy egyszerre nem lehet igaz az, hogy a fenomenális tulajdonságok a tudatos tapasztalatok esszenciális tartozékai, és az, hogy tévedhetetlenek vagyunk azzal kapcsolatban, milyen típusú tudatos tapasztalatban van részünk, nem kell, hogy azzal is egyetértsen, hogy a tudatos tapasztalatoknak soha nincsenek fenomenális karakterei. Elképzelhető, hogy csupán nem vagyunk tévedhetetlenek azzal kapcsolatban, hogy milyen fenomenális karakterrel rendelkezik egy tudatos tapasztalat, illetve elvben az is lehetséges, hogy tudatos tapasztalatainkat nem azoknak a fenomenális karaktere individuálja.

Úgy vélem, hogy ha fenomenális eliminativizmus alatt szigorúan azt az elméletet értjük, amely szerint még tudatos tapasztalataink sem rendelkeznek fenomenális karakterrel, mert nem léteznek fenomenális tulajdonságok, akkor ki

sem lehetne találni olyan elméletet, amelynek határozottabban ellentmondanak a bizonyítékok. Ugyanis közvetlenül megjelennek számunkra a tudatos fenomenális események. Nem akarom ezt a jól ismert megfigyelést túlságosan részletezni: tudatos tapasztalataink *valamilyenek*, valamilyen őket átélni, és *közvetlenül tudatában* vagyunk annak, hogy ez így van, még akkor is, ha fogalmainkkal nehéz leírni, hogy tudatos tapasztalatainknak ez a valamilyensége pontosan milyen. Minden egyes tudatos pillanat, amit átélünk, igen erős bizonyítékot szolgáltat arra, hogy a fenti értelemben vett fenomenológiai eliminativizmus hamis.

Persze a fenomenális eliminativisták is tudatában vannak ennek a – hogy egészen finoman fogalmazzak – nehézségnek. Frankish a következőképpen válaszol rá:

Amennyiben a realisták továbbra is kitartanak amellett, hogy a fenomenális tudatosság egy magyarázatra szoruló adat, akkor azt kell mondaniuk, hogy episztemikus hozzáférésünk a fenomenológiai tudatossághoz különleges, és kizár bármiféle tévedést. És mivel nincs olyan oksági folyamat, amely ilyen bizonyosságot szolgáltathatna, azt kell mondaniuk, hogy ez a hozzáférés nem okságilag közvetített. Néhány realista tényleg ezt is állítja. Chalmers szerint közvetlen ismeretségben vagyunk a fenomenológiai tulajdonságokkal [...]. [...] Ez a nézet megvédi a fenomenális tudatosság státusát mint magyarázatra szoruló adatot, de ennek nagy ára van. Először is, az ismeretségnek nem lehet fizikai jelentősége. Ahhoz, hogy beszéljünk vagy gondolkodjunk saját fenomenális tulajdonságainkról, arra van szükség, hogy mentális reprezentációkat alakítsunk ki róluk, és mivel a reprezentacionális folyamatok fallibilisak, az ismeretség által biztosított bizonyosság soha sem adható tovább – amennyiben kognitív rendszereknek tekintjük őket – másoknak vagy akár magunk számára. Az ára annak, hogy a tudatosságot egy magyarázandó adattá tesszük, az, hogy ez az adat pszichológiai értelemben nem csinál semmit. Másodsor, az ismertség-elmélet azt feltételezi, hogy a reakciók és asszociációk, amelyeket egy érzékelt esemény idéz fel, nem befolyásolja azt, hogy milyenek érezzük az érzékelt eseményt, mivel nem vagyunk közvetlen ismeretségben sem e reakciókkal és asszociációkkal, sem ezek hatásaival. Ám jó okunk van azt gondolni, hogy reakcióink és asszociációink befolyásolják azt, hogyan érzékeljük saját tapasztalataink milyenségét [...].

[...]

Az ismertség elmélet ráadásul még nagy metafizikai teherrel is jár. Nehéz látni, hogy fizikai tulajdonságok mégis hogyan mutatkozhatnak meg közvetlenül számunkra, így az elméletnek valószínűleg szüksége van a fenomenális tudatosság valamilyen antifizikalista felfogására. Ráadásul talán még a tapasztaló szubjektum antifizikalista értelmezésére is szüksége lehet. Amennyiben a megismerők összetett fizikai organizmusok, hogyan is kerülhetnének közvetlen ismertségbe a fenomenális tulajdonságokkal? (Frankish 2016. 15.)

Tulajdonképpen Frankish azzal válaszol arra a felvetésre, hogy a fenomenológiai karakterrel rendelkező tudatos események olyan tények, amelyeket nem tagadhatunk le és magyarázat után kiáltanak, hogy rámutat arra, mennyi teoretikus nehézséget okoz, ha elfogadjuk a fenomenológiai karakterrel rendelkező tudatos mentális események létezését. Aggályai így foglalhatóak össze: amennyiben elfogadjuk a fenomenális karakterrel rendelkező tudatos események létezését, akkor nehezen tudjuk magyarázni, hogy ezek hogyan viszonyulnak más mentális állapotainkhoz, vagy hogy ezek létezése hogyan volna összeegyeztethető az ember fizikalista felfogásával. Fogadjuk el, hogy Frankishnek igaza van: ezeken a területeken nincsenek kielégítő elméleteink. Az azonban, hogy nincsenek kielégítő elméleteink valamiről, nem ok arra, hogy tagadjuk azt a jelenséget, ami ott van az orrunk előtt (vagy mögött): tudatos tapasztalataink valamilyenek.

Frankish – miként a fenomenológiai eliminativizmus minden képviselője – arra gondol, hogy bármennyire is nyilvánvalónak tűnjék az, hogy vannak olyan tudatos tapasztalataink, amelyek valamilyenek, vannak még biztosabb igazságaink, amelyek bizonyosan kizárják a fenomenális karakterrel rendelkező tudatos események létezését. De hogyan lehetne bizonyosabb a fenomenológiai karakterek létezésénél az, hogy – teszem azt – a fizikalizmus igaz, és hogy a fizikalizmus kizárja a fenomenális karakterrel rendelkező események létezését? Erre ezt lehetne válaszolni: nem zárhatjuk ki, hogy Frankishnak vagy valaki másnak egyszer csak sikerül ezt megmutatni. Ezért végső soron nem egészen nyilvánvaló, hogy a fenomenológiai eliminativizmus szembemegy a bizonyítékokkal. És ezért lehetséges az, hogy a fenomenális eliminativizmussal szemben a tudományos közösség toleráns. És a tudományos közösség azért intoleráns az olyan elméletekkel szemben, mint a fiatal Föld-elmélet, mert egészen bizonyos, hogy ezek nem egyeztethetőek össze a bizonyítékokkal.

Ez a fajta mentegetése a tudományos közösség intoleranciájának elhibázott. Gondoljunk csak bele: elvi szinten bármelyik elmélet összhangba hozható a bizonyítékokkal, ha előállunk olyan erős érvekkel, amelyek döntően kikezdi bizonyos alapvető háttérfeltevéseinket. Ha például valaki előállna egy olyan nagyon erős érveléssel, amely már-már teljesen bizonyossá tenné, hogy létezik egy olyan istenség, amelyik a Bibliát diktálta, és amely a világot úgy rendezte be, hogy a tudósok szokásos módszereiket követve hamis következtetésekhez jussanak, akkor – miután sikerült ez a bizonyítás – *többé nem mondhatnánk*, hogy a fiatal Föld-elmélet ellentmond a bizonyítékainknak. Csakhogy joggal feltételezhetjük, hogy soha senki sem fog olyan erős érveléssel előállni egy ilyen istenség létezése mellett, amely felülírhatná a józan ész és a tudomány által szolgáltatott bizonyítékainkat. Ugyanígy joggal feltételezhetjük, hogy nem várható olyan erős filozófiai érv a fenomenális tulajdonságokat kizáró fizikalizmus mellett, amely képes lenne felülírni azt a közvetlen bizonyosságunkat, hogy léteznek fenomenális karakterrel rendelkező tudatos tapasztalatok.

Ezért úgy vélem, más magyarázat után kell néznünk akkor, ha azt akarjuk megérteni, hogy a fenomenológiai eliminativizmus mellett miért lehet érvelni Q1-es besorolású lapokban, a fiatal Föld-elmélet mellett viszont miért nem. A tudományos közösség intoleranciáját nem magyarázhatjuk főként és pusztán azzal, hogy a fiatal Föld-elmélet szemben áll a rendelkezésünkre álló bizonyítékokkal, hiszen a fenomenológiai eliminativizmus ugyanannyira, ha nem jobban szemben áll velük.

### III. A LAPOS FÖLD-ELMÉLET, A NORMATÍV HIBAELMÉLET ÉS AZ EPISZTEMIKUS ELVEK

Talán azzal lehetne érvelni, hogy a fiatal Föld-elmélet és a fenomenológiai eliminativizmus között az igazi különbség a következő: a fiatal Föld-elméletet lehetetlen anélkül képviselni, hogy ne kellene elfogadnunk valamiféle összeesküvés-elméletet. A fiatal Föld-elmélet híveinek, látva az erős tudományos konszenzust, végül – így vagy úgy – arra kell utalniuk, hogy a tudományos közösség szándékosan nyomja el őket vagy juttatja őket hátrányos helyzetbe. Ez még inkább megfigyelhető a lapos Föld-elmélet hívei esetében. Szerintük a tudományos közösség nem egyszerűen háttérbe szorítja a lapos Föld-elméletet és az elmélet mellett szóló erős érveket, de szándékosan teszi ezt, azzal a céllal, hogy valamilyen uralkodó elit és a saját érdekeit kiszolgálja. A fenomenológiai eliminativizmus esetében viszont fel sem merül az, hogy az elmélet koherens kifejtéséhez szükség van valamiféle összeesküvés-elméletre. A fenomenológiai eliminativizmus hívei szerint elméletük viszonylagos népszerűtlenségét nem kollégáik összeesküvése, hanem az agy által létrehozott (nem fenomenális természetű) illúzió okozza.

De miért is annyira problematikusak az összeesküvés-elméletek? E kérdésnek egy külön tanulmányt (vagy akár könyvet is) lehetne szentelni. Hadd éljek egy saját javaslattal: az összeesküvés-elméletek fő problémája, hogy *szerkezetük* miatt eleve lehetetlen velük szemben indokokat felhozni. Tulajdonképpen nem is mindegyik összeesküvés-elméletet tartjuk önmagában episztemikusan problematikusnak. Ha például valaki történeti bizonyítékok alapján érvel amellett, hogy Kennedy egy összeesküvés áldozata lett, s közben nagyjából világos, hogy az elmélet körülbelül miféle történeti bizonyítékok fényében omlana össze kártyavárként, akkor sem az elmélet, sem az elmélet képviselője nem tűnik patológikusnak. Ha viszont az elmélet kifejtése során nyilvánvalóvá válik, hogy nincs olyan bizonyíték, amely akár elviekben megingathatná az elméletet, mivel egyértelműen kiderül, hogy az elmélet minden ellenbizonyítékot eleve úgy kezel, mint amellyel kapcsolatban eleve bizonyos, hogy csupán *látszólagosan szólhat* az elmélet ellen, akkor az elmélet és a képviselője már valóban patológikusnak látszik. Az összeesküvés-elméletek esetében ez akkor válik általában nyilván-

valóvá, amikor kiderül: az elmélet olyan nagy hatalommal rendelkező ágensek összeesküvését feltételezi, akiktől éppen arra számíthatunk, hogy a lehetséges összes *látszólagos ellenbizonyítékot* megtevesztő szándékkal a valóságba plántálják. Ezért a szóban forgó összeesküvés ellen felhozható bizonyítékok még sokszor meg is erősítik az elméletet – legalábbis az elmélet képviselői szerint. A lapos és a fiatal Föld hívei is, amikor kifogynak az érvekből, gyakran így viselkednek: felhívják a figyelmet arra, elméletük milyen sok érdeket sért, és sejtelmesen (vagy kevésbé sejtelmesen) utalnak arra, hogy a tudományos közösség legerősebb ellenérvei alighanem az adatok szándékosan torz bemutatásán alapulnak.

A fenomenológiai eliminativizmus nem immunizálja ilyen mértékben magát a lehetséges ellenvetésekkel szemben. Ha például az agykutatás valamiképpen bizonyítékokkal szolgálna arra, hogy az agy működése nehezen érthető nem anyagi jellegű entitások feltételezése nélkül, akkor a fenomenológiai eliminativizmus képviselői alighanem elvetnék az elméletüket. Innen nézve úgy tűnik, ígéretes lehet a következő felvetés: a fenomenológiai eliminativizmus irányában a tudományos közösség azért toleránsabb, mint a lapos és a fiatal Föld elmélettel szemben, mert az előbbi nem immunizálja magát olyan nagy mértékben a lehetséges ellenvetésekkel szemben, mint az utóbbiak.

Ám léteznek más olyan filozófiai elméletek, amelyek még az összeesküvés-elméleteknél is jobban immunizálják magukat bármilyen ellenvetéssel szemben. Ilyen elmélet az ún. normatív hibaelmélet (Husi 2013; Streumer 2013, 2017; Olson 2014; Cowie 2016). A normatív hibaelmélet (amely elméletet az angol nyelvű szakirodalomban olykor „reason eliminativism”-nek vagy „error theory”-nak is neveznek) a következőképpen definiálható:

Ahogy mondtam, a legtöbb kognitivistá realista. De létezik egy olyan változata a kognitivizmusnak, amely nem realista: a hibaelmélet, amely szerint a normatív ítéletek olyan vélekedések, amelyek normatív tulajdonságokat tulajdonítanak valaminek, miközben normatív tulajdonságok nem léteznek. (Streumer 2017. 6.)

Más szavakkal: nem léteznek normák, csak tények (amelyek – természetesen – nem normatívak). Így – az elmélet szerint – nem igaz az *simpliciter*, hogy „lopni helytelen”, hiszen akkor ez egy létező normát fejezne ki. Legföljebb – az elmélet értelmében – csak azt tekinthetjük igaznak, hogy „a jelenleg társadalmilag elfogadott normarendszer szerint lopni helytelen”, hiszen ez egy nem normatív tényt fejez ki. Ugyanis abból a tényből, hogy milyen normarendszert fogad el a társadalom, önmagában semmi sem következik arra vonatkozóan, hogy mit kellene tennünk, vagy hogy milyen normák léteznek valóságosan. Ami következik ebből a mondatból, az csupán annyi, hogy „lopni helytelen a jelenleg elfogadott normarendszer szerint”, és e kijelentés igazságát a normatív hibaelmélet képviselői készséggel el is fogadják. Amit tagadnak, az annak a mondatnak a *simpliciter* igazsága, hogy „lopni helytelen”.



Amiben a morális hibaelmélettől különbözik a normatív hibaelmélet az az, hogy nem csupán a morális normatív tulajdonságok létezését tagadják, hanem az összes normatív tulajdonság létezését. Streumer az alábbi módon fogalmazza meg mit ért normatív tulajdonságokon:

A „normatív tulajdonság” kifejezés az alábbiakban helyettesíthető a „helyes”, „helytelen”, „megengedhető”, „jó”, „rossz”, „racionális”, „valamire indokot szolgáltatni” vagy bármilyen egyéb tulajdonsággal, amely ekvivalens vagy magában foglalja ezeknek a tulajdonságoknak az egyikét [...] (Streumer 2017. 106).

Amit ehelyütt külön érdemes kiemelni: a normatív hibaelmélet tagadja, hogy léteznének episztemikus indokok, azaz tagadja, hogy léteznének olyan indokok, amelyek alapján valamilyen elméletet el kellene fogadni vagy el kellene vetni. Soha nem igaz az a mondat, hogy a „Rendelkezésre álló bizonyítékok fényében *E* elméletet el kell vetni”, legföljebb az, hogy „A jelenleg elfogadott episztemikus normák keretén belül a rendelkezésre álló bizonyítékok fényében *E* elméletet el kell vetni”. Ám ez egy paradoxonhoz vezet: ha a normatív hibaelmélet igaz, akkor nem léteznek olyan indokok, amelyek alapján a normatív hibaelméletet el kellene fogadni. Küzdve ezzel a paradoxonnal, van, aki azt állítja, hogy végső soron nem is igazán lehet a hibaelméletben hinni (Streumer 2013, 2017), és vannak, akik azt állítják, hogy bőven elég, hogy a jelenleg elfogadott episztemikus normák fényében el kellene fogadnunk a hibaelméletet, mert *de facto* mindnyájan elfogadjuk ezeket az episztemikus normákat (Cowie 2016; Husi 2013).

Ám most nem a normatív hibaelmélet ezen problémájára koncentrálok, hanem arra, hogy ha valaki elfogadja a normatív hibaelméletet, akkor az egyszersmind azt is jelenti: nem hiszi, hogy a bizonyítékok vagy bármilyen érv megdöntheti az elméletet. Az elmélet kizárja, hogy létezhessenek olyan megfontolások, amelyek fényében az elméletet fel kell adni, legfeljebb olyan megfontolások létezhetnek, amelyeket bizonyos körök episztemikus normái szerint fel kell adni. Ha jobban megnézzük, a normatív hibaelmélet bevallottan kizárja azt, hogy ellene döntő érveket lehessen felhozni, míg az összeesküvés-elméletek ilyen messzire nem mennek el. Az összeesküvés-elméletek szerkezetét és híveik gondolkodásmódját megvizsgálva csupán következtethetünk arra, hogy ezek az elméletek végeredményben eleve kizárják, hogy az elméletet keretén belül bármit is erős ellenvetésként azonosíthassunk. Ennyiben a normatív hibaelmélet még az összeesküvés-elméleteknél is magasabb szinten immunizálja magát a lehetséges ellenvetésekkel szemben.

Fel lehetne vetni a következő érvet: fontos különbség az összeesküvés-elméletek és a normatív hibaelmélet között, hogy az előbbi tipikusan szándékos cselekvésnek tekinti azt, hogy egyes csoportok normatív autoritással rendelkező normarendszernek igyekeznek beállítani saját csoportjuk normarendszerét,

míg a normatív hibaelmélet még csak nem is sejteti, hogy az egyes csoportok által elfogadott normarendszerek abszolutizálása szándékos lenne. Bár ez a különbség számos tekintetben fontos, azzal kapcsolatban irreleváns, hogy vajon a tudományos közösség miért intoleráns az összeesküvés-elméletekkel szemben, miközben helyet ad a normatív hibaelméletnek. A tudományos közösség megtűri a szándékos magyarázatokat, gondoljunk csak például a történelmi magyarázatokra, ahol igazság szerint a bizonyítékokkal alátámasztott összeesküvés-elméleteknek is van helyük.

Egy másik ellenérv szerint az a különbség a rosszfajta összeesküvés-elméletek és a normatív hibaelmélet között, hogy az utóbbit képviselői feladnák, amennyiben – például – erős érvekkel szolgálnánk a fizikalizmussal szemben, amely filozófiai keret az egyik legfontosabb motivációja a normatív hibaelméletnek. Ezzel szemben a paranoid összeesküvés-elméletek hívei semmilyen érv hatására nem adnák fel álláspontjukat. Ezzel az érveléssel a fő gond az, hogy azt feltételezi: az összeesküvés-elmélet hívei nem képesek feladni az álláspontjukat. De ez nem igaz. Képesek rá, ha valamiért hajlandóak végre nem a saját elméletük keretén belül tekinteni a bizonyítékokra. Ám sajnos csak ugyanebben az értelemben lehetnek képesek a normatív hibaelmélet védelmezői felismerni azt, hogy elméletüket valamilyen érv miatt fel kell adniuk. Hiszen amíg az elméletük keretén belül tekintenek bármilyen érvre, addig azokról azt gondolják: eleve nem is lehet olyan normatív ereje semmilyen érvnek vagy megfontolásnak, aminek alapján bármilyen elméletet fel kellene adni.

#### IV. ABSZURD JOGÉRTELMEZÉSEK, AUTONÓM GÉPEK MORÁLIS FELELŐSSÉGE ÉS A KÁROS HATÁSOK PROBLÉMÁJA

Még egy lehetséges felvetést vizsgálnék meg, mielőtt rátérnék arra a válaszra, amelyet ugyan problematikusnak tartok, de a legvalószínűbbnek. E lehetséges felvetés szerint a tudományos közösség ellenségessége a fiatal Föld- és a lapos Föld-elmélettel szemben arra vezethető vissza, hogy ezeknek az elméleteknek roppant káros társadalmi hatása lenne, ha elterjednének. Bár talán ezekben az esetekben nem egészen nyilvánvaló – érvelhetne valaki –, hogy a tudományos közösség intoleranciája erre vezethető vissza, vannak esetek, amelyek ezt az aspektust egészen nyilvánvalóvá teszik. Vegyük például azokat a jogi okoskodásokat, amelyek szerint az Amerikai Egyesült Államokban a *hatályos törvények* alapján szövetségi (vagy olykor más jellegű) adókat nincs joga szedni az államnak. Elég könnyen belátható, hogy ha ez a vélekedés elterjedne és uralkodóvá válna, akkor az katasztrofális következményekkel járna az Amerikai Egyesült Államok társadalmára és gazdaságára nézvést, és közvetve beláthatatlan következményekkel járna más társadalmakra és a világgazdaságra is. A tudományos közösség társadalmi felelősségének tudatában nem ad helyet sem jogtudományi

sem filozófiai szaklapokban olyan érveléseknek, amelyek ezt és más veszélyes jogi ötleteket propagálnának.

Fontos tisztázni: nem arról van szó, hogy e megközelítés szerint a tudományos közösség azokat az elméletekkel szemben intoleráns, amelyeknek várhatóan rossz hatásuk lenne, ha elterjednének. Hanem arról, hogy a tudományos közösség azon valószínűtlen vagy episztemikusan problematikus elméletekkel szemben intoleráns, amelyekre az is igaz, hogy valószínűleg rossz hatással lennének a társadalomra.

Ezzel a felvetéssel a legnagyobb probléma az, hogy végső soron nem ad jó magyarázatot arra, hogy a tudományos közösség miért nem intoleráns például a fenomenológiai eliminativizmussal vagy a normatív hibaelmélettel szemben. Mire számíthatnánk, ha a társadalomban nagymértékben elterjedne az a vélemény, hogy nem is létezik a fájdalomnak az a kellemetlen karaktere, amely összekapcsolódik a fájdalom tapasztalatával? Vagy ha az emberek nem hinnének többé abban, hogy léteznek olyan erős érvek, amelyek alapján valakinek egy nézetet a bizonyítékok fényében el kellene vetnie vagy valamilyen cselekvést véghez kellene vinnie? Józan ésszel gondolhatjuk-e azt, hogy ezeknek a vélekedéseknek az elterjedése nem járna negatív társadalmi hatással?

Egyesek talán azt válaszolnák egy ilyen felvetésre, hogy semmilyen döntő bizonyíték nincs arra, hogy e nézetek elterjedése végül rossz eredményekhez vezetne. Lehetne például úgy is érvelni, hogy e filozófiai irányzatoknak esélyük sincs általánosan elterjedni, mert annyira összeférhetetlenek alapvető késztetéseinkkel és gondolkodásmódunkkal. Ugyanakkor azt sem lehet teljesen kizárni, hogy az Egyesült Államokban az adóztatás jogszerűtlenségével kapcsolatos nézetek elterjedése egy olyan jogi és társadalmi átalakuláshoz vezethetne, amely hatása összességében pozitív lenne. Néhány igazán rettenetes tézist leszámítva *egészen biztosak* soha sem lehetünk abban, hogy egy-egy nézet elterjedése hosszú távon milyen lenne, ezért e vélekedések elterjedése kapcsán csupán valószínűségekről beszélhetünk.

Egy másik, az előbbivel rokon válaszlehetőség a következő. Ha a társadalom jelentős része elfogadná a fenomenológiai eliminativizmust vagy a normatív hibaelméletet, akkor a társadalom tagjai oly mértékben másképpen gondolkodnának, mint jelenleg, hogy merő spekuláció ezeknek a nézeteknek a társadalmi hatásaival kapcsolatban kijelentéseket tenni. Ezzel szemben – így az érv – a lapos Föld- és a fiatal Föld-elméletek elfogadása vagy az Egyesült Államok jogrendszerének (félre)értelmezése nem igényli az egyén és a társadalom teljes hitrendszerének revízióját, csupán azt, hogy egy specifikus területen változzon meg az emberek jelenlegi gondolkodásmódja. Ezért a lapos/fiatal Föld elméletek és az alternatív jogi interpretációk esetében joggal feltételezhető, hogy elterjedésük hatása negatív lenne, míg a fenomenológiai eliminativizmussal és a normatív hibaelmélettel kapcsolatban ez merő spekuláció.

Ez a válasz némileg jogos, de nem képes megvédeni azt a feltételezést, amely a tudományos közösség intoleranciáját egyes elméletekkel szemben azok (várható) társadalmi hatásával magyarázza. Léteznek ugyanis olyan filozófiai elméletek, amelyek nemcsak nehezen összeegyeztethetőek a bizonyítékokkal, de várható társadalmi hatásuk is valószínűleg negatív lenne, ráadásul elfogadásukhoz nem kellene olyan alapvető vélekedéseinket megváltoztatni, minthogy a tudatos tapasztalatoknak van fenomenális karaktere, illetve, hogy bizonyos indokoknak van a konvencióktól független normatív erejük. Egy ilyen elmélet az, amely azt állítja: olyan mesterséges autonóm ágensek is rendelkezhetnek morális felelősséggel, amelyek nem tudatosak. Nemrégiben Jaap Hage (2017) érvelt amellet a nézet mellett, amely szerint a morális és jogi felelősséggel összefüggésben teljesen lényegtelen, hogy a gépi autonóm ágensek *valójában* (a társadalmi gyakorlatoktól függetlenül) rendelkeznek-e ágenciával, szándékkal vagy szabad akarattal. Egy entitás (legyen ez gépi vagy biológiai) akkor számít morálisan felelős ágensnek, ha jó következményei lennének annak, ha morálisan felelősnek tartanánk őket (hasonló nézeteket vall Dennett 1998; Floridi-Sanders 2004; Stahl 2006; Sullins 2006; Hellström 2013). Hage helyenként egészen radikálisan fogalmazza meg ennek a megközelítésnek a következményeit:

Mivel [az ágenciának, a szándékoknak, valamint a morális és jogi felelősségnek] a tulajdonítása elmfüggő, ágencia és felelősség elvben [*theoretically*] bárminek tulajdonítható, akármilyen alapokon. Lehet eseményekre úgy tekinteni, mintha azok állatoknak, isteneknek vagy szervezeteknek a cselekedetei volnának, és tarthatjuk az állatokat, isteneket és a szervezeteket felelősnek és felelősségre vonhatónak [*liable*] ezekért a „cselekedetekért”. Az azonban csak történelmi szempontból igaz, hogy ágenciát ezeknek azzal analóg módon tulajdonítottunk, ahogyan az embereknek ágenciát tulajdonítunk. Ontológiai nézőpontból semmiféle különbség nincs ugyanis aközött, hogyan tulajdonítunk ágenciát az embereknek és más ágenseknek.

Amennyiben szervezeteknek tulajdoníthatunk ágenciát, valamint felelősnek és felelősségre vonhatónak tarthatjuk őket a cselekedeteikért, ezt megtehetjük az autonóm gépi ágensekkel is. A kérdés nem az, hogy lehetséges-e tulajdonítani nekik mindezeket, hanem az, hogy vajon kívánatos-e így tennünk. Az embereknek tulajdonítunk ágenciát és felelősséget, és az emögött rejlő megfontolások világosabbá tehetik a gondolkodásunkat az autonóm ágensek számára való ágencia- és felelősségtulajdonításra vonatkozóan. (Hage 2017. 261.)

Ezzel az elmélettel is az a helyzet, hogy nagyon nehezen egyeztethető össze a bizonyítékokkal, és jelentős mértékben immunizálja magát a lehetséges ellenérvekkel szemben, ráadásul valószínűleg kártékony lenne, ha ilyen alapokon tulajdonítanánk morális felelősséget a gépeknek.

Könnyen belátható, hogy ez az elmélet ebben a formában tarthatatlan. Például: ha elterjedne egy olyan vírus, amely arra kényszeríti az embereket, hogy va-

lahányszor egy fényes tárgyat látnak, lopják el, s valamiért a vírus befolyása csak akkor gyengülne le vagy tűnne el egészen, ha a vírus által irányított személyeket úgy kezelnék, mintha morálisan felelősek lennének, akkor nem lennének a vírus által irányított személyek morálisan felelősek a vírus által kikényszerített tetteikért. Annak ellenére ez a helyzet, hogy a fenti elméletből az következik, hogy ezek a fertőzöttek rendelkeznek morális felelősséggel, hiszen ilyen körülmények között haszonnal járna hibáztatni őket.

Még sok másik, hasonlóan erős ellenvetést lehetne felhozni az elmélet ellen, de annyiban felesleges, hogy az elmélet szerint – ahogy Hage ezt hangsúlyozza is – elvben bárminek lehet bármilyen alapon morális felelősséget tulajdonítani, ha az hasznos. Nem számít, hogy valaki milyen ellenpéldát hoz fel: az elmélet szerint a morális felelősség tulajdonítása egy teljesen önkényes gyakorlat, aminek csak az szab határt, hogy végül mely gyakorlatok bizonyulnak hasznosak. Hogyan is lehetne – akár elvben – egy ilyen elmélettel szemben olyan ellenvetést felhozni, amelyet az elmélet képviselője komoly ellenvetésként képes lehet felismerni anélkül, hogy előbb elvetné az elméletet?

Továbbá józan ésszel nehezen lehet azt feltételezni, hogy hasznos volna az, ha mindenki elfogadná ezt az elméletet, és ennek alapján próbálnánk megreformálni morális gyakorlatainkat. A társadalom tagjai – bármi is legyen az, ami valójában hasznos – nagyon-nagyon másképpen gondolkodnak arról, hogy mi hasznos és mi nem az. Hogy ne távolodjunk el egészen az autonóm gépek felelősségének a problémájától: egyesek szerint igencsak hasznos lenne, ha az autonóm gépek gyártói jelentős mértékben felelnének azokért a károkért, amelyeket az autonóm gépek okoznak, hiszen ezáltal a károsultaknak nagyobb reményük lenne arra, hogy valamilyen kárpótlást kapnak; mások szerint a gyártókat felelősnek tekinteni ezekért a károkért elhibázott volna, mert visszavetné ezeknek a cégeknek a termelékenységét és innovációs teljesítményét. Ha e két tábor és mindenki más is a fejébe venné, hogy nincsen semmilyen, a hasznosságtól függetleníthető releváns tény a morális felelősséggel összefüggésben, akkor az egyes társadalmi csoportok sokkal gátlástalanabban igyekeznének politikai erejüket kihasználni annak érdekében, hogy saját hasznosságról alkotott elképzeléseik szerint alakítsák át a morális felelősséggel összefüggő társadalmi gyakorlatokat. Ráadásul maga a társadalom is kevesebb ellenállást tanúsítana az ellen, hogy ne az egyes csoportok érdekérvényesítő erején és a hasznosságról kialakított nézetein múljon a morális felelősséggel kapcsolatos gyakorlatok sorsa, hiszen azok – legalábbis a gondolatkísérletünkben uralkodóvá váló elmélet szerint – nem nagyon függnak metafizikai vagy pszichológiai tényektől. Tekintettel a jelenlegi társadalmi helyzetre, ahol egy vagyonos és nagy politikai hatalommal rendelkező kisebbség aránytalanul nagy hatással van a törvények kialakítására, nagy valószínűséggel olyan törvények és gyakorlatok születnének az autonóm gépek által okozott károkra összefüggésben, amelyek aránytalanul kedveznek a vagyonos és nagy politikai hatalommal rendelkező kisebbség érdekeinek.

E gondolatmenet fényében nagyon valószínű: a morális felelősség kérdését a hasznosságra redukáló felelősségelmélet elterjedése nem társadalmilag hasznos, hanem társadalmilag roppant káros eredményeket szülne (még akkor is, ha az elmélet mégiscsak igaz lenne).

Bár legalábbis igencsak megfontolandó az, hogy a morális felelősség kérdésének ilyen szélsőséges redukciója a hasznosság kérdésre társadalmilag nagyon is káros lehet, *senkiben még csak fel sem merül* az – és ez így van jól –, hogy a tudományos közösségnek intoleránsan kellene viselkednie az ilyen megközelítésekkel szemben. Ám a most tárgyalt hipotézis szerint a tudományos közösség a lapos Föld, a fiatal Föld, és az adózás jogszerűtlenségét a hatályos jogszabályok alternatív értelmezése alapján hirdető elméletekkel szemben azért olyan intoleráns, mert valószínűleg ezen elméletek elterjedése jelentős társadalmi károkat eredményezne. Így e hipotézis értelmében a tudományos közösségnek legalábbis intoleránsan *kellene* fellépnie a tudattal nem rendelkező autonóm gépek morális felelősségét támogató, szélsőségesen konzekvencialista elméletekkel szemben – ami abszurd.

Természetesen vannak olyan elméletek, amelyekkel szemben a tudományos közösség etikai okokból intoleránsan lép fel. Ám ezek az elméletek közvetlen módon fenyegetik egyes emberek méltóságát vagy jólétét. A rasszista vagy más módon méltóságsértő, tudományosan is gyenge lábakon álló elméletekkel szemben részben etikai okokból kifolyólag intoleráns a tudományos közösség. Ugyanakkor a tudományos közösség által kiátkozott elméletek többsége ilyen közvetlen módon senkinek sem fenyegeti a méltóságát vagy a jólétét.

Ha azonban nem igazán tartható az, hogy a tudományos közösség tipikusan etikai okokból intoleráns egyes elméletekkel szemben, akkor valamilyen más okot kell találnunk. A következő szakaszban egy felvetést fogalmazok meg azzal kapcsolatban, hogy mi lehet ez az ok, s hogy e magyarázatot miért találok gyakorlati szempontból problematikusnak.

## V. A SZAKÉRTŐI TUDÁS MEGKÉRDŐJELEZÉSE ÉS A TUDOMÁNYOS INTOLERANCIA

Úgy vélem, a lapos Föld- és fiatal Föld-elméleteket, valamint az Egyesült Államok adóinak jogosultságát a hatályos jogszabályok alapján kétségbe vonó nézetet az különbözteti meg a fenomenológiai eliminativizmustól, a normatív hibaelmélettől, illetve a tudattalan gépek morális felelősségét elvben lehetségesnek tekintő szélsőséges konzekvencializmustól, hogy míg az előbbiek valamilyen tudományos szakértelmen alapuló tudást vonnak kétségbe, addig az utóbbiak olyan „hétköznapi” tudást, amely nem tudományos (vagy bármilyen más) szakértelemből származik. Nincs szükségünk tudományos készségekre ahhoz, hogy közvetlen (nem tanúságtételre épülő) tudásra tegyünk szert arra vonatkozóan,

hogy tudatosságunknak van fenomenológiai karaktere, hogy léteznek olyan normák, amelyeket figyelembe kell(ene) vennünk (függetlenül attól, hogy mi elfogadjuk-e ezeket a normákat vagy sem), és hogy a teljességgel tudattalan, a programuknak vakon engedelmeskedő gépeknek nincs morális felelősségük. Ezzel szemben ahhoz, hogy belássuk – és ne csak a tudomány tekintélyére támaszkodva fogadjuk el – a Föld korával és alakjával, valamint az amerikai jogrendszerrel kapcsolatos igazságokat, ahhoz olyan szakértelemre (és akár műszerekre) van szükségünk, amelyet csakis a tudományos közösség tagjaként szerezhethünk meg. A Földdel kapcsolatos igazságok esetében ez egészen nyilvánvaló, de nehéz volna tagadni, hogy az Amerikai Egyesült Államok jogszabályainak és azok normatív következményeinek szisztematikus feltárása nem csak jogi jártasságot, de jogtudományi képzettséget is igényel.

Nem zárhatom ki, hogy ezzel a hipotézissel szemben is lehet találni jó ellenpéldákat, de jómagam nem találtam egyet sem. Ezért zárásképpen nem ennek a feltételezésnek a tarthatóságát szeretném megvizsgálni, hanem arra szeretnék rávilágítani, hogy ez a feltételezés miért nyugtalanító.

Az első probléma: a tudományos intolerancia bevett retorikája – amennyiben az általam javasolt hipotézis igaz – *őszintétlen*. Amikor a nyilvánosság előtt védelemzik ezeknek az elméleteknek a kizárását a tudományos fórumokról, akkor tipikusan nem azzal érvelnek, hogy ezek az elméletek a tudományos szakértelem tekintélyét aláássák, hanem arra hivatkoznak, hogy ezek az elméletek mennyire gyenge lábakon állnak.

De miért nem őszintébb a tudományos intolerancia indoklása? Ez elvezet a második, súlyosabb problémához. Ugyanis, ha az intolerancia legfőbb oka nem episztemológiai, hanem – a szó lehető legtágabb értelmében – tudománypolitikai jellegű, akkor végül is igaz az, amit gyakorta hangoztatnak a tudományos fórumokról kizárt elméletek hívei. Nevezetesen: azért vannak kizárva, mert mélyen megkérdőjelezzik az uralkodó, a tudományos közösség tekintélye szempontjából releváns nézeteket. Azaz a tudományos közeg nem pusztán episztemikus jellegű, elvi fenntartásai miatt nem ad teret ezeknek a nézeteknek – hiszen láthattuk, hogy episztemikus értelemben más, roppant problémás nézeteknek viszont nagyon is helyet ad –, hanem azért, mert meg akarja őrizni a bevett módszerek és nézetek tekintélyét. Azt hiszem, szerencsésebb volna, ha a tudományos közösség mégiscsak elvi, episztemológiai jellegű elvekre tudna hivatkozni akkor, amikor megindokolja ezekkel az elméletekkel szembeni intézményes és szisztematikus intoleranciáját.

## IRODALOM

- Cowie, Christopher 2016. Good News for Moral Error Theorists: A Master Argument against Companions in Guilt Strategies. *Australasian Journal of Philosophy*. 94/1. 115–130.
- Dennett, Daniel C. 1998. When HAL Kills, Who's to Blame? Computer Ethics. In David G. Stork (ed.) *HAL's Legacy: 2001's Computer as Dream and Reality*. Cambridge/MA, The MIT Press. 351–365.
- Dennett, Daniel C. 1988. Quining Qualia. In Anthony Marcel – Edoardo Bisiach (eds.) *Consciousness in Modern Science*. New York, Oxford University Press. 381–414.
- Floridi, Luciano – Sanders, J. W. 2004. On the Morality of Artificial Agents. *Minds and Machines*. 14/3. 349–379.
- Frankish, Keith 2016. Illusionism as a Theory of Consciousness. *Journal of Consciousness Studies*. 23/11–12. 11–39.
- Hage, Jaap 2017. Theoretical Foundations for the Responsibility of Autonomous Agents. *Artificial Intelligence and Law*. 25/3. 255–271.
- Hellström, Thomas 2013. On the Moral Responsibility of Military Robots. *Ethics and Information Technology*. 15/2. 99–107.
- Humphrey, Nicholas 2011. *Soul Dust: The Magic of Consciousness*. Princeton/NJ, Princeton University Press.
- Husi, Stan 2013. Why Reasons Skepticism Is Not Self-Defeating. *European Journal of Philosophy*. 21/3. 424–449.
- Olson, Jonas 2014. *Moral Error Theory: History, Critique, Defence*. Oxford, Oxford University Press.
- Pereboom, Derk 2011. *Consciousness and the Prospects of Physicalism*. New York, Oxford University Press.
- Stahl, B. C. 2006. Responsible Computers? A Case for Ascribing Quasi-Responsibility to Computers Independent of Personhood or Agency. *Ethics and Information Technology*. 8/4. 205–213.
- Sullins, John P. 2006. When is a Robot a Moral Agent. *Machine Ethics*. 2006/6. 23–30.
- Streumer, Bart 2013. Can We Believe the Error Theory? *The Journal of Philosophy*. 110/4. 194–212.
- Streumer, Bart 2017. *Unbelievable Errors: An Error Theory about all Normative Judgements*. Oxford, Oxford University Press.