

Pécsi Tudományegyetem
Filozófia Doktori Iskola

DOKTORI ÉRTEKEZÉS

Egy érv a reduktív materializmus mellett

Herpai Csaba Sándor

Témavezető: Dr. Kocsis László, Pécsi Tudományegyetem

Pécs, 2022

Tartalomjegyzék

| | |
|---|-----|
| 1. Bevezetés..... | 4 |
| 1.1 Ellentétes intuíciók..... | 4 |
| 1.2 Revizionista elmefilozófia: a „tudat tagadói”..... | 12 |
| 1.2.1 Tudat és behaviorizmus..... | 13 |
| 1.2.2 Rey a fenomenális tudatosság ellen..... | 18 |
| 1.3 Antirevizionista elmefilozófia: Chalmers..... | 22 |
| 1.4 A laikusok intuíciói..... | 32 |
| 1.5 Összefoglalás: a két alapvető intuíció..... | 37 |
| 2. Elmefilozófiai viták..... | 41 |
| 2.1 Antirevizionizmus: Block a fenomenális tudatosságról..... | 42 |
| 2.1.1 Block kritikái..... | 52 |
| 2.1.2 Block válaszai..... | 62 |
| 2.2 Revizionizmus: Dennett..... | 68 |
| 2.2.1 Dennett – Kinsbourne: Time and the Observer..... | 69 |
| 2.2.2 Dennett: Consciousness Explained..... | 80 |
| 2.3 További eliminativista kísérletek..... | 88 |
| 2.4 Illuzionizmus..... | 102 |
| 2.5 Összefoglalás: az elméletek kétértelműsége..... | 113 |
| 3. További metafizikai megfontolások..... | 118 |
| 3.1 Chalmers a tudat metaproblémájáról..... | 118 |
| 3.2 Antirevizionizmus és intrinzikusság..... | 120 |
| 3.3 Rorty és Antipodea lakói..... | 127 |
| 3.3.1 Behaviorizmus..... | 132 |
| 3.3.2 Szkepticizmus..... | 133 |
| 3.3.3 Materializmus és Rorty alternatívája..... | 133 |
| 3.3.4 Rorty értékelése..... | 135 |
| 3.4 Az egyensúlyi elmélet keresése..... | 139 |
| 4. A fenomenális hozzáférhetetlensége..... | 144 |
| 4.1 Schlick, Wittgenstein, Farrell, Ryle..... | 150 |
| 4.1.1 Schlick..... | 150 |
| 4.1.2 Wittgenstein..... | 153 |
| 4.1.3 Farrell..... | 156 |

| | |
|---|-----|
| 4. 1. 4 Feyerabend..... | 161 |
| 4. 1. 5 Ryle..... | 163 |
| 4.2 Livingston a strukturális magyarázatokról..... | 166 |
| 4.3 Összefoglalás..... | 170 |
| 5. A reduktív materializmus mint egyensúlyi elmélet..... | 175 |
| 5.1 Place..... | 177 |
| 5.2 Smart..... | 182 |
| 5.3 Feigl..... | 185 |
| 5. 4 Armstrong és Lewis: érett reduktív fizikalizmus..... | 192 |
| 6. Konklúzió és kitekintés..... | 201 |
| Bibliográfia..... | 207 |

1. Bevezetés

1.1 Ellentétes intuíciók

Ha melegem van, elzárom a fűtést. Hasonlóan, ha a hőmérséklet túllép bizonyos meghatározott értéket, a termosztát lekapcsolja a fűtést. Látásom segítségével tárgyakat észlelek és reagálok ezekre. Hasonlóan, egy önvezető autó kamerájával vizuálisan észleli az útakadályokat és kikerüli őket. De látja-e az akadályokat? Idézőjelbe kell-e tennünk a „látja” szót, ha úgy döntünk, hogy alkalmazzuk az önvezető autóra? Ha azt mondjuk, hogy a termosztátnak „melege volt”, idézőjelbe kell-e tennünk, vagy valami furcsát állítunk? Empirikus tény, hogy vannak olyanok, talán az emberek többsége, akik számára ezek a kérdések érthetőek.

A csaliként horogra húzott giliszta nyilvánvalóan sérüléseket szenved, és heves mozgásokat mutat. Átéli-e ennek során fájdalmat (Feigl 1958: 424), és mondhatjuk-e, hogy szenved? Érez-e fájdalmat a homár, amikor forró vízbe dobják? Ezeket a kérdéseket látszólag mindenki érti, és talán még fontosnak is tartja (etikai jelentőségük miatt).

A termosztát-problémával (vagy a kamera „látásának” kérdésével) szembeesülve a laikusok és szakértők lényegében kétféle választ adhatnak abból a szempontból, hogy milyen különbség áll fenn szerintük az összehasonlított jelenségek között. Van, aki lényeges különbséget lát az „érzés” (látás, melegérzet) valamiféle csak az emberre (vagy legalábbis az élőlényekre) vonatkoztatható eredeti, és (a gépeknek, vagy akár bizonyos élőlényeknek való) pusztán metaforikus tulajdonítása között. Mások nem érzékelnek ilyen lényegi eltérést. Ahogy a termosztát és ahogy az ember érzékeli (vagy *átéli*) azt, hogy meleg van, eltérhet ugyan feldolgozás és a kiváltott reakciók komplexitásának tekintetében, de további, abszolút különbséget a kétféle helyzet között nem látnak.

Valószínű, hogy mindenkinek eszébe jutott már az a kérdés, hogy ugyanolyannak látják-e különböző személyek ugyanazokat a színeket (Farrell 1950: 177, Friedl 2013: 205). Számunkra ebben a pillanatban még nem a válaszlehetőségek a legfontosabbak, hanem a kérdés értelmezésének lehetőségessége. Van, aki számára nyilvánvaló, hogy mit kérdez, aki felteszi ezt a kérdést. Mások számára, úgy tűnik, nem nyilvánvaló: vagy nem értik, vagy – az előbbi tábor szerint – félreértik a kérdést.

Évszázados múltú kérdés a filozófiában, a pszichológiában és persze az irodalomban, hogy a „mesterséges intelligencia” és a „robotok” – tehát az emberhez hasonló viselkedési

tulajdonságokkal vagy kognitív képességekkel rendelkező mesterséges objektumok – élnek-e, gondolkodnak-e vagy éreznek-e, és személyként kell-e viszonyulnunk hozzájuk (Feigl 1958: 423). A felsorolt kérdések természetesen különálló problémák, és mindegyik pontosításra szorul, mielőtt filozófiai vagy tudományos válaszra egyáltalán esély lenne. A mindennapi életben sokféleképpen felteszik a kérdést: „valóban gondolkodnak-e” a gépek, „lehetnek-e érzései egy robotnak”; egy számítógépes játékban meglőtt karakter „valóban” szenved-e és így tovább.

Sokan kézenfekvőnek tartják, hogy az ember fájdalma, látásélménye, hőérzete stb. „valódi”, míg a gépé pusztán szimuláció vagy utánzat, és ez lényeges különbség. Egy zivatar számítógépes szimulációjától sem leszünk csuromvizesek (Dennett 1978: 416.) Mások számára azonban éppen ilyen kézenfekvő, hogy az emberi elme és annak a hozzá hasonló képességekkel rendelkező számítógépes szimulációja között nem lehet lényeges különbség – hiszen végső soron agyunk is csupán egyfajta számítógép (Searle 1980). A két alapvető választípus mögött intuitív meggyőződések lappanganak. Az egyik fajta meggyőződés szerint egyfelől az ember (és talán bizonyos állatok), másfelől a gép vagy szimuláció között valamilyen lényeges különbség van abban a tekintetben, hogy rendelkezhetnek-e élményekkel, átélhetnek-e szubjektív tapasztalatokat, „valamilyen”-e nekik lenni. A másik fajta meggyőződés szerint nincs ilyen lényeges különbség. Különösen a tudományokban jártas, a tudományos szemléletmódnak nagy autoritást adó gondolkodók hajlamosak elutasítani az ember „különlegességébe” vetett hitet mind tudományelőtti vagy vallási eredetű babonát. Ám ha egyfelől az ember és az állat, másfelől a termosztát vagy robot között nem látunk lényeges különbséget, azt többféleképpen is értelmezhetjük. Nem mindegy ugyanis, hogy azért nincs különbség az ember és a robot között, mert mindkettő rendelkezik szubjektív élménnyel, azzal az állítólagos többlettel, amivel a „valóban érző” ember rendelkezik a robottal szemben azok szerint, akik radikális különbséget látnak ember és robot között – vagy azért, mert egyik sem. Önmagában az ember „különlegességének” elutasítása ezt a dilemmát nem dönti el. Sokféle alapon lehet az ember „különlegességét” állítani, és mindezeket az álláspontokat el lehet utasítani, tagadva mindenféle különlegességet. Ezek a viták azonban nem feltétlenül érintik a tudatos élmény problémáját.

Egy nemrég megjelent elmefilozófiai bevezető szövegben a következőt olvashatjuk: „Természetes meggyőződésünk szerint bizonyos mentális állapotaink fenomenális-tudatosak, rendelkeznek fenomenális karakterrel, szubjektív minőségekkel. [...] van olyan, mint például megízlelni egy csésze zamatos kávé. Van olyan, mint átélni egy szörnyű fogfájást.” Ám a szöveg így folytatódik: „ha nem volna bizonyos mentális állapotainknak fenomenális karaktere, akkor

egyikünk sem kérne érzéstelenítést a fogorvosnál. Egyikünk sem hajszolná a szexuális örömeket.” (Tózsér 2008:10).

Mint nemsokára látni fogjuk, az első idézet viszonylag világossá tesz, kiemel egy bizonyos intuitív meggyőződést, ami a fenti problémakörben mindazon álláspontok mögött áll, amelyek az említett problémák valóságosságát, fontosságát vagy nehézségét állították. A szöveg folytatása azonban elbizonytalanító: nem látszik teljesen egyértelműnek, hogy csupán illusztrálja-e az első kiemelt szövegrészt, vagy megváltoztatja annak tartalmát. Ebben a dolgozatban számos hasonló esetet fogunk még vizsgálni.

Az olyan élmények, mint a fogfájások, a kávé íze, egy szivárvány látványa stb. kétségkívül rendelkeznek valamiféle „minőséggel”, ám mindig valaki számára rendelkeznek azzal. Valakinek lenni mindig „valamilyen”. Más szóval, elterjedt meggyőződés, hogy a kvalitatív tapasztalat szubjektív, azaz első személyű nézőponthoz kötött. Mindannyiunknak egy sajátos nézőpontja van, ahonnan nézve ilyen vagy olyan a világ. Egyesek számára ez rejtélyes, mert nem illik bele a világ fizikai leírásába, mely harmadik személyű nézőpontot feltételez (Nagel 1974/2004: 4). Mások szerint azonban ez semmiféle problémát nem vet fel: minden egyes ember valahol és valamikor tartózkodik a térben és időben, és első személyű nézőpontjának létezését ez a tény kielégítően magyarázza.

A felsorolt nézetkülönbségek esetében intuitív meggyőzésekkel, illetve az ezek mögött lappangó megérzésekkel, sejtésekkel van dolgunk. Ezek az intuíciók vagy azt a természetesnek tűnő meggyőződést támasztják alá, hogy a tudatosság egy fontos, ugyanakkor rejtélyes – nehezen megmagyarázható, a fizikai világban nehezen elhelyezhető – jelenség, vagy ennek ellenkezőjét: azt a talán ritkább, de sok ember számára nem kevésbé természetes meggyőződést, hogy tudatosságról nincs értelme beszélni, vagy ha van, akkor valamilyen könnyen megmagyarázható jelenségről van szó, ami nem vet fel semmiféle rejtélyt.

Felmerülhet, hogy az utóbbi sohasem természetes meggyőződés, hanem némelyeknek a tudományos magyarázatok sikerei hatására kialakított meggyőződése. Ez azonban kétségbe vonható, egyrészt empirikus felmérések alapján, melyeket a későbbiekben idézni fogunk.

A másik ok a következő. A tudományok számos területen magyarázati sikereket értek el. Csakhogy ennek jelentőségét a tudat kérdésének szempontjából többféleképpen lehet értelmezni. Akinek eleve az az intuitív meggyőződése, hogy a tudat rejtélyes, vagyis a tudományos magyarázatok (legalábbis az eddig ismertek) által megragadhatatlan jelenség, annak számára nem relevánsak a tudományok eddig elért magyarázati sikerei, sőt, mivel az eddigi tudományos

magyarázatok nem járulnak érdemben hozzá a tudatprobléma megoldásához, a tudománytörténet bizonyos szempontból a tudattal kapcsolatos magyarázati *kudarok* története. Csak azt töltik el optimizmussal a tudat tudományos megmagyarázásának lehetőségével kapcsolatban a tudományok eddigi sikerei, akinek eleve az az intuitív meggyőződése, hogy a tudományok módszereivel ez a probléma megoldható. Ha elfogadjuk, hogy az elmefilozófia fő feladata: értelmezni az elmével kapcsolatos természetes meggyőződéseinket és ezek viszonyát tudományos ismereteinkhez (Tózsér 2008: 11), akkor az elmefilozófia célja és tárgya egészen más azok számára, akik az említett intuitív meggyőzések közül az egyik vagy a másik típusba tartozókkal rendelkeznek. Ez pedig komoly problémát okoz az elmefilozófia számára, mert ha sem az explanandum, sem pedig a magától értetődőnek, biztosnak tartott kiindulópontok nem ugyanazok minden elmefilozófus esetében, akkor nemcsak, hogy nem várhatók konszenzuálisan elfogadott válaszok az elmefilozófia kérdéseire, hanem még értelmes vita sem mindig – a viták résztvevői elbeszélnek egymás mellett.¹ Így aztán nem csodálkozhatunk azon, hogy az elmefilozófiai vitákat követve félreértésekbe (illetve egymás félreértéssel vádolásába), patthelyzetekbe, egymás melletti elbeszélésekbe botlunk. Más esetekben pedig azt láthatjuk, hogy egy-egy probléma megoldását a filozófusok újra és újra megkísérlik, de az egyre invenciózusabb terminológia és egyre korszerűbbnek tűnő példák és analógiák mögött csak ugyanazok az alapvető elméleti döntések rejtőznek, és a problémamegoldásban nem történnek lényeges előrelépések.

E tanulmányban feltételezem, hogy a felsorolt filozófiai patthelyzetek és kölcsönös félreértések egyik oka az lehet, ha a filozófusok eltérő intuitív meggyőződéssel vágnak neki az elmefilozófiai problémák megoldásának. Hasonlóan a nem-filozófusokhoz, ők is alapvetően kétféle intuíció egyikének hatása alatt állnak.

A két intuíciót a továbbiakban revizionista („R”) és antirevizionista („A”) intuíciónak nevezem. A „revizionizmus” elnevezést Tartaglia nyomán alkalmazom (Tartaglia 2020: 24), aki a tudat létezésének tagadásának radikális álláspontját, amit ő Ryle-nak és Rortynak tulajdonít, úgy jellemzi, mint szélsőségesen implauzibilis álláspontot, mert az szerinte szemben áll természetes meggyőződéseinkkel. Az elnevezés azt a sejtést sugallja, hogy a revizionista intuíció hatásának jelentkezése, a revizionista nézetek megjelenése a filozófiában újabb keletű, a tudományos

1, „Az elmefilozófus dolga [...] az, hogy belegyömöszölje a tudatot a természettudomány által nyújtott történetbe [...] ennek során mindig úgy tűnik, hogy *újraredefiniálják a tudatot* [...] Hirtelen a tudat problémája máris egyszerűbbnek tűnik [...] a probléma az, hogy [...] úgy tűnik, témát váltottunk” (Goff 2016. 190-191)

gondolkodás térhódításával terjedt el, míg korábban az elterjedtebb álláspont a természetesebbnek tűnő antirevizionista meggyőződéseknek felelt meg.

Felvethető, hogy az „antirevizionizmus” elnevezés szerencsétlen: ha ugyanis az antirevizionista meggyőződések ugyanolyan természetesek, mint a revizionisták, ráadásul a filozófiában történetileg előbb adtak hangot antirevizionista téziseknek, mint revizionistáknak, mely utóbbiak csak a tudományok sikereinek hatására terjedtek el, akkor ügyletlenség ennek az elnevezésnek a revizionistaéba „anti-” előtaggal leszámaztatni. Ugyan a kérdésnek csekély a jelentősége, megjegyezhetjük, hogy történetileg az „antirevizionizmus” artikulálására valójában legfeljebb csak szórványosan került sor azelőtt, hogy a revizionizmus valamilyen formája kihívásként vagy fenyegetésként megjelent volna a filozófiában vagy a kultúra más területein. Amikor a revizionista kihívás megjelent, az antirevizionizmus védelmezői mindazonáltal álláspontjukra mint természetesre, revizionizmus előttire hivatkoztak.

Az A-intuícióval (és az ennek hatására kialakított antirevizionista intuitív meggyőződéssel) azok rendelkeznek, akik számára hosszabb magyarázat nélkül is nyilvánvaló, mi a giliszta-probléma, a mások színélményének problémája és más efféle kérdések tárgya és tétje; ők az első Tözsér-idézetet találónak és fontosnak érzik, míg a másodikat nem feltétlenül, sőt, talán kontraproduktívnak tartják. Az A-intuícióval rendelkező személyek úgy vélik, vagy legalábbis plauzibilisnek tartják, vagy legalábbis értelmesnek tartják azt a felvetést, hogy az emberi látás, érzékelés vagy gondolkodás „valami több”, mint a gépi látás, érzékelés vagy gondolkodás. Vagyis az előbbinek van egy olyan aspektusa, amely több, mint pusztán „struktúra és funkció” (Chalmers 1995/2004:24). Aki az A-intuícióval rendelkezik, az egyetért azzal, vagy legalábbis nem tartja abszurdnak és érthetetlennek azt a gondolatot, hogy az emberi agy működését szimuláló elektronikus (vagy éppen hidraulikus, vagy egészen bizarr módon megvalósított, lásd Block 1978 példáit) számítógép pusztán azért, mert az emberi agyműködést valamilyen szempontból eredményesen szimulálja vagy modellezi, még nem feltétlenül él át ugyanolyan gyönyöröket vagy kínokat, sőt, talán még azokkal a véleményekkel sem rendelkezik, mint a szimulált személy.

Az R-intuíciónak megfelelő revizionista intuitív meggyőződést hétköznapi nyelven, szlogenszerűen így lehetne megfogalmazni: „mihelyt megmagyaráztuk, hogyan történik az érzékelés, észlelés vagy hogyan hozunk döntéseket, nem marad semmi további kérdés”. Másként: „túl az érzékelés, észlelés, kogníció és viselkedés mechanizmusain, nincs semmi további tárgya a pszichológiának”. Tehát nem létezik semmiféle „speciális értelemben vett” érzés, „érzetminőség”, „élményminőség”, az „átélés milyensége” vagy más hasonló tárgy vagy tulajdonság, amiről az elme-filozófusok a „tudat rejtélyével” kapcsolatban beszélnek. Összevetve az antirevizionista

meggyőződéssel, a revizionista meggyőződés valamit tagad. A revizionista álláspont számos megfogalmazásában megtalálhatóak olyan kifejezések, mint „nem több, mint”. A revizionista tagadás tartalmát metafizikailag, episztemológiailag vagy módszertanilag is értelmezhetjük, filozófiai tézisekké fejleszthetjük. A revizionista tézis metafizikailag azt jelenti, hogy nem létezik az a fajtájú létező vagy tulajdonság (a tudat vagy a tudatosság, a tapasztalat „kvalitatív karaktere” vagy „milyensége”), amely az antirevizionista intuíció számára középponti jelentőségű. Episztemológiailag azt jelenti, hogy nincs olyan típusú tudásunk, amelynek ténylegesen a tárgya az említett fajta létező vagy tulajdonság. Módszertanilag pedig azt jelenti, hogy sem a pszichológiai, sem a filozófiai vizsgálódásnak nem lehet tárgya (és nem is kell, hogy tárgya legyen) az antirevizionisták által sui generisnek tartott létező vagy tulajdonság.

A revizionista intuitív meggyőződés módszertani következménye az elmefilozófia olyan képe, mely szerint az elmefilozófia a „pszichológia filozófiája” kell hogy legyen, olyan értelemben, ahogy lehetséges a „biológia filozófiája” vagy a „számítástechnika filozófiája”. Az elmefilozófiának e koncepció szerint nincs olyan sajátos tárgya, melyhez a pszichológia nem fér hozzá: az elmefilozófiának feladata lehet, hogy értelmezze a pszichológia eredményeit, összevetve azokat a mindennapi gondolkodásunk kategóriáival és hétköznapi meggyőződésünkkel – és önmagában ez sem jelentéktelen feladat, ld. pl. a népi pszichológia státusa körüli problémák burjánzását (Demeter 2008) – de a „tudat rejtélye” nem rejtély, csupán misztifikáció.

Az A-intuíció szerint azonban *létezik a tudatosság*, mint sajátos entitás vagy tulajdonság. Így az antirevizionista intuitív meggyőződésekkel rendelkezők szerint az elmefilozófiának van egy sajátos tárgya (a tudat, és különösen a fenomenális tudatosság), amely a pszichológiának és más tudományoknak nem tárgya (Farrell 1950: 173), ám mégis fontos filozófiai téma. Az antirevizionisták szerint a mentális állapotok természetének kérdése épp attól igazán izgalmas, mert a mentális állapotok egy része „fenomenálisan tudatos”, így „valamilyen” átélni őket. Ez a tulajdonságuk pedig (legalábbis ma még) nem tárgya egyik tudománynak sem, így a kutatása a filozófia számára jelent feladatot, vagy új utakra kell vezetnie a tudományos kutatásokat.

Az elhangzottak fényében felmerülhet, hogy az egymással versengő elmefilozófiai elméleteket legalább részben úgy tekinthetnénk, mint az A- és R-intuíciók, illetve ezek nyomán támadt A- és R-meggyőződések filozófiai kidolgozását. Az elmefilozófiai pozíciók, legalábbis részben, az intuitív meggyőződések következményeinek kifejezései a filozófiai elmélet nyelvén. Alátámasztanának-e az elmefilozófiai szövegek egy ilyen értelmezést? Szövegeik alapján tulajdoníthatjuk-e az A- vagy R-intuíciót az egyes szerzőknek? Sikerült-e az elmefilozófusoknak artikulálniuk és világossá tenniük a kétféle intuitív meggyőződésnek megfelelő filozófiai

pozíciókat? A kérdés nem mellékes, mert az elmefilozófia sikere vagy kudarca szempontjából nem lényegtelen, hogy helyesen méri-e fel egyáltalán, mi a sajátos tárgya, illetve hogy tulajdonképpen tárgyra nézve egyáltalán releváns tud-e lenni.

Könnyűnek tűnik az a feladat, hogy példákat találjunk az A-intuícióval összhangban lévő elmefilozófiára, hiszen plauzibilis, hogy az elmefilozófia a mai formájában nem született volna meg, ha soha senkinek sem lett volna A-intuíciója. Tankönyvek, ismeretterjesztő cikkek és komoly elmefilozófiai tanulmányok tucatjai kezdődnek azzal, hogy a tudatprobléma felvetését a természetesnek, intuitívnek tartott antirevizionista meggyőződés melletti világos elköteleződéssel motiválják.

Aki azonban az A-intuíciót meg akarja fogalmazni, annak nincsen egészen könnyű dolga. Így például Ned Block kénytelen Louis Armstrongnak a jazz mibenlétéről szóló bon motját idézni: „Ha meg kell kérdezned, sohasem fogod megtudni” (Strawson 2018). Mégis, mintha az efféle gesztusok gyakran elégségesek lennének. Lewis szerint például az olyan szavak, mint a *kválé*, a *fenomenális karakter* vagy a *szubjektív minőség* nem részei a hétköznapi nyelvnek, de a kválé fogalma, amire ezek utalnak, része népi pszichológiánknak, így amikor a filozófusok megkísérlik elmagyarázni a laikusoknak, mit értenek a kválékon, a laikusok elég könnyen, kevés szóból is megértik azt (Lewis 1995: 140).

Strawson (2018) szerint „ami a tudatos tapasztalatot illeti, egy mélyen alapvető értelemben teljes mértékben ismeretségben vagyunk vele, pusztán azáltal, hogy rendelkezünk vele. Rendelkezni vele maga a tudás róla. Így hát azok, akik azt mondják, hogy a tudat rejtély, tévednek... ez a legismertebb dolog, ami létezik – bármilyen nehéz is szavakba önteni.” Tudatosak vagyunk, és ezt tudjuk magunkról, szögezi le Dretske (2003), mielőtt hozzálátna megvizsgálni, honnan tudjuk, hogy nem vagyunk filozófiai zombik – eredménytelenül.

Strawson szerint az emberi gondolkodás történetének legfurcsább epizódja, a valaha megtett „legidétlenebb állítás” a tudat, a tapasztalat „milyenségének” tagadása, az az álláspont, melyet Strawson többek között Farrellnek, Feyerabendnek, Rortynak és Dennettnek tulajdonít. Tartaglia (2020: 15) egyetért: szerinte Rorty és a többi „ryleánus” álláspontja a „filozófiai implauzibilitás végső határait” feszegeti. E nyilatkozatok kiválóan szemléltetik, hogyan beszél egy A-intuícióval rendelkező gondolkodó az R-intuícióval rendelkező szerzők munkásságáról. Azonban a helyzet az, hogy amit Strawson és Tartaglia implauzibilis vagy idétlen álláspontnak nevez, korántsem csupán egy egzotikus, elszigetelt filozófiai pozíció, amely mellett hiába érvelnek, mindenképpen kontraintuitív marad, hanem egy nagyon is elterjedt, és sok ember számára természetes, intuitív

meggyőződéseikkel összhangban lévő álláspont. Bármelyik, a tudat kérdése iránt érdeklődő, azt Chalmers (1995/2004) kifejezésével élve „nehéz problémának” tekintő filozófus vagy laikus szembesülhet vele, hogy egy revizionista vitapartnernek nem mindig tudja elmagyarázni, mi is az a kérdés, ami foglalkoztatja. A hétköznapi példák sorolása – a rózsa illatáról, a hajnali ég színéről stb. szóló retorikai fordulatok – a revizionistából csak értetlenkedést váltanak ki. Csak rontja a helyzetet, ha az antirevizionista az érzelmek vagy a mentális képek témáját hozza fel, hogy kiszabadítsa a beszélgetést a „percepció” vagy „érzékelés” körüli diskurzusból. Csöbörből vödörbe kerül: minél képzetesebb a vitapartner, annál hosszabb – irrelevánsnak tűnő – monológokat kell végighallgatnia az érzelmek döntésben játszott szerepéről, reprezentációkról vagy a kognitív funkciók evolúcióbiológiájáról.

Az antirevizionista szemszögéből ebben az esetben vagy az történik, hogy vitapartner egyszerűen félreérti, vagy az, hogy a vitapartner a probléma nehézsége előtt egy olyan területre menekül, amiről a tudása biztos, és arról beszél – mert talán abban reménykedik, az antirevizionista szerint valószínűleg tévesen, hogy más területeken elért tudományos haladás előbb-utóbb a tudatprobléma megoldásában is segít. Így például a címében a tudatosság magyarázatát ígérő művének 280. oldalán Dennett megjegyzi: „Eddig szemérmes voltam a tudattal kapcsolatban. Óvatosan elkerültem, hogy elmondjam, elméletem mit mond arról, hogy mi a tudat” (Dennett 1991: 280). A kérdéses fejezetben Dennett egyébként kifejti, szerinte micsoda a tudat – és elmélete aligha elégít ki bárkit, aki antirevizionista meggyőződéssel vág neki a könyv olvasásának. Dennett azonban számít erre a reakcióra, és olvasója lelkiismeretére (vagy önérzetére) próbál hatni, amikor először is „a tibeti imamalom”-nak nevezi azt az ellenvetést, hogy pusztán a kognitív funkciók leírása nem elégséges a tudat magyarázataként, majd az ellenvetés abszurditását olyan módon próbálja bizonyítani, hogy párhuzamot von közte és egy olyan vitalista okvetetlenkedése között, aki szerint a DNS működésének megmagyarázása még nem magyarázza meg az életet. Revizionista szemszögéből az antirevizionista olyannak tűnik, aki egyszerűen nem hajlandó megérteni a tudományos magyarázatokat.

Mivel az antirevizionista intuíció gyakran annak a hiedelemnek a formájában jelentkezik az elmefilozófiában, hogy van valami (a tudatosság), ami a „struktúrákon és funkciókon” felül (Chalmers 1995/2004: 24) még magyarázatra szorul, ezért a pusztán a funkciókat megmagyarázó elméletek „kihagynak valamit”, a revizionista intuíció szerint azonban ilyen többlet nincs, ezért revizionista nézőpontból az antirevizionista további magyarázatot követel akkor, amikor már minden meg van magyarázva. Erre az egyetlen logikus magyarázat, hogy nem érti az elhangzott magyarázatokat, így azokat még alaposabban meg kell ismételni. Antirevizionista nézőpontból

természetesen az egyre alaposabb, de továbbra is csak struktúrákra és funkciókra vonatkozó magyarázatok teljesen irrelevánsak.

Bár a filozófusok közti viták rendszerint udvariasabbak, a következő idézet példa arra, milyen értetlenséggel fogadhatja egy antirevizionista a revizionista elmefilozófiát (Lanier 1995): „Úgy tűnik, lehetséges a zombik megkülönböztetése a személyektől. A zombinak más a filozófiája... Tehát, a zombikat csak akkor lehet detektálni, ha történetesen filozófusok. Dennett nyilvánvalóan zombi... Csak egy Dennetthez hasonló zombi képes írni egy olyan „Consciousness Explained” [A megmagyarázott tudat] című könyvet, amely a tudatról egyáltalán nem szól.”

Amikor a filozófiában makacs patthelyzetek és félreértések adódnak, akkor gyakran arról van szó, hogy a vitapartnerek számára más számít „nyilvánvaló ténynek” – mások az intuitív meggyőződések. Ezért felvethető a kérdés, hogy csoportosíthatók-e az elmefilozófiai elméletek azon szempont szerint, hogy antirevizionista vagy revizionista meggyőződéseket juttatnak-e kifejezésre.

1.2 Revizionista elmefilozófia: a „tudat tagadói”

A Strawson (2018) által a tudatot tagadóknak nevezett filozófusok a behaviorizmus, eliminativizmus és illuzionizmus különféle változatainak képviselői. A következőkben elsősorban két kérdést vizsgálok ezekkel a szerzőkkel kapcsolatban. Egyrészt azt, hogy elméleteik valóban leírhatók-e a tudat tagadásaként, másrészt azt, hogy a tudat tagadása valóban olyan implauzibilis nézet-e, mint Tartaglia és Strawson tartja.

Kifejtettek-e filozófusok olyan álláspontot, amely felér a „tudat tagadásával”? Strawson jelöltjei erre a behaviorista, illuzionista és eliminativista elméletek. Könnyen belátható, miért. A (filozófiai) behaviorizmus a mentális állapotokat és folyamatokat – így a tudati jelenségeket és magát a tudatot is – viselkedésekkel, viselkedési diszpozíciókkal és ezekre épülő elmélet konstrukciókkal azonosítja. Strawson megjegyzi, hogy a filozófiai behavioristák nem állították (talán Hull kivételével), hogy a tudatot „eliminálták” volna: csupán azonosították valamivel. Ám ha X „valójában nem több, mint Y”, jegyzi meg Strawson, akkor X-et mégiscsak elimináltuk ebben az esetben. Ha ugyanis a tudatállapotok „valójában nem mások, mint” bizonyos viselkedési diszpozíciók, akkor a tudat igazából nincs (mint valami szubjektív, a publikusan megfigyelhető tulajdonságoktól megkülönböztethető dolog, a tapasztalat intrinzikus minősége vagy kvalitatív

karaktere). Egyelőre fogadjuk el ezt az értékelést, bár később visszatérünk arra, jogos-e ez a megállapítás.

Az eliminativista elmefilozófiák esetében első ránézésre még egyértelműbbnek tűnik a helyzet, hiszen ezeket szerzőik kifejezetten „eliminatív” célkitűzésekkel fogalmazták meg. Az eliminativizmus az az álláspont, hogy mentális állapotok egyszerűen nincsenek – rájuk vonatkozó hétköznapi meggyőződéseink radikálisan tévesek. Azonban sokféle eliminativizmus van. A népi pszichológiával kapcsolatos eliminativizmusból (Churchland 1981) például nem következik triviálisan a „tudat tagadása”.

Ám vannak kifejezetten a fenomenális tudatosságot célba vevő eliminativista kezdeményezések is. Ugyanez a helyzet az illuzionizmussal, azzal az állásponttal, hogy „a tudat illúzió”. Az A-intuícióval rendelkezők számára az az álláspont, hogy a tudat illúzió, első hallásra paradox, sőt semmitmondó: az illúziók jellemzője, hogy valamit „valamilyennek” tapasztalunk, ám az a valóságban nem olyan, mint amilyennek tapasztaljuk. Az illuzionizmus egy ilyen olvasat szerint a tudatproblémát nem szünteti meg, csupán megváltoztatja: a tudat problémáját a „tudat illúziójának” problémájává alakítja. De talán nem ilyen egyszerű fogást találni az illuzionizmuson. Az „illúzió” jelenthet kognitív illúziót (talán jobb lenne delúzióknak nevezni) is: eszerint az a *hiedelmünk*, hogy tudatos lények vagyunk, hamis. Végül, a reduktív materializmusnak létezik egy eliminativista színezetű hagyománya (például: Rorty 1965), amely a mentális és fizikai állapotok azonosítását nem olyan módon fogalmazza meg, hogy „a mentális és fizikai állapot ugyanaz”, hanem úgy, hogy „igazából nincsenek mentális állapotok, csak fizikai állapotok”.

A következőkben néhány ilyen elméletet vizsgálok meg, mint lehetséges jelölteket a revizionista elmefilozófiára.

1.2.1 Tudat és behaviorizmus

Tagadták-e a behavioristák a tudat létezését? A legtöbb esetben pusztán azt állították, hogy a tudat nem lehet a tudományos kutatások tárgya, ezért a pszichológusoknak figyelmen kívül kell hagyniuk a tudattal kapcsolatos kérdéseket. Néha azonban, legalábbis a behaviorizmus egyes kritikusainak interpretációjában a tudat tagadásáig is valóban eljutottak. Pratt például Watsont idézi, aki szerint

A behaviorista álláspont súlyos félreértése azt mondani, hogy ‘Természetesen a behaviorista nem tagadja, hogy mentális állapotok léteznek; csupán inkább figyelmen kívül hagyja őket’. Abban az értelemben hagyja őket figyelmen kívül, ahogy a kémia az alkímiát, vagy a csillagászat a horoszkópot. (Pratt 1922: 603)

Valamint

A gondolkodás nem különbözik lényegesen a teniszezéstől, úszástól, vagy bármilyen más tevékenységtől, kivéve, hogy el van rejtve a szokványos megfigyelés elől és komplexebb. (Pratt 1922: 603)

Figyeljük azonban meg, hogy Watson a mentális állapotok és az alkímia vagy a horoszkóp között von analógiát. Az antirevizionisták szerint azonban a tudat nem pusztán olyasmi, aminek létezését más jelenségek megmagyarázásának céljával feltételezzük, hanem – Chalmers kifejezésével élve – maga az *explanandum*, és így nem dobható félre, mint egy elavult elmélet. Így Watson radikálisnak ható kijelentésének az antirevizionisták számára olyan olvasata is adódik, mely szerint a behaviorista pusztán azért utasítja el a mentális állapotok egész kategóriáját, mert a pszichológiai magyarázatok szempontjából pontosan olyan túlhaladottnak és terméketlennek tartja, mint a modern csillagászat a horoszkópokat. Ennek során az *explanansok* valamilyen fajtáját utasítja el, nem pedig az antirevizionisták *explanandumát*, amiről valójában nem is beszél. Ezt az olvasatot csak támogatja a második idézet, ahol a *gondolkodásról* van szó. A gondolkodás természetének és a fenomenális tudatosság mibenlétének kérdése ugyanis nem tartozik szorosan össze. (Sellars 1962/1963: 34-35). Ha Watson radikális álláspontja pusztán annyiban áll, hogy elutasítja azt, hogy a pszichológiai magyarázatok mentális állapotokra hivatkozzanak, valamint abban a meggyőződésben, hogy a gondolkodás „nem különbözik lényegesen a teniszezéstől”, ezzel még nem mondta ki világosan azt, hogy nincs tudat.

Pratt hivatkozik Singerre is, aki szerint a tudat nem valami olyasmi, amire a viselkedés alapján lehet következtetni, hanem *nem más, mint viselkedés*. Valóban a tudat tagadói közé sorolhatjuk Watsont és Singert? Pratt szerint a tudatot tagadó behavioristák tagadják „a tudat mint szubjektív entitás létezését”. A tudatot mint „szubjektív entitást” tagadni azonban nem jelenti egyértelműen a „tudat tagadásának” álláspontját. Ugyan William James nem tartozott a

behavioristák közé, de érdekes példát szolgáltat erre. A következőket olvashatjuk „Létezik a ‘tudat?’” című tanulmányában (James 1904): „Azt hiszem, hogy a ‘tudat’ [...] úton van afelé, hogy teljesen eltűnjön. Egy nonentitás neve [...] eljött az óra, hogy nyíltan és általánosságában kidobjuk.” E radikálisnak hangzó kijelentések után, melyeket akár Watson is mondhatott volna, James rögtön pontosít: „Csak azt szándékozom tagadni, hogy ez a szó egy entitást jelöl, de a leghatározottabban állítom, hogy igenis jelöl egy funkciót”. „Funkción” azonban a kor tudósai nem elvont gépállapotok terminusaiban felfogott funkciót értettek Putnam (1980) módjára, hanem konkrétabb biológiai funkciót (Klein 2020). Ez azt jelenti, hogy James számára a tudat nem entitás, hanem folyamat.

James az olvasói lehetséges ellenvetéseit anticipálva nem azt az intuíciót veszi célba, hogy „létezik tudatosság”, hanem azt a fajta gondolatmenetet, mely szerint a tudatosság szembeállítható a tárgyaival, és ez mindenképpen valamiféle metafizikai dualizmusra enged következtetni. Amikor James azt mondja, hogy „az az ‘én gondolom’, aminek Kant szerint minden [gondolati] tárgyat kísérnie kell, az az ‘én lélegzem’, ami ténylegesen kíséri őket”, akkor már csak azért sem azt kívánja állítani, hogy nincs gondolkodás, vagy nincsenek tapasztalatok, csakis légzés és más biológiai folyamatok (a kettőt szembeállítva, kizáró viszonyt tételezve fel közöttük), mert szerinte „a konkrét gondolatok teljesen valóságosak. De a konkrét gondolatok ugyanabból az anyagból vannak, mint a dolgok”. Ez még materialista olvasatban sem a szubjektív tapasztalat tagadása, csupán egy, a gondolatoktól és más mentális eseményektől különválasztható, különálló entitásként felfogott „tudat” tagadása. Klein (2020) szerint James sohasem tagadta, hogy a mentális állapotok fenomenális minőséggel rendelkezhetnek, állítása inkább az, hogy a mentális állapotok ilyen vagy olyan fenomenális-tudatos minőséggel rendelkeznek, de nincs egyféle tulajdonság, amivel rendelkezni azt jelentené, hogy az adott állapotot „tudatosnak érezzük”. Klein szerint „nem valamilyen (egyféle) érzés az, hogy milyen denevérek lenni, mondja James; sokféle érzés.”

Mi a helyzet Singerrel, Pratt említett kritikájának másik céltáblájával? Egyik híres tanulmányában Singer a következőképpen érvel. Közvetlenül csak saját mentális állapotaimról van tudomásom. Amikor más testek viselkedését megfigyelem, csak következtetek arra, hogy nekik is vannak elmeállapotaik (Singer 1911: 180). De van-e pragmatikus jelentősége annak, hogy egy másik testnek van elméje is, ezáltal állíthatjuk-e, hogy értelmes a megkülönböztetés az emberként viselkedő, elmével rendelkező lény és a hasonlóan viselkedő, de elmével nem rendelkező lény között? Az „automata szerető” (Singer 1911: 182) példája azt mutatja, hogy igen, mert a szeretőként viselkedő automatához nem ugyanúgy viszonyulnánk, mint egy valódi szeretőhöz. Csakhogy Singer szerint ez a válasz elsietett, mert igazából amikor azt feltételezzük, hogy szeretőnk „lelketlen”, akkor attól tartunk, hogy viselkedése nem mindig fog megfelelni várakozásainknak (Singer 1911:

183). Ha biztosak lennénk abban, hogy sohasem viselkedik „lelketlenül”, akkor az a hipotézis, hogy esetleg nincs lelke, nem lenne komolyabb, mint az a hipotézis, hogy egy háromoldalú síkidom eseteg nem háromszög.

Singer szerint „A tudatosság nem valami, amire a viselkedésből következtetünk, hanem viselkedés. Vagy pontosabban a tudatosságba vetett hitünk a valószínű viselkedéssel kapcsolatos várakozás a megfigyelt tényleges viselkedés alapján” (Singer 1911: 183). Miután tehát Singer kijelenti, hogy a tudatosság viselkedés, állítását azon nyomban részben vissza is vonja, és arról kezd beszélni, hogy az emberek hogyan viszonyulnak a tudatosság kérdéséhez, vagy hogyan kellene viszonyulniuk hozzá, ha jó empiristák lennének. Később pedig párhuzamot állít fel a „tudatosság viselkedés” és az „a hő: a testek bizonyos viselkedése”, valamint az „ahogy egy meleg test test plusz hő, úgy egy élő test: élettelen test plusz élet” állítások között (Singer 1911: 186).

Míg azonban James antirevizionista meggyőződéseire nem fér kétség (Flanagan – Wallace 2016), Singertől aligha vitathatjuk el a világos revizionista szándékokat.¹ Amellett igyekszik érvelni, hogy a tudatosság csupán valamilyen fajta viselkedés, amit szerinte az is alátámaszt, hogy a tudatosság fogalmát valójában éppen így értelmezzük. Ez pedig az antirevizionisták szerint nem így van: aligha győzné meg őket Singernek a „lelketlenséggel” kapcsolatos szójátéka arról, hogy igazából mindegy nekik, szeretőjük tudatos lény-e. Másrészt amikor Singer a tudatosság létezésé helyett a tudatosságba vetett hitről kezd beszélni, felveti azt a lehetőséget egy antirevizionista interpretátor számára, hogy Singer csupán amellet érvel, hogy akármi is legyen a helyzet a mechanikus szerető tudatosságával kapcsolatban, ez nem fontos, hiszen nem ez érdekel minket, hanem a viselkedése. Vagy nyitva hagyja (és értelmesnek tartja) a mechanikus szerető tudatosságának kérdését, vagy pedig végig elbeszél a téma mellett (korábbi ígéretei dacára).

A tudatról alkotott behaviorista elképzeléseket átfogó vizsgálatnak vetette alá Lashley (1923). A „tudat tényeitől” mint tudományosan nem vizsgálható (vagy vizsgálendő) tárgyaktól eltekintő módszertani behaviorista irányzatokkal szemben létezik a „tudat tényének alapvető tagadásának” radikális nézete is. Lashley tanulmányában amellet érvel, hogy a tudatnak nincsenek olyan, az introspekcióban adott „tényei”, amelyeket ne lehetne természettudományos fogalmakkal leírni, és ezért „megértésükhöz egy dualista interpretáció semmivel sem járul hozzá” (Lashley 1923:

¹ Amikor azonban James a tudat oksági hatékonysága mellett érvel (James 1879), ez paradox módon olyan álláspont, amit akár egy revizionista is elfogadhat, ha a „tudatot” valamilyen funkció ellátásának képességével azonosítja, például valamiféle „központi döntési mechanizmusként” fogja fel. A tudat e kettős – funkcionális és fenomenális – értelmezésének lehetőségéből fakadó nehézségekről a későbbiekben még lesz szó.

240-241). Lashley szerint az elme-test probléma a természettudományos módszereknek és posztulátumoknak a tudat „úgynevezett attribútumaira vagy elemeire” való alkalmazhatóságának problémája, és a vita a behaviorista és a dualista álláspont között zajlik (Lashley 1923: 246). Lashley megvizsgálja a tudat különböző jellegzetes attribútumait (mint amilyen a „kvalitatív sokféleség”, vagy „a tudat egysége”, és az érzéki minőségek, és megkísérli kimutatni, hogy fizikai, biológiai rendszerek is rendelkezhetnek olyan tulajdonságokkal, melyek párhuzamba állíthatók a tudat attribútumaival.

A tudat talán talán legnehezebb problémát felvető jellemzőjét, a tudattartalmak *érzéki minőséget* a következőképpen definiálja: „valaminek (a tartalom egy elemének) feloszthatatlansága valami máshoz (az introspekcióhoz) való viszonyában. Minden fizikai komplexum, amely egységként viselkedik egy másik fizikai folyamathoz való viszonyában, megfelel ennek a feltételnek” (Lashley 1923: 329). Ha ezek az értelmezések sikeresen meggyengítik is a dualizmus mellett szóló érveket, nem szólnak a „tudat tagadása” mellett, hiszen egy, a tudat létezését nem tagadó materialista vagy éppen neutrális monista is elfogadhatja őket. Lashley azonban éppen azért is ragaszkodik ahhoz, hogy elméletét „behaviorizmusnak” nevezze, nem elfogadva a „materializmus” elnevezést, hogy ezzel is hangsúlyozza, hogy nem metafizikát művel, hanem érvelése csakis tudományos. Végső soron emiatt álláspontja alig tér el a pusztán módszertani behaviorizmustól: az eltérést abban ragadhatjuk meg, hogy alaposan érvel az elme vagy a tudat dualista felfogása ellen.

Amikor Lashley a robot tudatosságának kérdését vizsgálja, akkor álláspontja végképp kétértelművé válik. Miután amellet érvel, hogy amikor a robot „piros fényt lát”, akkor a benne zajló folyamatok kielégítik a tudatosság összes szubjektív feltételét (megfelel a tudatosságnak a számunkra az introspekcióban feltáruló attribútumainak), megállapítja, hogy semmiféle „pszichikus anyagra” nem kellett hivatkozni ahhoz, hogy a gép viselkedését megmagyarázzuk (Lashley 1923: 332). Ez világos tagadása a dualizmusnak, de nem tekinthető a tudat tagadásának. Revizionista olvasatban „nincs tudat, csak viselkedés”, antirevizionista olvasatban pedig problematikus, de lehetséges feltételezés, hogy a (külső és belső) emberi viselkedést részletekbe menően pontosan utánzó gép hasonló tudatossággal bír, mint egy ember. Ez utóbbi lehetőség egy materialista, egy neutrális monista vagy éppen egy tulajdonságdualista számára sem feltétlenül kizárt. Lashley konklúziója: „gépünk... szerveződésének köszönhetően teljese mértékben tudatos” (Lashley 1923: 336). Lashley szerint a gép tehát *tudatos*, ám egy revizionista úgy értelmezheti ezt az állítást, hogy a tudatosság végső soron nem több, mint bizonyos fajta viselkedés. Úgy tűnik, ehhez a revizionista értelmezéséhez áll közel Lashley eredeti intenciója.

Revizionista intuitív meggyőződésekre következtethetünk ugyanis, amikor Lashley a „tudat ‘ineffábilis’ karakteréről” azt a megjegyzést teszi, hogy „semmit sem találok a tapasztalatomban, ami úgy tűnik, hogy kimaradna verbális jellemzéseimből. A tudatnak tehát vagy nincsenek ilyen kifejezhetetlen elemei, vagy nem vagyok tudatos, és a való életben példázom a ‘gondolkodó behaviorista paradoxonát’...” (Lashley 1923: 337)

Szándékaitól és intuíciótól függetlenül azonban tanulmányáról nem állítható egyértelműen, hogy tagadja a tudat létezését és meggyőzően érvel a tudat nemlétezése mellett. Az érvelés célpontja elsősorban a dualizmus, de a tudat létezését a dualizmus feltételezése nélkül is el lehet fogadni.

1.2.2 Rey a fenomenális tudatosság ellen

Vizsgáljunk meg egy hírhedt „eliminativista” szöveget (Rey 1988), amely klasszikus hivatkozási pont, amikor azt kell szemléltetni, hogy mit is jelent, ha egyes filozófusok tagadják a fenomenális tudatosság jelenségét. A rá válaszoló vitacikkkel (Smith 1988) együtt jól illusztrálják a félreértésnek és egymás melletti elbeszélésnek azt a módját is, amely az elmefilozófiában lépten-nyomon előfordul.

Georges Rey először is megállapítja: „Egyrészt a tudat állítólag... a legnyilvánvalóbb dolog a világon; másrészt úgy tűnik, senki sem tud róla különösebben megvilágító erejű dolgokat mondani” (Rey 1988: 5). Az első kijelentést, bővebben azt, hogy „akármi is az elméleted, ez nem lehet ok arra, hogy kételkedjem abban, hogy itt és most tudatos vagyok” (Rey 1988: 6), „karteziánus intuíciónak” nevezi, és célja kimutatni, hogy ennek ereje ellenére is kételkedhetünk abban, hogy tudatos lények vagyunk. Rey szerint valami akkor bír mentális étellel, ha engedelmeskedik bizonyos pszichológiai törvényszerűségeknek, bővebben: olyan regularitásokat mutat a viselkedése, melyek lehetővé teszik hiedelmek és preferenciák tulajdonítását – tehát intencionális rendszer. Csakhogy intencionális rendszer lehet egy számítógép is, ha néhány – meglehetősen egyszerű – képességet beleprogramoznak. Egy ilyen rendszer úgy viselkedik, hogy a Dennett által elemzett „intencionális hozzáállást” (Dennett 1971: 90) érdemes felvenni vele kapcsolatban, sőt szemben Dennettel, Rey úgy véli, hogy egy ilyen rendszernek az intencionális állapotokat „szó szerint” tulajdoníthatjuk (Rey 1988: 10). Ám mindebből nem következik, hogy egy ilyen rendszer tudatos is. A tudat ugyanis a bevett elképzelés szerint valami többről is szól, mint a gondolkodásról vagy az információfeldolgozó funkciókról (Rey 1988: 11-12). Az öntudatosság,

önreflexió (n-edrendű intencionalitás) kritériuma is triviálisan implementálható egy olyan számítástechnikai rendszerben, amely elsőrendűen már intencionális. Ezért Rey nem fogadja el azokat az elméleteket sem, melyek az emberi tudatosságot efféle önismeretre igyekeznek visszavezetni. Úgyszintén nem megfelelő kritérium a belső állapotokról való tudomással bírás, mely az azokról publikus nyelven való beszámolás képességében nyilvánul meg, hiszen ez is olyasmi, ami meglévő számítógépeinkben gond nélkül megvalósítható anélkül, hogy értenénk, miért is tulajdonítsunk e gépeknek tudatosságot. E ponton a kézenfekvő javaslat az, hogy a tudatosság feltétele az *érzet* legyen. Az érzetek – színek, fájdalmak – a fontosak, mondaná a karteziánus intuíció híve, nem pedig a puszta „funkcionális ekvivalenseik” (Rey 1988: 18). De honnan tudjuk, teszi fel a kérdést Rey, hogy mi ténylegesen átéljük például a piros színérzetet, a puszta funkcionális ekvivalensén felül? Egészen biztos, hogy nem onnan, hogy van egy jól alátámasztott elméletünk, mely lehetővé teszi ezt a megkülönböztetést.

Rey ezután röviden ismerteti azt a lehetőséget, mely szerint az érzelmi állapotaink szervezetünk nem kognitív, hanem fizikai, kémiai állapotain alapulnak. Ám ez a lehetőség Rey szerint egyelőre még homályos. Egy ellenérvvel is megpróbálkozik. Nevezzük a fájdalomérzet feltételül szolgáló nem-kognitív állapotot K-nak. Ekkor lehetséges, hogy valaki azt higgye, hogy fájdalmat él át, de K nélkül. Valaki így azt hiheti, hogy fájdalmat érez, anélkül, hogy ténylegesen érezné, ami Rey szerint abszurd (Rey 1988: 19). Ez az érv azonban nem különösebben erős, mert a fájdalom komplex jelenség, melyben megkülönböztethető egymástól a kauzális szindróma, a fenomenális aspektus és talán további aspektusok is. Mindennapi „fájdalom”-fogalmunk elfedi ezt az összetettséget, amelyet az olyan jelenségek révén fedeztünk fel, mint a fájdalom-aszimbólia (Klein 2015) vagy a különböző fájdalomcsillapítók hatásmechanizmusának sajátosságai (Dennett 1978: 431-432).

Rey döntőnek szánt érve azonban a következő. Egy számítógépet beprogramozhatunk nemcsak arra, hogy intencionális rendszer legyen, hanem arra is, hogy rendelkezzen a karteziánus intuícióval. Rey szerint erre nem válaszolhatjuk azt, hogy a „világosan látom, hogy semmi sem egyszerűbb számomra, mint saját elmém ismerete” felirat kinyomtatásának képessége csupán „funkcionális ekvivalense” a tudatosságnak, és nem maga a tudatosság. Hiszen aki ezt állítja, annak valami további rejtélyes feltételt kellene támasztania a tudatosság meglétével szemben – ez pedig olyan feltétel lenne, mint K, tudniillik a személy tudása ezzel kapcsolatban nem lenne tévedhetetlen. Ebben az esetben pedig saját tudatunk vonatkozásában nem vagyunk tévedhetetlenek – ami nyitva tartja annak lehetőségét, hogy nem is vagyunk tudatosak.

Mindez azt jelentené, hogy a tudatosság „funkcionális ekvivalenseit” felmutató gépek igenis tudatosak? Rey szerint erre a bevett nézet hívei szemszögéből kézenfekvő válasz, hogy nem, mert a gépekben nincs meg valamiféle „belső fény”, ami az emberben megvan – csak hogy egy ilyen válasz megalapozatlannak bizonyul, mihelyt megpróbáljuk meghatározni, hogyan viszonyul az a figyelemhez, érveléshez, emlékezethez és más kognitív funkciókhoz. Másként fogalmazva: ha már ismerjük a kognitív folyamatokat, *mit tenne hozzá* mindehhez a tudat? (Rey 1988: 22). Tudomásunk szerint semmit – ez pedig kétségeket ébreszt azzal kapcsolatban, hogy egyáltalán létezik a tudat. Ám ha a karteziánus intuíció igazságán múlik a „tudat” fogalmának alkalmazhatósága, akkor a karteziánus intuícióval rendelkező, ám tudattalan gép pusztán lehetőségéből következik, hogy nincs tudat.

Szemben a behavioristákkal és eliminativistákkal, Rey nem azt állítja, hogy a mentális terminusok ne lennének „tudományosan respektábilisak”, vagy fel kellene adni e terminusokat a köznap diskurzusban. Közvetlenül a tudatosság speciális értelmét veszi célba, azt, ami állítólag megkülönbözteti az embereket a gépektől: a fenomenális, élményszerű aspektust. Persze ha egy gép emberszerűen nézne ki és viselkedne, azt hajlamosak lennénk tudatosnak tekinteni (Rey 1988: 23) – amit megpróbálhatunk ugyan behaviorista érvekkel igazolni, ám a helyzet az, hogy semmilyen elméletileg jól megalapozott módon nem vagyunk megkülönböztethetők a „pusztán gépektől”.

Cikkében Rey felvonultatja a revizionista intuíció híveinek szokásos érveit. Nincs okunk rá, hogy a tudományos magyarázatokban szerepet nem játszó entitásokat vagy tulajdonságokat posztuláljunk; ha intuitív meggyőződésünk az, hogy tudatosak vagyunk, ez nem perdöntő, mert intuíciónk lehet téves. Az ember és a gép között pedig nincs ismert, jól megalapozott különbség.

Smith (1988) arra mutat rá, hogy Rey téved, amikor elvárja, hogy a tudat valamilyen specifikus mentális művelet korrelátuma kell legyen. Ehelyett előfordulhat, hogy a tudat „integráns része” minden mentális műveletnek, noha ő maga nem mentális művelet.

Ez a kritika lényegében ekvivalens a kétértelműségen alapuló kritikával. Rey mindössze annyit mutat meg, hogy az ember és a hozzá hasonló képességekkel rendelkező gép között a tudatosság szempontjából releváns különbség „rejtélyes”. Még ha ebből következne is, hogy ilyen különbség nem is áll fenn, abból viszont csupán az következik, hogy az ember és a pszichológiailag hozzá hasonló gép egyformán tudatos, vagy egyformán tudattalan. Annak ellenére, hogy Rey közel kerül ahhoz, hogy világosan és kompromisszumok nélkül képviselje a revizionista álláspontot, végül mégsem ér egészen célra. Az antirevizionista állásponton lévő olvasója olvashatja cikkét akár pánpszichista tanulmányként is (tekintettel az emberi kogníció funkcionálisan sajátos jellemzőinek

triviális implementálhatóságára, így valószínű prevalenciájára a természetben és az emberalkotta gépek körében), annak ellenére, illetve pontosan azért, mert Rey világosabban már nem is fejezhetné ki, hogy a „funkcionális ekvivalensek” és a „valódi” tudatos élmény közötti különbséget nemlétezőnek vagy legalábbis „rejtélyesnek” tartja.

Rey cikkét kritizálva Smith (1988: 25) erős A-intuíciónkra hivatkozik: a tudatot kétségbe vonó filozófus „elvesztette kapcsolatát az emberi tapasztalattal” a túlzott elméletgyártás közepette. „A tudat kínos a funkcionalizmus számára... bármi is egy mentális állapot kauzális szerepe, úgy tűnik, ez a funkció tudat nélkül is ellátható lenne”, állapítja meg, amivel persze ugyanabba a kétirányú utcába téved be, mint Rey. Vegyük észre, hogy ezt a mondatot Rey is elmondhatná, csak persze a „funkcionalizmus” szót „dualizmusra” vagy „karteziánus intuíció”-ra cserélve. Smith szerint Rey érvelése végül is azon a feltételezésen nyugszik, hogy (1) a pszichológiai képességek megléte nem implikálja a tudatosságot, de (2) magunkról csak azt tudjuk, hogy pszichológiai képességeink vannak – így lehetünk akár olyan lények is, amelyek pszichológiai képességekkel bírnak (de tudattal nem). Smith szerint azonban egyszerűen nem igaz (2), és ezt mindannyian „tudjuk valahogyan”. De honnan? Onnan, hogy tudatosak vagyunk – abból ugyanis már az is következik, hogy tudjuk, hogy azok vagyunk, ami persze a karteziánus intuíció része. De mi a tudat? Smith, kezdeti magabiztossága ellenére ebben a kérdésben cserben hagy minket (Smith 1988: 27-28). A tudatot saját mentális állapotainkról való tudomásunkkal próbálja magyarázni: „Amikor tudatosan gondolkodom... akkor tudomásom van arról, hogy gondolkodom... a tudat éppen ez a tudomással bírás”. Így tehát a tudat meghatározásakor olyan kognitív funkciót jelöl meg, amely semmilyen lényeges vonását tekintve nem különbözik Rey önmagukra reflektáló számítógépeinek bizonyos képességeitől, ami meglehetősen furcsa egy Rey gondolatmenete *ellen* érvelő cikkben.

Néhány sorral később Smith pontosít: „Nem az a lényeg, hogy arra következtetünk, és így azt hisszük vagy arról alkotunk elméletet, hogy tudatos mentális állapotaink vannak. A lényeg az, hogy az ilyen mentális állapottal rendelkezvén egyúttal *tudomásom* van arról az állapotról – *megtapasztalom* azt. És ez a tudomással bírás ad nekem tudást.” Ha Smith álláspontja tényleg eltér Rey-étől, akkor a különbséget ezekben a sorokban kellene megtalálni. Nyilván a megtapasztalás (experience) és a tudomás (awareness) / tudás (knowledge) kontraszt, valamint a „tudat általi tudás” elképzelése az, amin a különbség alapszik.

A kérdéses „tudomás” magának a tapasztalatnak a jellemzője, nem pedig valami külső hozzá képest (Smith 1988: 30). Ettől még azonban lehet ez pusztán a tapasztalat valamilyen funkcionális sajátossága. Smith ezt az ellenérvet úgy kísérli meg elhárítani, hogy kifejezetten elhatárolódik attól a „Quine-tól Rortyn át a Churchlandéig és Reyig” képviselt elképzeléstől, mely szerint az elméről

tett állítások tisztán teoretikusak: a tudat nem pusztán teoretikus entitás, hanem „bizonyos módon megfigyelhető” (Smith 1988: 31), amennyiben „megtapasztaljuk” azt. Nem *megfigyeljük*, persze (Ryle 1949/1999: 253), hanem a tudatosság magának a tudatosult elmeállapotoknak a tulajdonsága. Smith végkövetkeztetése tehát az, hogy a tudatot *megtapasztaljuk* (másokban is: az empátia révén), és a tudat valamilyen magyarázandó *adottság* (nem pedig magyarázó céllal feltételezett teoretikus entitás); nem hiedelem, nem ítélet. Nem tudunk meg azonban többet annál, mint hogy a tudat *adottság*, amit nem kérdőjelezhetünk meg. A Rey és Smith között fennálló „nézeteltérés” végső soron a két intuíció (revizionista intuíció: „ha a kognitív funkciókat megmagyaráztuk, nincs szükség a tudat hipotézisére”; antirevizionista intuíció: „tudatosságunk a legnyilvánvalóbb tény”) szembeállításában merül ki.

Rey és Smith deklarálják, hogy melyik intuíciónak adnak elsőbbséget, majd amikor nézeteiket részletesen kifejtik, veszélyesen közel kerülnek egymás álláspontjához. Rey nem tudja konkluzívan bizonyítani, hogy az ember és gép közti különbség „rejtélyessége”, illetve az ilyen lényeges különbség hiánya egyirányú utca, és nem lehetséges az, hogy az ember és a gép egyformán tudatos, nem pedig egyformán tudattalan. Smith pedig nem tudja egyértelműen megkülönböztetni saját „tudomását” a Rey által elemzett önreflexióra képes gépek n-edfokú intencionalitásától, csak úgy, hogy a karteziánus intuíciókra (vö. Ryle 1949/1999: 198) hivatkozik. Ezért kettejük vitája nem megy túl érdemben az ellentétes intuíciók felmutatásán.

1.3 Antirevizionista elmefilozófia: Chalmers

Chalmers 1995/2004 az antirevizionista elmefilozófia egyik leggyakrabban hivatkozott alapműve. A cikk által népszerűsített „nehéz probléma” kifejezés a filozófusok világán kívül is ismertté vált és elterjedt, ami mutatja a chalmersi problémafelvetés intuitív, közérthető jellegét. A nehéz problémát az élmény szubjektív minősége (*experience*) veti fel: Nagel (1974/2004) terminológiájával *valamilyen* tudatos lénynek lenni, és *valamilyen* tapasztalatokat átélni. A nehéz probléma Chalmers számára magyarázati jellegű: az, hogy „miért hoz létre a[z idegrendszerben zajló] fizikai feldolgozás gazdag belső életet?”. Chalmers megkülönbözteti a tudatot (consciousness) a valamiről való tudomással bírástól (awareness): az előbbi fenomenális jellegű, míg az utóbbi csupán a „központi feldolgozás számára rendelkezésre állást” vagy más hasonló funkcionális jellemzőt jelenti. A különbségtétel azért fontos Chalmers szerint, mert szerinte számos szerző a „tudat” szó többértelműségét arra használja, hogy annak funkcionális aspektusát (tehát a

tudomással bírást) tudományosan megmagyarázva a (fenomenális) tudatproblémát is megoldottnak nyilvánítsa.

Chalmers szerint a nehéz probléma azért nehéz, mert nem csupán a kognitív „kéességek és funkciók” (Chalmers 1995/2004: 16) magyarázatával kapcsolatos. A probléma ugyanis akkor is fennmaradna, ha az összes kognitív funkció ellátásának mechanizmusa meg lenne már magyarázva. A nehéz problémára való érzékenység tehát direkt tagadása a revizionista meggyőződésnek.

Chalmers hosszasan ecseteli, hogy a tudat megjelenése a pszichológiai funkciók ellátásán túli, további explanandum, amely „nem-triviális további kérdést” vet fel a funkciók magyarázatán túl. Attól azonban nem zárkozik el, hogy megengedje, a tudatos élmény is rendelkezhet pszichológiai funkcióval.

Ezek után Chalmers megvizsgál számos, a kortárs megismeréstudományban és neurológiában megfogalmazott tudatelméletet. Így például Crick és Koch kidolgozta a tudat „neurobiológiai elméletét” (Chalmers 1995/2004: 19), mely szerint a tudat neurális korrelátumai azok a 35-75 hertzes oszcillációk, amelyek elvégzik a „megkötést”, vagyis az ugyanarról a tárgyról szóló, különféle érzékelési modalitásokból származó információk egyesítését a központi információfeldolgozás számára. Baars ezzel szemben a tudat „kognitív elméletét” fogalmazta meg, mely szerint a tudat tartalmi azok, amelyek megjelennek a „globális munkaterületen”, az elme egy központi adatfeldolgozó rendszerében (Chalmers 1995/2004: 20).

Chalmers mindkét esetben ugyanazt a hiányosságot találja: még ha az egyes elméletek sikeresek is abban az értelemben, hogy a tudatosság általuk feltételezett neurális vagy funkcionális korrelátuma valóban korrelál a tudatossággal, hiányzik annak magyarázata, hogy a Crick és Koch által vizsgált oszcillációk hogyan és miért hoznak létre tudatos tapasztalatot, valamint miért válik tudatosan megtapasztalttá a globális munkaterületen lévő információ. E problémával szembesülve a kutatók különféle továbblépési vagy menekülési irányokat keresnek. Egy részük azt magyarázza meg a tudat helyett, amit meg tud magyarázni. Mások „tagadják a jelenséget” – Chalmers szerint például Dennett is – vagy nyílt tagadással, vagy a tudatnak valamilyen sikeresen megmagyarázott pszichológiai képességgel való azonosításával. A harmadik lehetőség a probléma ráolvasászerű megoldása: a kutatók egyszerűen azt állítják, hogy amit megmagyaráztak, az a tudat. „A magyarázat releváns lépésén gyorsan túlteszik magukat”, állapítja meg Chalmers (1995/2004: 21). A negyedik stratégia a tudatos tapasztalat struktúrájával analóg struktúrákat keres az agyműködésben, az ötödik stratégia pedig a „tapasztalat szubsztrátumát” keresi az agyban. A tudatos tapasztalat ugyanis valahogyan az agyban keletkezik, és nyilván azonosítható, milyen típusú agyi folyamatokban. A

teljes magyarázathoz azonban valamilyen „magyarázati hídra” is szükség van, érvel Chalmers, mivel nem elég azonosítani, hogy milyen folyamatok hozzák létre a tudatos tapasztalatot, hanem azt is ki kell fejteni, hogy miért és hogyan hozzák létre.

Chalmers szerint a tudat sikeres magyarázatához valamilyen „extra hozzávalóra” is szüksége van. Enélkül az újabb és újabb neurofiziológiai felfedezések hozzásegítenek az agyműködés mind teljesebb megértéséhez, de a kognitív funkciók mechanizmusainak feltárása nem teszi kevésbé rejtélyessé a tudatos tapasztalatot (Chalmers 1995/2004: 23).

A fizikai magyarázatok hozzásegítenek a fizikai struktúrák megértéséhez, valamint megmagyarázhatják funkciók ellátását, sőt, „amint adottak egy fizikai elmélet részletei, a strukturális és funkcionális tulajdonságokra úgyszólván automatikusan fény derül” (Chalmers 1995/2004: 24). Csakhogy a tudatos tapasztalatra vonatkozó tények nem lehetnek fizikai magyarázatok automatikus következményei ebben az értelemben, amíg elgondolható, hogy bármely fizikai folyamat megtörténhet anélkül, hogy tudatos tapasztalat kísérné: „A tapasztalat, meglehet, a fizikai szférából emelkedik ki, de nincs implikálva a fizikai szférában” (Chalmers 1995/2004: 24). A tudat redukálhatatlan, de ez nem jelenti azt, hogy a tudat „extra hozzávaló” státusza azonos lenne a vitalizmus elismerésével. A vitalizmust ugyanis az életfunkciók reduktív magyarázata megdöntötte. Ezzel szemben a tudatot a pszichológiai funkciók magyarázata nem magyarázza meg. „A kulcs az a *fogalmi* állítás, hogy a funkciók magyarázata nem elég a tapasztalat megmagyarázásához”, jelenti ki Chalmers. A tudatos tapasztalat és az *élan vital* közti különbség az, hogy a tudatos tapasztalat explanandum, míg az *élan vital* egy magyarázó célú feltételezés volt.

Mivel a reduktív magyarázat kudarcot vall, természetes módon adódik a nonreduktív magyarázat, szögezi le Chalmers. A „reduktív magyarázat kudarca” Chalmers szerint annyit tesz, hogy „fogalmi okokból” a reduktív magyarázat sohasem lehet magyarázat (mivel az azonosságelmélet Chalmers számára „a jelenség tagadása” lenne, ha pedig a tudat nem *azonos* valamilyen fizikai folyamattal, hanem *belőle keletkezik*, akkor ennek leírásához a fizikai és tudati jelenségeket összekapcsoló „magyarázati híd” szükséges). Chalmers azt javasolja, hogy tekintsük a tudatos tapasztalatot a világegyetem fundamentális jellemzőjének az eddig ismert fizikai jellemzőkön felül. Az így előálló elmélet dualizmus, és pedig Chalmers szerint „ártatlan”, naturalista dualizmus, mivel jelenlegi fizikai ismereteinknek nem mond ellent.

A továbbiakban Chalmers az elmülethez szükséges pszichofizikai elvekkel kapcsolatban kockáztat meg spekulációkat. Az általa javasolt első alapvető pszichofizikai elv a szerkezeti koherencia elve. Az említett koherencia a tudat és a tudomással rendelkezés szerkezete között áll

fenn, és egyik következménye, hogy „ahol információ áll rendelkezésre a közlés és a globális kontroll számára, ott van ezzel összefüggő tudatos tapasztalat is” (Chalmers 1995/2004: 29). Chalmers második javasolt pszichofizikai elve szerint két hasonló finom funkcionális szerveződésű rendszer kvalitatíve azonos tapasztalattal kell rendelkezzen. Ezt az elvet a „táncoló kválék” gondolat kísérletével igyekezik alátámasztani. Aki tagadja az elvet, annak azt kell állítania, hogy az agy valamely darabjának a neurális inputok és outputok szempontjából ekvivalens szilíciumchippre való cserélése megváltoztatja az élmény kvalitatív karakterét. Ekkor ha az említett chipet az eredeti agyterületet megtartva készítik el, és egy kapcsolóval lehet a régi agyterületet vagy ahelyett annak új szilícium-pótlását az agyhoz kapcsolni, akkor a kapcsoló ide-oda forgatásával az élmény kvalitatív karaktere is ciklikusan váltakozik – ez a „táncoló kválék” esete. Csakhogy a kísérleti alany ezt a változást nem veheti észre, mert a funkcionális szerveződés változatlan marad, miközben az egyik funkcionális egység fizikai implementációs részletei változnak. Márpedig ragaszkodnunk kell ahhoz, érvel Chalmers, hogy élményeink kvalitatív karakterének változásait észre kell vennünk, ellenkező esetben ki lennénk téve annak a szkeptikus lehetőségnek, hogy kváléink mindig is „táncolnak” anélkül, hogy ezt észrevennénk. Chalmers mondanivalóját egy mondatban is összefoglalhatjuk: a kválékat a funkcionális szerveződésnek kell meghatározni, ellenkező esetben nem tudnánk kognitíve hozzájuk férni. Két megjegyzés rögtön adódik a második chalmersi pszichofizikai elv vonatkozásában. Az egyik, hogy szerzőnk ezzel veszélyesen közel kerül a funkcionalizmushoz. Felmerül a kérdés, hogy ha az élmény kvalitatív karakterét meghatározza a funkcionális szerveződés, akkor milyen magyarázó szerep hárul az „extra összetevőre”, azon kívül, hogy egyszerűen formális értelemben alátámasztja a funkcionális szerveződésről a kvalitatív karakterre irányuló következtetéseket. A másik, hogy a chalmersi elvet alátámasztani szánt „táncoló kválék” gondolat kísérlet jelentősége igencsak homályos. Mivel egy szilíciumchipnek mások a fizikai jellemzői, mint az agy neuronjainak, ezért ha az agy egy részét helyettesítik vele, az így létrejött új agy funkcionális szerveződése már csak azért sem lehet azonos az eredeti agyéval, mert hiába rendelkezik a szilíciumchip ugyanazzal a számítási struktúrával, mint az általa helyettesített agyi régió, a chip és az agy kapcsolódási pontjainak megjelenése már önmagában megváltoztatja a funkcionális szerveződést.

Vagyis az átkapcsolás a chipre semmiképpen sem azt jelenti, hogy változatlan funkcionális szerveződés és megváltozó finom fizikai struktúra mellett megváltozik az élmény (ha megváltozik), hanem azt, hogy megváltozó finom fizikai struktúra és megváltozó funkcionális szerveződés mellett megváltozik (vagy nem) az élmény kvalitatív karaktere. Ráadásul Chalmers is elismeri annak

lehetőségét, hogy a releváns funkcionális szerveződési szint akár olyan finom is lehet, hogy egy bizonyos funkcionális szerveződés fizikailag történetesen csak egyféleképpen implementálható.

Végül Chalmers harmadik pszichofizikai elvnek az információ „kétaspektusos elméletét” javasolja, mely szerint az információnak van fizikai és fenomenális aspektusa – ami nyitva tartja annak lehetőségét, hogy minden információfeldolgozást, így egy termostátét is – tudatos élmény kísérjen. Annak az állításnak azonban, hogy akár egy termostát méréseit is tudatos élmény kísérheti, revizionista és antirevizionista értelmezése egyaránt lehetséges:

(TR) Mindaz, amit értelmesen „tudatos élménynek” nevezhetünk, csupán olyasvalami, ami akár egy termostátban is lezajlik: információfeldolgozás.

(TA) Amit tudatos élménynek nevezünk, az akár egy termostátban is lezajlik, hiszen információfeldolgozás esetében mindig tudatos élménnyel van dolgunk.

Az általa kidolgozott naturalista dualizmust érintő kritikai reakciókra Chalmers egy további tanulmánnyal felelt (Chalmers 1997), melyben a vita során kikristályosodó elmefilozófiai stratégiáknak értékes taxonómiáját nyújtja.

Chalmers *A-típusú materializmusnak* nevezi azt az álláspontot, mely szerint az idegrendszer által ellátott különféle kognitív és viselkedési funkciókon felül nincs magyarázandó jelenség – mihelyt megmagyaráztuk a funkciókat, mindent megmagyaráztunk. Más szóval nincs olyan nehéz probléma, mely elkülöníthető lenne a „könnyű” problémáktól. Ez az álláspont a tudat tagadásaként vagy valami mással, így például a tudományos módszerekkel a „könnyű problémák” körében megmagyarázható kognitív funkcióval vagy tulajdonsággal (ilyen lehet Chalmers 1995/2004: 15 szerint a mentális állapotokról szóló beszámolás képessége, a figyelem összpontosításának képessége, az ébrenlét) való azonosításaként is kifejezésre juthat. Mindenesetre nincs további érdekes tény az elmével kapcsolatban, túl a funkcionális tényeken. Az *A-típusú materializmus* nem egyszerűen az az állítás, hogy a tudat azonos egy funkcióval, vagy bizonyos funkcionális szerepet játszik, hanem az, hogy nincs is különálló tudatprobléma a pszichológia kérdésein túl. Chalmers ide sorolja többek között a behaviorizmust, eliminativizmust és az analitikus funkcionalizmust, de az a posteriori azonosságelméleteket és funkcionalizmusokat nem. Ez utóbbiak ugyanis szerinte elismerik, hogy van egy valóságos jelenség, ami magyarázatot kíván, és amely fogalmilag elkülöníthető a funkciók ellátásától. Az ilyen álláspontok neve *B-típusú materializmus*. A chalmersi *A-materializmus* kategóriájában felsorolt álláspontok ugyanazok, mint amelyek Strawsonnál a „tudattagadók” álláspontjainak listáján szerepelnek. Ezek azok az elméletek, melyek – első

közelítésben – az R-intuíció által ihletett elméletként sorolhatók fel. Hogyan érvelnek az A-típusú materialisták álláspontjuk mellett, mely Chalmers szerint „rendkívül kontraintuitív”, hiszen „egy manifeszt tény” (a tudatosság tényét) tagad velünk, emberi lényekkel kapcsolatban? Az egyik módszer a nehéz probléma deflációja különféle analógiák segítségével. Így Dennett egy, az élet nehéz problémájával viaskodó vitalistát képzel el, Paul Churchland a fény nehéz problémáján („valami kimarad a fény reduktív elméletéből, ti. maga a fény”) töprengő 19. századi filozófust, Patricia Churchland pedig a hő tudományos magyarázatának esetét hozza fel. Minden ilyen esetben végül is elégségesnek bizonyult a jelenségek mögötti struktúrák és funkciók megmagyarázása a jelenség megértéséhez. Chalmers azonban amellet érvel, hogy a tudatprobléma más jellegű, mint az élet problémája: az „élet nehéz problémája” sohasem létezett úgy, ahogy Dennett elképzelte, még a vitalisták számára sem.

Ami Churchlandék fény/hő analógiáit illeti, Chalmers szerint ez a kifogás azért szerencsétlen, mert az egyetlen dolog, amit „kihagy” a fény reduktív magyarázata, az a szubjektív fénylátás-élmény. Emiatt pedig a tudat-kérdés deflációja helyett azt csak újból felteszik. Ezt azonban nem veszik észre, tehetjük hozzá, mert az R-intuícióból kiinduló gondolkodók számára az, ami Chalmers szerint „nyilvánvaló”, korántsem az, hanem értelmetlenség. Egy revizionista számára a chalmersi „hiányzó összetevő” egy olyan összetevő, amely senkinek sem hiányzik.

Chalmers alig talál példát ellenfelei válaszaiban az érvelésre annak határozott állításán kívül, hogy a funkciók megmagyarázása márpedig elégséges a tudat megmagyarázásához. Chalmers szerint ennek során Dennett közel kerül a tudat tagadásához, amivel kapcsolatban azt a megjegyzést teszi, hogy Dennett talán a saját fenomenológiáját híven írja le, ha másokét nem is, amivel burkoltan arra céloz, hogy Dennett filozófiai zombi. Ugyan ezt a megjegyzést nyilvánvalóan nem szánja teljesen komolynak, az is elsietett lenne, ha ezt a fajta lehetőséget teljesen figyelmen kívül hagynánk. A későbbiekben erre a kérdésre még visszatérünk. Chalmers végső magyarázata a közte és Dennett között fennálló nézeteltérésre azonban nem Dennett fenomenológiájának az övétől való radikális eltérése, hanem az, hogy Dennett álláspontja az ő „harmadik személyű abszolutizmusának” következménye, míg a nehéz probléma éppen hogy az első személyű nézőpontból feltároló minőségekkel kapcsolatos. Dennett ragaszkodása a harmadik személyű nézőponthoz így meggátolja, hogy érdemben hozzászóljon a kérdéshez.

Chalmersnek ez a megjegyzése meglehetősen érdekes, mert eddig viszonylag keveset beszélt az első személyű nézőpont jelentőségéről. Hiába veszi ugyanis fel valaki a tudatosságot akár a természet fundamentális tulajdonságai közé: amíg csupán harmadik személyű nézőpontból írja le világot, addig nem tudja meggyőzően állítani azt, hogy az általa leírt dualista világ nem „zombi-

világ”, vagyis olyan világ, amiben nincsenek tudatos lények, akiknek „valamilyen” lenni, mert az utóbbi tulajdonság csak első személyű nézőpontból tárul fel.

Ezért Chalmers az első személyű nézőpont fontosságát is hangsúlyozza, és összeköti az első személyű nézőpont és a „milyenség” kérdését. De hogy pontosan hogyan, az végül is tanulmányából nem derül ki egyértelműen. Az alapvető probléma az, hogy a természettudomány harmadik személyű nézőpontból írja le a természetet, és ha *emiatt* nem lehetséges a tudat természettudományos eszközökkel történő megragadása, akkor ezen nem segít az sem, ha további fundamentális tulajdonságokat és alapvető természeti törvényeket posztulálunk, hiszen bármennyire is kibővítjük a természettudományt, az továbbra is harmadik személyű nézőponttal fog rendelkezni.

Így aztán természetes, ha egy „harmadik személyű abszolutista” filozófus vagy tudós fölöslegesnek ítéli ezeket a posztulátumokat, melyek semmiféle szerepet nem játszanak a tudományos magyarázatokban (melyek harmadik személyű nézőpontból jellemzik a világot).

Ha mármost feltételezzük, hogy létezik első személyű nézőpont, ahonnan feltárulnak a tudat fenomenális minőségei, akkor nem világos, hogy ennek megmagyarázásában miféle szerepet játszhat a (harmadik személyű nézőponttal rendelkező) tudomány ontológiájának sajátos nem-fizikai tulajdonságokkal való kiegészítése.

Ha az első személyű nézőpontból tárulnak fel a tudat fenomenális tulajdonságai, akkor ha nem magyaráztuk meg, micsoda az első személyű nézőpont, akkor a tudatosságot sem magyaráztuk még meg pusztán azzal, ha feltételeztük, hogy vannak nem-fizikai tulajdonságok (Hellie 2013, Levine 2001: 177).

Van-e tudatosság egy olyan világban, ahol vannak nem-fizikai tulajdonságok, de nincsen első személyű nézőpont, vagy egy olyan világban, ahol van első személyű nézőpont, de csak fizikai tulajdonságok vannak? Van-e „nehéz problémája” az első személyű nézőpontoknak, vagy megmagyarázható a létezésük pusztán struktúrák és funkciók terminusaiban?

Ezek a kérdések azért lehetnek fontosak, mert bár nyilvánvaló, hogy Chalmerst a dualizmus elfogadására az A-intuíció motiválta, korántsem egyértelmű, hogy a puszta dualizmus (nem-fizikai tulajdonságok vagy entitások feltételezése) feltevéséből egyértelműen következik a tudatosság létezésének feltételezése.

Amikor ugyanis Chalmers az első személyű nézőpont jelentőségére utal, és Dennett hajthatatlan álláspontját harmadik személyű abszolutizmusként jellemzi, akkor mintha ő maga is

érezelné, hogy a dualizmusnak létezik egy olyan „harmadik személyű abszolutista” olvasata, amelynek fényében a dualizmus feltételezése feleslegesnek tűnik.

Ha ez így van, akkor az antirevizionizmus meggyőző képviseléséhez nem dualizmusra, hanem az első személyű nézőpont antirevizionista elméletének kidolgozására lenne szükség. Ehhez persze nem elég kimondani, hogy „létezik első személyű nézőpont”, hiszen az első személyű nézőpont fogalmának könnyűszerrel adhatjuk például olyan funkcionalista magyarázatát, amely biztosan nem elégítené ki az antirevizionistákat.

Dennett elutasítja Chalmersnek azt a nézetét, hogy a tudatosság túl van a struktúrákon és funkciókon: szerinte „kivonva a funkciókat, semmi sem marad”, amivel szemben Chalmers (Rosenberg nyomán) a következő analógiát alkalmazza: a szín tulajdonságai a színárnyalat, teltség és fényesség. Ha „kivonjuk” a színárnyalatot, semmi látható nem marad, ám ez nem jelenti, hogy a szín csak a színárnyalat. Hasonlóan, még ha igaz is lenne, hogy a funkciók a tapasztalat szükséges összetevői, ebből nem következne, hogy a tapasztalat megértéséhez csakis funkciókat kellene megmagyarázni.

Chalmers számára azonban a legfontosabb törésvonal a tudatproblémával kapcsolatban azok között húzódik, akik szerint van egy magyarázatra szoruló jelenség, és azok között, akik szerint nincs, azaz az A-típusú materialisták és a többiek között. „A vita köztem és Dennett között”, állapítja meg Chalmers, „az első személyű fenomenológiával kapcsolatos bizonyos alapvető intuíciónkra vezethető vissza” (Chalmers 1997, 2.3 fejezet).

Dennett ugyanis felszólította Chalmerst, hogy hozzon fel „független bizonyítékokat” a tudatos tapasztalat „posztulátuma” mellett, amivel szemben Chalmers megállapítja, hogy számára a tapasztalat maga az explanandum, nem pedig más jelenségek megmagyarázásához felhasznált posztulátum. Szerinte Dennett csupán előfeltételezi, hogy az egyedüli explanandumok a funkciók. Ebben természetesen igaza van – csakhogy Dennettnak is igaza van abban, hogy a „tapasztalat” mellett semmiféle „független bizonyíték” nem hozható fel – ha létezne efféle bizonyíték, akkor az elmefilozófia művelése könnyebb lenne. Ezen a ponton beáll a patthelyzet Dennett és Chalmers vitájában, mely jól illusztrálja azt a bizarr helyzetet, hogy míg az antirevizionisták számára a tudat létezése a legnyilvánvalóbb tény, a revizionisták hiába kérnek erre bizonyítékot.

Az intuíciónk összecsapásával kapcsolatban Chalmers számos érdekes észrevételt tett. Így például felmérések eredményeire hivatkozik, melyek szerint a válaszolók nagy része szerint a funkciókon kívül valami más is magyarázatra szorul az elmével kapcsolatban – konszenzusról azonban nincs szó. Dennett maga is elfogadja, hogy álláspontja „radikális” és „kontraintuitív”.

Chalmers így arra a magyarázatra hajlik, hogy Dennett álláspontja inkább valamilyen elméleti axiómából – Dennett „harmadik személyi abszolutizmusából” – származik, mint intuitív meggyőződésből. Ugyanakkor Chalmers sem tagadja, hogy vannak, akik eleve az A-típusú materializmust tartják meggyőzőnek.

Chalmers a B-típusú materializmussal sem elégedett, ugyanis szerinte a tudat megmagyarázására ez sem alkalmas. Ugyan a B-materializmus elfogadja, hogy a funkciók ellátásán felül még van valami megmagyarázandó, de a magyarázatot materialista fogalmi keretben igyekszik megtalálni. Chalmers szerint azonban ez reménytelen. A fizikai elméletek ugyanis szerinte végső soron „struktúra és dinamika terminusaiban” vannak specifikálva. A fizikai elméletek alapvető fizikai struktúrákat, valamint ezek időbeli viselkedésének törvényszerűségeit írják le. És bár az egyszerű struktúrák bonyolultabb struktúrákat alkothatnak, és így végső soron az egyszerűbb struktúrák dinamikáját megértve összetettebb rendszerek viselkedése is megmagyarázható, Chalmers szerint „akármennyi struktúra és funkció is csak struktúrává és funkcióvá állhat össze”. Számos területen a struktúrák és funkciók alapvetőbb struktúrák és funkciók alapján történő magyarázata elégséges, a tudatosságot azonban a B-materialisták szerint ilyen módon nem lehet magyarázni, mert ott a struktúrán és funkción felül valami „extra összetevőre” is szükség van. Mivel a B-materialista szemben az A-materialistával elismeri, hogy a struktúrákon és a funkciók ellátásán felül valamilyen további tény a tudatosság, ezért a B-materialista számára minden funkció és struktúra (alapvetőbb struktúrák és funkciók alapján történő) megmagyarázása után is nyitott marad a kérdés, hogy *miért* kíséri bizonyos struktúra és funkciók működését tudatos élmény. A tudati és fizikai folyamatok *a posteriori* azonosítását Chalmers szintén elutasítja. Érvelésének lényege, hogy mivel a „tudattal kapcsolatos tények” (hogy melyek ezek, azokat nem részletezi) nem „esnek csak úgy ki” a neurális folyamatok struktúráival és funkcióival kapcsolatos tényekből, legalábbis ha elutasítottuk az A-materializmust, mint ahogyan a DNS struktúrájával és funkciójával kapcsolatos tényekből automatikusan adódnak a génekkel kapcsolatos tények. Ugyanis mindig föl lehet tenni a kérdést: *miért* kíséri ezeket a folyamatokat tudatos élmény?

Chalmers természetesen a standard elgondolhatósági érvelés (mely szerint elképzelhetőnek tűnik, hogy van agyi esemény mentális esemény nélkül, tehát a kettő megkülönböztethető, így nem lehet azonos) egy változatát alkalmazza, a hangsúllyal ezúttal a magyarázat fogalmán. A „tudattal kapcsolatos tények” magyarázatra szorulnak, a magyarázat igénye pedig azt jelenti, hogy a tudatnak olyan relációban kell állnia az agyi eseményekkel, amely lehetővé teszi a magyarázatot, de nem azonosság. Miért nem lehet azonosság?

Először is a B-materializmus feltételezése, hogy a tudat valamilyen, a funkciók ellátásán felüli további, azoktól legalább fogalmilag megkülönböztethető jelenség. Ezért a fizikai folyamatok és a tudat jelenségei között a B-materialista szerint nincs „fogalmi implikáció”, a releváns fizikai folyamatok meglétéből nem következik *a priori*, hogy a tudatosság is jelen van.

Másodszor, ha a funkciók ellátása és a tudatosság közötti viszonyt „empirikus azonosságnak” fogjuk fel, mint Clark és Hardcastle, akinek az elméletét Chalmers bírálja, akkor ez az azonosság a világ egy „magyarázatilag primitív ténye”. Bizonyos funkcionális vagy fizikai állapotok tudatos állapotok, ami tovább nem magyarázható, pusztán tény. A természetben azonban efféle magyarázatilag primitív relációkat csak a természet alapvető törvényei esetében találunk.

Az *a posteriori* azonosításokat felhasználó B-típusú materializmusnak Chalmers így azt veti ellene, hogy ez az elmélet elüt azoktól az empirikus azonosságtézisektől, amelyek a természettudományban előfordulnak, amennyiben egyszerűen állítják (és nem magyarázzák) a mentális-fizikai azonosságot. „De az azonosítások nem így működnek: minden más esetben meg kell értük dolgozni” (Chalmers 1997, 2.3 fejezet). Így például a gén azonosítása a DNS-sel olyan módon történik, hogy a tudósok kimutatják, hogy a DNS molekula mely tulajdonságai teszik lehetővé, hogy a DNS molekula géneként funkcionáljon. Egy ilyen magyarázattal dolgozik meg a tudomány az azonosításért. Chalmers szerint az elemfilozófiai azonosságelmélet kezdeményezői (Place és Smart) megértették ezt a gondolatot, és a tárgyközömbös elemzés manőverével igyekeztek biztosítani a tudatos tapasztalat és az agyi események azonosíthatóságát. A narancssárga színélményt például úgy elemezték, hogy az az állapot, amelyet bizonyos körülmények között narancssárga tárgyak idéznek elő bennünk. Chalmers szerint ez a stratégia „az A-típusú materializmus összes problémájában szenved, de legalább elismeri, mi kellene ahhoz, hogy a tézis igaz legyen”. Ezzel szemben Chalmers szerint a B-materialista „a magyarázat *helyett* tételez egy azonosítást”.

Chalmers e két tömör megjegyzését talán a következőképpen értelmezhetjük. Szerinte Place és Smart nyújt ugyan valamiféle beszámolót arról, hogy hogyan jön létre a narancssárga színélmény, csak éppen az nem biztos, hogy az általuk kauzális-funkcionális terminusokban meghatározott „narancssárga színélmény” csakugyan az a tudatos színélmény, ami Chalmers számára felveti a tudat nehéz problémáját. Place és Smart tehát megmagyaráznak valamit, csak éppen annak számára, aki elutasítja az A-materializmust, egy további lépés még hiányzik ahhoz, hogy ez a tudat teljes magyarázata legyen. A B-materialisták ezzel szemben egy „magyarázatilag primitív axiómával” próbálják összekötni a fizikai folyamatokat a tudat jelenségeivel.

Az ilyen típusú magyarázat nem olyan, mint a tudományos azonosítások (mint a gén azonosítása a DNS-sel). Nem is tekinthető a tudat reduktív magyarázatának. A tudatot és a fizikai jelenségeket összekapcsoló, magyarázatilag primitív „áthidaló elvet” hiába neveznénk azonosságnak, ezzel Chalmers szerint a materializmusnak csak a betűjét tartanánk meg, a szellemét nem. Így a B-materializmus végső soron Chalmers szerint közelebb áll az ő naturalista dualizmusához, mint egy standard materializmushoz, még akkor is, ha a fizikai és fenomenális között kapcsolatot teremtő áthidaló törvényeket „azonosságoknak” nevezi.

Úgy tűnik tehát, hogy Chalmers számára a mentális-fizikai azonosságelméletet valójában csak A-típusú materializmusként lehet képviselni, azt azonban kontraintuitívnek vagy a tudatprobléma melletti elbeszélésnek tartja. Így Chalmers szerint a tudat nehéz problémájára csak olyan elmélet lehet megfelelő válasz, amely nem reduktív: Chalmers szerint márpedig egy adekvát nonreduktív elmélet a tudatos tapasztalatot a természet fundamentális alkotóelemei közé veszi fel, a természet ismert alapvető törvényeit pedig pszichofizikai törvényekkel bővíti ki.

Ha az A-intuíció kielégítéséhez elég a revizionista meggyőződések határozott tagadása, akkor Chalmers célt ér írásában – szándéknyilatkozatai szintjén mindenképpen. Egyértelműen kinyilvánítja, hogy szerinte van valami, ami az elmével kapcsolatban megmagyarázandó, és több, mint struktúra és funkció. Az azonban korántsem biztos, hogy az A-intuíció tartalma ennyi és nem több. Chalmers a „nehéz probléma” bemutatásakor a „milyenség”-re hivatkozik, végső soron Nagelre. Nagel (1974/2004) problémafelvetése felől nézve azonban Chalmers kevesebb haladást ér el, mint amennyit a tudat kvalitatív aspektusára vonatkozó R-intuíció tagadásában. Nagel számára ugyanis az igazán fontos témát az első és harmadik személyű nézőpontok közti kontraszt jelenti. A harmadik személyű nézőpontból szemlélt fizikai univerzum kiegészítése egy további fundamentális természeti jellemzővel, mint Chalmers naturalista dualizmusában, nem feltétlenül érinti ezt a kérdést. Egy olyan elméletben pedig, ahol az első személyű nézőpontok eredete például valamilyen funkcionista magyarázatot kap, a tudatproblémát pedig az első és harmadik személyű nézőpont kontrasztja hozza létre, szükségtelen és a tudat kérdésének megoldásához nem feltétlenül járul hozzá a pszichofizikai törvények vagy fundamentális pszichikai tulajdonságok vagy entitások feltételezése.

1.4 A laikusok intuíciói

A korábbiakban szó volt a filozófusok intuitív meggyőződéseiről és az ezeknek megfelelő elmefilozófiai álláspontokról. De mik a laikusok elmefilozófiai intuíciói? A kérdés azért nem érdektelen, mert a filozófiai vitákban olykor felmerül, hogy a filozófusok által természetesnek vagy maguktól értetődőnek tartott bizonyos meggyőzések nem is annyira természetesek, hanem filozófiai újítások eredményei, és csak a filozófiailag iskolázottak számára tűnnek „természetesnek”. A közelmúltban lezajlott néhány „kísérleti filozófiai” vizsgálat, vagyis olyan felmérés, amely nem-filozófus megkérdezettek elmefilozófiai intuícióit kísérte meg feltérképezni. E vizsgálatok célja tisztázni, hogy voltaképpen mik is a természetes és nyilvánvaló meggyőzések a tudattal kapcsolatban.

A vizsgálatokat az motiválja, hogy szélsőséges egyet nem értés van a filozófusok között azzal kapcsolatban, ami elvileg „nyilvánvaló” (Sytsma 2010: 60): úgy tűnik, nagyon is lehet vitatni azt, ami egyesek szerint vitán felül áll (Sytsma 2010: 63). Annak ellenére, hogy sok filozófus szerint nyilvánvaló, sőt a legnyilvánvalóbb tény a fenomenális tudatosság, a kválék létezése, mégsem sikerül definiálni a fenomenális tudatosságot. Legfeljebb példák felsorolására szorítkoznak, abban bízva, hogy ha „rámutatnak a jelenségre”, akkor hallgatóságuk megérti, miről van szó, amivel természetesen az a baj, hogy nem világos, hogyan lehet egyértelművé tenni, hogy amikor rámutatunk egy paradicsomra („a paradicsom pirossága”) vagy egy csésze kávéra („a kávé illata”), akkor kválékra mutatunk rá, nem pedig fizikai tulajdonságokra vagy tárgyakra (Sytsma 2010: 64-66). A példák felsorolása sem megfelelő módszer, hiszen a túlságosan heterogén példák (a sör íze, Beethoven IX. szimfóniájának hangzása, a túsúrás érzete stb.) esetében nem magától értetődő, hogy mi bennük a közös.

Másfelől, állapítja meg Sytsma (2014: 3-4) az az állítás, hogy a fenomenális tudatosság valamilyen nehéz rejtvényt ad fel, maga is rejtélyes. Éppenséggel kételkedhetünk is e jelenség realitásában, már csak azért is, mert a minden magyarázatnak való makacs ellenállása felveti annak lehetőségét, hogy fogalmilag nem jól körülhatárolt explanandumról, rosszul feltett kérdésekről van szó. Miért indulnak ki tehát annyian abból, hogy „ez a rejtélyes jelenség” egyáltalán létezik? Azért, mert létezése evidensnek tűnik, és szemlátomást nagy szerepet játszik mindennapi életünkben. Sytsma Chalmerst idézi, aki szerint „semmi sincs, amit bensőségesebben ismernénk, mint a tudatos tapasztalatot”, és Stoljart, aki szerint „fenomenológiailag nyilvánvaló” a fenomenális tudatosság megléte. Sytsma szerint azonban kételkedhetünk benne, hogy nyilvánvaló. Ha nyilvánvaló lenne, akkor például azt várhatnánk, hogy tudunk róla. Mihelyt azonban feltesszük a kérdést, hogy tudunk-e akár saját fenomenális tudatosságunkról, és honnan tudunk róla, problémákba ütközünk. Dretske

(2003) például nem jut messzire annak megválaszolásában, honnan tudjuk, hogy nem vagyunk zombik, noha ő a maga részéről biztos ebben.

Ha a fenomenális tudatosság jelenléte nyilvánvaló, sőt „fenomenológiailag nyilvánvaló” tény, amellyel a legbensőségesebb ismeretségben vagyunk, akkor az utca emberének, aki filozófiai tanulmányok által még meg nem rontott intuíciókkal rendelkezik, minden bizonnyal rendelkeznie kellene a fenomenális tudatosság valamilyen fogalmával. Ennek eredtek nyomába a „kísérleti filozófiai” kutatások, melyeket Sytsma foglalt össze.

Így például Knobe és Prinz (Sytsma 2014: 6-9) felmérése szerint a laikusok „természetellenesnek”, furcsának tartották a fenomenális mentális állapotok (*fájdalmat érez* stb.) tulajdonítását például vállalatoknak, míg a nem-fenomenális mentális állapotok tulajdonítását (a cég *azt hiszi, azt várja* stb.) elfogadhatónak tartották. A kísérleti eredmény szerintük azt bizonyította, hogy a mentális állapotok között a laikusok is úgy tesznek különbséget, hogy az megfelel a fenomenális/nem fenomenális felosztásnak.

Hozzátehetjük, hogy ez az eredmény talán azt is jelenti, hogy a laikusok szerint a fenomenális mentális állapotok nem oszthatók fel tetszés szerint vagy tulajdoníthatók absztrakt entitásoknak, vagyis természetük valamilyen értelemben *konkrétabb*, mint a nem-fenomenális állapotoké.

Huebner (Sytsma 2014: 11-12) pedig úgy találta, hogy egy olyan szituációban, ahol embernek, illetve robotnak tulajdoníthatnak mentális állapotokat, a laikusok hajlamosak különbséget tenni az ember és a robot között bizonyos állapotok tekintetében, míg ez nem jellemző a többi állapotot illetően. Általában majdnem ugyanolyan készek voltak olyan, a filozófusok által nem-fenomenális állapotoknak nevezett állapotokat (így például hiedelmeket) tulajdonítani a robotnak, mint az embernek. Fájdalmat és boldogságot azonban jóval kisebb (20, illetve 40%-os) arányban tulajdonítottak robotnak, mint embernek. Nyilvánvaló a kontraszt aközött, ahogyan a laikusok a fenomenális, illetve nem-fenomenális állapotokat tulajdonítják, amiből azonban nem következik, hogy a felmérés egyértelműen azt bizonyította volna, hogy a laikusok természetes meggyőződéseik alapján különbséget tesznek a fenomenális és nem-fenomenális fogalmak között. A kutatások nem közvetlenül azt vizsgálták, rendelkeznek-e a laikusok a fenomenális mentális állapot fogalmával. Bármilyen „laikus fenomenális mentális tulajdonság fogalom”, aminek működésére ezekkel a felmérésekkel következtetni lehet, legfeljebb extenzióját tekintve tűnhet azonosnak a fenomenális mentális tulajdonság fogalmával. Az eredményeket, mint maga Huebner is megállapította, aztán többféleképpen lehet értelmezni, és talán ez ezeknek a felméréseknek a

legfontosabb eredménye. Így például a fájdalomnak és a boldogságnak hedonikus karaktere is van, nemcsak fenomenális, és lehet, hogy a laikusok erre gondolva nem tulajdonították ezeket az állapotokat a robotoknak, vagy egyszerűen a robot-technológia feltételezett korlátai miatt. (Sytsma 2014: 13).

Sytsma rámutatott, hogy a Knobe-Prinz felmérés eredményét úgy is lehet értelmezni, hogy a laikusok különbséget tettek a cégeknek is, és az azoknak nem tulajdonítható mentális állapotok között, de ennek nem feltétlenül van köze ahhoz, hogy a kérdéses állapotok fenomenálisak-e. Phelan és munkatársai kimutatták, hogy a laikusok a csoport-ágenseknek „disztributivista” értelemben tulajdonítják a mentális állapotokat, ezért a Knobe-Prinz típusú felmérések kérdéseire adott válaszok a példamondatok névmáshasználatától is függenek.

Sytsma és Machery felmérésében egy robotnak, illetve egy embernek tulajdoníthattak a kutatás alanyai olyan állapotokat, mint például „pirosat lát”, illetve „fájdalmat érez”. Míg a filozófiailag nem teljesen képzetlenek a kétféle állapotot hasonlóan kezelték (mindkettőt tulajdonították az embernek, egyiket se a robotnak), addig a filozófiát nem ismerők ettől abban tértek el, hogy a robotnak is tulajdonították a „pirosat lát” állapotot (Sytsma 2014: 20). Sytsmaék következtetése: mivel a laikusok nem veszik egy kalap alá a piros-látás élményét a fájdalomérzetével, holott mindkettő fenomenális állapot, ezért nyilvánvalóan nem tartják úgy, hogy valami közös ezekben az állapotokban – tehát nem használják a fenomenális állapot fogalmát. Sytsma és Machery arra következtetett, hogy a laikus állapot-tulajdonítások szempontjából a fenomenális minőség helyett a hedonikus érték lehet lényeges, ám későbbi kutatási eredményeik alapján ezt is kétségbe vonták.

Ennél a pontnál azonban érdemes elidőzni, mert szemlátomást, amikor a laikusok elvitatták a robottól a fájdalomérzés képességét, akkor ezt nem fenomenális tulajdonságokra hivatkozva tették, hanem arra, hogy egy robot nem képes kedvelni vagy utálni dolgokat (Sytsma 2010: 107). Ez a megfigyelés megmutatja, milyen elsietett dolog a hétköznapi fájdalomfogalmat lényegileg fenomenális minőségekkel kapcsolatos fogalomnak tekinteni. Ez pedig azért fontos, mert a materializmus elleni Kripke-típusú érvek támaszkodnak egy „nyilvánvalónak” tartott, lényegileg fenomenális mentálisállapot-, pontosabban fájdalom-fogalomra.

A fájdalom olyan állapot, amelynek van hedonikus, fenomenális és egyúttal kognitív, tágabban pedig viselkedési aspektusa is. Hasonlóan, a piros szín látásának észlelési és így kognitív aspektusa is van. Nehéz megmondani, hogy a laikus, szembesülve a kérdésekkel, milyen gondolatmenet alapján dönt. Talán a „látást” észlelési/kognitív siker-szóként értelmezi Ryle

(1949/1999:274) módjára? Talán a robot- és az emberi látás között képes különbséget tenni, de ezt az adott kontextusban nem tartja lényegesnek?

Éppen ez volt az a pont, ahol Huebner, Peressini, Fiala, Arico és mások megtámadták Sytsma és Machery kritikáját (Sytsma 2014: 23). Peressini egy felmérésében arra kérdezett rá, hogy a robotokat „tapasztalónak” (experiercer) tartják-e a megkérdezett laikusok. A fogalmat azonban be kellett vezetnie: „Ezzel szemben... a termosztátok, betörésjelző riasztók és kenyérvágógépek nem rendelkeznek belső étellel: nincs olyan, hogy amilyen ezeknek a tárgyoknak lenni... ezek nem tapasztalók”. A megkérdezettek a robotokat nem sorolták a tapasztalók közé, ám a kézenfekvő ellenvetés az, mint Sytsma is megállapította, hogy Peressini gyakorlatilag megtanította a vizsgálati alanyokat a „tapasztaló” fogalmára, nem pedig azt kutatta, hogy azok rendelkeznek-e eleve hasonló fogalommal.

Peressini feltett néhány, az észlelés fenomenális aspektusával kapcsolatos klasszikus kérdést is („biztosak lehetünk-e valaha abban, hogy ugyanolyannak látjuk a pirosat, mint mások?”). Az eredmények bizonytalanok és nagy szórást mutattak, noha Peressini számára elégségesen negatívak voltak ahhoz, hogy arra következtessen: az utca embere rendelkezik a fenomenális tudat fogalmával. Ez a következtetés is elsietett, hiszen más okból is lehetünk bizonytalanok abban, mások olyannak látják-e a pirosat, mint mi, például egyszerűen azért, mert a gyakorlatban nehéz pontosan összehasonlítani különböző személyeknek a pirosárnyalatok megkülönböztetésére való képességét, vagy azért, mert különböző személyeknek eltérő elképzeléseik lehetnek arról, mely színárnyalatok tartoznak a „piros” fogalmának terjedelmébe (Sytsma 2014: 26).

Sytsma egy jellegzetesen filozófiai tartott kérdést adott fel laikus vizsgálati alanyainak: lehetséges-e a spektrum inverzió? Úgy tűnik ugyanis, hogy sok filozófus szerint a spektrum inverzió lehetőségességebe vetett hit lényegében a szinkválék melletti elkötelezettséggel egyértelmű. (Sytsma 2010: 108). A mások által kéknak látott paradicsom esetére vonatkozó kérdésre az 1...7 skálán 3.33-as átlaggal és 2.2-es szórással érkeztek válaszok. A meglehetősen bizonytalan, és inkább az elutasító tartományban szóródó eredmények azt mutatják, hogy a megkérdezetteknek legalábbis egy része nem tartja a színeket kváléknak – a laikus felfogás a fenomenális minőségekről minden, csak nem nyilvánvaló és egyértelmű.

A felmérések középre húzódó, nagy szórást mutató eredményéből ennyit biztosan megtudhatunk, ha határozottabban nem is. Érdekes lett volna, ha a kísérleti filozófusok feltették volna azt a kérdést, hogy a megkérdezettek egyáltalán értik-e a spektrum inverzió problémáját. Erre a kérdésre talán határozottabb válaszok születtek volna, ugyanakkor ez a kérdés fontosabb, mint az,

hogy hogyan döntenek el a laikusok a spektrum inverzió lehetőségére irányuló kérdést: akár így, akár úgy válaszolnak erre, ha *értik*, akkor mindenesetre valamilyen értelemben birtokában vannak a kváléfogalomnak.

Sytsma végkövetkeztetése az, hogy a fenomenális tudatossággal kapcsolatban semmi sem olyan „nyilvánvaló”, mint számos filozófus feltételezi. Számunkra különösen fontos – és ez indokolja a „kísérleti filozófia” eredményeinek részletes áttekintését – az a mintázat, amit a fentebb említett viták követnek. Vagy közérthető, hétköznapi szavakkal megfogalmazott, természetes kérdéseket tesznek fel a kísérleti filozófia művelői a laikusoknak, ám ebben az esetben a laikusok válaszait értelmező filozófusok két táborra szakadnak: minden felmérési eredményt lehet úgy értelmezni, hogy az alátámasztja, hogy a laikusok rendelkeznek a fenomenális mentális állapot fogalmával, de ellentétesen is. Vagy pedig kellő mértékben pontosítják kérdéseiket – akár filozófiai terminus technicusok bevezetésével –, ám ekkor a válaszokkal kapcsolatban, noha tartalmilag egyértelműbbek, a kétértelműség visszaszökik a hátsó kapun: az a gyanú merül fel, hogy a filozófus nem a laikus intuícióit csalogatta elő, hanem egy bizonyos filozófiai elméletbe indoktrinálta, vagy olyan fogalomhasználatra szoktatta rá, amely befolyásolta a gondolkodását.

1.5 Összefoglalás: a két alapvető intuíció

Ebben a fejezetben olyan elmefilozófiai gondolatmeneteket vizsgáltam meg, amelyeknek a szerzői nem értenek egyet abban, hogy mik azok a „nyilvánvaló tények”, amelyeket minden filozófusnak el kell fogadnia. Chalmers szerint a fenomenális tudatosság létezése nyilvánvaló, ám a tudomány számára a megmagyarázása olyan *nehéz problémát* jelent, amely más jellegű, mint a tudományban általában felmerülő „könnyű problémák”.

Míg ugyanis a tudományos magyarázatokban, mint például az észlelés vagy az emlékezés neurofiziológiai magyarázatai esetében valamely funkció ellátásának módját magyarázzák meg a kutatók annak a struktúrának az alapos leírásával, amely ezt a funkciót ellátja, a tudatosság nem magyarázható ilyen módon, mert az valamilyen, a pszichológiai funkciókon túli „extra összetevő”.

Dennett elutasítja, hogy a tudatosság magyarázása ilyen értelemben nehéz problémát vetne fel, szerinte ugyanis az összes „könnyű probléma” megoldása választ adna az összes olyan kérdésre is, amit a tudatossággal kapcsolatban lehet feltenni. Ez azt jelenti, hogy a tudatosság nem a

funkciókon és struktúrákon felüli „extra összetevő”. Ami pedig a tudatosság létezésének nyilvánvaló voltát illeti, Rey szerint tudatosság nem is létezik.

Az eddig megtekintett példák alapján most megkísérelhetjük pontosabban megfogalmazni azokat az intuitív meggyőződéseket, amelyek az ilyen radikálisan eltérő elmefilozófiai koncepciók háttérében húzódnak. Összegezve az eddigieket, most pontosabban is megfogalmazom a revizionista és az antirevizionista meggyőződést, nemkülönben azokat a revizionista és antirevizionista téziseket, amelyeket a filozófusok hajlamosak az intuitív meggyőzések hatására elfogadni. Úgy vélem, hogy az elmefilozófusok gyakorta juttatják kifejezésre az alábbi revizionista, illetve az antirevizionista téziseket, és az elmefilozófiai elméletek jól csoportosíthatók azon szempont szerint, hogy melyik fajta intuitív téziseket képviselik.

Az antirevizionista meggyőzések kialakulásában az antirevizionista intuíció (AI), a revizionista meggyőzések kialakulásában pedig a revizionista intuíció (RI) játszik szerepet. Ezeknek a pontos tartalmát egyelőre nem határozom meg, noha első közelítésben az antirevizionista intuíciót lehetne jellemezni úgy, mint azt az intuíciót, mely szerint „létezik tudatosság”, a revizionista intuíciót pedig úgy, mely szerint „nem létezik tudatosság”. Egyelőre annak a kérdésnek is függőben hagyom, hogy aki revizionista intuitív meggyőzéssel rendelkezik, az valamilyen pozitív revizionista intuícióval rendelkezik-e, vagy pedig az ő esetében csupán arról van szó, hogy nem rendelkezik az antirevizionista intuícióval (és csupán ebben az értelemben rendelkezik a revizionista intuícióval, ami az antirevizionista intuíció tagadása).

A filozófiai elméletekben nem közvetlenül maguk az intuíciók, hanem azok az intuitív meggyőzések, illetve az ezek alapján kidolgozott filozófiai tézisek jutnak kifejezésre, amelyeket az adott intuícióval rendelkező filozófus alakít ki. A filozófiai elméletekben azokkal a különféle revizionista vagy antirevizionista tézisekkel van dolgunk, amelyeket az kapcsol össze, hogy háttérükben revizionista, illetve antirevizionista meggyőzések húzódnak meg, melyeket a filozófusok olykor explicit módon is kifejeznek.

Kezdjük a revizionista meggyőzéssel és három revizionista tézissel.

(RM) Amennyiben megmagyaráztuk az érzékelés fiziológiai mechanizmusait, vagy a kognitív funkciók működését, illetve azt, hogy hogyan áll elő a viselkedés, nincs további teendőnk az elme működésének megértését illetően. Csak a viselkedés megfigyelhető, így csak az

magyarázandó. Nincs független bizonyítékunk arra, hogy az emberi látás esetén valami nehezen meghatározható és magyarázható többlettel is számot kell vetnünk, túl a fényérzékelésen, alakok felismerésén stb. Ha egyszer elkészül az emberi viselkedést megbízhatóan bejósoló pszichológiai elmélet, abból nem fog hiányozni semmi, hiszen ez az elmélet azt is meg fogja jósolni, hogy némelyek miért beszélnek a „tudat rejtélyéről”. A tudomány útjában nem áll semmiféle leküzdhetetlen akadály, és nincsenek speciálisan „nehéz problémák”.

(RT1) Nincs „tudat” nevű entitás vagy „tudatosság” nevű tulajdonság, mely speciálisan „nehéz problémát” jelent a fizikalista filozófia, vagy a természettudományos megismerés számára, mert amit értelmesen „tudatnak” vagy „tudatosságnak” nevezhetünk, az maradéktalanul azonosítható az élőlény vagy egyéb tapasztaló rendszer valamilyen kognitív teljesítményével.

(RT2) Nincs a tapasztalatnak olyan „milyensége”, amely sajátos filozófiai problémákat vetne fel. Nem magyarázandó adottság az, hogy „valamilyen” átélni bizonyos tapasztalatokat. A tapasztalat „milyensége” ugyanis, amennyiben értelmes róla beszélni, a tapasztaló által megtapasztalt valóság fizikai tulajdonságaira vonatkozik, vagy pedig tapasztalatának olyan tulajdonsága, mely maradéktalanul elemezhető fiziológiai vagy funkcionális/kognitív terminusokban.

(RT3) Az elmefilozófia feladatai: tisztázni, mi az elme; tisztázni, milyen élőlények vagy egyéb entitások rendelkeznek elmével; mik az elmével rendelkezés feltételei; hogyan működik az elme. Mivel az elmével rendelkezés nem jelent többet, mint az élőlények (vagy gépek) bizonyos kognitív képességekkel, funkcionális tulajdonságokkal való rendelkezését, ezért az elmefilozófia nem egyéb, mint a „pszichológia filozófiája”, vagy annak egy részterülete, ahol a pszichológián mindenekelőtt a viselkedés leírásával, magyarázásával és előrejelzésével foglalkozó tudományt értjük. Sem a pszichológiának, sem az elmefilozófiának nem tárgya a tudat vagy „a tapasztalat milyensége” olyan értelemben, hogy ezek a vizsgálati tárgyak kívül esnének a struktúrák és funkciók (a kognitív funkciók és az ellátásukat lehetővé tevő struktúrák) körén.

És most foglaljuk össze az antirevizionista nézet alapjait, pontosabban meghatározva az antirevizionista meggyőződést és három antirevizionista tézist.

(AM) Még ha meg is magyaráztuk az érzékelés fiziológiai mechanizmusait, a kognitív funkciók működését, illetve azt, hogy hogyan áll elő a viselkedés, akkor sem számoltunk be a legszembevetőbb tényről: arról, hogy *valamilyen* érzékelni, gondolkodni és cselekedni. Noha kívülről csak a viselkedés megfigyelhető, *nyilvánvalóan* van valami más is a külsőleg megfigyelhető viselkedésen túl, amit meg kell magyarázni. Valamilyen *nyilvánvaló* többlet van abban, amikor látunk, mint a fényérzékelés és alakfelismerés puszta mechanizmusa, még akkor is, ha egyelőre semmilyen objektíve megfigyelhető sajátosságot nem ismerünk, ami ezt a többletet megmagyarázná. A viselkedésen túl az *élményt* is meg kell magyarázni, enélkül az emberképünk nem teljes. Ha egyszer elkészül az emberi viselkedést megbízhatóan bejósoló pszichológiai elmélet, abból még mindig hiányozni fog az egyik legfontosabb mozzanat, tudniillik az, hogy nemcsak cselekszünk, hanem *át is éljük* mindazt, ami velünk történik: az, hogy *valamilyen* nekünk lenni. A biológia és a pszichológia jelen formájában nem tud számot adni ezekről a tényekről, amit a „tudat rejtélyének” nevezünk. Ha egyszer olyan színvonalat ér el a robotika és számítástechnika, hogy az embert minden tekintetben elérő vagy meghaladó képességekkel rendelkező robotot tudunk építeni, ebből még akkor sem következik magától értetődően, hogy ez a gép képes lenne a szenvedés vagy gyönyör átélésére.

(AT1) Van az élőlényeknek (vagy bizonyos mentális állapotaiknak) egy tudatosság nevű tulajdonságuk, mely speciális, a többi tudományos problémától eltérő jellegű, ún. „nehéz problémát” jelent a természettudományos megismerés számára, mert nem azonosítható az élőlény biológiai, illetve pszichológiai, kognitív funkcióival, képességeivel.

(AT2) Az emberi tapasztalatnak van egy olyan sajátossága – az, hogy „valamilyen” nekünk lenni –, mely filozófiai problémák forrása. Ez a sajátosság ugyanis nem magyarázható meg egyszerűen a tapasztalatok átélőjének fiziológiai vagy funkcionális képességei és jellemzői alapján. Úgy tűnik ugyanis, hogy az élmény milyensége olyan adottság, ami további magyarázandóként merül fel az organizmus biológiai, funkcionális jellemzői mellett. Elgondolható lenne ugyanis az organizmus biológiai funkcióinak maradéktalan teljesülése, és kognitív képességeinek megléte anélkül, hogy „valamilyen” lenne számára létezni, vagy legalábbis nem közvetlenül nyilvánvaló, hogy miért ne lenne elgondolható.

(AT3) Az elmefilozófia feladata túlmutat a „pszichológia filozófiája” feladatkörén, mert meg kell kísérelnie megoldani a tudatosság filozófiai problémáit. Nem elég pusztán az intelligencia vagy az elme fogalmának meghatározása, vagy az intelligensnek levés vagy elmével rendelkezés feltételeinek feltárása, ezek ugyanis pusztán funkcionális terminusokban is elemezhető tulajdonságok. Az elmefilozófiának tárgya az is, ami túl van a „struktúrán és funkción” (Chalmers 1995/2004), illetve a „tapasztalat versus viselkedés” (Farrell 1950: 177) megkülönböztetéssel élve a tapasztalat oldalán helyezkedik el. Az elmefilozófiának tárgya lehet a tudatosság feltételeinek feltárása; annak meghatározása, hogy milyen élőlények (vagy gépek) élhetnek át tudatos állapotokat; vagy a tudatproblémával összefüggő olyan kérdések, mint az első személyi nézőpont, a személyi azonosság (diakronikus személyi azonosság és a tudat egysége) kérdése. Mindezek a kérdések túlmennek a pszichológia hatókörén, amennyiben a pszichológiát a viselkedés tudományának tekintjük. Ám ha a pszichológia régebbi felfogását részesítjük előnyben (mely szerint a pszichológia a viselkedés és a tudati jelenségek tudománya), akkor az elmefilozófia az antirevizionista szerint is tekinthető a „pszichológia filozófiájának”.

2. Elmefilozófiai viták

Ebben a fejezetben alaposabban megvizsgállok néhány elmefilozófiai álláspontot, illetve a körülöttük kialakult vitákat. Az előző fejezetben azt a sejtést fogalmaztam meg, hogy az elmefilozófiai álláspontok alapvető intuitív meggyőződések elméleti kifejezéseiként is értelmezhetők. Mivel a tudat létezéséről egymással radikálisan ellentétes intuitív meggyőződések léteznek, mind a filozófusok, mind pedig a laikusok körében, ezért felvetődött az a gondolat, hogy az elmefilozófiai műveket az alapján lehetne csoportosítani, hogy melyik intuitív meggyőződést juttatják kifejezésre. Amikor azonban megvizsgáljuk, hogy az antirevizionista („A”), illetve revizionista („R”) filozófia klasszikus művei képesek-e azoknak az intuitív meggyőződéseknek a képviselőjére, amelyek mellett szerzőik kiállnak, azt találjuk, hogy ebben az értelemben nem váltják be a hozzájuk fűzött reményeket: az elmefilozófiai műveknek rendszerint adódik egy ellentétes olvasata, mely nem felel meg a szerző által elfogadott intuíciónak, vagy ellentétes a mű bevett recepciójával. Ennek következtében arra sem alkalmasak, hogy hatással legyenek az olvasó intuitív meggyőződéseire, továbbá az egymással vitatkozó, de eltérő intuitív meggyőződéseket kifejezésre juttató elmefilozófiai művek közti viták a félreértések és egymás mellett elbeszélések rendszeres

mintázatait mutatják. A következőkben erre fogok példákat hozni, majd megkísérlem röviden összefoglalni a legjelentősebb elmefilozófiai pozíciók általános A- és R-olvasatát.

2.1 Antirevizionizmus: Block a fenomenális tudatosságról

Block (1995) a hozzáférési tudatosság és a fenomenális tudatosság közötti különbségtétel kidolgozásával, legalábbis tanulmányának *prima facie* olvasata és deklarált szándéka szerint, arra törekedett, hogy a filozófiai tudatprobléma (a tudat „rejtélye” vagy Chalmers kifejezésével a „nehéz probléma”) szempontjából releváns tudatfogalmat – a fenomenális tudatosság fogalmát – pontosabban meghatározza annak az egyéb tudatfogalmaktól való megkülönböztetése révén. Block szerint a tudatosság „hibrid, vagy még inkább korcs fogalom” (Block 1995: 227), és azok a szerzők, akik a tudatról beszélnek, olykor különböző fogalmakat kezelnek egyetlen fogalomként, ami tévutakra viszi őket.

A fenomenális tudatosság sajátos jellegének megvilágítása érdekében Block a vaklátás jelenségére hivatkozik. Amikor a vaklátó páciens látóterének vak részét érik fényingerek, a páciens arról számol be, hogy nem lát semmit, ám az inger bizonyos jellemzőit mégis ki tudja találni alkalmas kísérleti körülmények között. Úgy tűnik, az ilyen esetekben a látási inger bizonyos értelemben nem *tudatosul*, ám más értelemben mégiscsak *tudomásunk* van róla. Az efféle kísérleti eredmények alapján többen arra következtettek, hogy a (fenomenális értelemben vett) tudat funkciója az, hogy ez teszi hozzáférhetővé az agyban reprezentált információkat a racionális cselekvés és az érzetokről való beszámolás céljaira. Block ezt a fajta következtetést támadja, mely szerinte a tudat kétféle fogalmának összemosásán alapszik.

A kétféle tudatfogalom a *fenomenális* (P-) és a *hozzáférési* (A-) tudatosság fogalma. Egy állapot akkor fenomenálisan tudatos, ha „valamilyen” ebben az állapotban lenni (Block 1995: 228), és akkor hozzáférési tudatos, ha tartalmát a megfelelő információfeldolgozó funkciók a viselkedés racionális irányítását végző (központi végrehajtó) rendszer rendelkezésére bocsátják.

A fenomenális és hozzáférési tudatosság közti kapcsolatot sokféleképpen el lehet képzelni, a Block által elemzett Schachter-féle modellben például a fenomenális tudatosságnak külön alrendszere van az elmén belül, és ez közvetít az érzeteket feldolgozó rendszerek és a viselkedést szabályzó végrehajtó rendszer között, így a fenomenális tudatosság részt vesz a hozzáférési

tudatosságot megalapozó információfeldolgozó funkció implementációjában, de attól legalábbis fogalmilag megkülönböztethető.

Block hangsúlyozza, hogy a fenomenális tudatosság az, amely „tudományos rejtélynek” tűnik, és elismeri, hogy ő sem tudja nem-körkörösén definiálni; „a legtöbb, amit tehetünk, ha rámutatunk a jelenségre” (Block 1995: 230). Hogyan lehet rámutatni a fenomenális tudatosságra anélkül, hogy másra (például a hozzáférési tudatosságra) is rámutatnánk? Voltaképpen ez annak a problémának a lényege, amellyel Block birkózik, és amely szerinte sok szerzőt tévutakra vezetett. Block szerint az egyik módszer szinonimák felsorolása lehet: a fenomenális tudatosság tapasztalat; egy állapot akkor fenomenálisan tudatos, ha tapasztalati tulajdonságai vannak; egy állapotunk tapasztalati tulajdonságainak összessége pedig megadja, „milyen” abban az állapotban lenni. A szinonimák után példák következnek: fenomenálisan tudatos állapotokba kerülünk, amikor látunk, hallunk, ízlelünk vagy fájdalmat érzünk. Block szerint azonban a legmerészebb állítása az, hogy „a fenomenális tudatos tulajdonságok különállnak minden kognitív, intencionális vagy funkcionális tulajdonságtól” (Block 1995: 230). A fenomenális tudatosság egyik (Block szerint talán csak kontingens) jellemzője a magyarázati rés, vagyis annak állítólagos érthetlensége vagy rejtélyessége, hogy hogyan eredhet egy olyan jelenség, mint a tudat – a színes, szagos, fájdalmas stb. megélt élmény – olyasmiből, mint amilyen az idegszövetek működése (Block 1995: 231). Hasonló magyarázati részek nem merülnek fel az agy kognitív funkcióinak tudományos magyarázataival kapcsolatban, ezért a fenomenális tudatosságot akár az alapján is meg lehet határozni (legalábbis rá lehet mutatni), hogy ez az a valami, amire a magyarázati rés vonatkozik (Block 1995: 231).

Ez a magyarázati réssel kapcsolatos megjegyzés egyébként kétségbe vonható. Vajon csakugyan igaz-e, hogy nem rejtélyes vagy nehezen érthető az, hogy az agyműködés – a neurotranszmitterek, neuronpotenciálok stb. tevékenysége – miként lát el olyan összetett kognitív funkciókat, mint amilyen az emlékezés vagy az érvelés? Abban az értelemben biztosan nem könnyen érthető és rejtélyek nélküli problémával állunk szemben, hogy az átlagember (vagy akár a filozófus, sőt, egy mai agykutató) számára *részleteiben* könnyen megmagyarázható lenne ez a funkcionalitás. Inkább arról van szó, hogy nem vet fel különleges fogalmi nehézséget az, hogy egyáltalán elképzelhessünk egy ilyen magyarázatot, hiszen Chalmers kifejezésével, bár a részletek nagyon bonyolultak, ám végül is csak kellően összetett „struktúrák és funkciók” kérdése egy sikeres magyarázat. Ha a rejtélyesség és érthetlenség feloldásához elegendő, hogy ilyen absztrakt módon vázoljuk fel egy elképzelhető magyarázó elmélet jellemzőit (például azt gondolva, hogy „mivel a kogníció funkcionális, ezért elég, ha kellően bonyolult struktúrákat és funkciókat találunk, amelyek

megmagyarázzák”), akkor valóban igaz, hogy nincs filozófiai érdekes magyarázati rés az idegsejtek funkcionalitása és az elme összetett kognitív teljesítményei között. Ezzel szemben viszont fennáll a magyarázati rés, ha nem valami kognitív (tehát funkcionális) jelenséget magyarázunk valamilyen neurofiziológiai (tehát funkcionális) jelenségekkel, hanem ettől eltérő módon „különműeket” kívánunk a magyarázatunkban összekapcsolni. De pontosan mi szerepel ennek a relációnak a két oldalán a fenomenális tudatossággal kapcsolatos magyarázati rés vonatkozásában? Ha elfogadjuk Blocknak a funkcionalizmus elleni érvelését (Block 1978), akkor el kell ismernünk, hogy a fenomenális tudatosság *explanansa* nem valamilyen funkcionális tulajdonság lesz a tudat befejezett elméletében. A reduktív materialisták szerint például a fenomenális tulajdonságok valójában fizikai tulajdonságok, így nem kellene fenn magyarázati résznek fennállnia a fenomenális tulajdonság és bizonyos fizikai tulajdonságok között, hiszen a kettő azonos.²

A tudat másik, nem-fenomenális fogalma a hozzáférési tudatosság fogalma. Egy állapot akkor hozzáférési tudatos, ha inferenciálisan termékeny, azaz premisszaként használható a gondolkodásban, és rendelkezésre áll a viselkedés racionális vezérlése céljára. (Block 1995: 231). A tudatosságról szóló mindennapi diskurzusban a hozzáférési tudatosság fogalma fontos szerepet játszik, ami hozzájárul az általában vett „tudat” fogalmának meghatározatlanságához, illetve jelentése bizonytalanságához.

Block szerint a fenomenális (vagy P-) tudatosság és a hozzáférési (vagy A-) tudatosság között három fő különbség van:

(1) A P-tudatos tartalom fenomenális, az A-tudatos tartalom reprezentacionális, illetve egy állapot annyiban és amiatt P-tudatos, amennyiben és mivel fenomenális, és annyiban A-tudatos, amennyiben tartalma reprezentacionális.

Egy állapot P-tudatos tartalma nem más, mint az állapot tapasztalati (*experiential*) tulajdonságainak összessége, vagyis az, hogy *milyen* abban az állapotban lenni.

Egy állapot P-tudatos tartalma az adott állapot „tapasztalati értéke”, A-tudatos tartalma pedig „reprezentacionális értéke”. Egy tapasztalat tartalma tehát egyszerre lehet A- és P-tudatos is, a reprezentacionális tulajdonságai, illetve fenomenális minőségének megfelelően (Block 1995: 232).

² Vannak ugyanakkor filozófusok, akik szerint a magyarázati rés akkor is fennállhat, ha igaz a fenomenális/fizikai azonosság, lásd: Levine 1983.

(2) Az A-tudatosság funkcionális fogalom, szemben a P-tudatossággal, mely nem az. Az A-tudatos tartalom annyiban A-tudatos, amennyiben bizonyos módon használható fel valamely rendszeren belül (vagyis rendszer-relatív). Ha létezik valamiféle „P-tudatosság modul” az elmében (amelyben P-tudatossá válnak az e modul által feldolgozott tartalmak), akkor a tartalom annak köszönhetően válik tudatossá, hogy mi történik a „P-tudatosság modul”-ban, ám hasonló módon (különálló „A-tudatosság modul” meghatározásával) nem elemezhető az A-tudatosság. Egy tartalom ugyanis attól A-tudatos, hogy meghatározott módon áll rendelkezésére például az elme központi végrehajtó rendszereinek, nem pedig önmagában.

(3) Míg létezik olyan, hogy P-tudatos típusú állapot, ez nem mondható el az A-tudatosságról. A fájdalom érzett *minősége* P-tudatos típus: minden fájdalomnak rendelkeznie kell ezzel a sajátos minőséggel. Ezzel szemben egy bizonyos gondolat, mely A-tudatos egy bizonyos pillanatban, egy másik időpontban hozzáférhetlenné válhat (éppen nem jár az eszünkben, így nem befolyásolja döntéseinket, viselkedésünket). Egy információs szempontból egy bizonyos pillanatban termékeny állapot nem biztos, hogy későbbi időpontokban hasonlóan termékeny.

Ennek megfelelően Block szerint a P-tudatos állapotok paradigmaticus esetei az érzetek, míg az A-tudatosakéi a propozicionális attitűdök, noha Block szerint a gondolatok olykor P-tudatosak, az érzeteknek pedig lehet reprezentacionális tartalmuk. Így például sok filozófus értelmezte a fájdalmat reprezentacionális állapotként.

A kétféle tudatfogalom közötti különbség megvilágítására Block gondolatkísérleteket használ, ilyen például a szuper-vaklító esete. A szuper-vaklító olyan személy lenne, aki a valóságos vaklatókhoz hasonlóan látótérkiesésben szenved, ám a látótér „vak” részén is képes bizonyos mértékű vizuális észlelésre, ám a vaklatók esetétől eltérően a szupervaklító vaklátása ugyanolyan részletes, mint a normál látóké. Amikor a szupervaklátót arról kérdezik, hogy milyen tárgy van a látótere „vak” részében, akkor erre válaszolni tud úgy, hogy egyszerűen az eszébe jut a kérdéses tárgy, ám ezt nem kíséri látási élmény. A szupervaklító ugyanakkor arról számol be, hogy látótere ép részében „valamilyen” látni a tárgyakat, míg a vak részében semmilyen: ha a vak látóterébe kerül egy tárgy, akkor egyszerűen tudja ugyan, mi az, de nem „valamilyen” ilyen módon látni a tárgyakat.

Block a szupervaklátót úgy értelmezi, mint egy „korlátozott részleges zombit”, aki a „vak” látótér vonatkozásában a látott tárgyakról A-tudatos, de nem P-tudatos tudomást képes szerezni (Block 1995: 233).

Mivel – mint azt Block is elismeri – a szupervaklítás a valóságban jelenlegi tudományos ismereteink szerint nem fordul elő, ezért csak úgy foghatjuk fel, mint gondolat kísérletet, mely segít megérteni az A- és P-tudatosság közötti különbséget. Mint elhangzott, az A-tudatosság Block számára funkcionális fogalom, a P-tudatosság azonban nem. Ennek megfelelően elképzelhetőnek tűnnek fenomenális zombik, akik a normális emberek funkcionális duplikátumai, ám nem fenomenális duplikátumai. A szupervaklító Block szerint részleges fenomenális zombi (zombi a látótere egy részének tekintetében).

Ez a jellemzés azonban nem tűnik helyesnek, mert Block szerint a szupervaklító képes beszámolni arról, hogy látótere normális, illetve „vak” részében látni ugyanazt a tárgyat más minőségű élmény (ti. az egyik esetben „valamilyen”, a másik esetben „semmilyen”; az egyik esetben van látási élmény, a másik esetben „pusztán eszébe jut” az, hogy milyen tárgy van előtte), vagyis nem pusztán fenomenális, hanem kognitív-funkcionális eltérés is fennáll a szupervaklító és a normál látó között.

Látni valamit nyilvánvalóan másmilyen élmény, mint az, ha valami pusztán az eszünkbe jut, vagy az, ha valamit egyszerűen tudunk. Gondoljuk el a következő helyzetet. Tegyük fel, hogy egy „totális szupervaklító” (akinek nincs normális látótere, csak vaklító látótere – elképzelhetjük az egyszerűség kedvéért azt is, hogy egy egyik szemére vaklító, a másik szemére normálisan látó személyről van szó, akinek most valamilyen okból el van takarva a normálisan látó szeme) körülnéz egy eddig számára ismeretlen könyvtárszobában vagy képgalériában.

Az élmény *ex hypothesi* különbözni fog annak élményétől, amikor valakinek (a szó megszokott mindennapi értelmében) pusztán eszébe jutnak dolgok. A totális szupervaklítóknak ugyanis nem „egyszerűen az eszébe jutnak” a polcokon sorakozó könyvek címei vagy a festményeken ábrázolt tájak vagy arcok jellemzői, hanem bizonyos rend szerint jutnak az eszébe.

Tehát a pusztán tudás és az „eszünkbe jutás” hétköznapi életből ismerős példái félrevezetőek, mert nem jó modelljei annak, hogy mi lenne egy totális szupervaklító tapasztalata. Ez viszont azt jelenti, hogy nagyon nehéz elképzelnünk egy totális szupervaklító tapasztalatait.

Míg azt meglehetősen könnyű megérteni, hogy a gyakorlatban megfigyelt vaklítás miért nagyon más, mint a normális látás, hiszen a tényleges vaklítás sokkal kevésbé részletgazdag, mint a normális látás, tehát jelentős kognitív hátrányban van azzal szemben, ami természetesen a páciensek

verbális beszámolóiban és behaviorisztikus teszteken elért teljesítményeikben is tükröződik, a totális szupervaklítás esetében nem sok támpontunk marad azon kívül, hogy azt mondjuk: szemben a normális látóval, a totális szupervaklító számára „semmilyen” látni a tárgyakat (azt ugyanis nem mondhatjuk, hogy olyan, mint amikor mi egyszerűen tudunk valamit, vagy mint amikor nekünk eszünkbe jutnak dolgok). A totális szupervaklító és a normális látó élménye között a különbség persze Block szerint a P-tudatosság. Mivel tudjuk, milyen látni, és tudjuk, ehhez képest milyen az, amikor nem látunk, csak az eszünkbe jutnak dolgok, ezért ha a P-tudatosság az e két eset közötti különbséget jelentené, akkor értenénk, mi a P-tudatosság. Csakhogy a totális szupervaklító élményéről nincs világos elképzelésünk, mert a mindennapi életben nincs ehhez fogható.

Segíthet valamit a helyzetünkön, ha Block példájához hívebben nem totális, hanem részleges szupervaklítást próbálunk elképzelni?

A részleges szupervaklító Block szerint „nagyon korlátozott részleges zombi”, azonban nem nyilvánvaló, hogy ez miért lenne így. Egy ember zombi-ikre az illető funkcionális (vagy fizikai, viselkedési – az érvelés kontextusától függően) duplikátuma, mely ugyanakkor nem fenomenális duplikátuma. Egy részleges zombi csak részleges fenomenális duplikátum, miközben funkcionálisan továbbra is tökéletesen ekvivalens egy normális emberrel. Mint ahogyan a teljes zombi viselkedése sem tér el a nem-zombiétól, ennek a részleges zombira is igaznak kell lennie. A részleges szupervaklító tehát vagy abban a helyzetben van, hogy mint Block állítja, szembe tudja állítani azt a kétféle élményt, amit a „látás” jelent, amikor a látótere ép, illetve vaklító részével történik, vagy pedig részleges zombi – egy részleges zombi ugyanis semmilyen, a normálistól eltérő tapasztalatról vagy élményről nem számol be, hiszen funkcionálisan duplikálja a normális tapasztalat átélőjét.

Más szóval, ha a szupervaklító valóban részleges zombi, akkor nem számolhat be sajátos (a normálistól eltérő) szupervaklítói tapasztalatairól. Mindez pedig azért problematikus, mert rövid úton elvezet ahhoz a szkeptikus kérdéshez, hogy honnan tudhatja magáról bárki is, hogy nem szupervaklító (vagy zombi). A valóságban megfigyelt vaklátók persze nem szupervaklátók, és vaklításuk sokkal korlátozottabb, mint a normális látás vagy a szupervaklítás. Ezért nagyon könnyű félreérteni, ha valaki a fenomenális tudatosság fogalmát a vaklátáson keresztül próbálja elmagyarázni.

De erre nem megoldás az áttérés a vaklítás példájáról a szupervaklítására, mert mihelyt a vaklítás „szuper”, vagyis semmiféle kognitív hátrányt nem jelent a normális látással szemben, vagyis a szupervaklító valóban részleges fenomenális zombi (ám nem tér el egyéb módon a normál

látótól), egyúttal kimutathatatlaná is válik részleges zombi volta. Sem objektív képességeiben, sem pedig szubjektív beszámolóiban nem lesz többé megfigyelhető semmiféle abnormalitás, vagyis ő maga sem tudhat róla, hogy részleges zombi vagy szupervaklító.

Csak hogy az antirevizionista vagy tudat-realista intuíció szerint elszórt lenne azt állítani, hogy a P- és A-tudatosság (vagy a tudatosság valamilyen kognitív és kvalitatív aspektusa) között nem tehető különbség pusztán csak azért, mert a funkcionális vonatkozások végső soron kimerítik annak teljes tartalmát, amit a tudatosságról *mondani lehet*. Más szóval az antirevizionistának még akkor is ragaszkodnia kell az A- és P-tudatosság közötti ontológiai különbséghez, ha elfogadja, hogy episztemikus különbség nem tehető. Természetesen az olyan erőfeszítések, mint a Block (vagy a tudatfilozófiai bevezető tankönyvek és cikkek) által felsoroltakhoz hasonló példák felmutatása, az antirevizionista intuíció megnyilvánulásai, azonban önmagukban ezek a gesztusok még nem filozófiai érvek. Egy biztos: ha az ontológiai különbségtétel fenntartható, a vaklítás példája nem segít megérteni, hogy pontosan hogyan.

A vaklítás a P-tudatosság nélküli A-tudatosságra volt példa. Block megvizsgálja a fordított esetet is. Az egyik ilyen eset szerinte az, amikor valaki csak egy bizonyos időpontban lesz figyelmes egy olyan háttérzajra, ami már régóta jelen volt (Block 1995: 234).

Ugyan a példa meggyőzően illusztrálja annak egyik lehetséges rekonstrukcióját, hogy mit ért a köznyelv azon, hogy valamiről „tudomásunk van”, de „nem tudatosan” (*consciously aware of*), itt sem egyértelmű, hogy közelebb kerülünk az A/P distinkció megértéséhez. Emlékezzünk ugyanis arra, hogy Block a P-tudatossággal kapcsolta össze a tapasztalat „milyen”-ségét („milyen” egy bizonyos helyzetet átélni), márpedig ha valaki annyira elmerül egy beszélgetésben (mint Block példájában), hogy meg sem hallja az utcán csörömpölő légalapácsot, akkor plauzibilis értelmezése ennek a helyzetnek, hogy az illető számára „másmilyen” tapasztalat adódik, mint annak számára, akit ez a zaj zavar, például mert nem képes igazán ráhangolódni egy érdektelen beszélgetésre. Egészen biztosan a tapasztalat „milyen”-ségéhez tartozik az, ha egy beszélgetés annyira lebilincselő, hogy egy légalapács zaját is kiszorítja tudatunkból.

A kétféle értelemben vett tudatosságnak kétféle filozófiai zombi felel meg: a fenomenális zombi A-tudatos, ám nem P-tudatos. A fenomenális zombi viselkedése természetesen nem tér el a normális emberétől, a vudu-zombik vagy a zombifilmekben szereplő zombik azonban funkcionálisan károsodottak, míg fenomenálisan akár épek is lehetnek (Block 1995: 234). A különbség világos, ám az a nehézség megmarad, hogy a szupervaklító nem funkcionális mása a normális látással rendelkező embernek, így nem tekinthető részleges P-zombinak. Ennek

következtében Block igazából nem tudott előállni a P-tudatosság jelentőségét szemléltető olyan példával, amely nem kaphatna az eredeti szándékaival ellentétes kognitív olvasatot (mint a szupervaklítás példájában), vagy pedig közvetlenül a tudattal szemben eliminatív materialista olvasatot (mint a légkalapács példájában).

Block szerint a P-tudatosságnak valamilyen kognitív képességgel való azonosítása, mint amilyen az önmegfigyelés, saját mentális állapotaink belső „pásztázása” (Block 1995: 235, vö. Armstrong 1970: 79), végső soron a P-tudatossággal kapcsolatos eliminativizmushoz vezet. Hogy hogyan, arra Rey már vizsgált gondolatmenete a példa. Ha viszont a P-tudatosságot megkülönböztetjük a „belső pásztázási tudatosságtól”, akkor fenntartható az az állítás, hogy egy számítógépnek lehet „belső pásztázási tudatossága”, miközben P-tudatossággal nem rendelkezik. (Block már korábban érvelt amellett, hogy miért nem lehet a fenomenális tulajdonsággal szemben realistának lenni olyan módon, hogy a fenomenális tudatosságot valamilyen funkcionális szerveződéssel azonosítjuk: Block 1978).

Dennett többször is hangot adott annak a nézetének, hogy a tudat kulturális termék (Block 1995: 238), ám Block szerint nyilvánvaló, hogy a P-tudatosság nem az, így Dennett tudatelmélete nem vonatkozhat a P-tudatosságra. Amikor Block azt állítja, hogy a P-tudatosság nyilvánvalóan nem lehet kulturális konstrukció, akkor ezt azzal szemlélteti, hogy az ellentétes nézet azzal az abszurd következménnyel járna, hogy fel kellene tételeznünk, hogy volt olyan korszak, amikor az emberek ettek, ittak és szeretkeztek, ám anélkül, hogy ezt átélni *valamilyen* lett volna számukra. Ezért Dennett elmélete csak az A-tudatosságra vonatkozhat, és valóban, amikor Dennett azt fejtegeti, hogy az a tartalom tudatosul a különféle agybeli események közül, amely kellően nagy jelentőségre tett szert ahhoz, hogy tipikus viselkedési és kognitív hatásokat váltson ki, Block szerint úgy tűnik, hogy itt valóban az A-tudatosság elemzéséről van szó.

Ezzel kapcsolatban azonban megjegyezendő, hogy semmi akadályja nincs annak, hogy Dennettnek a tudattal kapcsolatos álláspontját úgy fogjuk fel, hogy ugyan a mentális tartalmak globális hozzáférhetővé válása vagy a viselkedés irányításában meghatározó szerep megkaparintása terminusaiban felfogott tudatossá válás csakugyan az A-tudatosságról szól, eközben azonban a centrális szerepért „versengő” mentális tartalmak maguk P-tudatosak. Az pedig, hogy „milyen” nekünk lenni, ebben a Dennett-interpretációban végső soron attól függ, hogy milyen a globális feldolgozásra hozzáférhető P-tudatos eseményekből összeálló (A-tudatos) állapot.

Mivel Block sem zárja ki azt a lehetőséget, hogy például egy lábfájás (vagy egy háttérzaj) P-tudatos legyen számunkra, ám figyelmünk perifériájára szorul (már nem A-tudatos), nem

egyértelmű, hogy igaza van Blocknak abban, hogy Dennett elmélete nem szól a P-tudatosságról. Tény, hogy Dennett elsősorban azzal foglalkozik, hogy mely ingerek, agybeli információszállító folyamatok válnak A-tudatossá. Ám ha például reduktív materialistaként feltételezzük, hogy ezek a folyamatok önmagukban rendelkeznek fenomenális tulajdonságokkal (hiszen azok fizikai tulajdonságokkal azonosak), csupán nem mindig válnak A-tudatossá, akkor Dennett elmélete alapján magyarázható, hogy milyen mechanizmusok határozzák meg, „milyen” nekünk lenni egy adott szituációban, hiszen ez attól függ, mely (P-tudatos) események A-tudatosak.

Block tulajdonképpeni célja a P- és A-tudatosság közti különbségtétellel az, hogy rávilágítson egy olyan tévkövetkeztetésre, melyet szerinte több szerző is elkövet. Így például Searle bizonyos epilepsziás rohamokon áteső páciensekkel kapcsolatban azt állítja, hogy elveszítik tudatosságukat, és ez olyan módon hat ki a viselkedésükre – azt kevésbé rugalmassá és kreatívvá téve – ami azt illusztrálja, hogy a tudat funkciója többek között az, hogy az emberi viselkedést kreatívabbá, rugalmasabbá teszi (Block 1995: 239). Block azonban rámutat, hogy a kreativitás és rugalmasság hiánya a gondolkodási folyamatok hiányosságait jelenti, tehát kognitív funkciók sérülését, ami azt sejteti, hogy ebben az esetben a tudatosságot, ami hiányzik, A-tudatosság értelmében kell venni, addig Searle arról beszél, hogy a P-tudatosság hiányzik, de anélkül, hogy ezt különösebb érveléssel alátámasztaná. Ez a fajta tévkövetkeztetés az A-tudatosság funkcióit ruházza át a P-tudatosságra, és oka az, hogy szokásos „keverék” tudatfogalmunkból indul ki, mely nem ismeri az A/P különbségtételt (Block 1995: 240).

A gyakorlatilag megfigyelhető (normális vagy patológikus) esetekben az A- és P-tudatosság rendszerint együtt jár vagy mindkettő hiányzik (Block 1995: 242). Ez a „keverék” tudatfogalom kialakulásának oka. Az olyan esetekben, mint a vaklátás, a pácienseknél az A- tudatosság és P-tudatosság egyaránt sérült, és ez tükröződik abban, hogyan birkóznak meg a kísérletek során különböző feladatokkal. Nehézségeiket könnyen meg lehet magyarázni az A-tudatosság hiányosságaival, ugyanakkor ha a páciensek állapotát úgy fejezzük ki, hogy a „tudatuk” szenved rendellenességben, majd pedig feltételezzük, hogy a „tudat” a „milyenséggel” áll szoros összefüggésben – vagyis tudaton P-tudatosságot értünk – akkor könnyen el lehet jutni arra a téves megállapításra, hogy a P-tudatosság szenved csorbát ezeknél a betegeknél. Ám az áttérés az A-tudatosság funkcióiról a P-tudatosságéira, csupán azért, mert a gyakorlatban a P- és A-tudatosság egyszerre jelentkezik vagy hiányzik, jogtalan. Hogy mi a P- és A-tudatosság viszonya, az nem eldöntött kérdés. Megfogalmazható az a hipotézis, természetesen, hogy az A-tudatosság függő viszonyban van a P-tudatossággal, és a P-tudatosság hiánya okozza az A-tudatosságét, mint ahogy a fordítottja és más további hipotézisek is felvázolhatók. (Block 1995: 242)

Sőt Block szerint az a helyzet, hogy azt is csupán feltételezzük, hogy a vaklátó esetében hiányzik a P-tudatosság. Block csupán úgy döntött, hogy elhiszi a vaklátók beszámolóit arról, hogy nincsenek a „vak” látóteret érő látási ingereknek P-tudatában. Block szerint a vaklátásban mind a P-, mind az A-tudatossággal gondok vannak, „szupervaklátót” pedig a valóságban még nem figyeltek meg.

Ha megkülönböztettük a P- és A-tudatosságot, akkor ezeknek különböző funkciókat is tulajdoníthatunk, így a P- és A-tudatosságért felelős különböző rendszerek viszonyainak eltérő felfogásai az elme funkcionális-kognitív felépítésének különféle modelljeihez vezetnek attól függően, hogy a P-tudatosságnak a cselekvés központi vezérlésében jut-e valamilyen funkció, vagy a P-tudatosság efféle funkció nélkül, csupán kísérőjelensége információfeldolgozási folyamatoknak, vagy pedig az A- és P-tudatosságért felelős funkcionális egységek az elmében ugyanazok vagy szorosan összetartoznak (Block 1995: 243).

Block két példát hoz fel, amelyek szerinte bizonyítékok lehetnek arra, hogy a gyakorlatban előfordulhat a P-tudatosság A-tudatosság nélkül. Az egyik eset a hipnotikus fájdalomcsillapítás, a másik pedig az ún. aerodontalgia jelensége.

A hipnotikus fájdalomcsillapítást Block úgy értelmezi, hogy ekkor fennáll egy P-tudatos állapot (például fájdalom egy végtagban), ám ez nem hozzáférhető a személy számára, akin a hipnózist végezték, azaz nem A-tudatosul számára. Mindamellettt közvetett bizonyítékok utalnak arra, hogy a fájdalmas állapot (mint P-tudatos állapot) igenis fennállt. Előfordulhat tehát különleges körülmények között, hogy egy személy ne szenvedjen fájdalmat, miközben fennáll egy olyan állapota, amely normális körülmények között fájdalmas lenne számára. Block szerint nincs különösebb okunk kételkedni abban, hogy az ilyen fájdalomállapot nem P-tudatos (Block 1995: 244).

A másik példa, az aerodontalgia lényege az, hogy altatásban (de helyi érzéstelenítés nélkül) végzett fogászati műtétek fájdalmas állapotok emléknymait hagyják maguk után, és ezek a fájdalmak bizonyos körülmények között (légnyomásváltozás hatására) később „reaktiválódhatnak”. Ha a fogászati kezelés helyi érzéstelenítéssel történt, akkor a reaktiváció nem következik be. Block a jelenséget úgy értelmezi, hogy a fogászati kezelés helyben P-tudatosuló fájdalommal jár, ha a kezelést nem helyi érzéstelenítés mellett végzik (Block 1995: 244-245).

Amennyiben elfogadjuk a P-tudatosság létezését, az A-tudatosságtól való megkülönböztettségét, továbbá plauzibilisnek tartunk valamilyen fenomenális-fizikai vagy fenomenális-neurofiziológiai azonosságelméletet, akkor valóban kézenfekvő elképzelés, hogy P-

tudatos fájdalmak (tehát megfelelő fizikai, neurofiziológiai típusba tartozó, és ennek megfelelően a fájdalom fenomenális karakterével bíró események) előfordulhatnak lokálisan, még akkor is, ha nem válnak A-tudatossá, vagyis a fájdalmat elszenvedő lény elméje számára központilag feldolgozottá. Egy, az A/P különbségtétellel vagy a P-tudatosság létezésével, vagy a P-tudatosság és lokális fizikai-neurofiziológiai események azonosíthatóságával szemben szkeptikus filozófus azonban alternatív magyarázatokkal is előállhatna, és nem világos, hogy Block milyen bizonyítékok alapján tudná saját értelmezését előnyben részesíteni.

Az egyik ilyen alternatív magyarázatnak Block is a közelébe kerül (Block 1995: 244), amikor felveti egy lokális A-tudatosság lehetőségét.

Fel lehetne vetni a következőt: még ha létezik is lokális P-tudatosság, nem emiatt állnak elő azok a jelenségek, amelyek révén Block arra következtet, hogy az A-tudatosság távollétében mégiscsak jelen van a P-tudatosság, és ez magyarázza a hipnotikus fájdalomcsillapítás közben megfigyelt sajátos jelenségeket vagy az aerodontalgia jellegzetességeit. Felvethető ugyanis az a lehetőség is, hogy minden lokális P-tudatosságot lokális, töredékszerű A-tudatosság kísér, amely azonban nem normál módon kapcsolódik a szervezet globális A-tudatos állapotához, akkora mértékben azonban tud rá hatni, hogy a leírt jelenségek előálljanak. Mivel kétséges, hogy egyáltalán bármikor előfordul P-tudatosság A-tudatosság nélkül (vagy fordítva), továbbá kétséges, hogy ha előfordulna P-tudatosság A-tudatosság nélkül, akkor arról tudomásunk lehetne (így bizonyítékul hivatkozhatnánk rá például egy elmefilozófiai tanulmányban), ezért komolyan kell venni azt a lehetőséget, hogy a Block által leírt esetekben a (globális A-tudatos állapotokhoz csökkent mértékben kapcsolódó) lokális A-tudatosság felelős azokért a jelenségekért, amelyekért Block szerint az A-tudatosság nélküli P-tudatosság. Ha viszont ez így van, akkor továbbra sem tudunk felsorolni olyan eseteket, amikor a P-tudatosság előfordul A-tudatosság nélkül, ekkor pedig felmerül a kérdés, hogy nem redundáns-e a P-tudatosság fogalma, minthogy a P-tudatosságra való hivatkozás soha, semmilyen körülmények között nem bír magyarázóerővel.

2.1.1 Block kritikái

Block tanulmányát kommentálva Armstrong (1995: 247-248) elfogadja az A- és P-tudatosság közti különbségtételt, de a kétféle tudatosságot némileg másképpen jellemzi: „a P-tudatosság középpontjában a percepció áll, és a percepció szerintem lényegileg propozicionális” (Armstrong 1995: 247). A hangsúly azon van, hogy a percepció valamilyen ismeretszerző

mechanizmus. Armstrong szerint a percepció során információkhoz jutunk a környezetünk (és saját testünk) állapotáról. Armstrong tehát nem ért egyet Blockkal, aki „tagadja, hogy a P-tudatosság egy kognitív, intencionális vagy funkcionális tulajdonság” (Armstrong 1995: 247). Természetesen az A-tudatosság is kognitív, funkcionális tulajdonság, csak éppen az A- és P-tudatosság különböző funkciókat lát el, így világosan megkülönböztethetők (Armstrong 1995: 248), ugyanis Armstrong felfogásában a P-tudatosság révén kapunk információkat a világról, míg az A-tudatosság magáról az elméről. (Armstrong korábban a tudatot az elme „önpásztázó” tevékenységeként fogta fel, mely révén tudomást szerzünk saját mentális állapotainkról: Armstrong 1970: 78-79). Ezért Armstrong számára nem jelent nehézséget a vaklátás megértése, csak éppen annak legkézenfekvőbb modellje az ő számára az, hogy vaklátásban fennáll a P-tudatosság, viszont korlátozott az A-tudatosság. Ennek megfelelően, állapítja meg Armstrong, félrevezető az A-tudatosság neve („hozzáférési tudatosság”), hiszen a világhoz való hozzáférés közvetlenül a P-tudatosság (Armstrong számára: „perceptuális tudatosság”) révén történik; az A-tudatosságot lehetne akár egyszerűen csak introspekciónak is hívni.

Az egyetértés Armstrong és Block között tehát mindössze arra korlátozódik, hogy meg lehet különböztetni a tudatosság többféle fogalmát, ám arra nézve, hogy mik ezek, és milyen viszonyban állnak, a két filozófus jelentősen különböző álláspontra kerül, ennek pedig az a nyilvánvaló oka, hogy Armstrong a P-tudatosság kognitív értelmezéséhez, információszerző mechanizmusként való felfogásához ragaszkodik, és teljesen elutasítja annak nem-kognitív, nem-funkcionális felfogását, amely Block distinkciója szempontjából kulcsfontosságú. Ez pedig éppen a revizionista és antirevizionista elmefilozófusok álláspontjának különbsége úgy, ahogyan azt Chalmers megfogalmazta: az egyik tábor szerint a tudatosság valami olyasmi, ami más, mint puszta struktúra és funkció, míg a másik tábor szerint nem, hiszen ha a struktúrákat és funkciókat leírtuk és megmagyaráztuk, azzal mindent leírtunk és megmagyaráztunk, amit az elme működésével kapcsolatban le lehet írni és meg lehet magyarázni.

Érdeemes figyelni arra, hogy Block a vaklátók beszámolóit és képességeit úgy értelmezi, hogy A-tudatossággal rendelkeznek (mivel birtokába jutnak bizonyos tudásnak, amiről csak látás révén tudhatnak), de P nélkül (mivel élményük minősége olyan, mintha nem látnának), míg Armstrong úgy gondolja, hogy ha a vaklátók vakságról számolnak be (arról, hogy élményük minősége olyan, mintha nem látnának), ez annak bizonyítéka, hogy károsodott az A-tudatosságuk (ezért hibásan arról számolnak be, hogy nem látnak), miközben P-tudatos hozzáférésük van a tárgyakhoz (hiszen képesek azokkal kapcsolatos feladatokat megoldani). Ez egyszerűen szemlélteti a tudatosság minden „nem kognitív, nem funkcionális” felfogásának azt az alapvető problémáját,

hogy ha képesek vagyunk beszámolni valamiről, akkor ennek előzménye (vagy magának ennek a képességnek a fennállása) valamilyen kognitív, funkcionális állapot. Így ha valaki arról számol be, hogy „nem lát”, miközben egyébként képes a nem látott tárgyakkal kapcsolatos feladatokat megoldani, míg egy másik személy arról számol be, hogy „lát”, miközben ugyanezeket a feladatokat el tudja végezni, akkor a két személy között fenn kell állnia valamilyen kognitív-funkcionális különbségnek. Ha az „A-tudatosság” kognitív, funkcionális tulajdonság, a „P” pedig nem, akkor a két személy között az A-tudatosságban kell fennállnia a különbségnek.

Ha valaki a páciensek beszámolóit vagy képességeit alapján a P-tudatosság károsodására vagy hiányára akar következtetni, miközben fenntartja, hogy a P-tudatosság *nem* kognitív vagy funkcionális tulajdonság, annak feltételeznie kell, hogy a normális P-tudatosság valamilyen módon feltétele a normális módon működő A-tudatosságnak, de maga nem kognitív-funkcionális tulajdonság. Erre a lehetőségre utal kommentárjában Atkinson és Davis (1995: 248-249), akik elfogadják a P-tudatosság nem kognitív vagy funkcionális jellegét: a P-tudatosság a relacionális A-tudatosság kategoriális és intrinzikus alapja lehet. Miközben ez a lehetőség teljesen megfelel például egy, a tudattal szemben határozottan realista (tehát antirevizionista) materialista azonosságelmélet képviselője számára, ezzel kapcsolatban is adódnak nehézségek: mindenekelőtt az, hogy hogyan szerzünk tudomást olyan tulajdonságainkról, amelyek bizonyos kognitív állapotainkkal nem funkcionális kapcsolatban állnak (tehát nem elménk funkcionális blokkábrájának két olyan „dobozaként”, melyet a lehetséges átmenet nyilai kötnek össze), hanem e kognitív állapotok kategoriális bázisaként szolgálnak.

Kommentárjában Atkinson és Davies felveti, hogy a P- és A-tudatosság úgy viszonylik egymáshoz, hogy a P-tudatosságnak szerepe van az A-tudatosság kauzális magyarázatában (Atkinson – Davies 1995: 248). Az elképzelés az, hogy az A-tudatosság fogalma diszpozicionális tulajdonságot jelöl. Márpedig ahol jelen van egy diszpozicionális tulajdonság, ott „természetes dolog egy intrinzikusabb tulajdonságot keresni”, amely annak alapjául szolgál. Atkinson és Davies szerint például természetes népi pszichológiai meggyőződésünk szerint egy fájdalomérzetről azért tudunk beszámolni, mert a fájdalom egy fenomenálisan tudatos állapot. A fájdalomviselkedésre irányuló diszpozíció kategoriális alapja a fájdalom mint fenomenálisan tudatos állapot. A fájdalomról való tudomásszerzés és beszámolás (tehát a fájdalomra vonatkozó A-tudatosság) diszpozíciója és az annak alapjául szolgáló fenomenálisan tudatos állapot közti viszony nem fejezhető ki teljes mértékben információfeldolgozással kapcsolatos fogalmi keretek között, hiszen „a P-tudatosságról nem adható teljes magyarázat információfeldolgozási terminusokban” (Atkinson – Davies 1995: 249).

Atkinson és Davies szerint az A- és P-tudatosság viszonya úgy írható le, hogy a P-tudatosság a „diszpozicionális és *relatív relációs* A-tudat kategoriális és *relatív intrinzikus* bázisa”. Ezért lehetséges a P-tudatosság A-tudatosság nélkül, de nem fordítva. Ehhez hasonlóan egy másik kommentátor azt veti fel, hogy a (fenomenális) tudat talán betölthet olyan szerepet, mint a víz egy hidraulikus számítógépben (Farah 1995: 255). Ez a hasonlat azt jelenti, hogy az, hogy egy információfeldolgozó rendszer tudatos-e, az implementációjának bizonyos aspektusaitól függ (vagyis az információfeldolgozó funkcionalitás alapjául szolgáló fizikai szubsztrátum tulajdonságaitól). Ezt a fajta tulajdonságot egy funkcionalista elmélet nem tudja megragadni, de ha a tudatosság az elméhez így viszonyul, akkor a fenomenális tudatosságnak sincs külön alrendszere az emberi elme blokkábráján.

Church (1995: 251) szerint egy állapot csak akkor rendelkezhet fenomenális tulajdonságokkal, vagyis akkor számíthat „tapasztalatnak”, illetve akkor lehet „valamilyen” a kérdéses állapotban lenni, ha ez az állapot „valamilyen egy szbjektumnak, vagy a szbjektum számára”, azaz valamilyen relációban kell állnia a szbjektummal, vagy – karteziánus szbjektum híján – a szbjektum bizonyos további *állapotaival*.

Mármost ha egy mentális állapotnak intrinzikus jellemzője a fenomenális karaktere, mint azt a fenomenális tulajdonságok számos híve állítja Church szerint, akkor a fenomenális karakter nem lehet relációs tulajdonság. Ezért Church elutasítja a fenomenális karakter ilyen felfogását. Az alternatíva az, hogy a fenomenális karakter relációs tulajdonság. Ekkor viszont „plauzibilis feltételezni, hogy a releváns relációk hozzáférési relációk” (Church 1995: 252), vagyis olyan relációk, melyeknek szerepük van az organizmus egységes szbjektummá szervezésében. Church valószínűnek tartja, hogy a fenomenális tudatosság fogalmának, ha az egyáltalán értelmes, akkor ugyanaz a tartalma, mint a hozzáférési tudatosságnak. A kétféle tudatosság valójában ugyanaz a tulajdonság, amit alátámaszt az is, hogy a gyakorlatban a kétféle tudatosság mindig egyszerre lép fel, az igazán meggyőzőnek szánt ellenpéldák (mint a milyen a szupervaklítás esete) pedig csupán spekulatív, és nem valódi példák.

Dennett szerint a P- és A-tudatosság megkülönböztethető, de csak mennyiségi jellegű, nem pedig minőségi különbség van köztük. A P-tudatos állapotok tartalomgazdagok, ezért teljes komplexitásukban leírhatatlannak tűnnek, és *fenomenálisan impresszívek*. Más mentális epizódok szegényesebb tartalmúak (Dennett 1995: 252), ezért az sem mindig nyilvánvaló, hogy a szó mindennapi értelmében tudatosak-e egyáltalán. Mivel azonban ezek között csak fokozati különbség van, valójában nincs példa „tisztá” (A-tudatosság nélküli) P-tudatosságra vagy P-tudatosság nélküli A-tudatosságra.

Dennett emlékeztet arra, hogy Block egyszerűen elhiszi a vaklátó páciensnek, hogy az ő esetében hiányos a látási P-tudatosság. Csakhogy ha verbális beszámolóknak hiszünk a P-tudatossággal kapcsolatban, akkor, mint Dennett megállapítja, az „A-tudatosság legjobb kritériumát használjuk a P-tudatosság egyetlen bizonyítékaként”, vagyis maga Block állítja, hogy a hozzáféréstől független tudatosság fogalma inkohérens.

Dennett szerint ha valaki olyan különleges vaklátásról számolna be, amellyel meglehetősen részletgazdag érzékelés jár, akkor sokkal nehezebb lenne neki elhinni, hogy egyszerűen kitalálja, milyen tárgyak esnek a látótere vak részére. Block nem indokolja meg, miért is hinnénk el a szupervaklátónak, hogy nem normál látó. Valóban elhinnénk egy szupervaklátónak, kérdezi Dennett, hogy *semmilyen* számára érzékelni egy öt centiméter magas narancssárga X-et egy kék-zöld háttér előtt? (Dennett 1995: 253) Megfordítva, ha valaki azt állítja, hogy nem vaklátó, hanem normál látó, ám egy betűre nézve csupán annyit tud megállapítani, hogy egy X betűt lát, nem pedig O-t, akkor aligha értenénk, mire céloz az illető, amikor arról beszél, hogy, hogy P-tudatos látási élménye van. Dennett számára tehát a P- vagy A-tudatos állapotok különbsége pusztán fokozati, és az állapotok tartalomgazdagságától függ.

Harman arra mutat rá, hogy hogy a tapasztalat rendszerint reprezentacionális, intencionális tartalommal rendelkezik. A perceptuális tapasztalat például olyan módon reprezentálja számunkra a környezetünket, ahogyan az az adott környezetben bizonyos látószögéből látszik. Harman megjegyzését a könyvtárra vonatkozó korábbi példát felhasználva úgy értelmezhetjük, hogy ha megállunk egy könyvtár egy bizonyos pontján, és egy meghatározott irányba nézünk, majd körbefordulunk, akkor látási élményeink (tehát hogy „milyen” abba az irányba nézni) változása jellemző arra, hogy milyen módon változott a tekintetünk iránya, és az egyes könyvsorokhoz képest milyen pozícióban voltunk a körülnézés során. Mivel azonban Harman vitatja a minden reprezentacionális tartalomtól mentes „nyers érzetek” létezését (bővebben Harman 1990), ezért számára nyilvánvaló, hogy a nyers érzetek megkülönböztethetők a perceptuális érzet minőségétől (tehát az, hogy „milyen” egy bizonyos helyről egy bizonyos irányba nézve vizuális percepció révén tapasztalatokra szert tenni, megkülönböztethető a nyers érzettől, mert az előbbinek van intencionális tartalma, míg az utóbbinak nincs). Ekkor viszont a P-tudatos tapasztalat, tehát ami „valamilyen” az átélője számára, intencionális és reprezentacionális tartalommal rendelkezik, miközben „hozzáférni egy tapasztalathoz egyszerűen annyi, mint tudatában lenni” (Harman 1995: 257), ezért az A-tudatosság elnevezés Harman szerint redundáns.

Levine egyetért a P- és A-tudatosság megkülönböztetésének szükségességével. „A szubjektivitás jelensége, mely a fenomenális tapasztalat magvát alkotja, lényegileg hozzáférést

involvál, és ez az a tény teszi lehetetlenné, hogy világosan megkülönböztessük a P- és A-tudatosságot.” (Levine 1995: 261). Levine szerint „végső soron kétféle hozzáféréssel van dolgunk”. Az egyik a kognitív tudomány számára is jól elemezhető fajta, vagyis információáramlások terminusaiban elemezhető hozzáférés. A másik pedig az, amely „összefonódik a szubjektív élménnyel”.

Lloyd (1995: 262) számára az alapvető tudatfogalom a fenomenális tudatfogalom, a hozzáférési tudatosság fogalma ezen élőködik: csak azért plauzibilis tudatfogalom, mert hozzáférési tudatosság mindig csak ott lép fel, ahol fenomenális tudatosság is. A blocki megkülönböztetés ezért félrevezető: a tudatosság fogalma egyértelmű, alapjelentése a fenomenális tudatosság. A tudatos elme „tapasztalati állapotokból” épül fel. Ezen állapotok némelyike centrálisan hozzáférhetőbb, vagy inferenciálisan termékenyebb, mint némely másik (Block ezeket nevezné A-tudatos állapotoknak), de nem ettől válnak tudatossá, hanem eleve azok (tudniillik fenomenális értelemben).

Lloyd elképzelése így sajátos módon hasonlít a Dennettéhez: csupán fokozati kérdés, hogy valami csak P-tudatos, vagy A-tudatos is. Elképzelésének gyengesége viszont az, hogy mivel a P-tudatosságot tartja alapvetőnek, nem pedig a funkcionális terminusokban viszonylag könnyen definiálható A-tudatosságot, ezért végső soron közvetlenül definiálnia kellene, mi is a P-tudatosság (hiszen arra sincs lehetősége, hogy az A-tudatossággal szembeállítva, az A-tudatosságból kiindulva definiálja). A P-tudatosság közvetlen definíciói pedig, mint már maga Block esetében is láttuk, rendszerint körkörös megfogalmazásokra, illetve olyan gesztusokra szorítkoznak, melyek csak annál az olvasónál érik el a megfelelő hatást, aki a megfelelő intuitív meggyőződésekkel rendelkezik.

Revonsuo (1995: 266) Lloydnál is továbbmegy, szerinte az A-tudatosság fogalmának létjogosultsága megkérdőjelezhető, hiszen A-tudatossággal egy zombi is rendelkezhet. Tudatfogalmunk lényegi tartalma a P-tudatosság.

Rey, akit a fenomenális tudattal kapcsolatos eliminativizmus egyik klasszikus képviselőjeként szokás számon tartani, felteszi a kérdést, hogy létezik-e egyáltalán olyan, „erős értelemben vett” P-tudatosság, amelyre – Block meghatározásaival összhangban – rá lehet mutatni, és nem funkcionális tulajdonság. Rey szerint azonban Blocknak nincs jó oka arra, hogy feltételezze, létezik ilyesmi (Rey 1995: 266). A probléma az, hogy mivel a P-tudatosság Block szerint talán „funkció nélküli”, illetve „nincs hatással későbbi mentális folyamatokra”, ezért nincs okunk feltételezni a létezését – még saját esetünkben sem. Ha a P-tudatosság szerepe analóg a víznek a

hidraulikus számítógép működésében betöltött szerepével, akkor felmerül a kérdés, hogy a tulajdonságaihoz hogyan férünk hozzá mentális életünkben. Rey nagyon fontos problémát vet fel. Érdekes ezt egy példán szemléltetni, mely azt is segíti, hogy jobban megértsük a „relatív intrinzcitás” (Atkinson és Davies által is alkalmazott) fogalmát, illetve Farah gondolatát az implementációs részletek szerepéről.

A víz szerepe a hidraulikus számítógépben fontos lehet a gép komputációs funkcióinak realizálása szempontjából, de a gép komputációs állapotainak tartalmát nem határozza meg vagy befolyásolja az, hogy víz implementálja őket. A gép a komputációi során különféle komputációs állapotok között végez szabályszerű átmeneteket. Ezeknek egymáshoz való viszonya ábrázolható lehetne egy absztrakt folyamatábrán, illetve a gép komputációs struktúráját implementálni lehetne egy másfajta közegben is, például szilíciumchipben vagy egy mechanikus elven működő gépben. Ez a többszörös realizáció közismert elve, amiből azonban az is következik, hogy ha a hidraulikus gépnek az S1 absztrakt komputációs állapotból S2 állapotba kell kerülnie, akkor ugyanezt kell tennie a vele azonos funkcionális elrendezésű elektronikus gépnek is. Ez még akkor is így van, ha történetesen S1 az az állapot, amibe a gép akkor kerül, ha azt a kérdést teszik fel neki, hogy „milyen fizikai közegben implementáltak?”. A gép az absztrakt állapotok terében mozog normális működése során, és így nem tudhat az absztrakt állapotainál alacsonyabb vagy alapvetőbb szintekről információt szerezni (kivéve az olyan speciális eseteket, amikor a komputert felszereltük saját maga fizikai jellemzőinek kutatására alkalmas szenzorokkal és szerszámokkal – ám ennek tárgyalása messzire vezetne).

Másik példával érzékeltetve: ha egy sakktáblán sötét éppen mattot adott, akkor akkor is mattot adott, ha a sakktábla üvegből van, és akkor is, ha műanyagból, hiszen a sakkjáték is többszörösen realizálható. Éppen ezért viszont a játék során a játék megnyerése szempontjából senki sem tudja előnyére fordítani azt a tényt, hogy a sakk-készlet üvegből készült, mert a készlet anyaga nem befolyásolja a mindenkori sakkszituációt. Ám mindez nem jelenti, hogy a számítógép, mely nem tud számot adni saját implementációs részleteiről, nincs implementálva, vagy azt, hogy a sakk-készletnek semmiféle fizikai természete (mérete, színe, anyaga) nincs.

Ha a P-tudatosság valóban nem valamilyen információfeldolgozási vagy kognitív jellemző, hanem az információfeldolgozási folyamatokat realizáló neurofiziológiai vagy fizikai szubsztrátum tulajdonságaihoz kötődik szorosan, (például mint a magasabbrendű diszpozicionális tulajdonságok kategoriális bázisa), akkor Reynek természetesen igaza van abban, hogy nincs kognitív hozzáférésünk a P-tudatosság tulajdonságaihoz. Egy reduktív materialista számára, aki a P-tudatosságot azonosítja az agy bizonyos fizikai tulajdonságaival, ez egyszerűen azzal az állítással

egyenértékű, hogy introspekció útján nem tudjuk megismerni agyunk anatómiáját vagy biokémiáját. De ezt nem is várhatjuk, hiszen az introspekció, a számítógép példájára utalva, lehetséges „komputációs állapotaink” (esetünkben ezek pszichológiai állapotok) terében mozog, hiszen *mentális* funkciónk. Következik mindebből, hogy el kell utasítanunk a P-tudatosság fogalmát? Nem, hiszen agyunk mentális funkciói valamilyen módon implementálva vannak, agyunk rendelkezik fizikai tulajdonságokkal, melyek egy része az azonosságelmélet hívei szerint egyébként azonos a fenomenális élményeink tulajdonságaival. Az azonban világos, hogy ezeket a tulajdonságokat nem ismerhetjük meg introspekcióval. De mióta feltétele valami létezésének, hogy megismerhető-e introspekcióval?

Rey mindazonáltal helyesen mutat rá, hogy az introspekció „hozzáférés”, vagyis az egymásból elérhető funkcionális állapotok hálózatának része, de természetesen nem alkalmas arra, hogy hozzáférjen, ahogy Rey fogalmaz, „hidraulikánkhöz” (Rey 1995: 267). Rey konklúziója, hogy elutasítja az „erős értelemben vett” P-tudatosságot annak introspektív megismerhetetlensége miatt, ami oda vezet, hogy egy olyan szubjektum is tulajdoníthatna magának P-tudatosságot, amely ezzel nem rendelkezik, és nincs módunk eldönteni, hogy mi is ilyenek vagyunk-e vagy sem. Rey tehát Block-kritikájában is következetesen képviseli azt a nézetet, ami miatt a „tudat tagadói”, sőt, a kifejezett „fenomenális eliminativisták” közé szokás sorolni.

Annak bizonyítása azonban, hogy a fenomenális tudatosság nem megismerhető az introspekcióban olyan módon, hogy megismerése funkcionális/kognitív különbségeket idézzen elő az elme működésében, nem ugyanaz, mint bebizonyítani, hogy nincs tudatosság. Egy reduktív materialistának például nem kell azt mondania, hogy introspekció alapján vagy az érzékelés során „közvetlen ismeretségbe” kerülünk a tapasztalat fenomenális tulajdonságaival és „közvetlen ismeretségből vett tudás” alapján tudjuk, hogy nem vagyunk zombik. A reduktív materialista *következtet* arra, hogy nem vagyunk zombik, abból kiindulva, hogy a *normál módon implementált* (vagyis *fizikailag normális*) emberi elme *fenomenológiailag normális*. Egy funkcionalista, aki szerint „ugyanolyan embernek lenni”, mint az emberi elme működését funkcionális szerveződése szempontjából híven utánczó mechanikus gépezetnek (Block 1978) szintén feltételezheti, hogy a nem „standard módon” implementált emberi elme szintén *fenomenológiailag normális*. Egy realizáció-funcionalista vagy reduktív materialista erre persze további bizonyítékokat várhat, hiszen szerinte ha sikeres is az emberi elme funkcionális utánczása egy másik közegben, ebből nem következik, hogy fenomenális utánczása is megtörtént. Természetesen ha introspekció útján hozzáférhetnénk élményeink fenomenális tulajdonságaihoz, akkor el kellene fogadnunk, hogy

funkcionális duplikátumunk egyúttal fenomenális duplikátumunk is. Ugyanis ha Kovácsot megkérdezik, hogy zombi-e (vagy: milyen érzés neki lenni), akkor ennek hatására alaposan „magába néz”, és „V1” állapotba kerülve meghatározott választ ad. Kovács robot-utánzata, funkcionális duplikátuma ugyanazt a választ fogja adni, vagyis (másként implementált) „V1” állapotba kerül, aminek előzménye mindkét esetben „F” fenomenális állapot. Ezt a konklúziót csak akkor kerülhetjük el, ha kivonjuk a fenomenális állapotokat a funkcionális állapotok összefüggéséből, éppen, ahogy Block javasolta. Ekkor viszont első ránézésre kísérletileg eldönthetetlen kérdés lesz, hogy a Kovács-robot fenomenális élményminősége ugyanolyan-e, mint a Kovácsé. Rey ebből arra következtet, hogy nincs értelme feltételezni a funkcionálisaktól különálló fenomenális tulajdonságokat.

Aki azonban természetes meggyőződése következtében inkább elfogadja a tudat létezését, vagyis az antirevizionizmusra hajlik, az a fenomenális tudatosság megismerhetetlenségéből azt a következtetést vonja le, hogy nehéz megítélni, készíthető-e például tudatos számítógép, nem pedig azt, hogy nehéz eldönteni, zombik vagyunk-e. Ha nem az introspekcióna hivatkozva, akkor egyéb okból tartják lehetségesnek, hogy higgyenek a fenomenális tudatosság létezésében. Vagyis Rey itt sem tud olyan érveléssel előállni, ami egy antirevizionistának komoly fejtörést okozna.

Shepard (1995: 268) felteszi a kérdést, hogy függ-e a P-tudatosság a komplexitástól oly módon, ahogy vélhetőleg az A-tudatosság (mely komputációs természetű) igen. Feltételezhető, hogy egyszerűbb rendszerekben (azaz alsóbbrendű élőlényekben) az A-tudatosság már nincs, vagy kevésbé van jelen; a komplexitás csökkenésével az A-tudatosság valahol „elfogy”. Függ-e ilyen módon a komplexitástól a P-tudatosság? Ha a P-tudatosság fennállhat anélkül is, hogy róla való beszámolásra készen állna, vagy hogy célirányos cselekvés alapjául szolgálna, akkor ennyi erővel akár egyetlen neuron tüzelését is kísérhetné valamilyen privát érzetminőség. Sőt, az sem kizárt, hogy egy elektronnak legyenek fenomenális érzetminőségei. Ezzel Shepard azon túl, hogy felidézi a neutrális monista álláspontot, érintőlegesen egy olyan kérdést is feszeget, amely a *fenomenális hozzáférhetetlenségének* irodalmában már sokszor megjelent. Míg Dennett és sok más szerző úgy találja, hogy a P-tudatosság ineffabilitásának oka, hogy túlságosan információgazdag ahhoz, hogysem szavakba lehessen öltetni a P-tudatos állapotok tartalmát, és meglehetősen régi elképzelés, hogy az érzet és a gondolat különbsége ezek komplexitásában rejlik az előbbi javára (kritikáját lásd: Sellars 1956/1963: 156), sok szerző ellenkezőleg úgy látta, hogy a fenomenális megragadhatatlanságának oka éppen ellenkezőleg, a fenomenális állapotok információszegénysége, tartalmatlansága. (Farrell 1950, Feyerabend 1970, Smart 1959).

Az eddigiek alapján úgy tűnik, hogy azok a szerzők, akik a fenomenális tudatosság önállóságát vitatják, vagyis a fenomenális tudatosságot a hozzáférési tudatosságra vagy más funkcionális, kognitív tulajdonságokra redukálják, elfogadhatják a fenomenálisnak tartott állapotok információgazdagságának téziséét, hiszen a fenomenális állapot az információtartalma révén foglalja el a helyét az egymás előzményeiként és következményeiként kapcsolatban álló különféle mentális állapotok funkcionális terében. Ha nem vesz részt információfeldolgozási és kognitív folyamatokban, akkor az információtartalmat – legalábbis az elme kognitív folyamatai számára hozzáférhető információtartalmat – el kell vitatni a fenomenális állapotoktól. Ez azonban rögtön elvezet annak rejtélyéhez, hogy hogyan szerzünk egyáltalán tudomást fenomenális tudatosságunkról. Shepard a kérdést úgy veti fel, hogy egyáltalán hogyan lehetséges, hogy beszélni tudunk a fenomenális tudatosságról.

Amikor Dennett azt mondja, hogy a látási élmény „gazdag” részletekben és információkban, és pedig annyira, hogy szavakban nehéz kifejezni, akkor ezzel jól leírja azt a különbséget, ami például egy híres festmény megpillantása és aközött van, ha csak olvasunk róla. A látási élmény „gazdagabb”. Másfelől Harman (1990) felhívta a figyelmet arra, hogy az a „gazdag” élmény, ami ilyenkor elementáris erővel hathat ránk, egyszerűen egy lenyűgöző tárgy vizuális tulajdonságainak percipiálásában áll. Nem az élményünk, hanem a *tárgya* az, ami érdekesebb, mint a róla szóló puszta leírások. Sőt, ha a fenomenális tudatos állapotok mentesek minden funkcionális és információs szereptől, és nem töltenek be semmiféle szerepet a funkcionális és kognitív állapotaink ökonómiájában, akkor megismerhetetlenek, és az sem világos, egyáltalán miért beszélünk róluk. (Rey ebből nemlétezésükre következtetett.)

Shepard csupán félig komolyan szánt felvetése szerint ha lennének az emberek között olyanok, akik rendelkeznek fenomenális tudatossággal, és olyanok, akik nem (minden egyéb jellemzőt változtatlanul hagyva), akkor a fenomenális tudatosság tesztjének lehetne használni a mások színlátásával kapcsolatos rejtvény (Farrell 1950: 177) megértésére való képességet. Aki nem rendelkezik fenomenális tudatossággal, az egyszerűen nem érti, mire gondol, aki azon töpreng, hogy „mások is ugyanolyan pirosnak látják-e a piros színt, mint én”, míg aki rendelkezik ezzel, az érti.

Mindennek pedig Shepard szerint etikai jelentősége van, ha elfogadjuk, hogy egy zombinak okozott szenvedés erkölcsileg közömbös. Ekkor persze – hiszen kognitív képességeinket, viselkedésünket az evolúció során a túlélés szempontja formálta – a zombik várhatóan olyan értelmű válaszokat adnak a kérdéseinkre, melyek alapján nem-zombinak tűnnek – saját jól felfogott érdekükben. A gondolatmenet nyilvánvaló problémája, hogy ha a zombi és a nem-zombi között

tényleg csak a fenomenális tulajdonságokban áll a különbség, akkor ezek a lények elkerülhetetlenül ugyanúgy fognak válaszolni a tesztkérdésre, ha viszont a fenomenális tulajdonságokban mutatkozó különbség elégséges ahhoz, hogy olyan funkcionális különbségekhez is vezessen, melyek megnyilvánulásaként a zombi saját bevallása szerint nem érti a mások színjátására irányuló kérdését, míg a nem-zombi igen, akkor azok, akik szkeptikusak a nem-funkcionális tulajdonságként felfogott fenomenális tudatossággal szemben, joggal feltételezhetik, hogy a zombi és a nem-zombi közötti állítólagos fenomenális tulajdonságokban mutatkozó különbség igazából funkcionális különbség. Ha viszont van funkcionális tesztje a „fenomenális tudatosságnak”, az nyitva tartja annak a lehetőségét, hogy az, amit „fenomenális tudatosságnak” neveznek, valójában csupán valamilyen funkcionális tulajdonság. Ezt csak akkor zárhatjuk ki, ha Blockkal egyetértve ragaszkodunk hozzá, hogy a fenomenális tudatosság mint olyan nem bír funkcionális vonatkozásokkal. Ekkor viszont nem lesz viselkedési, verbális, kognitív, tehát funkcionális tesztje.

2.1.2 Block válaszai

Block (1995: 273) a fenomenális élmény információgazdagsága vagy -szegénysége tárgyában rámutat, hogy az ízlelés kevésbé gazdag, mint a látás, de ettől még nem kevésbé fenomenális. Hasonlóan, az orgazmus különösen információszegény, mégis erőteljes fenomenális élmény.

Dennettnek válaszolva Block elismeri, hogy nehéz elhinni a szupervaklátónak azt, hogy nincs vizuális élménye, és plauzibilis, hogy bizonyos szintű információgazdagság *elégséges* feltétele lehet a fenomenalitásnak az emberekben, de azt nem, hogy *szükséges* feltétele lenne (lásd az ízlelés vagy az orgazmus példáját), ezért Block azt továbbra sem fogadja el, hogy a fenomenális tudatosság mindössze az információgazdag tartalommal bíró mentális állapotokkal rendelkezéssel lenne azonos.

Blocknak teljesen igaza van abban, hogy még ha el is fogadjuk például Tye (1995: 268)-nak azt az álláspontját, hogy még az orgazmus is értelmezhető reprezentacionális állapotként, mégiscsak *prima facie* ellenpéldának tűnik az orgazmus (vagy éppen a fázás, hőérzet, egy tompa nyomásérzet stb.) arra az elméletre, hogy a fenomenális állapotok információgazdag állapotok. Ezért elvethető az az állítás, hogy a fenomenalitás: információgazdagság.

Ezzel kapcsolatban megjegyezhetjük, hogy az „információgazdagságot” sokféleképpen lehet definiálni és mérni. Egy fontos üzenet lehet néhány betű hosszú, míg egy bonyolult zaj lehet, hogy nem hordoz hasznos információt, miközben (például egy számítógépben) az utóbbit sokkal több byte-on lehet csak reprezentálni.

Vegyük azt az esetet, amikor fázunk. Ennek az állapotnak a cselekvésünk meghatározása szempontjából releváns információtartalma meglehetősen szegényes, hiszen vagy fázunk (ekkor feljebb csavarjuk a fűtést), vagy nem. Másfelől amikor fázunk, akkor nem egyszerűen tudunk erről az egy bitnyi információról, hanem számtalan idegvégződésünk egyidejűleg „üzeni” számunkra, hogy hideg van. Így egy viszonylag „komplex” állapotba kerülünk, még ha hasznos információtartalma alacsony is.

Ez alapján tehát egyetérthetünk Blockkal abban, hogy a fenomenalitás azonosítása az információgazdagsággal, első közelítésben, elcsúsztattnak tűnik, és az olyan esetekben, mint amilyen az orgazmus vagy a fázás, könnyen megérthető, hogy a kérdéses érzetállapot „információgazdagságának” sokféle meghatározása is lehetséges. A hozzáférési tudatossághoz közelebb eső meghatározások a viselkedés gyakorlati irányítása szempontjából állapítják meg az információgazdagságot, de könnyen belátható, hogy az élmény minősége szempontjából, ami a fenomenális tudatossággal látszik szorosabb kapcsolatban lenni, egy érzet „gazdagabb” lehet, mint a másik szempontból.

Block szerint az orgazmus „nem információgazdag, de fenomenálisan impresszív”. Talán ennek az az oka, hogy az orgazmus azért impresszív, mert úgy érezzük, hogy testünk számos részében lép fel és sokféle hatást gyakorol ránk egyidejűleg, és pedig meglehetősen intenzíven. De az impresszivitás nem információgazdagság. Ha egy kórus énekel, akkor a szöveg nem lesz bonyolultabb, de az előadás esetleg impresszívebb, mintha egyvalaki énekel.

Most már értjük, hogy miközben Dennett, Armstrong és Farah hajlamosak a fenomenálisan tudatos állapotokat információgazdag állapotoknak tekinteni, addig komoly táborra van annak az álláspontnak is, hogy az érzetek „tulajdonságnélküliek” vagy „színtelenek” (mint Farrell 1950, Smart 1959 és Feyerabend 1970 szerint). A „fenomenális impresszivitás” az, ami miatt a fenomenális állapotok információgazdagnak tűnnek. Mivel ez bizonyos fajta komplexitáson alapul, ezért bizonyos értelemben a fenomenális állapotok valóban információgazdagak, de más szempontból (a hozzáférési tudatosság számára érdekes gyakorlati szempontokból) nem azok. Így aztán amennyiben kivonjuk a látási élményből annak funkcionális szempontból fontos vonatkozásait, ami visszamarad, a póre fenomenális karakter funkcionális és kognitív tartalom

nélkül. Erről *ex hypothesi* semmit sem tudunk mondani, és nem befolyásolja a viselkedésünket. Mivel ennyire nincs szerepe az önmagában vett fenomenális karakternek, Dennett szerint nehéz elhinni valakinek, hogy „vak”, ha aprólékosan beszámol arról, milyen képek vannak előtte.

Nem kerülünk közelebb ahhoz, hogy bizonyítottnak tekintsük egy teljességgel nem funkcionális, nem kognitív fenomenális tudatosság létezését és megkülönböztethetőségét a hozzáférési tudatosságtól. Ugyanis éppen az a tény, hogy egyáltalán fel tud merülni az élmények „gazdagsága” gyakorlati információtartalmukon túl, arról árulkodik, hogy van valami a tapasztalataink megélésében, ami mintha túlmutatna azok gyakorlati jelentőségén. Ám ez önmagában nem elég a radikális P/A megkülönböztetéshez.

A „kognitív olvasat” ugyanis továbbra is rendelkezésünkre áll, sőt, felületesen szemlélve mintha ki is kényszerítené a kognitív olvasatot (tehát a P- és A- tudatosság olyan értelmezését, amiben a fenomenális tudatosságnak is van kognitív-funkcionális szerepe) az a tény, hogy képesek vagyunk, legalábbis az antirevizionisták számára *prima facie* úgy tünik, hogy képesek vagyunk megérteni a mások színjátására vonatkozó kérdést, a zombi-problémát vagy éppen Block cikkét. Csakhogy ha a kognitív olvasatot végigvisszük, akkor ezzel a blocki megkülönböztetés revizionista olvasatát állítottuk elő. A revizionista olvasat alapján Block kérdésére, mely szerint „létezik-e a P?” (Block 1995: 274), a válasz határozott „nem”, vagy legfeljebb: „igen, de nem abban az értelemben, ahogyan azt az antirevizionisták remélik”.

Rey bizonyítékokat követelt, amelyek szükségessé teszik a mindenféle komputacionális jelenségtől megkülönböztethető P-tudatosság „posztulálását”. Block Rey álláspontját „zavarbaejtőnek” nevezi (Block 1995: 274), és „saját tapasztalatunk bizonyítékaira” hivatkozik. Enélkül Block szerint figyelmen kívül hagynánk egy olyan „bizonyítékforrást, amellyel mindannyian rendelkezünk” saját tapasztalataink révén. Ezzel azonban a kör bezárult, és ismét azt látjuk, hogy a filozófiai vita kimerül abban, hogy a filozófusok közlik egymással, mi tünik számukra nyilvánvalóan igaznak.

Nem különösebben meggyőző azonban olyan bizonyítékokra hivatkozni, amelyek teljesen irrelevánsak. R-olvasatban saját tapasztalatunk nem más, mint a percepció és minden más információszerzési képességeink gyakorlása. Kognitív, funkcionális fogalmakban teljes egészében leírható. Mint már láttuk, a „fenomenális tapasztalat gazdagsága” azokban az esetekben, amikor egyáltalán valóban információgazdag perceptuális állapotokról van szó, információgazdagságot jelent gyakorlati szempontból (a viselkedésre gyakorolt hatások sokfélesége szempontjából). Más, Blocktól vett kifejezéssel, inferenciálisan termékeny állapotokról van szó. Egy vizuális állapot

„gazdag”, mert sok részlet fedezhető fel a látómezőben. Egy könyvekkel teli könyvtár vizuális képe például számos könyvet tár elénk. Bármelyikhez odamehetünk, és leemelhetjük a polcra, vagyis a vizuális tapasztalat számos gyakorlati következtetést tesz lehetővé, azaz számos lehetséges cselekvési alternatívát nyit meg előttünk, azaz információgazdag, komplex.

Mi a helyzet a fenomenális impresszivitással? A tikkasztó hőség, a csontig ható fázás, az ólmos fáradtság vagy éppen az orgazmus mind impresszívek. Kétféle információtartalommal bírnak. Egyrészt, ezek révén regisztrálunk egy olyan típusú értesülést, amelyekhez e modalitások révén jutunk tipikusan hozzá. Például megtudjuk azt, hogy a helyiség számunkra túlságosan hideg, és ideje becsukni az ablakot vagy bekapcsolni a fűtést. Másrészt, amikor beszámolunk arról, hogy milyen erős hidegérzetünk van, akkor maga az impresszivitás válik információvá, az impresszivitás pedig azt mondja meg, hogy milyen erős (mennyire intenzív, kiterjedt és hosszan tartó) hatást gyakorolt szervezetünkre a releváns inger. Így például minél intenzívebb a fázás, annál jellemzőbb, hogy további testi érzeteket vált ki, mint például a végtagok elgémberedését, fájdalmakat, remegést. Az információtartalom nem nő meg, hiszen a „nagyon hideg van” nem feltétlenül több információ, mint a „hideg van”, azonban egy komplex, kiterjedt érzés fenomenálisan impresszívebb, mert számos testi érzetre tagolódik, melyek mindegyike rendelkezik bizonyos reprezentacionális tartalommal, még akkor is, ha ezek együttesen mind csak a „nagyon hideg van” információt hordozzák a cselekvés irányítása szempontjából.

A fenomenális impresszivitás végső soron információ, számos párhuzamos, azonos tartalmú információ áramlásaként érthető meg. Semmiféle akadálya nincs annak, hogy információáramlási, funkcionális, kognitív módon értelmezzük a fenomenális impresszivitást, miközben továbbra is világosan megkülönböztethetjük a hozzáférési információgazdagságtól.

Ennek lehetősége miatt aki fenn akarja tartani, hogy *létezik tudat*, abban az értelemben, hogy van valami, túl az információfeldolgozáson, ami meghatározza, hogy milyen egy bizonyos tapasztalatot átélni, és abban az értelemben, ahogy ezt szemlátomást az emberek jelentős része érti, annak közvetlenül meg kellene mutatnia a fenomenális tudatosság létezését. Ennek fontosságát maga Block is elismeri Rey-nek adott válaszában (Block 1995: 274). Természetesen Block éppen ezt kísérelte meg tanulmánya bevezető szakaszaiban, amikor megkísérelt „rámutatni” a fenomenális tudatosság szerinte nyilvánvaló eseteire. Ez azonban olyan gesztus, ami a szkeptikusokat nem győzi meg.

Atkinson és Davies szerint P „relatív intrinzikus és kategoriális” A-hoz képest, mint ahogy az oldhatóság kémiai alapja az oldhatóság diszpozíciójához képest. De ha ez így is van, ez

önmagában nem szünteti meg az R-olvasat lehetőségét. A P-tudatosság az R-olvasatban egyszerűen az A-tudatosság kategoriális bázisa, és *semmi több*. Megfigyelni a viselkedést lehet és azokat a mentális állapotokat, melyek funkcionális szerepet játszanak. Ami közvetlenül megfigyelhető, az az A-tudatosság (illetve a viselkedés, amiből az A-tudatosságra következtetünk, vagy ami alapján az A-tudatosságot definiáltuk), a P-tudatosságot mint ennek a kategoriális bázisát posztuláltuk.

Az antirevizionisták gyakran úgy gondolják, hogy „közvetlen tudásuk” van fenomenális tudatosságukról; legalapvetőbb tapasztalatuk, hogy *tapasztalnak*, ami A-olvasatban azt jelenti, hogy tudatos élményeken mennek keresztül. Bármennyire is nehéz ezt fogalmilag megragadni, az antirevizionisták számára mégsem fér kétség a kognitív és funkcionális tulajdonságokra nem redukálható tudatosság létezéséhez, amit úgy értelmezhetünk, hogy számukra a fenomenális tudatosság veendő primitívnek, mint kiindulópont (lásd például Revonsuo álláspontját).

Amikor a fenomenális impresszivitás és a hozzáférési információgazdagság viszonyát elemeztük, akkor megállapítottuk, hogy egy élmény impresszivitása és információtartalma között az a viszony áll fenn, hogy az élmény impresszivitásáért felelős mechanizmusok közlik vagy implementálják az információtartalmat – meglehetősen, redundánsan.

Ezt legalább kétféleképpen lehet értelmezni. Az egyik lehetőség egyfajta „jelzőberendezés-modell”, amelyben az impresszív információközlést úgy foghatjuk fel, mintha egymással párhuzamosan számos csatornán történne ugyanannak az információnak a közlése, vagy a közölt információ tartalma egyszerűen az lenne, hogy „egyidejűleg sok csatornán érkezik jelzés”, a kevésbé impresszív információközlést pedig úgy, mintha ez kevesebb csatornán menne végbe. Ekkor az impresszivitás megkülönböztethető a hozzáférési információtartalomtól, noha bizonyos, az impresszivitásra vonatkozó információk végső soron hozzáférhetők.

Egy másik lehetőség a „hidraulikai modell” a hidraulikus komputer blocki példájából kiindulva. Az információközlő és -feldolgozó folyamatok valamilyen közegben vannak implementálva, valamilyen fizikai hordozójuk van. Ez a fizikai hordozó (pontosabban annak aktuális állapota) rendelkezhet olyan sajátosságokkal, amelyek fenomenálisan impresszívek lehetnek, például a komplexitás vagy intenzitás bizonyos fokán. Ebben a modellben nem válik közvetlenül hozzáférhető az impresszivitásra vonatkozó információ.

Akármelyik modellt is vesszük szemügyre, egy olyan filozófus, aki szkeptikus a fenomenális tudatossággal kapcsolatban (legfeljebb annak olyan deflacionárius értelmezéseit engedve meg, amelyek visszavezetik a fenomenális tudatosságot funkcionális jellemzőkre), képes produkálni e modellek revizionista olvasatát. A jelzőrendszer-modell esetében csak annyit állít,

hogy a „fenomenális impresszivitás” egy olyan további információ a „hozzáférési” tudatban, amely kiegészíti a gyakorlati szempontból lényeges információtartalmat.

A hidraulikai modell esetében a revizionista annyit mond, hogy a „fenomenális impresszivitás” valamilyen fizikai jellemző, amely bizonyos funkcionális állapotok alapjául szolgáló fizikai állapotokra és folyamatokra jellemző, de fizikai jellemzőin túl nincs értelme még további, fenomenális tulajdonságokat posztulálni, melyekkel ezek rendelkeznek. Ha erre válaszul az antirevizionista arra hivatkozik, hogy introspekciója vagy érzetei „közvetlen bizonyítékkal” szolgálnak arról, hogy minden informatív testi érzetét vagy perceptuális állapotát valamilyen fenomenális minőség jellemzi vagy kíséri, akkor a revizionista a hidraulikai modell feltételezése esetén rámutat, hogy ilyen bizonyítékkal mentális állapotaink biztosan nem szolgálhatnak, hiszen a fenomenális impresszivitásért felelős folyamatok „relatívén intrinzikus” folyamatként csak megalapozzák a mentális állapotokat, de nem fordulnak elő a mentális állapotok logikai terében. Viszonyuk tehát olyan a mentális állapotokhoz, mint a sakktábla anyaga vagy mérete a sakktáblán kialakuló játéksituációhoz (például döntetlen vagy matt) képest. A hidraulikai modellben a fenomenális impresszivitás revizionista értelmezése az, hogy a „fenomenális impresszivitás” végső soron része a gyakorlati szempontból lényeges (hozzáférési) információnak, például a fenomenálisan impresszív fázis egyszerűen az intenzív fázis, aminek információtartalma a különösen alacsony testhőmérséklet.

Visszatérve Block tanulmányához, Block elképzelhetőnek tartja, hogy az emberi elme működését szemléltető blokkábrán Schachter modelljének megfelelően külön „doboz” legyen a fenomenális tudatosságnak. Ekkor a fenomenális tulajdonságok és az információtartalom viszonya nem a relatívén intrinzikus és az általa megalapozott funkcionális tulajdonságok közötti viszonyra emlékeztet, hanem az előbb vizsgált modellek közül közelebb van a jelzőcsengő-modellhez. A probléma természetesen az, hogy mihelyt megengedjük, hogy a P-tudatosság külön dobozt kapjon az elme funkcionális szerkezetét szemléltető blokkábrán, kitesszük elképzelésünket azon revizionista olvasat lehetőségének, hogy a P-tudatosság az ábrán szemléltetett funkcionális szerep betöltője (töltse be akármi ezt a szerepet) és *semmi más*.

Block szerint „puszta józan ész” alapján hihető az az elgondolás, hogy a fenomenális élményben valami több is van, mint annak puszta reprezentacionális tartalma. (Block 1995: 280). Block ezzel szembeszáll Tye, Harman és Lycan nézetével, akik szerint a fenomenális tartalom is teljes egészében reprezentacionális. Block a különböző fenomenális karakterű (például égető, szűrő, éles, lüktető) fájdalmak példáját hozza fel, amire azonban azt lehetne válaszolni, hogy egy égető vagy szűrő fájdalom reprezentacionális tartalma is más, mert (élve a Smart által bevezetett

tárgyközömbös fordítások módszerével) egy szúró fájdalom az ujjam hegyében olyan módon reprezentálja az ujjamat, hogy az olyan állapotban van, mintha megszúrták volna (és nem olyanban, mintha megégették volna).

Reynek arra a kérdésére, hogy ha a P-tudatosság semmi funkcionálisnak, kognitívnek vagy intencionálisnak nem felel meg, akkor ugyan mi okunk lenne létezését posztulálni, Block a következő érdekes választ adja: noha egy számítógép belsejében lévő vezetékek színe epifenomenális „pszichológusi értelemben” (vagyis nincs hatással a komputer komputacionális állapotaira, hasonlóan ahhoz, ahogy sakktáblás példáinkban a sakktábla mérete nem befolyásolja a sakkállapotot), nem epifenomenális „filozófusi értelemben”, mert hatással van a megfigyelőkre (megfigyelhető a számítógépbe belenéző szerelő által). A „fenomenálistudat-modul”, mely elménkben van – ha van ilyen – akkor hasonló módon megfigyelhető fiziológusok által akkor is, ha pszichológiai szerepe nincs (Block 1995: 282). Ez igaz, persze a fiziológusok csak akkor tudják megkeresni az agyban a fenomenális tudatért felelős régiót vagy tulajdonságokat, ha feltételezzük valamilyen fenomenális-fizikai azonosságelmélet igazságát.

Block szerint Dennett korábban eliminativista volt, a 90-es években azonban már a tudattal kapcsolatban realistiként határozta meg pozícióját, bár hozzátéve, hogy az eliminativizmus és a redukcionista realizmus közötti különbség „gyakran tisztán taktikai” (amely megjegyzés nem segít tisztázni, hogy voltaképpen *mi* is Dennett álláspontja egészen pontosan). Block helyesen jegyzi meg, hogy nem minden redukcionizmus áll közel az eliminativizmushoz (Block 1995: 283). Block úgy értelmezi Dennett pozícióját, például ahogyan az Dennett (1991)-ben megjelenik, hogy mivel olyan szubsztantív állításokat tesz a tudatról, mint például hogy a tudat kulturális konstrukció, Dennett „messze maga mögött hagyta” a tudattal kapcsolatos eliminativizmust. Igaz, mint azt Block is elismeri, az a tudat, amivel kapcsolatban Dennett realista, valamilyen deflacionárius értelemben vett tudat. Nem is gondolja minden filozófus, hogy Dennett valóban realistává vált volna a tudattal olyan értelemben, ami átsorolná őt az antirevizionisták közé.

2.2 Revizionizmus: Dennett

Daniel Dennett egyike azoknak a szerzőknek, akiket Strawson a tudat tagadói között említett. Strawson nincs egyedül Dennett nézeteinek ezzel az értelmezésével. A Chalmerssel és Blockkal folytatott viták alapján nem nehéz erre a megállapításra jutni. A következőkben

megvizsgáljuk Dennett néhány olyan művét, amelyek miatt Dennettnek más filozófusok a tudat tagadásának álláspontját tulajdonították, és az ennek hatására kibonakozó viták jó példákkal szolgálnak a tudatfilozófiai álláspontok értelmezésének nehézségeire.

2.2.1 Dennett – Kinsbourne: Time and the Observer

Idegtudományi kísérleti eredményekre hivatkozva Dennett és Kinsbourne saját „több pizkogat” (multiple drafts) elmémmodelljük mellett érveltek a „kartezianus materializmussal” szemben, mely az az álláspont, hogy a tudatos elmék létezése egy nézőpont fennállását feltételezi, és ez a nézőpont, a szubjektivitás székhelye, pontosan lokalizálható az agyban. Descartes feltételezte, hogy az agynak van egy olyan középponti helyzetű része (a tobozmirigy), amely „átjáróként” szolgált a tudatos elméhez (Dennett – Kinsbourne 1992: 184). A tobozmirigy tehát olyan, mint egy színpad, amelyen bemutatásra kerülnek az elme számára az információk. A kartezianus materializmus az az álláspont, mely szerint létezik az agynak egy olyan centrális része, ahol a tapasztalat számai „mind összefutnak” (Dennett – Kinsbourne 1992: 184): az az információ válik a tudatos tapasztalás részévé, amely elérte ezt a központot vagy „főhadiszállást” (Dennett – Kinsbourne 1992: 185).

A kartezianus materializmus szerint amikor egy inger az érzékszervek felől eléri az agyat, legalábbis elvileg meg lehetne határozni, hogy milyen határvonalat kell ahhoz átlépnie, hogy a tudatosuljon. A tudatos tapasztalat annak folyamata, ahogy a „kartezianus színházban” egymás után színpadra lépnek a különféle érzetek és gondolatok, érkezésük sorrendjében. Ez a folyamat teszi ki szubjektumunk történetét.

Ennek alternatívája, a „több pizkogat” modell párhuzamos információfeldolgozási folyamatokkal és eseményekkel számol, melyek az agy különböző pontjain mennek végbe. Ha már valamilyen ingerfeldolgozás, megfigyelés vagy megkülönböztetés végbement valahol az agyban, az így létrejött információtartalomnak nem kell még ezen felül megjelennie valamilyen központi szintéren, valamiféle centrális újrarendelés céljából, hanem maguk ezek a lokális feldolgozási események járulnak valamilyen módon hozzá a teljes agyállapothoz, így tehát a viselkedéshez (Dennett – Kinsbourne 1992: 185).

A tanulmányban az eseményeknek a szubjektív tapasztalatban fennálló időrendje és az e tapasztalatot megelőző agybeli diszkriminációs, adatfeldolgozási események objektív időrendje

közi összefüggés kérdése nagy hangsúlyt kap. Az erre vonatkozó kísérleti eredmények ugyanis alátámasztják a szerzőknek azt az állítását, hogy tudatfolyamunk nem tekinthető „egyetlen, jól meghatározott narratívának”, hanem esetenként egymásnak is ellentmondó, párhuzamos, állandó revízió alatt álló tartalmak folyama, melyből egyetlen szál sem lehet a tudatos élmény „igazi verziójaként” kiemelni (Dennett – Kinsbourne 1992: 185). Ha ugyanis nincs világos érkezési sorrend, akkor nincs egyértelmű határvonal sem. Ekkor viszont egységes narratíva sincs, csak párhuzamos történet-töredékek.

Az elmélet jelentőségét azon kívül, hogy számos kritikusa a „tudat tagadásának” álláspontját rótta fel neki, az is adja, hogy a *nézőpont* vagy az egységes, térben-időben jól körülhatárolható *szubjektum* elképzelésének elutasítása olyan lépés, amely az illuzionizmus védelmezésekor is szerepet kaphat. Azzal az állítással szemben ugyanis, hogy a tudat illúzió, kézenfekvő módon felmerül, hogy az illúzió *valakinek* az illúziója: valami valahogyan *valaki számára* tűnhet másnak, mint amilyen valójában. Ha ez a valaki tudatos megismerő szubjektum, mert illúziója csak tudatos megismerő szubjektumoknak lehet, akkor az ellenvetés szerint az illuzionizmus önellentmondásba keveredik. Ha azonban szubjektumról vagy nézőpontról valójában nem is beszélhetünk, akkor ez az ellenvetés érvénytelen. (Sokat vitatott kérdés, hogy tudatosságról beszélhetünk-e akkor is, ha a nézőpontok vagy a szubjektumok létezését elvetettük.)

Dennett és Kinsbourne négy kísérletet ír le, melyek arra engednek következtetni, hogy az embert érő ingerek időbeli tulajdonságai a tudatos élményben nem tükrözik híven az ingerek objektív időbeli struktúráját: így például a phi-jelenség esetében a kísérleti alanyt felváltva érő pillanatnyi ingerek például folytonos mozgásként tudatosulnak, ráadásul az észlelés késleltetett, a folytonos mozgássá való „átszerkesztése” utólag, visszamenőlegesen zajlik le. Libet híres kísérlete pedig megmutatta, hogy a szándékos cselekvés esetében a szándék csak a cselekvés megkezdése után tudatosul, ami megkérdőjelezi azt a karteziánus képet, mely szerint egy tudatos én kezdeményezi a cselekvéseket, miután a releváns információk birtokába jutott (Dennett – Kinsbourne 1992: 187).

A végkövetkeztetés az, hogy amikor az agy megkísérli rekonstruálni a külvilágból érkező ingerületek alapján az események objektív sorrendjét, akkor ez a sorrend nem függ közvetlenül attól, hogy milyen sorrendben érkeznek be az ingerületek az agyba, vagy annak valamely meghatározott részébe. Ehelyett az agyon belül egy utólag megszerkesztett történet jön létre, mely a későbbiekben tovább is változhat (Dennett – Kinsbourne 1992: 189).

Ha egy ilyen elbeszélés másképp ábrázolja az eseményeket, mint ahogy azok a valóságban történtek, azt kétféleképpen lehet elképzelni. A „hamisítás” vagy az aktuális átélés jelenében történik meg, és már így kerül rögzítésre, vagy az emlékezetbe vésett benyomások utólagos manipulációjára kerül sor. Egy sosemvolt bírósági tárgyalásról történő beszámoló az *orwelli*, egy gondosan megkoreografált ál-tárgyalásról szóló beszámoló a *sztálini* revízió esete. Noha az újságolvasó számára a két eset ekvivalens, és a dezinformáció egyformán sikeres lehet, a két módszer között képesek vagyunk fogalmilag éles különbséget tenni. E gondolat a „kartezianus materializmus kognitív illúziójának” forrása (Dennett – Kinsbourne 1992: 191).

Ha a szubjektum nézőpontja „térben és időben szétkent”, akkor bizonyos esetekben az „orwelli vagy sztálini?” kérdésre nincs egyértelmű válasz. (Dennett – Kinsbourne 1992: 192). A kartezianus materialisták számára azonban létezik egy térbeli-időbeli „nagy választóvonal”, (Dennett – Kinsbourne 1992: 194), ami előtt az ingerek még nem tudatosultak, és ami után igen.

A Dennették által vizsgált kísérleti jelenségeket, mint például a szín-phi jelenséget orwelli vagy sztálini mechanizmus alapján is lehetne értelmezni. A sztálini magyarázat szerint a kísérletben szereplő színérzetek a tudatosulásuk előtt szerkesztésre kerülnek, és a tudat elé már egy kész, koherens, de csalóka „film” kerül, és ez kerül az emlékezetben is elraktározásra. Az orwelli magyarázat szerint a tudatos észlelés minden pillanatban pontosan megfelel ugyan a pillanatnyi vizuális ingernek, azonban mire az észlelő alany eljut vizuális tapasztalatának verbális kifejezéséig, eltelik annyi idő, hogy ekkorra a koherens narratíva megteremtése érdekében az agy meghamisította az emlékeket. Vagy meg van hamisítva az észlelés már akkor, amikor tudatosul, vagy pedig akkor még nem, de az emlékezetbe hamis emlék rögzül. Mindkét magyarázat konzisztens azonban mindennel, amiről a kísérleti alany beszámol vagy amire emlékszik (Dennett – Kinsbourne 1992: 193).

Az orwelli és a sztálini elképzelés máshol húzza meg a „nagy választóvonalat”, aminek átlépésétől az függ, hogy tudatosulnak-e az információk. Ha viszont empirikusan semmiféle különbséget nem jelent, hogy hol van ez a választóvonal, akkor a kartezianus materializmus empirikusan alátámaszthatatlan. A szerzők konklúziója: „Az egymásra következők reprezentációja a tudatfolyamban az agy interpretatív folyamatainak terméke, nem az e folyamatokat felépítő események egymásra következésének közvetlen tükröződése... A tudatban lévő időbeli egymásra következők... tisztán a reprezentált tartalomtól függ, nem a reprezentálás időzítésétől” (Dennett – Kinsbourne 1992: 200).

A tanulmány, hasonlóan Dennett nem sokkal korábban megjelent, nagy sikerű könyvéhez (Dennett 1991), nagy vitát kavart, melynek során egyebek mellett Dennettnek a fenomenális tudatosság létezésének tagadását is felrötták.

Így például Block szerint Dennett és Kinsbourne érvelése azon az „elhallgatott és bizonyítatlan feltevésen nyugszik, hogy nincs olyasmi, mint fenomenális tudatosság” (Block 1992: 205). Block szerint éles szakadék választja szét azokat, akik elfogadják egy mindenféle kognitív, információfeldolgozási vagy funkcionális fogalomtól független tudatfogalmat, és azokat, akik efféle fogalmat nem fogadják el. A Block által említett éles szakadék két oldalán azok vannak, akiket mi antirevizionistáknak, illetve revizionistáknak nevezünk. A fenomenális tudatosság a Block által megkövetelt erős értelemben tehát valami nem funkcionális, nem kognitív és nem információfeldolgozással kapcsolatos értelemben vett tudatosság.

Kritikájában Block világosan kifejti antirevizionista meggyőződését. Szerinte „a szóban forgó tudatfogalom olyan leírásokat vonz, mint ‘nyers érzetek’, ‘közvetlen fenomenális minőségek’, és Nagel (1974/2004) ‘milyen’-je. Ez az a tudatfogalom, mely miatt ‘magyarázati rés’-ről beszélünk: a releváns tudományok mai állapotában fogalmunk sincs, hogy fájdalmam neurális szubsztrátuma hogyan tudná megmagyarázni, miért *ilyenek* és nem másmilyenek vagy semmilyenek sem érződik.”

Block szerint a fenomenális tudatfogalom ihleti a spektrum inverzióval, a hiányzó kválékkal vagy a zombikkal kapcsolatos gondolat kísérleteket, ám ezek a rejtélyek, teszi hozzá, csak „utak a fenomenális tudatossághoz – nem konstituálják azt. Elfogadhatjuk a fenomenális tudatosságot anélkül, hogy ezek bármelyikét elfogadnánk, mert a fenomenális tudatossághoz való alapvető hozzáférésünk a vele való ismeretségünkől származik.”

Block állítása az, hogy Dennett és Kinsbourne érvelése különösebb indoklás nélkül feltételezi a fenomenális tudatosság nemlétezését, miközben a karteziánus materializmus elleni érvelés önmagában nem érv a fenomenális tudatosság ellen. Dennett és Kinsbourne szerint az általuk vizsgált helyzetekben az „orwelli” vagy „sztálini” rekonstrukció mind az introspekcióban, mind kívülről, a megfigyelt viselkedés alapján megkülönböztethetetlen, ezért önkényes lenne az agybeli reprezentációk feldolgozásának valamelyik (és nem egy másik) stádiumát mint tudatos felcímkézni (Block 1992: 205). Block szerint azonban ha némelyik agyi reprezentáció fenomenálisan tudatos, némelyik pedig nem, akkor éppenséggel nem önkényes azt mondani, hogy bizonyos agyi események tudatosak. Block értelmezése szerint csak annak számára tűnik önkényesnek bizonyos reprezentációk tudatosként való megjelölése, aki eleve elfogadja Dennett és

Kinsbourne „kimondatlan irrealizmusát” a fenomenális tudatossággal kapcsolatban. Hiszen ha a fenomenális tudatosság valamilyen valódi tulajdonság, akkor nem lehet tetszés szerint „tulajdonítani” ennek vagy annak a reprezentációnak.

Aki elfogadja a fenomenális tudatosság létezését, annak nem kell elfogadnia a karteziánus materializmust, ezért maga a karteziánus materializmus elleni érvelés nem működik a fenomenális tudatosság irrealitását kimutató érvként, hanem Block szerint fordított a gondolatmenet: az orwelli és sztálini rekonstrukció csak akkor ekvivalens, ha eltekintünk a fenomenális tudatosság létezésétől.

Block ugyanis elgondolhatónak tartja azt a lehetőséget, hogy a tapasztaló szubjektum élménye, az, hogy „mit érez”, az orwelli és sztálini rekonstrukció szerinti esetben nem ugyanolyan. A szubjektum *ítéletei* tekintetében nincs különbség a kétféle eset között, de ezek nem merítik ki a szubjektum tapasztalatát azok szerint, akik hisznek a fenomenális tudatosságban. A Dennett és Kinsbourne által vizsgált esetek Block szerint legfeljebb annyit bizonyítanak, hogy az orwelli és sztálini rekonstrukció között nem tudunk dönteni introspekció alapján, mivel élményeinkhez ugyan privilegizált hozzáférésünk van, ám nem vagyunk tévedhetetlenek a róluk szóló emlékeinkkel kapcsolatban, amit nyugodtan elismerhet az is, aki hisz a fenomenális tudatosságban. (Blocknak ez a megjegyzése egyébként rámutat arra, hogy Dennett és Kinsbourne tézisei anti-karteziánusak abban az értelemben is, hogy ellentmondanak annak a saját élményünkkel kapcsolatos tévedhetetlenség karteziánus tételének.)

A sztálini és orwelli eset ráadásul Block szerint elvileg kívülről nézve is megkülönböztethető, hiszen az orwelli esetben a memóriában olyan utólagos változás megy végbe, amilyen a sztáliniban nem, amit pedig az agy vizsgálatával elvileg ki lehetne mutatni.

Blockhoz hasonlóan Antony (Antony 1992: 201) is úgy értelmezi, hogy Dennett és Kinsbourne álláspontja a tudattal szembeni eliminativizmus: a tudat e szerint nem létezik. Antony egyébként Blockhoz hasonlóan a tudat funkcionalista elméleteinek kritikusa. (Lásd: Antony 1994). Antony szerint háromféleképpen értelmezhető Dennett és Kinsbourne álláspontja. Ha állításuk az, hogy a tudatos élményeknek nincs időbeli helyük, akkor ez materialista előfeltevések mellett azt jelenti, hogy tudat nem létezik (hiszen a materialisták szerint minden, ami létezik, térben létezik). Ha állításuk az, hogy a tudatos élménynek nincs *pontos* időbeli helye, ez az az állítás, hogy a tudatos állapotok időben elmosódtak, körülhatárolásuk nem lehet tetszőlegesen pontos. Ennek az értelmezésnek azonban az a gyengesége, hogy ez az állítás triviálisan igaz, hiszen minden esemény időbeli határai elmosódtak kellően finom skálán. A harmadik lehetséges értelmezés Antony szerint az, hogy noha a tudatos élmények az időben lokalizálhatók, a neurofiziológiai események

meghatározásának időskálája túl finom ahhoz, hogy ezek alapján nem-önkényes időbeli határvonalakat lehessen megállapítani a tudati események számára. Ez akkor lehetséges, ha a tudatos események nem azonosíthatók funkcionális állapotokkal vagy folyamatokkal (hiszen azok meglehetősen pontosan körülhatárolhatók időben). Ha viszont elfogadjuk ennek azt a következményét, hogy a tudatos élménynek nincs szerepe az agy funkcionális szerveződésében, annak folyományaként azt is el kell fogadnunk, hogy – mivel a tudatos élmények ezek szerint nincsenek kauzális kapcsolatban az agy funkcionális állapotaival – a tudatos élményekről nem tudunk beszámolni, a tudatos élményeket nem is idézhetik elő érzéki ingerek stb. Ebben az esetben nehéz felfogni, egyáltalán hogyan tudunk gondolkodni a tudatos élményről (hiszen gondolataink nincsenek vele kauzális kapcsolatban). E megfontolások miatt a harmadik olvasat Antony szerint abszurd következményekkel jár. A három értelmezés közül csak az nem triviális vagy inkohérens, amely szerint Dennett és Kinsbourne eliminativisták a tudattal kapcsolatban (Antony 1992: 202). Cikkükben azonban Dennett és Kinsbourne nem szállnak síkra nyíltan az eliminativizmus mellett, és időnként úgy használják a „tudat” szót, mintha mégiscsak elfogadnák a tudat létezését. Tagadja-e Dennett és Kinsbourne a tudatot, csak éppen tagadja azt, hogy tagadja – mint ahogy az Strawson szerint sokszor előfordul – vagy sem?

Aronson, Dietrich és Way szerint (1992: 202-203) fenomenológiailag a tudat egységesnek *tűnik*, és tudatfolyamunk „egyetlen narratíva”. Ahhoz, hogy megmagyarázzuk, miért tűnik egységesnek, mégiscsak feltételezni kell valamilyen egységesítő folyamatot, ha viszont erre válaszul Dennették tagadják, hogy a tudat egységes, akkor a tudatot „nem megmagyarázták, hanem kidobták”. Az ilyen típusú érvelés problémája az, hogy egyrészt a tudat egységének tagadása nem azonos a tudat létezésének tagadásával. Még fontosabb azonban, hogy nem egészen világos, hogy a tudat valóban *egységesnek tűnik-e*.

Fontoljuk meg a következő ellenvetést. Aronsonék azzal a példával támasztják alá, hogy a tudat „egyetlen narratíva”, hogy amikor valaki fejben hangosan gondolkodik, mint egy Joyce-regényhős, akkor egyetlen szöveget mond magában (nem például többet szimultán). Találhatnának kevésbé meggyőző példát is. Aki egész nap cipőt visel, az nem érzi minden pillanatban, hogy cipő van a lábán (kivéve talán, ha a cipő annyira szűk). Ha azonban az illető a figyelmét arra irányítja, hogy mit érez a lábával, esetleg mozgatja is a lábát, akkor hamarosan érzékeli, hogy nincs mezítláb. Amikor éppen nem tudatosul benne, hogy cipő van a lábán, vagyis a pillanatnyi élményének, „hogylétének” nem része a lábát érő nyomás, feltételezhetően akkor sem az a helyzet, hogy átmenetileg megszűnt az érzékelés a lábában, csupán az, hogy a szóban forgó érzékelési információ nem válik annyira „centrálissá”, hogy részét képezze a szubjektum „önmagáról szóló

narratívájában”. Ha megkérdeznék az illetőt, hogy hogyan érzi magát, akkor például beszámolna arról, hogy éhes vagy unatkozik, esetleg elfoglalt, de azt, hogy cipő van a lábán, csak speciális esetekben hozná szóba (például a szorító cipő esetében). Ennek a közismert ténynek az ortodox karteziánus értelmezése az lenne, hogy a karteziánus színházban nem mindig tűzik műsorra a cipő által a lábra gyakorolt nyomás által felkeltett érzetet, csak időnként (ami ahhoz az érdekes kérdéshez vezet, hogy ki és hol tevékenykedik az a rendező, aki eldönti, színpadra léphet-e ez az érzet). Egy továbbfejlesztett karteziánus elképzelés elismerheti, hogy a cipő által okozott érzet a tudatban van (vagyis a karteziánus színpadon), de nem mindig válik centrálissá (vagyis a beszámoló és visszaemlékezés szempontjából hozzáférhetővé), vagyis nem irányul rá a reflektor fényköre. Aronsonék mintha úgy gondolnák, hogy ha egyetlen ilyen fénykör van, akkor úgy tűnhet, mintha csak az ez által megvilágított szereplők lennének a színpadon, de ha így is tűnik, abból nem következik, hogy a fénykörön kívül nincs semmi. Ennek a nézetnek a Dennett számára minden bizonnyal szimpatikusabb alternatívája az, hogy nincs fénykör (mint ahogy karteziánus színház sincs), és a tudat látszólagos egysége abból ered, hogy verbális beszámolási, történetmesélési képességeink olyanok, illetve emlékezetünk olyan módon működik, hogy azok a potenciálisan tudatosulni képes különféle agyi eseményeket például csak meghatározott sorrendben tudják történetekbe fűzni, verbálisan kifejezni vagy az emlékezetben rögzíteni.

Elhagyva a képes beszédet, amikor valaki hangosan beszél, akkor a beszéd (és beszédszervek) sajátosságai miatt egyszerre csak egy szöveget tud mondani, és talán nem lehetetlen képzeletünkben egyszerre két hangon megszólalni, a gyakorlatban ezt nehéz kivitelezni és nincs praktikus haszna. Hasonlóan, kezeink és szemeink száma (és egymástól mért távolsága) korlátozott, aminek – és egyéb korlátaink – következtében praktikus korlátai vannak annak, hogy egyszerre hányféle dologgal foglalkozhatunk (még ha a történelem fel is jegyzett olyan rendkívüli személyeket, akik két kezükkel két különböző szöveget tudtak egyszerre írni – de olyan személy például biztosan nincs, aki három kezével három szöveget tud egyszerre írni).

Akármilyen szeszélyesen is viselkedjünk, csak egyetlen testünk van, mely egyszerre csak egy helyen tud tartózkodni, egyszerre csak egy szöveget tud elmondani, egyszerre csak egy látószögéből látja a világot. Magától értetődő, hogy egy ilyen lény valamilyen bizonyos mértékig egységes narratívát fog mondani magáról.

Clark arra mutat rá, hogy lehetségesek tények akkor is, ha nincs róluk utólagos verbális beszámoló. Gyakorlatilag kivitelezhetetlen, hogy egy fél másodpercig tartó eseménysor kellős közepén több kérdést is feltegyünk valakinek, amire válaszol is. Ha el akarjuk dönteni, hogy az orwelli vagy a sztálini modell-e a helyes, akkor csak utólagos verbális beszámolókra

hagyatkozhatunk, vagy legfeljebb a kísérlet közben adott egy-egy (megkésett) reakcióra, de ez nem jelenti azt, hogy nem lehet valójában igaz az egyik vagy másik modell. Csak dönteni nem lehet közöttük. Ahhoz, hogy ez kizárható legyen, tagadni kell azt, hogy a tudatos tapasztalattal kapcsolatban lehetnek egyáltalán tények, tehát tagadni kell a tudatosságot. Csakhogy Dennették sehol sem érvelnek explicit módon egy ilyen pozíció mellett (Clark 1992: 207-208).

Hasonlóan, Glymour et al. (1992: 209-210) és Van Gulick (1992: 229) felveti annak lehetőségét, hogy a verbális beszámolók korlátai átléphetők neurobiológiai megfigyelések által, tehát ha sikerülne megismernünk a tudat neurális mechanizmusait, akkor rendelkezhetnénk a tudatosság pontosabb, megbízhatóbb tesztjeivel. Egy, a tudattal kapcsolatban realista és materialista kutató számára nem okoz gondot elképzelni, hogy létezhetnek tudati tények a róluk való verbális beszámolóktól függetlenül is.

Farah szerint Dennett és Kinsbourne tanulmánya egyáltalán nem szól a tudatosságról. (Farah 1992: 209), hiszen elemzésük akár egy számítógép tudattalan reprezentációjáról is szólhatna. Miközben más kritikusok azt róják fel Dennettéknek, hogy tagadják a tudat létezését (és ezt nem támasztják kellőképpen alá), Farah a másik fajta tipikus értelmezést alkalmazza: az ő számára Dennették elbeszélnek a tárgyuk mellett. Ugyanígy látja Lloyd (1992: 215-216), hozzátéve, hogy ha úgy értelmezzük Dennett munkásságát, hogy az antirealista a tudattal kapcsolatban, akkor azzal nem érthetünk egyet, mert „a tudatos tapasztalat realitása tagadhatatlan”. Ezért Dennett és Kinsbourne eredményeit úgy kell hasznosítani, hogy „a tudat azonosságelmélete” (mely szerint „a tudatos feldolgozás azonos a kognitív feldolgozással”) alapján interpretálandó az az eredményük, hogy nem vonható éles határvonal a tudatos és nem tudatos kogníció között. Van Gulick szintén azt állapítja meg, hogy Dennették elmémmodellje elhanyagolja a fenomenális és nem-fenomenális különbségét, ezt azonban csak akkor teheti meg, ha e különbségtétel lehetőségét elutasítja, emellett azonban nem érvel (Van Gulick 1992: 229). A szerkesztői kommentár (Dennett & Kinsbourne 1992: 233) ugyancsak arra mutat rá, hogy az „objektív és szubjektív minőségek összemérhetetlenségét” demonstráló tények nem az utóbbiak nemlétezését vagy meghatározatlanságát bizonyítják.

Kritikusainak adott válaszában Dennett és Kinsbourne kijelenti, hogy álláspontjuk „realista és materialista”: a tudatos események „valós események”, amelyekre valós időben és térben kerül sor, azonban bizonyos kérdések nem tehetők fel róluk, mert ehhez nem elég élesek az időbeli és térbeli határaik, pontosan úgy, ahogyan a Brit Birodalomról is fel lehet tenni olyan kérdést, amire nincs válasz. Mivel egy személy nem több, mint különféle idegrendszeri folyamatok összessége, ezért az a kérdés, hogy „mikor szereztem tudomást én (és nem pedig agyam valamelyik része) valamilyen eseményről?”, Dennették szerint „csapda”, mert a tudatos tapasztalatot „egymást

követő, agyban történő állapotok *konstituálják*, és nem olyasmi, ami más vagy több, mint ezek a folyamatok, melyek *okoznák*”. (Dennett – Kinsbourne 1992: 236)

Ebben a pillanatban Dennették a tudatos tapasztalat folyamáról beszél, melyet különböző állapotok konstituálnak, bár arra a lehetőségre nem térnek ki, hogy ezek az állapotok esetleg tudatos állapotok lennének. Miért nem? A következő módon magyarázzák a tudattal kapcsolatos „realizmusukat”:

A tudattal kapcsolatos „realistaként” abban hiszünk, hogy kell lennie valaminek – a K tulajdonságnak – mely megkülönbözteti a tudatos eseményeket a nem-tudatos eseményektől. Vegyük fontolóra a következő jelöltet K tulajdonság szerepére: egy tartalommal bíró esemény akkor válik tudatossá, ha és amikor része lesz egy *átmenetileg domináns aktivitásnak* az agykéregben [...] K tulajdonságnak ilyen definíciója kielégíti a „realizmus” minimális követelményeit, de fenyegeti az orwelli és sztálini revíziók közötti feltételezett különbségtételt. (Dennett – Kinsbourne 1992: 236)

Az egyik kulcsmozzanat ebben a magyarázatban, hogy a tudatosuló és nem tudatosuló elemek tulajdonságai között olyan különbségek vannak (például hogy meddig volt része a domináns aktivitásnak, de akár arra is gondolhatunk, hogy mit jelent a „domináns aktivitás”), amelyek inkább fokozati különbségek, mint éles kategoriális különbségek. A másik, hogy Dennett számára a „realizmus” elég, ha kielégíti annak „minimális követelményeit”, úgy, ahogy realisták lehetünk a Brit Birodalommal kapcsolatban is annak ellenére, hogy nem képzeljük, hogy a Brit Birodalom több vagy más, mint az általa uralt területek az azokon élő emberekkel stb. együtt.

Akik Dennettet a tudattal kapcsolatos „eliminativizmussal” vádolják, nyilvánvalóan nem elégszenek meg az így értett realizmussal. Számukra azonban nyitva áll az a lehetőség, hogy Dennettet úgy olvassák, hogy ő a tudaton valami mást ért, mint ők (talán olyasmit, mint Ned Block a hozzáférési tudatosságon), és az általa hivatkozott kísérleti eredmények, illetve érvek nem zárják ki a tudatnak egy olyan felfogását, mely szerint egy másik értelemben tudatos események (például: „fenomenális karakterrel” vagy „kvalitatív jellemzőkkel” bíró események) konstituálják tudatfolyamunkat – azok a fenomenális vagy kvalitatív karakterrel rendelkező események, melyek bekerültek a „tudatos események elit klubjába” (Dennett – Kinsbourne 1992: 238). Úgy tűnik, hogy a választóvonal ott van a dennetti „minimális” realizmus és a Dennett kritikusai által megkövetelt

szubsztantívabb realizmus között, hogy az utóbbiak megkövetelik, hogy a tudatosság valami konkrétabb létező legyen, mint amilyen a Brit Birodalom.

Lloydnak arra a kritikájára válaszolva, hogy elméletük nem a tudatról szól, Dennették kifejtik:

„Inkább azt állítjuk, hogy nincs frappáns módja annak, hogy megállapítsuk, pontosan melyek tudatosak a párhuzamos áramlatok (streams) közül. Az áramlatok mindegyike néha hozzájárul a tudathoz, máskor pedig nem. Egyik áramlat sem szükségképpen tudatos saját természetéből adódóan. És ez a mi tudatelméletünk. Rossz hír a „realizmusnak”, ha csak az olyan jelenségek tekinthetők valósnak, amelyeknek élesen körülhatárolt szükséges és elégséges feltételei vannak. [...] A tudat úgy tűnik, mint valami, ami vagy van, vagy nincs, de [...] ez nem jogosít fel arra a konklúzióra, hogy a tudat vagy van, vagy nincs” (Dennett – Kinsbourne 1992: 239)

Aki szerint a tudatosság olyan tulajdonságokkal azonos, vagy olyan tulajdonságoktól függ (például „kvalitatív” vagy „fenomenális” tulajdonságoktól), melyek a teljes tudatfolyamot alkotó egyes kisebb eseményeket is megillettek, az elfogadhatja, hogy a tudatosság ezen események vonatkozásában „vagy van, vagy nincs” még akkor is, ha nem tudjuk pontosan megmondani, hogy ezek az események *elégé centrálissá* válnak-e egy adott pillanatban az agyban, hogy hozzáférhessük őket az egyén élettörténetéhez, ami Dennett számára az igazán fontos.

Így aztán miközben Dennett tesz egy határozott állítást, mely szerint a tudatosság nem vagy van, vagy nincs, és nem a tudatfolyamot alkotó egyes események intrinzikus tulajdonsága, amellel nem érvel közvetlenül, hogy ezek az események nem is rendelkezhetnek olyan tulajdonságokkal, például „fenomenális” vagy „kvalitatív” jellemzőkkel, vagy „protofenomenális”, „protokvalitatív” jellemzőkkel, amelyek döntőek lehetnek a tudatos élmény meghatározottsága szempontjából. Dennett antirevizionista olvasói, akik elégedetlenek az ő „realizmusával”, úgy olvashatják Dennettet, mint aki egyszerűen nem foglalkozik a(z érdekesebb értelemben vett) tudatossággal. Block terminológiájával élve, Dennett nem foglalkozik a fenomenális tudatossággal, csak a hozzáféréssel, és éppenséggel azt sem bizonyítja be, hogy az előbbi nem létezik – csupán azt, hogy logikai tulajdonságai mások, mint az utóbbiéi.

Dennett számára azonban ebben az esetben nincs nagy különbség aközött, hogy valamivel nem foglalkozik, és aközött, hogy tagadja a létezését – Dennett ugyanis tipikus példa a revizionista filozófusra. Block kritikájára reagálva, mely szerint Dennett különösebb érvelés nélkül utasítja el a fenomenális tudatosság létezését, Dennett (1992: 239-240) rámutat, hogy Block nem tud komoly bizonyítékot felhozni a fenomenális tudatosság létezése mellett.

Block szerint nyilvánvaló, hogy van fenomenális tudatosság, még akkor is, ha nincs rá olyan bizonyíték, mint amit általában bizonyítéknak fogadunk el a tudományban, míg Dennett szerint nincs okunk feltételezni, hogy létezik, hiszen nincs rá bizonyíték. A két filozófus számára más a „nyilvánvaló” kiindulópont, amit úgy jellemeztünk, hogy egyikük (Dennett) revizionista, másikuk (Block) antirevizionista. A revizionista Dennettnek bizonyítékokra lenne szüksége ahhoz, hogy elfogadja a fenomenális tudatosság létezését, míg az antirevizionista Blockot csak komoly érvekkel lehetne meggyőzni arról, hogy a fenomenális tudatosság „nyilvánvaló” létezésébe vetett hite téves.

Dennett mindazonáltal inkább ott látja a fő választóvonalat az ő és a vele egyet nem értők között, hogy vitapartnerei elfogadják ama „kartezianus örökségből” eredő feltételezést, hogy a tudatosság *intrinzikus* tulajdonsága a tudatos mentális állapotoknak. Dennett ezt a tézist utasítja el. (Dennett – Kinsbourne 1992: 243)

Dennett álláspontja tehát a tudattal kapcsolatban realistának nevezhető, ha nem követeljük meg a realizmustól, hogy az általa reálisnak tartott tulajdonságok intrinzikusak legyenek. És miért követelnénk meg ilyesmit? – kérdeznék a revizionisták. Dennett tehát úgy gondolja, hogy realista lehet a tudattal kapcsolatban ugyanabban az értelemben, ahogy a Brit Birodalommal kapcsolatban is realisták lehetünk bizarr politikai metafizikai előfeltevések nélkül is.

Dennett ellenfelei azonban kitartanak amellett, hogy Dennett „eliminativista”, vagy legalábbis realizmusa nem arról szól, amit a tudaton érteni kellene. Úgy tűnik, az antirevizionisták számára a tudatosság csakugyan intrinzikus tulajdonsága mentális eseményeknek, illetve tapasztaló szubjektumoknak. Az, hogy *milyen* zöld színt látni vagy rózsaillatot érezni, az antirevizionista intuíció szerint nem megállapodás kérdése, nem egy elméleti elemzés valamelyik szintjén jelentkező eredmény, hanem *konkrét valóság*. Ez az alapvető intuíció gyakran szerepet játszik az elmefilozófiában, például a funkcionalizmus elleni érvelésekben (Block 1978), hiszen ha a tudatos állapotok funkcionális állapotoknak feleltethetők meg, akkor problémát okoz, hogy egy-egy rendszernek többféle különböző funkcionális interpretációja is lehetséges, márpedig az, hogy „milyen” egy bizonyos élményt megtapasztalni, nem tűnik a minket megfigyelő tudósok értelmezésétől függő ténynek. (Ennek megfelelően R-olvasatban a tudat funkcionalista elméletét a

fenomenális tudatosság elvetéseként értelmezhetjük, míg A-olvasatban akár a számos egymást átfedő, eltérő minőségű tudatos élményt is megengedő bizarr elméletté válik, ha csak ezt nem korlátozzák olyan kritériumok, melyek az elméletet nonredukcionistaává teszik.)

De okvetlenül karteziánus-e az, aki nem revizionista? Karteziánusok-e azok a reduktív materialisták, vagy éppen a neutrális monisták vagy a pánpszichisták, akik mindannyian elismerhetik, hogy a fenomenális tudatosság lehet egyes mentális események intrinzikus jellemzője? Dennett pozíciójának egyik értelmezése szerint Dennett a tudatosságon azt érti, mint Block a hozzáférési tudatosságon, vagyis számára a tudatosság az a tulajdonság, amely azokat a mentális eseményeket illeti meg, melyek egy bizonyos időpontban kellően prominens szerepre tettek szert az elme működése során („bejutottak az elit körbe” Dennett kifejezésével). Természetese ha erről a tudatosságról gondoljuk, hogy az egyes mentális események intrinzikus tulajdonságának kell lennie, az a mentális események „foszforeszcenciájának” karteziánus elképzelését idézi (Ryle 1949/1999: 199-200). Másfelől, ha az így értett tudatosságtól megkülönböztetjük a fenomenális tudatosságot (ahogyan ezt Block is teszi), akkor egy másik doktrínához jutunk, aminek karteziánus jellege legalábbis nem nyilvánvaló.

2.2.2 Dennett: Consciousness Explained

Dennett a *Consciousness Explained* című művében bevezeti a heterofenomenológia módszerét: a heterofenomenológus fikciós világként (mint amilyen egy irodalmi mű világa) kezeli a vizsgált személyek fenomenális tapasztalatának világát, semlegesnek maradva abban a kérdésben, hogy valóban léteznek-e azok az entitások, amelyekről a vizsgált alanyok beszámolnak. Mások heterofenomenológiai világának ismerete lehetővé teszi az első személyi beszámolóik értelmezését, és esetleg olykor a viselkedésük megértését is. Dennett másik hasonlata a heterofenomenológust az antropológushoz hasonlítja, aki egy őserdei törzs vallását vizsgálja, igyekezve elsajátítani a törzs hiedelmeit, ám anélkül, hogy hinne benne (Dennett 1991: 81-82). Később lehetségessé válhat a heterofenomenológiai világot benépesítő entitások azonosítása agyi történésekkel, ami annak felfedezésével egyenértékű, hogy miről beszélnek az emberek *valójában*, amikor tudati eseményeikről számolnak be (Dennett 1991: 85). Ez az álláspont úgy hangzik (egészen a primitív törzs vallási hiedelmeire vonatkozó hasonlatig bezárólag), mint Rorty (1965) „eltűnési típusú” azonosságelmélete, amelyet az eliminativizmus korai formájának vagy előfutárának (Tözsér 2008: 67) szokás tekinteni. De vajon tényleg a „tudat tagadói” közé kerül Dennett az e könyvben kifejtett nézeteivel?

A zombiproblémáról például azt mondja, hogy a zombiknak nyugodtan tulajdoníthatunk heterofenomenológiai világot (Dennett 1991: 95): a heterofenomenológiai objektumok azonban absztrakt entitások, „mint az Egyenlítő vagy a tömegközéppontok”.

Dennett számít arra az ellenvetésre a heterofenomenológiai objektumok elméleti fikciós státuszával kapcsolatban, hogy az emberek várhatóan ragaszkodnak hozzá, hogy autofenomenológiai objektumaik viszont valóságosak, nem fiktívek. Erre Dennett válasza az, hogy empirikus kérdés, hogy van-e az agyban olyan entitás vagy folyamat, amely valamilyen módon megfeleltethető ezeknek az objektumoknak (Dennett 1991: 98), ami nem hangzik egyértelműen eliminativista megállapításnak, hiszen ezt egy, a fenomenális tudatossággal kapcsolatban realista materialista is mondhatná.

Másfelől Dennett elmélete a tudatról Dennett 1991-ben az, hogy minden olyan rendszer tudatos „a szó legteljesebb értelmében”, amelynek „irányító rendszere [joyce-i] virtuális gép”. Ez egy tisztán funkcionalista elméletnek hangzik. Arra a kézenfekvő ellenvetésre, hogy igen könnyűnek tűnik elgondolni azt, hogy egy gép megfelel bizonyos funkcionális leírásnak, és mégsem tudatos, Dennettnek két válasza van. Az egyik, hogy kételkedik abban, hogy valóban könnyű lenne elképzelni egy nem tudatos, joyce-i gép által vezérelt lényt vagy robotot. A másik pedig, hogy egy elképzelt vitalistához hasonlítja vitapartnerét, aki nem akarja elfogadni, hogy valami, ami úgy néz ki és viselkedik, és olyan anyagból van, mint egy macska, csak ettől már *él* is, nemcsak egy szerves robot (Dennett 1991: 281). A vitalista álláspontja meglehetősen abszurd, és éppen ilyen abszurd Dennett szerint azé is, aki nem funkcionalista, ez pedig azt mutatja, hogy Dennett számára a tudatosság teljes egészében funkcionális fogalom, és semmiképpen sem valami olyasmi, ami Chalmers kifejezésével „túl van struktúráján és funkción”.

Dennett szerint a funkcionalizmusban való kételkedés annak a karteziánus érvek a leszármazottja, amely a test és a lélek különbségére következtet azok világos és egyértelmű különálló megkülönböztetőségéből. Dennett azonban egyáltalán nem tesz különbséget kétféle „zombi-hipotézis” között: aközött, amiben egy élő ember pontos fizikai másolatáról vélik elgondolhatónak, hogy nem tudatos lény, és aközött, amiben egy élő ember funkcionális másolatáról (például Block példájában egymással rádióon kapcsolatot tartó kínaiak által implementálva). Az, hogy számára nem jelent nehézséget egyenlőségjelet tenni a kétféle hipotézis közé, azt mutatja, hogy számára a tudatosság nyilvánvalóan funkcionális tulajdonság, ami az R-intuícióban osztozók számára nem problematikus állítás, azonban mások számára cseppet sem nyilvánvaló igazság.

A *Consciousness Explained*ben Dennett azt állítja, hogy nincsenek kválék, noha úgy tűnik, hogy vannak:

Úgy tűnik, hogy vannak kválék, mert tényleg úgy tűnik, mintha a tudomány bebizonyította volna, hogy a színek nem lehetnek odakint, tehát idebent kell lenniük. Továbbá úgy tűnik, hogy ami idebent van, az nem lehet csupán mindaz az ítélet, amit akkor teszünk, amikor a dolgok színesnek tűnnek számunkra. (Dennett 1991: 372)

A következőkben Dennett kifejti a színérzékeléssel kapcsolatos koncepcióját, melynek eredménye, hogy a másodlagos minőségek, mint a szín, jól érthető okokból nehezen definiálhatók. A gondolatmenet egyik érdekes pontján Dennett világossá teszi:

A CADBLIND Mark I [egy kamerával és manipulátorral felszerelt robot] biztosan nem rendelkezik semmiféle kváléval (legalábbis várakozásom szerint a kválék barátai ezen a ponton így vélik), ezért következik az összehasonlításomból, hogy azt állítom, hogy nekünk sincsenek kváléink. Az a fajta különbség, amit az emberek a bármilyen gép és emberi tapasztaló között képzelnek el[...] az, amit határozottan tagadok: nincs efféle különbség. Csak úgy tűnik, mintha lenne. (Dennett 1991: 375)

Meg kell jegyeznünk, hogy a gép és az ember látása közötti, a közvélemény (egy része) által feltételezett radikális különbség többféleképpen is nem létezőnek bizonyulhat: R-olvasatban úgy, hogy egyik sem rendelkezik fenomenális látásélménnyel, A-olvasatban pedig úgy, hogy mindkettő rendelkezik azzal. Dennett természetesen nem siklik el e probléma mellett, és úgy tűnik, megkísérel közvetlenül is választ adni a felmerülő kérdésekre: erre látszik utalni, hogy a következő fejezetének az „Élményeink élvezése” címet adja. (Az „élvezet” Samuel Alexander óta az a terminus, amivel az angol nyelvű irodalomban a (tapasztalattal szembeállított) élményt igyekeznek kifejezni, mivel az angol nyelv híján van az élmény/tapasztalat (Erlebnis/Erfahrung) különbségtételnek).

„De Dan, a kválék azok, amik miatt érdemes élni!”, idézi Dennett Wilfrid Sellars egy megjegyzését (Dennett 1991: 383). Úgy tűnik, valóban eljutottunk arra a pontra, amikor Dennett közvetlenül megvizsgálja azt a kérdést, hogy létezik-e fenomenális tudatosság, bár az előjelek nem

kedvezőek, hiszen a kérdés nem az, hogy miért látszik a rózsaszín olyannak, amilyenek látszik (mint azt Dennett elképzelt vitapartnere, Otto kérdezi), hanem az, hogy van-e egyáltalán a tapasztalatnak fenomenális minősége, intrinzikus „valamilyen”-sége, élményjellege, vagy sem.

Az élvezetről Dennett tisztán funkcionális-behaviorisztikus beszámolót nyújt. Amikor finom bort iszunk, akkor amit akarunk, az az, hogy „kedvenc módunkon informálódjunk annak kémiai tartalmáról” (Dennett 1991: 384). Ez kétségkívül igaz, noha nem visz sokkal közelebb annak eldöntéséhez, hogy Dennett mit gondol az élmény fenomenális minőségének a bor élvezetében betöltött szerepéről. A probléma az, hogy Dennett elismeri, hogy nem pusztán „informálódunk a bor kémiai összetételéről”, amikor bort kóstolunk, hanem „kedvenc módunkon informálódunk” róla, kedvenc módunk viszont az, hogy egészen konkrétan megkóstoljuk a bort egészen konkrét nyelvünkkel, és noha erről a folyamatról adható funkcionalista beszámoló, az antirevizionisták viszont könnyen úgy értelmezhetik, hogy akár a fenomenális élményminőségnek is szerepe lehet benne, mely akár azonos is lehet a bor és az ízlelőszervek kölcsönhatásának bizonyos fizikai jellemzőivel. Más szóval, a „kedvenc módunkon informálódunk” mintha nyitva hagyná azt a kiskaput, hogy az elméletet máshogy is érthetjük, mint egy szigorúan funkcionalista, egyben a fenomenálissal szemben eliminativista elméletet.

Később Dennett megkísérli az álláspontját egyértelműbbé tenni, amikor azt fejtegeti, hogy a kválének az élvezet okaként való felfogása *virtus dormitiva* típusú magyarázat. Egy *virtus dormitiva* típusú magyarázat persze ideiglenesen elfogadható, mint ami kijelöl egy problémát. Később egy megfelelő mechanikus, reduktív magyarázat válthatja fel ezt az üres magyarázatot.

Ennek megfelelően kezdhethetjük úgy, hogy a kválé definíció szerint az élvezetünk és szenvedésünk közvetlen oka, majd pedig ezt idővel felváltjuk egy valódi tudományos magyarázattal. Tehát az élvezet és szenvedés kap egy (például evolúciós biológiai) magyarázatot, és ez szükségtelessé teszi a kválékra hivatkozást. A „kvalofilek” (azok, akik „még mindig hisznek a kválékban”) azonban, állapítja meg Dennett, ezt elutasítják: „ragaszkodnak ahhoz... hogy azok a kválék, melyeket pusztán mechanikusan megvalósított diszpozíciók komplexumaivá „redukáltunk”, nem azok a kválék, melyekről ők beszélnek” (Dennett 1991: 386).

A „kvalofilek” azért nem érik be Dennett evolúciós magyarázataival, mert számukra az élvezet tárgya, vagy maga az élvezet valamilyen okkurrens tulajdonság (nem pedig diszpozicionális), illetve mert a kválét az élvezet *tárgyának* tekintik és nem az *okának*. (Az élvezet szót ebben a mondatban az alexanderi értelemben értem. Ha az élvezet a szó hétköznapi értelmében értendő ezekben a mondatokban, tehát mint egy tapasztalatra adott értékelő reakció, akkor

természetesen a kvalofil számára a kválé az élvezet oka lesz, de akkor nem világos, miben is nem ért egyet a kvalofillel azon kívül, hogy a kvalofil szerint az élvezet kiváltó okának van valamilyen intrinzikus természete, míg Dennett szerint talán nincs.) Röviden, például egy materialista kvalofil azt mondhatja, hogy amikor S fizikai állapotba kerül a tapasztaló szubjektum, akkor „valamilyen” neki lenni, vagyis valamilyen fenomenális állapotot „él át” vagy műszóval „élvez”. Az S fizikai állapotba valamilyen M mechanizmus működése eredményeként kerül. Dennett szerint ha M-et megmagyaráztuk, azzal megmagyaráztuk a kválé természetét, a kvalofil szerint viszont S állapot átélése jelenti azt, hogy a tapasztaló szubjektum rendelkezik a kérdéses kváléval.

Amikor a kvalofil azt állítja, hogy akármennyit is tud a publikusan hozzáférhető és közölhető információkból arra vonatkozóan, hogy milyen lehetett (Dennett példájában) 1725-ben lipcei lakosként Bachot hallgatni, marad valami kimondhatatlan maradék, ami ilyen módon nem ismerhető meg, vagyis bizonyos értelemben „sohasem tudjuk meg, pontosan milyen lehetett ez”, akkor Dennett szerint téved, mert ha valóban elégséges információval rendelkezünk arról, milyen lehetett 1725-ben lipceinek lenni, akkor nem maradna semmi rejtély, el tudnánk képzelni, milyen lehet az. Dennett és az ő optimizmusában kételkedő kvalofil vitája azonban másként is értelmezhető.

Ezen értelmezés szerint amikor a kvalofil azt mondja, hogy „sohasem tudjuk meg, igazából milyen 1725-ben lipceinek lenni”, akkor igazából arról beszél, hogy akármennyit is olvasunk erről a kérdéstről, vagy akármennyit is erőltetjük a fantáziánkat, konkrétan nem éljük át sohasem, hogy lipceiek vagyunk 1725-ben, abból az egyszerű okból, hogy nem vagyunk lipceiek 1725-ben. A kvalofil ezt fenntarthatja, miközben elfogadhatja azt, hogy végül is nincs elvi akadály annak, hogy különféle gyakorlatok hatására (melyben képzelőereje gyakorlása, történelmi olvasmányok és más szokványos módszerek játszhatnak szerepet, vagy egészen más megoldások, például „virtuális valóság”-gépezetek vagy más sci-fibe illő technikai újítások) pontosan olyan állapotba kerüljön, mint az egy lipcei polgár 1725-ben. A kvalofil fenntarthatja, hogy pusztán előadások, olvasmányok vagy a képzelete révén sohasem kerülhet ebben az állapotba, vagyis az „első kézből való tudás” idiomatikus kifejezés, ami a közvetlen *átélést* jelenti, olyasmit jelölhet, ami világosan megkülönböztethető a hétköznapi értelemben vett, „másodkézből” való tudástól (amit „leckék” révén sajátítunk el: Lewis 1988/2008: 351). Ez nem jelenti, hogy el kell utasítania a tudományos magyarázatokat, vagy bármiféle obskurantizmusba kellene menekülnie. A különbség Dennett és a kvalofil álláspontja között ekkor az, hogy míg Dennett számára a kválé „diszpozíciók komplexuma” (Dennett 1991: 389), addig a kvalofil szerint a közvetlenül átélt okkurrens állapot tulajdonsága vagy része.

A zombi-hipotézisekkel és az általuk alátámasztani szándékozott dualista elméletekkel kapcsolatban Dennett kifejti:

„noha talán van valamilyen mód arra, hogy megkülönböztessük a zombikat a valódi emberektől (ki tudja, talán a zombiknak zöld az agya), a különbség nem jelentkezik a megfigyelők számára funkcionális különbségként. [...] [Ez] veszélyesen idétlen, mert azokat a fajta teljesen motiválatlan előítéleteket idézi, melyek elvitatták a teljes értékű személy voltot egyes emberektől a bőrszínük alapján. [...] Más módon is fontolóra vehetjük a zombik lehetőségességének kérdését [...] Lehetségesek a zombik? Nem csupán lehetségesek, hanem valóságosak. Mind zombik vagyunk. [Dennett lábjegyzete: „Elkeseredett intellektuális becstelenség lenne ezt az állítást a kontextusa nélkül idézni!"]. Senki sem tudatos – mármint olyan szisztematikusan rejtélyes módon, hogy ez olyan doktrínáknak ágyazna meg, mint az epifenomenalizmus! Nem tudom bebizonyítani, hogy nem létezik olyanféle tudatosság. Azt sem tudom bebizonyítani, hogy nem léteznek gonosz manók. A legtöbb, amit tehetek, hogy megmutatom, nincs respektálható indítékom hinni benne.” (Dennett 1991: 405-406)

Az idézet első felében Dennett a funkcionalizmus mellett érvel, Block (1978) kifejezésével a *sovinizmus* veszélyére hivatkozva: ha egy lény rendelkezik mindazokkal a kognitív képességekkel és viselkedési tendenciákkal, mint egy ember, akkor indokolatlan elvitatni tőle a tudatosságot. Ennek az érvelésnek ismert problémája, hogy amilyen meggyőző abban a formában, ahogy Dennett tálalja (ahol az ember és az idegen lény közötti megkülönböztető vonás az agyszín a bőrszín mintájára), annyival kevésbé nyilvánvalóan igaz a Block által emlegetett különféle abszurd példákban. A sovinizmussal ellentétes *liberalizmus* abszurditása kételyt ébreszthet a funkcionalizmus igazságában, de a funkcionalizmus problémáinak kimerítő tárgyalása ebben az értekezésben nem lehet cél. (Persze a funkcionalizmus is válhat sovén állásponttá, ha például a kognitív képességek hiánya miatt megtagadjuk a teljes értékű személyvontot egy csecsemőtől vagy egy demenciában szenvedő személytől.)

Akárhogy is alakuljanak a funkcionalizmus körüli viták, a fenti idézet érdekessége, hogy első felében Dennett érve az, hogy a robot (vagy az úrlény) sem zombi, míg a második felében az,

hogy mindannyian zombik vagyunk. Világos, hogy miért nem lenne helyes az utóbbi megállapítást kiragadni a kontextusából. Dennett álláspontja az, hogy nem lehet radikális különbséget tenni egy ember és egy intelligens gép (vagy egy másik fajba tartozó értelmes lény) között tudatosság tekintetében, mert semmiféle „rejtélyes” tulajdonsága nincs az embernek, ami egy ilyen különbségtételt megalapozna, tehát a zombi-gondolatkísérletek értelmetlenek. Ezt az álláspontot olvashatjuk revizionista módon a „tudat tagadásaként”, vagy pedig antirevizionista módon egy „liberális” álláspontként, amely szerint a tudatosság mindenhol előfordulhat, ahol valamilyen rendszer bizonyos funkciókat ellát. Revizionista olvasatban amikor Dennett tagadja, hogy a kválék okkurrens állapotok vagy a tapasztalat intrinzikus minőségei lennének, akkor ezzel tagadja, hogy létezne a tudat abban az értelemben, ahogy az antirevizionisták szerint „nyilvánvaló tény” a létezése. Amikor funkcionális jellemzőkkel azonosítja a tudatot, akkor revizionista olvasatban csupán újrahasznosít egy szót, de anélkül, hogy azt a „rejtélyes” tulajdonságot nevezné meg vele, aminek létezését már nincs okunk feltételezni. Antirevizionista olvasatban viszont vagy azt a tudatot igenli, amiben az antirevizionisták hisznek, és abban hisz, hogy az csakugyan előáll, valahányszor egy rendszer rendelkezik a szükséges funkcionális tulajdonságokkal, vagy pedig egyszerűen hallgat a tulajdonképpeni tudatról – hallgatólagosan előfeltételezve, hogy nyilvánvalóan létezik – és filozófiája másról (például Block kifejezésével a hozzáférési tudatosságról) szól elsősorban.

Így például amikor Dennett azt fejtegeti, hogy a szenvedés nem azt jelenti, hogy a szenvedő egy „ineffábilis, de intrinzikusan szörnyű állapotot” él át, hanem céljainak, reményeinek megghiúsulását (Dennett 1991: 449), akkor a revizionista olvasat szerint itt tényleg arról van szó, hogy Dennett szerint a szenvedés intencionális, illetve diszpozicionális terminusokban maradéktalanul megérthető, hiszen „hogylétünknek” semmiféle intrinzikus, okkurrens minősége nincs. Antirevizionista olvasatban viszont a hangsúly oda kerül, hogy Dennett sem állítja, hogy ha beüti a lábujját a küszöbbe, akkor nem él át „szörnyű állapotokat” – ezeket azonban mindig meg lehet magyarázni (vagy legalábbis jellemezni lehet, körül lehet írni) célok megghiúsulásának terminusaiban. Például aki beütötte a lábujját a küszöbbe, azt egyelőre körülményei megakadályozzák abban, hogy folytassa tevékenységeit, melyek elvégzését nemrég elhatározta, valamint új célkitűzései és vágyai is megjelennek (az, hogy ne „informálódjon” tovább arról, amiről jelenleg a fájdalomérzet révén „informálódik”, és pedig Dennett kifejezésével nem éppen a „kedvenc módján”).

Amikor Dennett amellet érvel, hogy a fájdalom maradéktalanul magyarázható funkcionális képességekre hivatkozva, akkor azonban nem teljesen meggyőző. Etológiai kutatások alapján megállapítja, hogy valószínűleg azokban az állatokban alakul ki a szenvedés képessége, amelyek

tudnak tenni azon körülmények elkerülésére, amelyek a szenvedést okozzák, így egy, a veszélyforrások elől elmenekülés képességét ki nem fejlesztő élőlény minden bizonnyal nem alakítja ki a szenvedés képességét sem. „Nem lenne evolúciós értelme egy olyan fának, ami képes lenne arra, hogy csendben szenvedjen, ha levágják az ágait” (Dennett 1991: 451). Dennett azonban elkerüli azt a kérdést, hogy például mi volt az evolúciós értelme annak, hogy az ember *szenvedjen* veszélyforrás vagy sérülés jelenlétében ahelyett, hogy pusztán tudomást szerezne az elkerülendő körülményekről. Evolúciós érvelés alapján nehéz megérteni, miért olyan fájdalmas egy vesegörcs, amilyen. Revizionista olvasatban, mivel nincs semmiféle „fájdalom” túl a sérülések érzékelésén, ez a probléma fel sem merül, míg antirevizionista olvasatban Dennett azt a szubsztantív állítást látszik tenni, hogy a veszélyforrások elől menekülni képes állatok képesek a szenvedésre (amely ismeretlen okból ráadásként jár az elkerülő magatartások és a sérülések konstatálása mellé), míg az erre nem képes élőlények nem.

Dennett kevés egyértelműbb bizonyítékát adja, hogy tudatában van az ellentétes intuitív meggyőződések összecsapásának, mint 1991-es műve utolsó lapjain, amikor azt a vádat anticipálja, hogy könyvében a tudatot – a mű címével ellentétes módon – nem megmagyarázta, csupán „kimagyarázta” (Rorty 1998: 98-99). Válasza meglehetősen árulkodó, hiszen lényegében ismét egyfajta vitalizmusvádat fogalmaz meg:

Amikor megtudjuk, hogy az egyetlen különbség az arany és ezüst között az atomjaikban lévő szubatomi részecskék számának eltérése, akkor dühöseket lehetünk vagy átverve érezhetjük magunkat – ezek a fizikusok valamit kimagyaráztak: az arany eltűnt az aranyból; kihagyták az ezüsből éppen azt az ezüstösséget, amit nagyra becsültünk. [...] valami valóban „kimaradt” [...] Kihagyni valamit: ez nem a sikertelen, hanem a sikeres magyarázatok jellemzője.” (Dennett 1991: 454)

Csak az az elmélet tarthat igényt a „tudat magyarázata” elnevezésre, amely nem-tudatos események alapján magyarázza meg a tudatosságot, pontosan úgy, ahogy az élet nem vitalista tudományos magyarázatának is „ki kell hagynia” az *élan vital*-t, hiszen különben nem lenne magyarázat: „Ha a modelled arról, hogy a fájdalom hogyan áll elő az agytevékenység termékeként, még mindig tartalmaz egy „fájdalom” feliratú dobozt, akkor még el sem kezdted megmagyarázni, mi a fájdalom”, írja Dennett. Azok, akik azt gondolják, hogy a tudat funkcionalista magyarázata

nem teljes, hanem kihagy valamit, a kihagyottnak vélelmezett tulajdonságot „kválé”-ként csempészik vissza elméleteikbe. (Dennett 1991: 455).

2.3 További eliminativista kísérletek

A revizionista elmefilozófia feltárását olyan elméletek vizsgálatával folytatom, melyek kifejezetten a kválék létezését próbálják megcáfolni, vagy éppen azt az álláspontot, hogy „valamilyen” a tapasztalatainkat átélni.

A fenomenális tudatossággal kapcsolatos eliminativizmus egyik klasszikus szövege szintén Dennett műve (Dennett 1988). Úgy tűnik, ez a tanulmány valóban közel kerül a „tudat tagadásának” álláspontjához, legalábbis bizonyos mondatai alapján (és a tanulmány recepciója alapján) lehet ilyen benyomásunk. Egy jellemző részlet a tanulmány legelejéről:

Az, amilyenek a tej ízét érzed, egy másik, ízlelési kválé, és az, amilyenek hangzik, miközben lenyeled, egy hallási kválé[...] Semmit sem ismerhetsz, úgy tűnik, bensőségebben, mint a saját kváléidat [...] Descartes azt állította, hogy mindenben kételkedik, amiben kételkedhet, de sohasem kételkedett abban, hogy tudatos tapasztalatainak vannak kváléi, olyan tulajdonságai, amelyek által ismerte vagy felismerte őket. (Dennett 1988)

Amikor Dennett a tej ízét és hangját említi, akkor követi az elmefilozófusoknak azt a szokását, hogy amikor a kválék témáját vezetik be, példákat sorolnak fel. A bevett megfogalmazás szerint azért *valamilyen* átélni a velünk történő eseményeket, mert tapasztalatainknak valamilyen *milyensége* van, vagyis fenomenális tulajdonságai – kváléi. Amikor a filozófusok ehhez a manőverhez folyamodnak, akkor abban bíznak, hogy olvasójuk nem érti félre a mondanivalójukat, mert a megfelelő intuitív meggyőződésekkel felvértezve olvassa az efféle sorokat, és mint Lewis megállapította, kevés szóból is megérti, miről van szó (Lewis 1995).

A kísérleti filozófiai felmérések azonban megmutatták, hogy nem lehetünk biztosak abban, hogy a filozófiai tanulmányok olvasói egységesek az elmefilozófiai intuícióik vonatkozásában. A filozófusok közti viták pedig azt mutatják meg, hogy a filozófusok sem ugyanazokból az intuitív

meggyőződésekből indulnak ki, és ezért – vagy más okokból – eltérően értelmezhetik egymásnak a fenomenális tudatossággal kapcsolatos állításait.

Ha alaposabban megnézzük azt a módot, ahogyan Dennett bevezeti a kválék témáját, akkor két sajátosságra figyelhetünk fel. Az egyik nem specifikusan Dennett megközelítésére jellemző, hanem a példák felsorolásán alapuló bevezetésekre általánosan: amikor a tej ízéről vagy lenyelésének hangjáról van szó, akkor mindig lehetséges az ilyen példák említésének tisztán kognitív vagy fizikai olvasata, mindenféle, a fenomenális tudatosságra való vonatkozás nélkül. Nem segít az sem, ha *minőségről* vagy *milyenségről* beszélünk, hiszen a revizionista – akinek számára korántsem magától értetődő, hogy a szituáció kognitív vonatkozásain túl még további, számára megfoghatatlan és motiválatlan fenomenális jellemzők is magyarázatra szorulnának – ezt úgy értelmezheti, hogy a *tej tulajdonságairól*, illetve a tejről alkotott hiedelmeinkről, tejjel kapcsolatos attitűdjeinkről beszélünk. A tejnek természetesen vannak tulajdonságai – kémiai, fizikai tulajdonságok, a vele kapcsolatos attitűdjeink és hiedelmeink pedig azok a fiziológiai és pszichológiai reakciók, melyeket a tej ivása kivált belőlünk. A revizionistát az sem hozza zavarba, ha a „tapasztalat tulajdonságairól” beszélünk, hiszen értelmesen lehet beszélni arról, hogy egy adott helyzetben például milyen érzelmi állapotban vagyunk, milyen propozicionális attitűdökkel rendelkezünk, illetve látóterünkben milyen tárgyak vannak, és ízlelőérzékünk milyen ingerületet fog fel. Ha két különböző szituációban tejet iszunk, akkor könnyen lehetséges, hogy a két szituáció között lesznek olyan hasonlóságok, amelyek abból adódnak, hogy mindkét esetben történetesen tejet iszunk. Például mindkét esetben teljes tapasztalatunk része lehet valamilyen T állapot fennállása (a tej ízének érzése). Ez esetben a tej ízének érzését a „tapasztalatunk tulajdonságának” nevezhetjük. Továbbra sincs semmi okunk arra, érvelne egy revizionista, hogy ezzel összefüggésben másról is beszéljünk, mint hiedelmekről vagy érzékszervünk állapotairól. Csak az antirevizionista számára nyilvánvaló, hogy a tej megízlelése esetében a kognitív vonatkozásokon felül (az íz érzékelésén, az ital színének konstatálásán stb. felül) a tapasztalatunknak még valami egyéb, fenomenális tulajdonságáról is beszámolót kellene nyújtani.

Mindezt pedig azért fontos hangsúlyozni, mert Dennett már tanulmánya bevezető soraiban igen közel kerül ahhoz, hogy a kválé fogalmának, illetve a kválé fogalmát bevezető jellegzetes példának kognitív olvasatot adjon, amikor kváléink bensőséges *ismeretéről* beszél, vagy arról, hogy Descartes nem kételkedett abban, hogy tapasztalatainak kváléi vannak.

A továbbiakban azonban Dennett igyekszik hangsúlyozni, hogy ő valóban *annak* a létezését szándékozik tagadni, ami az antirevizionisták szerint nyilvánvalóan létezik: „Céлом felforgató. Egy

olyan elképzelést szándékozom megdönteni, amely ilyen vagy olyan formában, „nyilvánvaló” a legtöbb ember – tudósok, filozófusok, laikusok számára.”

A továbbiakban azonban Dennett kifejti, hogy nem tagadja a tudatos tapasztalat realitását, állítása csupán annyi, hogy amikor valamit „valamilyennek és nem másilyennek” tapasztalunk, akkor ez ugyan valamilyen, az adott pillanatban a tapasztalóra jellemző tulajdonságoknak köszönhetően történik, ám e tulajdonságok „annyira nem olyanok, mint a tudatnak hagyományosan tulajdonított tulajdonságok, hogy súlyosan félrevezető lenne a régóta keresett kválékkal azonosítani bármelyiküket is. A kváléknak különleges tulajdonságoknak kell lenniük egy bizonyos nehezen meghatározható értelemben. Állításom... hogy a tudatos tapasztalat nem rendelkezik olyan tulajdonságokkal, amelyek olyan módon különlegesek, ahogy a kváléknak állítólagosan különlegesnek kellene lenniük.”

Dennett tehát önértelmezése szerint nem tartozik egyértelműen a tudat tagadói közé, sőt, még a „tapasztalat milyenségének” tagadói közé sem. Más kérdés, hogy a tudat olyan tisztán kognitív és funkcionális fogalmával dolgozik, amely aligha elégti ki az antirevizionistákat.

Dennett azt szándékozik megmutatni, hogy kválékról beszélni ugyanolyan elavult és nehezen védhető, mint az életerőről beszélni a modern biológiában. Arra a kérdésre, hogy mik a kválék, Block *Louis Armstrong bon motjával* (a továbbiakban LABM) felelt. Dennett szerint ha érvelése sikeres, akkor ez a válasz, melyet sokan meggyőzőnek tartanak, ugyanolyan elfogadhatatlanná válik, mint egy olyan elképzelt érv, hogy egy élőlény nem kételkedhet az életerő létezésében.

E megjegyzések alapján felvázolható, hogy mit gondol Dennett az elmefilozófiai pozíciók mögött megbúvó intuitív meggyőződésekről és ezeknek az elmefilozófiai elméletekhez való viszonyáról. A kválékat egy rossz, elavult elméletben feltételezett tulajdonságoknak tekinti. Meghaladhatók, éppen úgy, mint az életerő.

Továbbá Dennett úgy véli, hogy a kválék a bevett nézet szerint „különleges tulajdonságokkal” rendelkeznek, amely megfoghatatlanná teszi őket, és ez magyarázza LABM elfogadottságát. Ha tehát a rossz elmélet helyébe egy jobb elmélet kerül (mint ahogyan az életerő-elméletet felváltja a modern molekuláris biológia), akkor LABM elveszti vonzerejét.

Dennett szerint tehát a tudatos tapasztalat általa bírált felfogása olyan tulajdonságokra hivatkozik (a kválékra), melyek különleges tulajdonságaiknál fogva megfoghatatlanok, ezért csak olyan talányos módon lehet róluk beszélni, mint amilyen LABM. A tudatos tapasztalat jobb

elméletében nincsen szó különleges tulajdonságaikból eredően megfoghatatlan tulajdonságokról, csak olyan tulajdonságokról, amelyekre a bon mot nem igaz.

Már ebből a rövid rekonstrukcióból kitetszik az a potenciális nehézség, hogy nem egyértelmű, hogy a kválék hagyományos felfogásának megdöntéséből következik, hogy *LABM* elveszíti plauzibilitását az egész fenomenális tudatosság vonatkozásában is.

Dennett helyesen látja, hogy amikor egy filozófus kénytelen *LABM*-t idézni, akkor gondban van: valamilyen „megfoghatatlan” tulajdonságra próbál utalni. Azért kell a bon mot-t segítségül hívnia, mert az összes többi példa (a tej íze és színe stb.) félreérthető lenne. Dennett szerint ezeknek a filozófusoknak a kínjait a kválé fogalmának inkohérens volta okozza. Ha a kválé fogalmát kiküszöböljük, akkor a bon mot komolyan vehetőségének látszata is megszűnik. Ám mi van akkor, ha a bon mot valamilyen változata (például nem a kváléval, hanem egyszerűen a „tudatossággal” vagy a „tapasztalat milyenségével” kapcsolatban) akkor is hihető marad, ha a kválé fogalmát sikerült eltörölni? Ekkor a kválé eltörlése nem vezet a „tudat tagadásához”.

Ha nem a kválé fogalmának sajátosságai vezettek oda, hogy a filozófusok zavarukban *LABM*-t voltak kénytelenek idézni, hanem fordítva, a kválé fogalmát azért vezették be, hogy megpróbálják explikálni, mi az a megfoghatatlan valami, ami a tapasztalatra jellemző annak kognitív-funkcionális vonatkozásain túl az antirevizionista intuíció szerint, akkor a kváléfogalom furcsaságai csak tünete bizonyos filozófusok és laikusok alapvető meggyőződésének, nem pedig okai annak.

Ha viszont ez így van, akkor még ha sikerül is Dennettnek kiküszöbölnie a kválékat, ezzel nem tudja eloszlatni azt az intuitív meggyőződést, ami végül is a kválék „feltalálásához” vezetett, és az antirevizionista filozófusok a kválékkal szembeni kritikáját legfeljebb olyan kihívásként értelmezik, amely felhívta a figyelmet arra, hogy az antirevizionista intuíció elméleti formában történő kifejezésének nem megfelelő módja, ha ez a kválék terminusaiban történik. A kválék kiküszöbölésével tehát Dennett nem lesz egyben a tudat eredményes „tagadója” is.

Másfelől, amikor Dennett a tudattal kapcsolatos realizmusáról beszél a 80-as és 90-es évekbeli műveiben, akkor mindig világossá teszi, hogy tudaton kizárólag valamilyen funkciók ellátását érti. Ebben az értelemben Dennettet a „tudat tagadói” közé sorolhatnánk, de egy antirevizionista számára Dennett funkcionalista gondolatmenetei nemcsak az antirevizionista intuíciókkal szembeni állásfoglalásnak tűnhetnek, hanem a téma elkerülésének is.

Visszatérve Dennett 1988-as tanulmányára, Dennett tizenöt gondolatkísérlet (pontosabban *intuáció-szivattyú*) segítségével igyekszik bemutatni, hogy tapasztalatainknak semmilyen

tulajdonsága nem felel meg a kválék négy klasszikus ismérvének: annak, hogy a kválék *intrinzikusak, ineffábilisak, privátak és közvetlenül hozzáférhetőek*.

Így például a harmadiktól a hatodikig terjedő példákban Dennett megállapítja, hogy nem lehetséges a kválék interszjektív összehasonlítása, sőt, még azt sem tudjuk megmondani, hogy saját kváléink megváltoztak-e egy rajtunk végzett műtét hatására (vagy kváléink nem változtak meg, csak a rájuk adott reakcióink), vagyis introspekcióink számára korántsem jól hozzáférhetőek a kválék. A kávékóstolók híres példájának (7-8. gondolat kísérlet) tanulsága, hogy mivel a kválék a viselkedésünkre csak a velük kapcsolatos ítéleteink közvetítésével hatnak, ezért a viselkedésből mindig csak két tényező (a kválé és a rá adott, emlékezettől is függő tipikus reakciók) eredőjére lehet következtetni, így magáról a kváléről soha nincsen viselkedési bizonyíték. A kválék tehát vagy teljesen megismerhetetlenek számunkra, vagy legfeljebb következtetni lehet rájuk viselkedési és pszichológiai reakciómintázatok alapján. A tapasztalt sörívő példája (9. gondolat kísérlet) kétségbe vonja, hogy a sör íze intrinzikus tulajdonság lenne (hiszen a tapasztalt sörívő számára másmilyen élmény sört inni, mint annak, aki először kóstolja meg). Dennett további három példát hoz fel a kválék intrinzikussága ellen. Rámutat továbbá, hogy az intrinzikusság fogalma egyébként is tisztázatlan a filozófiában, illetve arra, hogy olyan tulajdonságok is, mint a színek, relacionális tulajdonságok.

Ami a kválék ineffábilisát illeti, Dennett szerint elfogadhatjuk, hogy tudatos tapasztalatainknak „gyakorlatilag ineffábilis” tulajdonságaik vannak, vagyis meg tudunk különböztetni, illetve újra tudunk azonosítani tapasztalatokat anélkül, hogy szavakba tudnánk foglalni, pontosan milyen tulajdonságok figyelembevételével, milyen módon voltunk képesek erre, ahogy azt már Smart (1959) is megállapította.

A tapasztalat ineffábilis, intrinzikus, privát és közvetlenül hozzáférhető tulajdonságai helyére tehát „gyakorlatilag ineffábilis”, publikus tulajdonságok kerülnek, melyekhez nincs közvetlen, tévedhetetlen hozzáférésünk, sőt a találgatásnál több autoritásunk csak relacionális, extrinzikus tulajdonságokkal kapcsolatban van. Ha tehát a kválék a tapasztalat ineffábilis, intrinzikus, privát és közvetlenül hozzáférhető tulajdonságai, akkor nincsenek kválék.

Dennett tanulmánya terjedelmes vitákat váltott ki; a tipikus ellenérveket a következőkben csak nagyon röviden fogjuk bemutatni. Mielőtt ezekre rátérünk, két megjegyzést érdemes tenni. Az egyik, hogy az intrinzikusság kérdését Dennett végső soron függőben hagyta, noha néhány gondolat kísérletét érvnek szánja amellet, hogy az ízek, színek stb. tekinthetők relációs tulajdonságnak és így nem intrinzikusak.

A második, hogy Dennett számos gondolat kísérletében szerepel egy kválé és egy vele kapcsolatos ítélet, vagy rá adott reakció, vagy vele összehasonlított emlék, melyek együttesen határozzák meg az élmény minőségét, a tapasztaló ítéleteit, vagy a viselkedést. Dennett bebizonyítja, hogy ezekben a szituációkban, melyeket úgy jellemezhetünk, hogy jelen van egy kválé és még valami (ítélet, reakció, viselkedés, emlékezet), a kválé hozzájárulása a szituáció egészéhez annyira bizonytalan, hogy meg kell állapítanunk, a kválé fogalma hasznavehetetlenül üres vagy zavaros. Azt azonban Dennett nem állítja, hogy *semmilyen* átélni az adott szituációt. Más szóval, aki realista a tudattal kapcsolatban, az állíthatja azt is, hogy az, hogy *milyen* például első vagy sokadik alkalommal sört kóstolni, természetesen nem független attól, hogy mennyire tapasztalt sörivők vagyunk. Dennett gondolat kísérlete rámutat, hogy hiba lenne a sörivés élményét felosztani a sörnek az általunk érzett, de rá adott reakcióinktól mégis függetlenül tekintett ízére egyfelől és az ízre adott reakciónkra másfelől. Egy intrinzikus, ineffabilis stb. *söríz-kválé* fogalma, meglehetősen inkohérens. Lehet, hogy egyáltalán nincsenek kválék. Ettől még értelmesen állíthatjuk, hogy számunkra *valamilyen* sört inni. Természetesen szó sincs arról, hogy tapasztalatunk e minősége valamilyen különleges (a dennetti „gyakorlati”-nál erősebb) értelemben ineffabilis vagy privát lenne, hiszen bárki a maga számára reprodukálhatja az élményt, ha sört iszik. A sörivés tapasztalatának minősége *közvetlenül hozzáférhető* abban az értelemben, hogy sör ivásával képesek vagyunk instanciálni a sörivés tapasztalatának minőségével rendelkezés tulajdonságát. Nem közvetlenül hozzáférhető abban az értelemben, hogy introspekcióval belátást szerezhethetnénk abba, hogy a sör ízének élvezetének mi a *természete* (például milyen agybeli folyamatokkal azonos). Továbbá, bár a sörivés aktusa és a sörivés által a sörivőben előidézett neurofiziológiai állapot tekinthető relacionálisnak vagy extrinzikusnak abban az értelemben, hogy összetett fizikai rendszerek egymással való kölcsönhatásából áll elő, azok a materialisták, akik elutasítják a funkcionalizmust, feltételezhetik, hogy a szituáció tipikus tulajdonságai olyan neurofiziológiai vagy fizikai tulajdonságok terminusaiban határozhatók meg, melyek végső soron legalább részben a sörivő szervezetének – például agyának – intrinzikus fizikai tulajdonságain alapulnak. Röviden, a sörivő olyan S állapotba kerül, mely még ha több elemből áll is, ezek némelyike intrinzikus fizikai vagy neurofiziológiai tulajdonság.

A fenti megjegyzések alapján Dennett cikkéről megállapíthatjuk, hogy bár Dennett szerint nincsenek kválék, ezzel Dennett nem feltétlenül iratkozott fel a „tudat tagadó” közé, hiszen a kválék hagyományos felfogásának elutasítása – a kválék dennetti kritikájának elfogadása – mellett is állítható, hogy *valamilyen* átélni a tudatos tapasztalatokat, és ennek során a tapasztalat relatíve

ineffabilis, gyakorlatilag privát, bizonyos értelemben közvetlenül hozzáférhető, és talán intrinzikus (vagy intrinzikus tulajdonságoktól nem független) tulajdonságáról beszélünk.

Ugyanakkor a kválékkal szembeni dennetti eliminativizmust nem mindenki fogadta el. Így például Seager a kválék Dennett által megfogalmazott négy jellemzőjének olyan „minimalista” interpretációját (Seager 1993: 347) próbálta megfogalmazni, amelyek nem esnek áldozatul a dennetti példának. Az *ineffabilitás* és *privátság* például pusztán gyakorlati; az *intrinzikusság* a kválék vonatkozásában tartható, mivel az objektumok alapvető fizikai tulajdonságai intrinzikusak, és ezen nem változtatnak a Dennett által leírt szituációk. A *közvetlen hozzáférhetőséggel* kapcsolatban Seager úgy véli, hogy ez mindössze annyit jelent, hogy tudatállapotainkról „noninferenciális módon tudomásunk van”.

Seager szerint Dennettnek azok a gondolatkísérletei, melyek az emlékezet és a pillanatnyi kválék összevetésén alapulnak, indokolatlanul verifikacionisták: Dennett abból, hogy valamilyen dilemmát nem lehet eldönteni (például: a jelenleg érzékelt kválé változott meg, vagy az emlékeinkben tárolt kválé?), arra következtet, hogy az ilyen dilemmák alapjául végül is nem szolgálnak szubsztantív kérdések. Seager szerint azonban a gyakorlatilag eldönthetetlen kérdésekből ez nem következik. Általánosabban: tapasztalatunknak lehet valamilyen fenomenális tulajdonsága akkor is, ha ezzel kapcsolatban végül nem tudunk bizonyos kérdésekre válaszolni.

Egy másik választípus szerint a kváléknak mindössze fenomenális karakterrel kell rendelkezniük, de nem kell intrinzikusnak, privátnak, ineffabilisnak vagy közvetlenül hozzáférhetőnek lenniük. Frankish kifejezésével ezek a „diétás kválék” (Irvine – Sprevak 2020: 7).

Irvine és Sprevak továbbfejleszti Dennett 1988-as érvelését úgy, hogy a diétás kválékkal szemben is működjön. Az érvelés első lépcsőjében azt kell kimutatni, hogy sohasem tudjuk, hogy pillanatnyi tapasztalatunk melyik (diétás) kválét instanciálja. A második lépésben pedig ebből kiindulva érvelnek a kválék eliminációja mellett (Irvine – Sprevak 2020: 9).

Az első lépcsőben a szerzők egy sci-fi operációt írnak le, amelynek szenvedő alanya ébredés után az őt körülvevő világot megváltozott színűnek tapasztalja, de nem tudja eldönteni, hogy az emlékei vagy a pillanatnyi látási kváléi változtak-e meg. Ezt a közismert fajta gondolatkísérletet a szerzők továbbfejlesztik olyan módon, hogy abban a dilemmát az agy vizsgálatával (objektív vizsgálati módszerekkel) sem lehet eldönteni. Végkövetkeztetésük az, hogy ha a példabeli személy nem tudja eldönteni, hogy a kváléi megváltoztak-e, ha tehát sem introspektív, sem objektív módszerekkel nem tudjuk eldönteni, hogy változnak-e a kváléink, akkor saját kváléink is akár bármikor változhatnak, anélkül, hogy ezt észrevennénk (Irvine – Sprevak 2020: 12).

A második lépcsőben a kválék megismerhetetlenségéből arra következtetünk, hogy azokat, mivel teljesen funkció nélkül maradnak, eliminálhatjuk ontológiánkból. A kválék Wittgenstein dobozba zárt „bogaraihoz” hasonló szerepet játszanak. A kválé-realista ragaszkodhat ahhoz, hogy a kválék továbbra is jelen legyenek ontológiánkban annak ellenére, hogy nem megismerhetők számunkra, ám ez az álláspont motiválatlan (Irvine – Sprevak 2020: 13).

Irvine és Sprevak két lehetséges ellenvetést említ meg gondolatmenetével kapcsolatban. Az egyik az, hogy mindazonáltal úgy tűnik számunkra, hogy léteznek a kválék; ez az „illúzió-probléma”, mely elmefilozófia „nehéz problémája” helyére kerül.

A másik, hogy abból, hogy nem tudjuk eldönteni, mai kváléink ugyanolyanok-e, mint a tegnapiak, nem következik, hogy a jelenlegi kváléinkról semmiféle tudásunk nincs. Így például minden pillanatban egyszerre több kválé instanciálódik tapasztalatunkban, és ezeket össze tudjuk hasonlítani egymással.

Az Irvine – Sprevak-féle érvelés első fázisával kapcsolatban felvethető az a probléma, hogy a szerzők szerint a futurisztikus agysebészeti beavatkozás alanya arra ébred, hogy megváltoztak látási élményei. Azt ugyan nem tudja eldönteni, hogy az emlékezete vagy a pillanatnyi látási tapasztalatai változtak-e meg, ám valamilyen változást tapasztal. Ha ez így van, akkor a gondolatkísérlet noha olyan szituációt ábrázol, amiben valaki nem tudja eldönteni, hogy változtak-e a *kváléi* vagy sem, azt viszont tudja, hogy látási *tapasztalatai* változtak.

Másmilyen az illetőnek lenni az agyműtét előtt és után, mert a tárgyak máshogyan néznek ki a műtét után, mint előtte (vagy az emlékek változtak meg?). A tapasztalat milyensége nem eliminálható azon az alapon, hogy nem játszik szerepet, hiszen itt szemlátomást szerepet játszik; akár szerepet játszik a látási kválék megváltozása a tapasztalat változásában, akár nem. Seager érvét idézve pedig hozzátehetjük, hogy abból, hogy nem tudjuk elválasztani a látási kválék hozzájárulását az egyéb hatásoktól, még nem következik a kválék nemlétezése.

Irvine és Sprevak mondhatnák azt is, hogy az agyműtét hatására megváltoztak a látási kválék, de ebből a műtét szenvedő alanya semmit sem vesz észre, hiszen a pusztán kválék minden mástól elszigetelt változásait sohasem vesszük észre. Ezt azonban nem mondhatják, hiszen a céljuk éppen amellet érvelni, hogy nincsenek kválék.

Ezért azt kell mondaniuk, hogy történik egy agyműtét, aminek hatására a tárgyak máshogyan néznek ki a kísérleti alany számára. Miután a kísérleti alany felébred, konstatálja, hogy vagy a kváléi, vagy az emlékei változtak meg. Holott konstatálhatná azt is, hogy megváltozott a

tapasztalata, aminek részét képezhetik akár az intellektuális érzelmek is, például az a meghökkenés, amit az vált ki, hogy a citromok kékek, az ég pedig sárga.

Összegezve, az Irvine-Sprevak-érv nem perdöntő a látási kválék létezése ellen, továbbá az sem egyértelmű, hogy ha sikerül is bizonyítani egyes kválék megismerhetetlenségét, az egyenértékű-e az egész tudattal kapcsolatos eliminativizmussal.

A Dennett és Irvine-Sprevak-féle kválé-eliminativista érvekkel kapcsolatban megfogalmaztuk azt a kritikát, hogy az egyes kválék megismerhetetlenségével, vagy a kváléfogalom problémáival kapcsolatos érvek nem feltétlenül elégségesek a „tudat tagadásához”. De lehet-e közvetlenül érvelni amellett, hogy *semmilyen* átélni tapasztalatainkat, nincs olyan, hogy *amilyen* nekünk lenni?

P. M. S. Hacker (2002) valóban felteszi azt a kérdést, hogy egyáltalán értelmes-e arról beszélnünk, hogy *valamilyen* nekünk lenni vagy *valamilyen* számunkra különféle tapasztalatokat átélni. Ha azt kérdezik egy megvakult, de a látását egy műtét eredményeként visszanyerő embertől, hogy milyen érzés újra látni, akkor természetesen az a válasz várható, hogy „csodálatos” vagy „furcsa”. Amikor azonban egyszerűen az a kérdés, hogy milyen látni a körülöttünk lévő hétköznapi tárgyakat, akkor nem világos, hogy ez a kérdés voltaképpen mire irányul. Természetesen a különféle tárgyakat látni más-más fajta tapasztalat, annak függvényében, ahogy ezek a tárgyak eltérnek egymástól. Arra a kérdésre, hogy milyen érzés a migrénes fejfájás, értelmesen lehet válaszolni, mert ennek az élménynek negatív hedonikus karaktere van, negatív érzelmi reakciók kapcsolódnak hozzá. Egy hétköznapi színérzékeléshez azonban semmi ilyesmi nem kapcsolódik, ezért Hacker szerint „nincs kvalitatív karaktere – nem kellemes vagy kellemetlen”.

Hacker tehát úgy látja, hogy a tapasztalatokkal kapcsolatban két értelemben lehet arról beszélni, hogy azok milyenek. Egyrészt a tapasztalatok tárgyai eltérnek egymástól, és ebben az értelemben a tapasztalatok is: egy barna vagy zöld folt látása másféle tapasztalat, bár a „milyen?” kérdésre adott válasz ezekben az esetekben az érzékelt tulajdonságokat írják le. Másrészt egyes tapasztalatoknak további – negatív vagy pozitív – érzelmi értékük is van, és amikor a „milyen volt ezt és ezt átélni?” kérdés értelmes, akkor az erre irányul; ugyanakkor a legtöbb tapasztalatunk, például közönséges, mindennapi érzékelésünk nem bír semmiféle említésre méltó érzelmi vagy hedonikus értékkel.

Mivel ebben a két értelemben vethető fel, hogy milyen egy tapasztalat, nyilvánvaló, hogy magára a tudatos tapasztalás egészére vonatkozóan nem lehet értelmesen megkérdezni, hogy milyen azt átélni vagy azzal rendelkezni.

Hacker szerint tehát az az állítás, hogy csak a tudatos lények számára „valamilyen nekik lenni”, zavaros. Ehelyett annyit lehet állítani, hogy „csak a tudatos, érző lények tapasztalatok szubjektumai és tapasztalataikat ... csak ők találhatják kellemesnek vagy kellemetlennek, érdekesnek vagy unalmasnak” (Hacker 2002). Azt azonban Hacker nem fejti ki bővebben, hogy mit jelent ebben a mondatban a *tudatos* és az *érző*. Lehet, hogy azt jelenti, amire azok a szerzők próbálnak utalni, akik ezeket a lényeket úgy jellemzik, mint amelyek számára *valamilyen* nekik lenni, de lehet, hogy nem. Talán a „tudatos” és „érző” Hacker megfogalmazásában valami tisztán funkcionális tulajdonságot jelöl, ebben az esetben viszont Hacker és az általa kritizált szerzők végső soron két malomban örölnék: még ha Hacker be is bizonyítja, hogy a „számukra valamilyen nekik lenni” rossz explikációja annak a tulajdonságnak, amit ők elemezni próbálnak, problémáikat, amennyiben antirevizionisták, nem oldja meg, ha ezt helyettesítik egy csupán funkcionális fogalommal.

Hacker arra is rámutat, hogy az olyan kifejezések, mint a „milyen denevérnek lenni” vagy „milyen embernek lenni”, eltérnek az olyan szokványos és könnyen megválaszolható kérdésektől, mint például „milyen volt számodra katonának lenni a II. világháborúban” vagy rövidebben: „milyen orvosnak lenni”. Ezekre a kérdésekre könnyű válaszolni, mert a kérdezett leírhatja feladatait és tipikus élményeit, egy átlagos napját. Ezek a kérdések „milyen X számára Y-nak lenni” alakúak. A „milyen embernek lenni?” kérdés nem ilyen alakú, ezért túl általános ahhoz, hogy ilyen módon válaszolni lehessen rá. A „milyen orvosnak lenni?” kérdésre például úgy lehet válaszolni, hogy a válaszoló leírja, milyen feladatai vannak, hogyan telik egy tipikus napja, milyen emlékezetes sikerei és kudarcai vannak. Meglehetősen furcsa kérdés azonban az, hogy valakinek *emberként* vagy emberi *minőségében* hogyan telik egy napja, vagy emberként milyen sikerei vagy kudarcai vannak. Ezért Hacker szerint Nagel tévesen állítja, hogy „tudjuk, milyen nekünk lenni”, és Edelman és Tononi tévesen állítják, hogy mindannyian tudjuk, „milyen nekünk lenni”.

Hacker érvelése azonban csak azt mutatja meg, hogy a tudatosság egészére vonatkozóan nem jogosultak a „milyen ...-nek lenni” típusú állítások. Nincs sajátos minősége annak, hogy milyen tudatos lénynek lenni, vagyis annak, hogy „valamilyen nekünk lenni”. Egyszerűen az a helyzet, hogy a tudatos, érző lények, mint mi, „valamilyennek” (például kellemesnek vagy kellemetlennek) találunk némely helyzetet, amibe belekerülünk. Ezt azonban nem vitatja Hacker, ezért noha elutasítja azt az elképzelést, hogy a tudatosság problémáját jól meg lehet fogalmazni arra hivatkozva, hogy egyes lényeknek „valamilyen” lenni, pusztán ennek alapján aligha lehet besorolni a „tudat tagadói” közé.

Mi a helyzet a denevérekkel? Hacker szerint nem értelmesek a „milyen X számára X-nek lenni” típusú kérdések, ezért a „milyen egy denevér számára denevért lenni?” kérdés sem. Másfelől, a „milyen számokra embernek lenni?” kérdést átfogalmazhatjuk olyan módon, hogy értelmes és megválaszolható legyen, például: „Hogy telik egy napod?” vagy „milyen dolgoknak örülsz vagy mik bosszantanak fel általában?”

Farrell (1950) mélyreható elemzését nyújtotta az olyan típusú kérdéseknek, mint amilyen a „milyen lehet denenevért lenni” vagy a „milyen lehet marslakónak lenni” kérdése. Abban a kérdésben, hogy „milyen lehet ópiumszívónak lenni?”, szerinte nincsen semmi megválaszolhatatlan. Bárki megtudhatja, milyen ópiumszívónak lenni, mindössze el kell kezdenie ópiumot szívni. Ez a „sekély” vagy „triviális” kérdés látszólag eltér az olyan filozófiailag „mélyebb”, problematikusabb kérdésektől, mint a „milyen lehet denevért lenni”. Ez azonban nincs így, hiszen megtudhatjuk, milyen denevért lenni, amennyiben például denevérré változtat egy boszorkány. Más kérdés, hogy visszaváltoztatásunk után milyen fajta emlékeink maradnak denevértként eltöltött időnkéről, és be tudunk-e számolni róluk. Csak emberi nyelven, emberi fogalmakkal tudunk beszámolni róla másoknak, illetve emberként csak ilyen módon tudunk gondolkodni a kérdéses időszakról. Ez azonban nem különbözik lényegileg attól az esettől, amikor valaki ópiumot szív, majd pedig egy nem-ópiumszívónak próbálja leírni az ópiumszívás közben keletkezett látomásait, vagy saját maga próbálja felidézni kijózanodása után ezeket a víziókat. Hacker formuláit felhasználva tehát ha a „milyen X számára X-nek lenni” típusú kérdések nem értelmesek, a „milyen X számára Y-nak lenni” típusúak viszont igen, akkor értelmesen felvethető a kérdés, hogy „milyen volt számokra denevért lenni”, a „milyen volt számokra katonának lenni” analógiájára.

Farrell szerint pontosan annyira lehet (vagy lehetetlen) beszámolni arról, milyen volt denevért lenni (miután a boszorkány denevérré változtatott, majd vissza), mint arról, hogy milyen volt ópiumszívónak lenni. Bizonyos értelemben semmilyen efféle beszámoló nem tökéletesen kielégítő, mert az ilyen beszámolók hallgatója maga „nem éli át első kézből” azt, amiről a beszámolót hallgatja. Ebben az értelemben igaz az, hogy „sohasem tudhatjuk meg, milyen denevért lenni”, de csak annyira, amennyire igaz az, hogy „sohasem tudhatjuk meg, milyen érzés ópiumot szívni”. Természetesen megtudhatjuk, hogy az utóbbi milyen (az „első kézből való megtapasztalás” értelmében), ha történetesen ópiumot szívunk.

Milyen embernek lenni? Farrell terminológiájában „első kézből tudni”, milyen embernek lenni azt jelenti, hogy emberek vagyunk. Mivel mindannyian „első kézből tudjuk”, milyen

embernek lenni, ezért nem szoktuk feltenni azt a kérdést, hogy „milyen embernek lenni”, egy olyan típusú élménybeszámolót várva, mint amilyen az ópiumszívóé.

Nagel (Nagel 1974/2004) szerint csak akkor tudhatnánk meg, milyen denevérnek lenni, ha képesek lennénk felvenni a nézőpontját; Farrell kifejezésmódjában ezt úgy lehet megfogalmazni, hogy „nem tudjuk első kézből”, milyen denevérnek lenni, amíg nem vagyunk denevérek. Másfelől viszont amit *mondani* lehet arról, vagy emberként vagy más kívülállóként *tudni* lehet arról, hogy milyen denevérnek lenni (például ha egy varázslat hatására egy ideig denevérként röpködtünk, majd visszaváltoztunk), az nem több vagy kevesebb, mint az, amit mondani lehet arról, vagy kívülállóként tudni lehet arról, hogy milyen ópiumszívónak lenni.

Ha egy boszorkány valakit denevérré változtat, majd bizonyos idő után visszaváltoztatja, akkor az illető esetleg emberi nyelven, emberi fogalmakkal be tud számolni arról, milyen volt számára denevérként röpködni a sötétben, milyen sikerélménnyel járt a szúnyogok eredményes levadászása, vagy milyen kudarcélménnyel a függönybe való beleütközés. Amiről beszámol, az valószínűleg semmivel sem részletesebb vagy érzékletesebb, mint amilyen egy igen kompetens denevérkutató hipotetikus beszámolója lenne. Ez a beszámoló pontosan olyan, mint valakinek a beszámolója arról, hogy milyen volt számára katonának lenni a II. világháborúban (vagy egy történész hipotetikus beszámolója arról, hogy milyen lehetett katonának lenni a II. világháborúban). A „milyen egy denevér számára denevérnek lenni” helyére az kerül, hogy „milyen volt számokra szúnyogokra vadászni ebben a szobában, amikor denevérré voltál változtatva”. Amennyiben a válasz tartalmas, akkor a különféle átélt eseményekre (és az ezekkel kapcsolatos érzelmi attitűdökre) koncentrálnak, de tartalma sohasem az, hogy „milyen denevérként egyáltalán létezni” vagy milyen denevérként „tudatos lénynek lenni”.

Hacker és Farrell elemzése alapján tehát behatárolhatjuk, hogy mennyit tudhatunk arról, hogy milyen denevérnek lenni. Erről a témáról meglehetősen sokat tudhatunk, ugyanakkor sohasem élhetjük át „első kézből”, hogy milyen denevérnek lenni, ha csak nem változunk denevérré. De ez nem különbözik olyan más esetektől, mint a „milyen lehetett katonának lenni a II. világháborúban” vagy a „milyen lehet ópiumszívónak lenni”.

Akiket azonban ez a válasz nem elégít ki, nyilván arra kíváncsiak, hogy „valamilyen-e egyáltalán” denevérnek lenni, vagyis amikor a denevér éhes vagy jóllakott, és vadászata vagy a tárgyak sötétben való kikerülése közben sikerei és kudarcai vannak, akkor az ezekkel kapcsolatos gyönyörei és szenvedései, pozitív és negatív érzelmi reakciói ugyanúgy rendelkeznek-e valamiféle „minőséggel”, mint az ember hasonló állapotai.

Ugyan Hacker rámutatott arra, hogy a „milyen embernek lenni?” vagy a „milyen egy denevér számára denevérré lenni?” kérdésekre nem várhatunk értelmes választ, Farrell pedig arra, hogy az ilyen esetekben megkülönböztethető az „első kézből való tudás” – mely valójában nem tudás, hanem csupán az *annak levés*, vagy Nagel megfogalmazásában egy *nézőpont felvétele* – attól a tudástól, ami közölhető és semmiféle rejtélyt nem vet fel, ám látszólag mindig tökéletlen, mert „másodkézből való tudás” – ám akik azt szeretnék tudni, hogy a denevér valóban érző, tudatos lény-e, ezt a kérdést továbbra is értelmesen feltehetik, és az sem világos, hogy be kellene-e érniük a tudatosság vagy érzés egy tisztán funkcionális elemzésével. A tapasztalat „kvalitatív karakterét”, a kválékat illetően Hacker álláspontja az, hogy amikor filozófusok az érzeteink sajátos minőségi karakteréről beszélnek, és ezen biztosan nem az adott érzettel kapcsolatos érzelmi attitűdöt értik (mert például a kérdéses érzettel kapcsolatos attitűdünk semleges), akkor igazából az érzet tartalmáról beszélnek csupán. Egy piros-színérzet természetesen másmilyen, mint egy kék-színérzet, és pedig azért, mert az egyik tartalma a piros szín, a másiké a kék szín. „Abban az illúzióban szenvedünk, hogy amikor egy piros almát látunk, akkor... képesek vagyunk a figyelmünket a látásunkra irányítani” (Hacker 2002), holott amire rá tudunk mutatni vagy amire figyelmünket irányítani tudjuk, az csak az alma (vagy annak pirossága). Ez egy transzparencia-érv (vö. Harman 1990). Noha Hacker konklúziója az, hogy a tudatnak nincs semmiféle rejtélye, „csak empirikus tudatlanság és fogalmi misztifikáció”, és élesen kritizálja mind a kválékkal kapcsolatos elterjedt diskurzust, mint a tudatosságnak a „valamilyen számára lenni” kifejezéssel való körülírását, mégsem sorolhatjuk a „tudat tagadói” közé. Hacker szerint is vannak tapasztalataink, továbbá ezek élvezetesek vagy kellemetlenek lehetnek. Revizionista olvasatban Hacker tanulmánya alkalmas arra, hogy a „kvalitatív karaktert” és a „valamilyen neki lenni” beszédmódot tisztán funkcionista terminusokban értelmezzük. Antirevizionista olvasatban Hacker tanulmánya azt mutatja meg, hogy az antirevizionista intuíció megfogalmazására alkalmatlan az a megfogalmazás, hogy némely lényeknek „valamilyen lenni”, ám ebből nem következik, hogy ne léteznének kellemetlen vagy kellemes élmények, a kérdés csak az, hogy az élőlények mely körének vannak ilyenjei, és mitől függ az, hogy egy lény „érez”, illetve tudatos-e.

Hacker egy másik tanulmányában (Hacker 2012) közvetlenebb támadást intéz a „tudatkutatók” (a Nagel és Chalmers nyomán a tudat „nehéz problémáját” elismerő filozófusok és tudósok) ellen. A tudatnak az általuk használt modern felfogását Nageltől eredezteti, és e felfogás középpontjában az az elképzelés áll, hogy a tudatos lényeknek „valamilyen” lenni, valamint a tudatos állapotok azok, amelyekkel „valamilyen” rendelkezni. Az ezekkel az elképzelésekkel

kapcsolatos problémákat már másik tanulmánya kapcsán áttekintettük; érdekesebb a tudatfilozófia helyzetéről adott általánosabb jellemzése.

A nageli tudatfelfogás népszerűsége Hacker véleménye szerint négy tényezőre vezethető vissza. Először is, meggyőzőnek tűnhet az az állítás, hogy semmilyen téglának vagy tintasugaras nyomtatónak lenni, de valamilyen denevérnek lenni. Másodsor, a „valamilyen valaminek lenni” alakú állítások szokatlan szerkezete elaltatja nyelvi érzékenységünket és nem tűnik fel, amikor ezeket használva tulajdonképpen értelmetlenségeket beszélünk. Harmadsor, a filozófusok „meg szeretnék menteni emberségünket” a redukcionizmus és funkcionalizmus támadásaitól. Végül pedig a filozófusok szeretnék úgy érezni, hogy valami rejtélyes, nagyon nehéz problémával foglalkoznak. Csakhogy a filozófiában – így Hacker – nincsenek rejtélyek, csak misztifikációk.

A következőkben Hacker az általa támadott tudatfelfogással szemben három antitézist védelmez. Az első szerint a tapasztalatokat általában nem a „milyenségük” individuálja, hanem a tartalmuk, és a legtöbb tapasztalatnak nincs is kvalitatív karaktere (ami Hackernél a tapasztalatokra adott érzelmi reakciókat vagy hedonikus értéket jelenti). A második szerint nincs olyan, hogy „milyen” egy bizonyos tapasztalattal rendelkezni, a harmadik szerint pedig nincs olyan „milyen” embernek (vagy denevérnek) lenni. Az érvelés az előzőekben vizsgált tanulmánybelihez hasonlít.

Hacker az érvelés során megállapítja, hogy „igaz, hogy a súlyos fájdalom átélése borzasztó, rózsák illatának érzése kellemes, és megcsonkított testek látványa rettenetes. Ezek bizonyos tapasztalatok kvalitatív karakterei”. Kétségtelen, hogy mivel Hacker elutasítja a „kvalitatív karakternek” azt a felfogását, ami az általa támadott filozófusokra általában jellemző (ez valamilyen fenomenáliskarakter- vagy kváléfogalom, amely valami olyan fenomenális tulajdonságra irányul, mely elválasztható az érzelmi attitűdöktől vagy a hedonikus karaktertől), de ezzel a kválék filozófiai problémája helyébe csupán a fájdalmak és élvezetek problémáját helyezi. Sokak számára továbbra is értelmes és érdekes filozófiai kérdésnek tűnhet, ha nem is az, hogy „valamilyen”-e téglának vagy tintasugaras nyomtatónak lenni, de az a kérdés, hogy egy tégl (vagy ha nem is tégl, egy giliszta vagy csiga) átélhet-e kellemetlen vagy kellemes érzéseket, és ha nem, akkor miért nem.

Természetesen Hacker állíthatná azt, hogy ez a kérdés sem vethető fel értelmesen, mert a téglához hasonlóan az ember sem él át fájdalmat vagy kellemes érzéseket, de Hacker szemlátomást nem állítja ezt. Ha pedig a kérdést Hacker oly módon intézi el, hogy rámutat arra, hogy a tégl (vagy csiga) nem tud olyan funkcionális állapotokba kerülni, amilyenekbe az ember akkor, amikor éppen kellemes érzései vagy fájdalmai vannak, és ezen az alapon vitatható el a téglától és tulajdonítható az embernek a fájdalom vagy a kellemes érzések átélésének képessége, akkor ennek a

válasznak lehetséges revizionista olvasata (a téгла nem képes fájdalmat érezni vagy kellemes élményeket élvezni, az ember viszont igen, ám ez csupán annyit jelent, hogy az ember képes F funkcionális állapotba kerülni, míg a téгла nem), de antirevizionista olvasata is (a téгла nem képes fájdalmat érezni vagy kellemes élményeket élvezni, az ember viszont igen, erre ugyanis az olyan lények képesek, melyek képesek F funkcionális állapotba kerülni, mint például az ember).

2.4 Illuzionizmus

A 2010-es években népszerűvé vált az elmefilozófiában az illuzionizmus, mely szemlátomást határozottan kiáll a „tudat tagadása” mellett: a fenomenális tudatosság az illuzionizmus szerint nem létezik, csupán úgy tűnik számunkra, hogy létezik. Így például Dennett szerint:

robotok vagyunk, amelyek robotokból állnak, amelyek robotokból állnak... melyeknek együttesen sikerül létrehozniuk a Tudatos Személy felhasználói illúzióját, egy egységes ágensét, a selfét mint narratív tömegközpontét. Nekünk kívülállóknak arra kell a felhasználói illúzió, hogy leírjunk másokat[...], egymást[...], és önmagunkat.”(Dennett 2019: 55)

Az irányzat bemutatása érdekében a következőkben röviden összefoglalom Keith Frankish híres tanulmányának gondolatmenetét, mely nagy vitákat váltott ki az elmúlt években (Frankish 2016).

Frankish szerint a tudat filozófiai elméletei vagy valóságosnak tartják a fenomenális tudatosságot és célul tűzik ki megmagyarázását, vagy pedig illúziónak tartják, és azt tartják céljuknak, hogy megmagyarázzák, miért tűnik úgy, hogy létezik a fenomenális tudatosság. Ez utóbbi megközelítést lehetne (régebbi kifejezéssel) eliminativizmusnak nevezni, Frankish azonban inkább az illuzionizmus nevet javasolja. Az illuzionizmus szerint a fenomenális tudatosság feltételezése nem pusztán elméleti tévedés, hanem illúzió; továbbá az illuzionizmus nem feltétlenül osztja az eliminativizmus bizonyos törekvéseit, például azt, hogy a fenomenális tudatosság realitásának elutasítására a népi pszichológia elutasításának részeként kerüljön sor, vagy azt, hogy a

fenomenális tulajdonságokról szóló beszédet ki kellene küszöbölni. Frankish szerint az „irrealizmus” sem jó elnevezés, mert az illúziók nagyon is valóságosak. Az illuzionizmus célja megmagyarázni, hogy miért tűnik úgy, hogy tapasztalatainknak fenomenális tulajdonságai vannak – ez az *illúzió-probléma* kerül a „nehéz probléma” helyére.

Mielőtt közelebbről szemügyre vennénk Frankish nézeteit, két megjegyzést lehet tenni. Az egyik, hogy mint Frankish is jelzi, Place és Smart is illuzionistának vagy az illuzionizmus előfutárának nevezhető, mert tagadták a fenomenális mezőt és fenomenális objektumokat. A másik, hogy mint Sytsma és az általa idézett kutatók kísérleti filozófiai felmérései rámutattak, nem annyira egyértelmű, hogy fenomenálisnak tűnik a tapasztalat legalábbis azon nem-filozófusok számára, akik nem jártasak az elmefilozófiai irodalomban (Sytsma 2010: 115). Nem olyan egyszerű tehát a helyzet, hogy mindenki számára nyilvánvaló, égető problémára keresné a választ az illuzionizmus.

Frankish szerint amikor valamilyen anomáliával találkozunk, mely nem illeszkedik tudományos világgépünkbe, akkor erre háromféleképpen reagálhatunk. Elfogadhatjuk a jelenség realitását és megkísérelhetjük a tudományos világgépünk kibővítését, hogy ezt befogadja, érvelhetünk amellet is, hogy az új jelenség nem is anomália, és illeszkedik a meglévő tudományhoz, végül pedig minősíthetjük a jelenséget illuzórikusnak és megpróbálhatjuk kikutatni, hogyan keletkezik az illúzió. (Megjegyzendő, hogy mivel az illúzió megmagyarázása azt jelenti, hogy a jelenséget a meglévő tudományos ismeretek segítségével magyarázzuk, ezért a két utóbbi opció közötti különbség nem teljesen tűnik élesnek.)

Frankish most ezt a sémát alkalmazza a tudatos tapasztalat kérdésére. A tudatos tapasztalatnak „szubjektív aspektusa” van, azaz „valamilyen” például színeket látni. Ezt Frankish úgy értelmezi, hogy a tudatos tapasztalatnak introspekcióban hozzáférhető kvalitatív tulajdonságai vannak, röviden, „fenomenális tulajdonságok”. Ezek Frankish szerint anomalózusak, mert úgy tűnik, nem fizikai tulajdonságok, és nem magyarázhatók meg fizikai terminusokban. Háromféleképpen lehet viszonyulni hozzájuk Frankish szerint.

A *radikális realizmus* szerint a fenomenális tulajdonságok valóságosak, és megmagyarázhatatlanok jelentős elméleti innováció nélkül. Ilyen a dualizmus, neutrális monizmus, vagy az új fizikát javasoló irányzatok. A *konzervatív realizmus* elfogadja a fenomenális tudatosság realitását, és fizikai magyarázatot keres rá. A legtöbb fizikalista ebbe a táborba tartozik. Végül, a harmadik lehetőség az *illuzionizmus*. Az illuzionista szerint a tapasztalatnak nincsenek fenomenális tulajdonságai, és azt kell megmagyarázni, miért tűnik úgy, mintha lennének, hogyan keletkeznek hamis ítéleteink a létezésükkel kapcsolatban.

Az *erős illuzionizmus* szerint a tapasztalatnak igazából nincs kvalitatív, „milyen” (what it is like) karaktere. A gyenge illuzionizmus nem teszi meg ezt a radikális állítást, csak azt, hogy a fenomenális tudatosság némely tulajdonsága illuzórikus.

Frankish szerint a gyenge illuzionizmus (és a konzervatív realizmus) nem stabil pozíció, mert ha a konzervatív realisták nem tudnak előállni a fenomenálisnak olyan erős fogalmával, ami erősebb, mint a kvázi-fenomenálisé (kvázi-fenomenális az a tulajdonság, amit az introspekció fenomenálisként reprezentál), de még nem ütközik a konzervatív realizmus korlátaiba, akkor opciókként a radikális realizmus és az erős illuzionizmus maradnak a porondon.

Ismét felmerül a kérdés, hogy tényleg magától értetődő-e az, hogy introspekciónk fenomenálisként reprezentál némely tulajdonságokat, vagy csak az elmefilozófiai diskurzusban járatos filozófusok számára nyilvánvaló ez. Még lényegesebb probléma, hogy a fenomenális tulajdonság kellően erős fogalmának szerepére jó jelöltnek tűnnek a fizikai tulajdonságok. Egy reduktív fizikalista (tehát nem funkcionalista), aki a fenomenális tulajdonságokat fizikaiakkal azonosíthatja, kitarthat a konzervatív realizmus mellett. Természetesen a reduktív materialistának is sok dolga lehet azzal, hogy megmagyarázza, miért éppen olyan a tapasztalat minősége, amilyen (abban az értelemben, hogy miért úgy strukturálódik, ahogyan).

Frankish felsorol néhányat az illuzionisták által elméletük illusztrálására alkalmazott analógiák közül. Dennett például a felhasználói illúzió metaforáját alkalmazta már 1991-es könyvében is. Az introspektív világunkat benépesítő tárgyak olyasmi státusszal bírnak, mint egy számítógép „asztalán” lévő „mappák” és „ikonok”. A lényeg az, hogy az introspekció saját bonyolult fizikai tulajdonságainkat egyszerű fenomenális tulajdonságokká torzítva reprezentálja.

De mi a helyzet a külvilágból érkező ingerekkel? Frankish szerint a tapasztalat transzparenciájának tézise felveti a „kifelé néző illuzionizmus” lehetőségét, mely szerint tapasztalatunk a külvilági ingereket is látszólag fenomenális tulajdonságokkal rendelkezőkké torzítja.

Tanulmánya 1.7 fejezetében Frankish felteszi és megválaszolja a kérdést, hogy az illuzionisták azt állítják-e, hogy fenomenális zombik vagyunk. Frankish válasza:

Ha az egyetlen dolog, ami a zombikból hiányzik, a fenomenális tudatosság... akkor... zombik vagyunk. Csakhogy a zombikat tőlünk nagyon különböző lényekként szokták bemutatni – nincsen belső életük[...] ahogy Chalmers fogalmazott, „Nincs olyan, hogy

amilyen zombinak lenni[...] minden sötét belül”[...] az illuzionisták [szerint] ez nem jó leírás rólunk. Inkább [...] azt fogják mondani, hogy az azzal a belső étellel való rendelkezés, ami nekünk van, egyfajta introspektív öntudatban (self-awareness) áll, amely megteremti a gazdag fenomenológia illúzióját. De nem éppen a fenomenális tulajdonságok azok, amelyek a tapasztalást *valamilyenné* teszik? [...] Az illuzionisták azt mondhatják, hogy tapasztalataink akkor *valamilyenek*, ha tudomásunk van róluk egy funkcionális értelemben[...] (Frankish 2016: 9)

Vagyis Frankish szerint a tapasztalat úgy *valamilyen* egy lény számára, ahogy a külvilágbeli tárgyak is *valamilyenek* lehetnek számára. Ebben az értelemben *valamilyen* számunkra a tapasztalat, állapítja meg Frankish, és ebben az értelemben *valamilyen* zombinak lenni.

Vajon szegényesebb-e vagy más módon más az a *milyenség*, amivel Frankish zombijainak (vagyis nekünk a) tapasztalatai rendelkeznek, mint amilyeneknek általában elképzeljük az emberi tapasztalatot? A Frankish-féle zombik egy „*valamilyen*” külvilág tárgyairól szereznek ingereket és különféle funkcionális mechanizmusok révén ezeket feldolgozzák. Vajon nem kellene-e feltételeznünk ahhoz, hogy azt hihessük, semmilyen Frankish zombijának lenni, hogy az őt körülvevő világ is „*semmilyen*”, bármit is jelentsen ez? Más szavakkal: ha a világ „*valamilyen*” (rendelkezik tulajdonságokkal, van természete), akkor a benne zajló információáramlások, mozgások, tevékenységek is „*valamilyenek*”, és ha egy komplex felépítésű élőlény különféle bonyolult megismerő és célirányos cselekvéseket végez ebben a világban, akkor nehéz megérteni, miért ne lenne „*valamilyen*” az ő nézőpontjában lenni. Más szóval, amikor Frankish esetleg eltakarít az útból néhány nehezen kezelhető, avítt filozófiai fogalmat, biztosan valami radikálisat állít arról, hogy milyen érzés élni? Frankish nem mond olyat álláspontja világossá tétele érdekében, hogy „*olyan embernek lenni, mint egy darab kőnek lenni*”, hiszen feltehetően ezt nem is így véli.

Különös módon Frankish nem sok szót veszteget arra a lehetőségre, hogy éppúgy, ahogy ha a külvilág tárgyai *valamilyenek* (ugyanis rendelkeznek *valamilyen* fizikai természettel), akkor számunkra is *valamilyenek* tudnak lenni (kölcönhatásokba lépve érzékszerveinkkel), „*belső*” állapotaink (testi érzeteink, agyállapotaink) is „*valamilyenek*”, hiszen rendelkeznek *valamilyen* fizikai természettel. Hogy hogyan szerzünk ismereteket környezetünkről és önmagunkról, az különféle funkcionális magyarázatok tárgya lehet, de ezekben a magyarázatokban szerepelni fognak a környezet, illetve az emberi szervezet fizikai jellemzői, vagyis „*milyensége*” is. Egy materialista

vagy akár neutrális monista számára ebben a képben elfér a tapasztalat „milyensége” anélkül, hogy ez bármiféle rejtély forrása lenne.

Frankish az illuzionizmus mellett a tudományos konzervativizmus elveit hozza fel: minél kevésbé radikális lépéseket kell tennünk a tudattal kapcsolatos problémák megválaszolása érdekében, noha nem világos, hogy például egy egyszerű fizikai-mentális vagy fizikai-fenomenális azonosságelmélet, ami semmilyen új feltételezéssel nem bővíti a fizikát, mennyiben tekinthető radikálisabb elképzelésnek, mint az az egyébként érthetetlen tézis, hogy „a világ végső soron *semmilyen*”.

A radikális elméletekkel szemben az epifenomenalizmus problémájával érvel Frankish, de érdekesebb a vitája a konzervatív realizmussal, mely azonosítja a fenomenális tulajdonságokat fizikaiakkal.

Frankish szerint a konzervatív realizmus nem stabil álláspont, hanem hajlamos beleolvadni az illuzionizmusba. Ennek oka, hogy a fenomenális tulajdonságok Frankish szerint „túlságosan furcsának tűnnek” ahhoz, hogy fizikai magyarázatot kaphassanak, ugyanakkor ellenállnak a funkcionális elemzéseknek is és függetlennek tűnnek az őket megmagyarázni hivatott fizikai mechanizmusoktól.

Erre válaszul azt lehetne mondani, hogy akár a fordítottja is igaz lehet: az erős illuzionizmus hajlamos beleolvadni a gyengébe, mert az illuzionista tézis „túl furcsának tűnik”, hogy elfogadják, ugyanakkor a fizikai és fenomenális nyilvánvalóan csak akkor azonosítható, ha elfogadunk gyenge illuzionista téziseket arról, hogy a fenomenális tulajdonságok nem olyanok, mint amilyenek azok az állítólagos „antifizikalista intuíciók” mutatják őket, melyek egyáltalán a tudat-agy problémát okozzák. E mögött az inkább csak retorikailag érdekes válasz mögött azonban felsejlik az a mélyebb probléma, hogy Frankish meglehetősen sokat beszél, az illuzionizmust motiválandó, a fenomenális tulajdonságok anomáliás, „furcsa” jellegéről. Sőt, mint látni fogjuk, azt is elfogadja, hogy természetes „antifizikalista intuícióink” vannak.

Természetesen a filozófusok kiindulhatnak „antifizikalista intuíciókból” mint premisszákból, de ha ígéretesebb egy másik kiindulópont, akkor nem kötelező ragaszkodniuk ezekhez. Az „antifizikalista intuíciókból” való kiindulás mellett azonban nem jó érv, hogy ezek az intuíciók természetesek, noha kétségkívül hihetőnek tűnik, hogy némelyik filozófiailag képzetlen ember természetes meggyőződése antifizikalista, például Lewis egy alkalommal az „utca emberének dualizmusáról” beszél (Lewis 1966:25). Másrészt sem a filozófusok, sem a laikusok intuitív meggyőződései nem egységesek (hiszen vannak revizionisták és antirevizionisták, amire

felfigyelhetünk a szakirodalomban vagy éppen a már idézett „kísérleti filozófiai” felmérések eredményei között), és az pedig végképp nem igaz, hogy a természetes meggyőződések kifejlett filozófiai doktrínák lennének. Ezért érdemes megfontolni, hogy azok az intuíciók, amelyeket gyakran antifizikalistának vagy dualistáknak neveznek, esetleg máshogyan is interpretálhatók, tehát nem antifizikalista, hanem például antifunkcionalista intuíciókként, vagy más módon. Továbbá azt sem szabad szem elől téveszteni, hogy nagyon sok a természetesnek tartott meggyőződések közül modern elméleti újítások eredményeként jött létre, ha nem is közvetlen filozófiai „indoktrináció” hatására, de minden esetre egy olyan kultúrán belül, amire hatással vannak a filozófiai elméletek. Ez nem különösebben merész feltevés, hiszen számos filozófus (például Dennett) kárhóztatja más filozófusok vagy tudósok „kartezianus” elképzeléseit, melyek olykor úgy kerülnek az elméletekbe, hogy szinte észrevétlenül lopódnak be. Ryle „hivatalos doktrínának” (Ryle 1949/1999: 17) nevezte a hiedelmeknek azt a Descartes hatására kialakult komplexumát, amely sok téren meghatározta szerinte az elméről való gondolkodást a filozófusoktól a laikusokig. Dennett lényegében Ryle tevékenységét folytatja, amikor feltárja, hogy számos széles körben magától értetődőnek gondolt elképzelés igazából a „hivatalos doktrína” maradványa.

Antifizikalista intuíciókról beszélni azonban gyakran problematikus: kétségtelenül vannak bizonyos erős intuícióink (korántsem egységesen), amelyeket azután egyes filozófusok antifizikalistaként interpretálnak, de sok esetben ez az interpretáció vitatható. Az „utca emberének” egyik klasszikusan emlegetett reakciója az agyat az elmével azonosító materializmusra, hogy „lehetetlen, hogy az, hogy egy rózsát látok, azonos azzal, hogy másfél kilogramm szürke fehérjetömegben valami láthatatlan, színtelen elektrokémiai folyamatok zajlanak”. Ezt azonban túlzás lenne antifizikalista intuíciónak nevezni, mert ez a reakció antifizikalista érvként azon a teljesen elemi hibán alapul, hogy két, eltérő szemléletes kép azonosításából következtet arra, hogy nem lehet azonos a fizikailag leírt folyamat annak valamelyik szemléletes képével, ha viszont nem antifizikalista érvként olvassuk, akkor teljes mértékben megállja a helyét az, hogy a rózsá látványa valóban nem egy szürke fehérjetömeg látványával azonos, persze a fizikalisták nem is ezt állítják (Feigl 1958: 455).

E kitérő után folytassuk Frankish gondolatmenetének vizsgálatát. Sok filozófus szerint antifizikalista intuícióink forrása az a mód, ahogyan a fenomenális tulajdonságokat konceptualizáljuk introspekciónkban. Vannak „fenomenális fogalmaink”, amelyek nem kapcsolódnak a priori a fizikai fogalmakhoz, vagyis fogalmilag izoláltak, emiatt, noha a fenomenális tulajdonságok fizikai tulajdonságok, ez nem tűnik így, hanem antifizikalista intuícióink keletkeznek. Ez a *fenomenális fogalom stratégia*.

Frankish szerint a stratégia nem oldja meg a konzervatív realizmus problémáit, mert a kérdéses fogalmak továbbra is *fenomenálisak*, nem pedig pusztán „introspekcióban megjelenő *valamik*”. Ha *úgy* konceptualizálnánk a tapasztalat tulajdonságait, akkor nem éreznénk semmi ellenállást azzal szemben, hogy azok fizikaiak.

Ismét jelezni kell, hogy egyrészt az olyan szerzők szerint, mint például Smart (1959), éppen arról van szó, hogy a tárgyközömbös fordítások kiiktatnak bármilyen, nem-fizikai tulajdonságokra referálni látszó fenomenális fogalmakat, másrészt pedig nem kell túlzott jelentőséget tulajdonítani az élmények tulajdonságait fizikai tulajdonságokkal való azonosítással szembeni „ellenállással”, mert könnyen lehet, hogy egyszerű félreértésen alapul (lásd korábban a Feigl által elemzett félreértést a szemléletességgel kapcsolatban).

Frankish szerint ha viszont a fenomenális fogalmak arra vonatkoznak, hogy „hogyan érződik” az élmény (nem „pusztán valamikre”), akkor a konzervatív realistáknak vagy meg kellene magyarázniuk, hogyan lehetnek ezek az érzések fizikaiak, vagy el kell fogadniuk, hogy a fenomenális fogalmak tévesen reprezentálják a tapasztalatot, ahogyan azt az illuzionisták állítják. Frankish szerint az eddig javasolt reduktív magyarázatok problémája, hogy a gyenge illuzionista nézetnek, mely szerint a fenomenális tulajdonságok ugyan nem privátak, intrinzikusak és ineffábilisak (mint a kválék Dennett (1988) által kritizált bevett felfogásában), de mégis valóságosak és valóságosan kvalitatívak, a fenomenalitás olyan fogalmára van szükségük, amely erősebb a kvázi-fenomenalitásnál. A fenomenális tulajdonságoknak tényleg „éreződniük” kell valahogyan, nem elég, hogy a fenomenalitás reprezentációit okozzák. Nem világos azonban, állapítja meg Frankish, hogy ez hogyan kivitelezhető. „Milyen fenomenális maradvány marad, ha már eltávolítottuk az olyan jellemzőket, mint a privátság, intrinzikalitás, és ineffabilitás?”

Frankish egyébként úgy véli, hogy a fizikalisták ellenállhatnak az illuzionizmus csábításának, de csak azon az áron, hogy bele kell nyugodniuk abba, hogy, amint Levine megállapította, magyarázati rés marad a fenomenális tulajdonságok és a nekik megfelelő agyi események között. (A „magyarázati rés” a következőt jelenti: ha elfogadjuk, hogy a fájdalom a C-rostok tüzelésével azonos, akkor ezzel ugyan kaptunk egy neurofiziológiai magyarázatot a fájdalom előállításának mechanizmusáról, és azt is mondhatjuk, hogy kiderítettük, igazából mi a fájdalom, de a C-rostok felfedezése nem magyarázza meg, hogy miért *olyannak* érződik a fájdalom, amilyennek érződik (Levine 1983: 357)). Frankish szerint azonban jobb egy megmagyarázható illúzió, mint egy megmagyarázhatatlan realitás.

A konzervatív realizmus elleni érvelés után Frankish rátér az illuzionizmus melletti pozitív érvelésre (tanulmánya 2.3. fejezetében). Az első érve, mely az antifizikalista elméletek ellen irányul, az, hogy a fenomenális zombik ugyanazokat az állításokat tennék a fenomenális tapasztalatról és a tudatosságról, mint mi. Ezért a mi állításaink a tudatról nem képeznek bizonyítékot a fenomenális realizmus mellett, ezért az hamis. Ez az érv nem meggyőző az olyan fizikalista elméletekkel szemben, amelyek nem ismerik el a zombik lehetőségességét.

Frankish második érve szerint „általában véve az anomáliás jelleg az illuzórikusság bizonyítéka. Ha egy tulajdonság ellenáll a fizikai terminusokban való magyarázatoknak vagy csak bizonyos perspektívából érzékelhető, akkor a legegyszerűbb magyarázat erre az, hogy illuzórikus”. Továbbá, ha esély van arra, hogy tévedhetünk a fenomenális tudatosság létezésével kapcsolatban, akkor ez erős érv amellest, hogy ténylegesen tévedünk vele kapcsolatban, mert nincs introspektív eszközünk, amivel ellenőrizhetjük introspektív reprezentációink helyességét. Tehát lehet illuzórikus, ráadásul anomáliás, amiből arra következtethetünk, hogy illuzórikus.

Az illuzionizmus előnye Frankish szerint, hogy bizonyos értelemben komolyabban veszi a tudatot, mint a realisták, mert nyugodtan elismerheti, hogy „mennyire mágikusak és földöntúliak” a fenomenális tulajdonságok és tárgyak. Lehet, hogy így van, de a fenomenális tudatosság valószínűleg csak annak mágikus és földöntúli, akinek anomáliás – és nem mindenkinek az, csak annak, aki megfelelő – például karteziánus – filozófiai „neveltetést” kapott.

Végül, Frankish szerint az illuzionizmus előnye az is, hogy adaptív tulajdonságként képes megmagyarázni a tudatosságot. Lehetséges, hogy a tudatosság funkciója éppen az, hogy azt a benyomást adja nekünk, hogy „mágikus, nem-fizikai belső életünk van”.

Ez utóbbi fejtegetések jelentősége az, hogy az illuzionisták nagyon komolyan veszik azokat a különféle intuitív meggyőződéseket, amelyek alapján az emberek kialakítják elvárásaikat az elmefilozófiai elméletekkel szemben. Akinek alapvető meggyőződése a tudatosság létezése (tehát antirevizionista), annak számára az elképzelések szerint az illuzionizmus nem irreleváns vagy hiányos elméletnek tűnik, hanem olyannak, amely alternatív magyarázatot ad arra, amit az illető megtapasztal. Akinek nincs ilyen meggyőződése (tehát revizionista), annak számára az illuzionizmus egy érdekes elmélet az elmeműködés bizonyos részletkérdéseiről. Csak az marad kérdéses, hogy az antirevizionisták számára valaha is hihető lesz-e egy olyan elmélet, amely a tudatosságot tévedésnek állítja be. A tipikus antirevizionista reakció ugyanis erre az, hogy a tudatosságot „közvetlenül” tapasztaljuk meg, „átéljük”, nem pedig feltételezzük vagy gondolkodunk róla.

Az illuzionizmus védelmében (tanulmánya 3. fejezetében) Frankish sorra veszi az illuzionizmussal szembeni gyakori ellenvezéseket. Az első jellegzetes kifogás az, hogy az illuzionista egyszerűen letagadja azt, amit magyaráznia kellene. A fenomenális tudatosság nem valami elméleti feltevés, mely más jelenségek megmagyarázására szolgál, hanem maga az adottság, ami magyarázatra szorul. Ám ha fenomenális tudatosság tényleg ilyen módon adott, akkor, mint Frankish rámutat, valamilyen episztemikus hozzáféréssel kell bírunk hozzá, ami lehetőleg a tévedést is kizárja. Mivel pedig kauzális folyamat ilyen bizonyosságot nem biztosíthat, a realistáknak Frankish szerint azt kell mondaniuk, hogy a fenomenálishoz való episztemikus hozzáférés nem kauzálisan közvetített, hanem például a „közvetlen ismeretség” viszonya.

Frankish szerint nagy árat kell fizetni a fenomenális tudatosságnak az így megvédett „adottság” státuszáért. Először is, a „közvetlen ismeretségben” megismert tulajdonságoknak nem lehet pszichológiai jelentőségük, például nem kommunikálhatók. Másodsor, az ismeretség-elmélet feltételezi, hogy az érzetek fenomenális minőségét nem befolyásolják a rájuk adott reakciók, ami ellentmondani látszik az ezzel kapcsolatos tudásunknak. Harmadsor, az ismeretség-elmélet valószínűleg összeegyeztethetetlen a fizikalizmussal.

Az, hogy a közvetlen ismeretség elmélete nem fér be a fizikalizmus keretei közé, igen elterjedt elképzelés (lásd például Feyerabend 1970). Frankish idézi Chalmersnek egy állítását, mely szerint a fenomenális tapasztalat „intrinzikusan episztemikus”, ami a karteziánus „hivatalos doktrínának” azt a tételét idézi, hogy a tudattartalmak „önélárulók” vagy „foszforeszcensek” (Ryle 1949/1999: 200), bár rövidesen megvizsgáljuk a közvetlen ismeretségnek az olyan deflacionárius értelmezéseit, amivel kapcsolatban ez az ellenvetés nem áll meg, legfeljebb az vonható kétségbe, hogy a felvázolt reláció valóban méltó-e a „közvetlen ismeretség” névre.

Ami a fenomenális kauzális inertségét és közölhetetlenségét illeti, nem teljesen világos, hogy ez önmagában miért lenne probléma. Egy reduktív materialista mondhatja azt, hogy amikor beszámolunk aktuális mentális állapotunkról, akkor egy kauzális-funkcionális állapotúnak egy olyan példányáról számolunk be, melyet aktuálisan éppen egy bizonyos fizikai állapotunk realizál.

Az ezzel a fizikai állapottal való rendelkezést, vagyis azt, hogy ez a fizikai állapot ténylegesen realizálja mentális állapotunkat, tekinthetjük az ennek az állapotnak a fizikai tulajdonságaival való tényleges „közvetlen ismeretség” egy minimálisan, deflacionáriusan felfogott esetének.

Ha mentális állapotunkat másfajta fizikai állapot realizálná (az egyszerűség kedvéért most feltesszük ennek lehetőségét), akkor is ugyanezeket a szavakat mondanánk ki, ennyiben tehát a

realizáció pontos mikéntje, tehát az, hogy egy meghatározott funkcionális állapotot milyen fizikai állapot realizál, nem kommunikálható, és a viselkedésünkben sem érhető tetten ³

Ennek ellenére igaz maradna az, hogy csak az első személyi nézőpontból hozzáférhető az élménynek az az aspektusa, amivel a „közvetlen ismeretség” viszonyában vagyunk – abban a minden misztikumtól mentes értelemben, hogy ha a „közvetlen ismeretség” egyenlő az „átéléssel” (vagyis a tulajdonsággal való tényleges rendelkezéssel), akkor sehogyan máshogyan nem lehet egy élménnyel ismeretségbe kerülni, mint az átélésével, abban a triviális értelemben, ahogy „Lengyelország sem rendelkezhet Bulgária történetével” (Ryle 1949/1999: 258).

Végső soron mindegy, hogy elfogadjuk-e a „közvetlen ismeretség” egy lehetséges modelljének azt a deflacionárius koncepcióját, amit akár egy reduktív materialista is elfogadhat: ha a funkcionális szerepeknek fizikai állapotok általi realizációját nem is nevezzük az ezekkel az állapotokkal való közvetlen ismeretségnek, akkor is elfogadhatja egy reduktív materialista a fenomenális pszichológiai inertségét és kommunikálhatatlanságát (mint azt az 1950-es évek reduktív materialista elmefilozófiai irodalma és ennek közvetlen filozófiai behaviorista előzményei mutatják).

Tehát ha a reduktív materialistákat a konzervatív realisták közé soroljuk, hiszen egy (mégoly deflacionárius) közvetlenismeretség-felfogás végső soron tulajdonítható nekik, akkor elmondható róluk, hogy készek megfizetni a Frankish által túl magasnak tartott árakat. Míg ha álláspontjukat (legalábbis gyenge) illuzionistának tekintjük (mint ahogy Frankish közéjük is sorolja Place-t és Smartot), akkor persze nincs értelme a „közvetlen ismeretség” problémájáról beszélni az esetükben, de a fenomenális inertségét és kommunikálhatatlanságát ekkor is elfogadják.

Figyeljük meg, mi történik, ha túl magas árak itéljük a fenomenális oksági hatástalanságát és kommunikálhatatlanságát, vagy fenomenális minőségük elválaszthatóságát a rájuk adott reakcióktól, és ezért az illuzionizmus mellett döntünk. A fenomenális helyére a kvázifenomenális kerül, amelynek bőségesen van funkcionális szerepe és nagyon is kommunikálható. Ez egyszerűen funkcionalizmushoz vezet, de azzal a kikötéssel, hogy ennek a funkcionalizmusnak az A-olvasata nem lehet az, hogy ahol a funkcionális specifikáció ki van elégítve, ott mindig „megjelenik a tudat”, hanem legfeljebb csak az, hogy az elmélet elsiklik a tudat kérdése mellett. Más szóval Frankish illuzionizmusát úgy is felfoghatjuk, mint a funkcionalizmus olyan finomítását, amely elzárja az

³ Feltéve, hogy a *viselkedést* olyan módon értelmezzük, hogy ne legyen tetten érhető benne. Ha a *viselkedés*be beletartozik a mikroszkopikus viselkedés is, és/vagy a koponyán belüli történések is, akkor a realizáció különbsége megmutatkozik valahol a viselkedésben.

egyik lehetséges antirevizionista értelmezés útját, vagyis valódi erőfeszítés a kétértelműség elkerülésére és a revizionista olvasat kikényszerítésére.

Amikor Frankish azt elemzi, hogy mit jelent az, hogy *valamilyen* nekünk lenni, erre kétféle értelmezést ad. Utalhat ez a kifejezés az introspekción szubjektívására (mely egy funkcionális tulajdonság). Frankish szerint az illuzionisták elfogadják az introspektív szubjektívást, csak „radikálisan félrevezetőnek” tartják. A „valamilyen valakinek lenni”-nek azonban van Frankish szerint egy erősebb értelmezése is, ez az *intrinzikus szubjektívás*, ami „abból ered, hogy azok a dolgok vagyunk, amik”. Frankish szerint a fenomenális tulajdonságokkal való „közvetlen ismeretség” is az intrinzikus szubjektíváshoz kapcsolódik: a fenomenális tulajdonságok fizikai felépítésünk „korrelátumai”. Frankish szerint azonban az intrinzikus szubjektívás „mélyen rejtélyes”, mivel a fizikai rendszerekhez „árnyékszerűen járuló társ”, amellyel bármilyen objektum rendelkezhet, ahogy a pánpszichisták is gondolják. Frankish konklúziója az, hogy az intrinzikus szubjektívás fogalma talán nem inkoherens, de ahhoz elég rejtélyes, hogy előnyben részesítsünk egy illuzionista magyarázatot rá.

Am ha a közvetlen ismeretség tárgyai nem a fizikai felépítésünk „korrelátumai”, hanem maguk a fizikai tulajdonságaink, és az intrinzikus kvalitatívás egyszerűen minden fizikai rendszerre jellemző (amíg elfogadjuk, hogy a fizikai állapotoknak és folyamatoknak vannak intrinzikus vagy legalább az általuk realizált, például funkcionális állapotokként felfogott mentális állapotokhoz képest „relatív intrinzikus” jellemzői), akkor aligha kell rejtélyes létezőkben és tulajdonságokban hinnünk ahhoz, hogy elfogadjuk az „intrinzikus kvalitatívás” lehetőségét.

Frankish szerint az illuzionizmus szerte tudja foszlatni a chalmersi „nehéz problémát”, pontosabban azt lecseréli az illúzióproblémára: már nem azt kell megmagyarázni, hogy miért kíséri életünket fenomenális élmény, és az miért éppen olyan, amilyen, hanem azt kell megmagyarázni, hogy hogyan és miért keletkezik a fenomenalitás illúziója, és miért ilyen erős. Frankish szerint bár az illúzióprobléma sem éppen könnyű, nem olyan különlegesen nehéz, mint amilyenek a nehéz problémát szokták tartani.

Ez azonban csak akkor szól igazán az illuzionizmus mellett, ha egyáltalán elismerjük a nehéz problémát és különlegesen nehéznek is tartjuk. Ha a nehéz probléma elutasítása vagy könnyebbekkel való helyettesítése a cél, akkor az illuzionizmusnak vannak alternatívái.

2.5 Összefoglalás: az elméletek kétértelmősége

Ebben a fejezetben olyan elmefilozófiai elméleteket vizsgáltunk meg, melyeket a revizionista vagy az antirevizionista elmefilozófia paradigmaticus eseteinek tarthatunk: többek között az egyik leghíresebb revizionista elmefilozófus, Dennett, valamint a funkcionalizmus híres kritikusa, az antirevizionista Block bizonyos érveit.

Mindegyik vizsgált elmélet esetében azonosítható volt, hogy szerzőjük inkább revizionista, vagy antirevizionista meggyőződésű-e. Ám érvelésük vizsgálata során fény derült arra, hogy mindegyik elméletet lehetséges kétféle olvasatban (revizionista és antirevizionista olvasatban) interpretálni, vagyis a szerzők nem tudják kikényszeríteni egyik vagy másik olvasatot. Ennek következménye, hogy az antirevizionista filozófiai elméletek semmiképpen sem tudják meggyőzni a revizionista olvasókat az antirevizionista meggyőződés helyességéről, és fordítva. A revizionista olvasó sohasem fog egy antirevizionista filozófiai tanulmányban olyan érveket találni, melyek megingatják intuitív meggyőződéseit, mert a revizionista olvasó revizionista olvasatban értelmezi az antirevizionistának szánt szöveget. Megfordítva, az antirevizionista olvasó antirevizionista módon interpretálja a revizionistának szánt szövegeket, így a revizionista érvelések nem fogják meggyőzni a revizionizmus helyességéről.

A következőkben röviden összefoglalom, hogy mi a főbb elmefilozófiai elméletek A- és R-olvasata, majd pedig azt a kérdést vizsgálom meg, mi történik akkor, amikor antirevizionista és revizionista filozófusok vitatkoznak egymással.

A-dualizmus: Elismeri a fenomenális tudatosság létezését, és azt a célt tűzi maga elé, hogy megmagyarázza ezt. Ennek érdekében nem-fizikai tulajdonságok vagy entitások létezését feltételezi.

R-dualizmus: Nem a fenomenális tudatosság megmagyarázása céljából, hanem a tudományos tudás hiányosságai miatt (például azért, mert el sem tudja képzelni, hogy egy olyan szerv, mint az agy, képes lehet a gondolkodásra) feltételezi egy titokzatos nem-anyagi entitás vagy tulajdonság létezését, hogy ezzel betömje a magyarázati hézagokat (Cottingham 2016: 67). A „nehéz problémával” szemben megfogalmazott vitalizmus-analógia a dualizmusnak efféle törekvést tulajdonít.

R-eliminativizmus: „nincsenek mentális állapotok” vagy „nincsenek kválék”, tehát elme vagy tudat sincs.

A-eliminativizmus: „nincsenek mentális állapotok” vagy „nincsenek kválék”, tehát mentálisállapot- és kváléfogalmaink radikálisan inadekvátak annak leírására, ami valójában *van*. Az eliminativizmus nem eliminálja a tudatot vagy az elmét, hanem *demitologizálja*.

A-materializmus: Place (1956), Smart (1959) és Feigl (1958) egy lehetséges olvasata (ami addig tűnik egyértelműnek, amíg nem térnek rá a fenomenális mezővel és fenomenális tulajdonságokkal kapcsolatos eliminativizmusra): a tudat létezik: egy folyamat az agyban; az érzetek léteznek: agybeli események.

R-materializmus: Rorty (1965) és Feyerabend (1970) prima facie olvasata (ami addig tűnik egyértelműnek, amíg nem a „demitologizáció” eseteként értelmezzük a „fenomenológia” elleni támadásukat): nem tudat van vagy érzetek vannak, *hanem* agybeli események.

A-misztérianizmus: a tudat olyan különleges jelenség, amit az emberi kognitív képességek korlátai miatt gondolkodásunkkal nem tudunk adekvátan megragadni

R-misztérianizmus: a tudatprobléma és a tudat fogalma olyan súlyos fogalomzavarok eredménye, amelyekből, minthogy az emberi kognitív képességek korlátozottságából fakadnak, nem tudunk kikeveredni.

R-behaviorizmus: nincs tudat, elme vagy érzet, csak viselkedés és beszéd – többek között beszéd a „tudat” nevű mitikus entitásról.

A-behaviorizmus: természetesen nyilvánvaló, hogy van tudat, de a tudományos vizsgálódások tárgya csak a viselkedés lehet, a tudatot figyelmen kívül hagyjuk.

R-illuzionizmus: Frankish álláspontja a saját értelmezése szerint: valójában csak a tudat illúziójával és ersatz kválékkal rendelkezünk, nem pedig tudatossággal és kválékkal.

A-illuzionizmus: minthogy rendelkezünk a tudatosság illúziójával és az ersatz kválékkal, van első személyi nézőpontunk és *valamilyen* nekünk lenni, csak éppen azok a kválék nem léteznek, amelyekről a „kvalofil” filozófusok beszélnek

Funkcionalizmus a tudatossággal kapcsolatban: Egy S mentális állapot akkor tudatos, ha kielégít egy meghatározott fajta funkcionális leírást.

A-funkcionalizmus a tudatossággal kapcsolatban: Léteznek tudatos állapotok abban az értelemben, ahogyan az antirevizionisták értik ezt a fogalmat. Ezek nem mások, mint azok az állapotok, melyek kielégítenek bizonyos fajta funkcionális leírásokat, melyek nem redukálhatók alapvetőbb szintű (például fizikai) leírásokra. Ha a megfelelő funkcionális szerepet valamilyen agyi (vagy mikrochip-beli) állapot betölti, akkor megvalósul a tudatos tapasztalat. Az A-funkcionalista, amennyiben fizikalista, non-reduktivista lesz abban az értelemben, hogy számára a tudatos állapotok bizonyos funkcionális leírásokat kielégítő agyállapotoknak vagy gépállapotoknak felelnek meg. Az A-funkcionalista a tudatos állapotok létezésének „tényéből” indul ki, és megkeresi azt a funkcionális állapotot, amely ennek megfelel, valamint azokat a fizikai állapotokat, melyek ezt implementálják.

R-funkcionalizmus a tudatossággal kapcsolatban: Nem léteznek tudatos állapotok azon túlmenően, hogy az elme képes bizonyos funkciók ellátására. Akik a tudatról, a tudat problémájáról vagy a tudat rejtélyéről beszélnek, valójában arról beszélnek, hogy milyen nehéz a tudomány mai állása szerint megmagyarázni, hogy ezeket a funkciókat hogyan látja el az agy (vagy egy számítógép). Az R-funkcionalista számára nem létezik a tudatos állapotok „ténye”, csupán abban az értelemben, hogy elfogadja, sokan (például számos filozófus) beszélnek ilyesmiről, melyen szemlátomást az elme valamilyen funkcionális teljesítményét értik.

Kissé kisarkítva akár azt is mondhatjuk, hogy a Block által elképzelt szituációra, amikor egymással rádió adó-vevőn kapcsolatot tartó kínaiak szimulálják egy emberi agy működését, egy A-funkcionalista azt fogja mondani, hogy a „Kína-agy” is tudatos, bármennyire hihetetlennek tűnik ez, míg az R-funkcionalista azt fogja mondani, hogy (antirevizionista értelemben) sem a Kína-agy, sem az emberi agy nem tudatos, ezért Block példájában semmi rendkívüli nincs.

A revizionista és antirevizionista filozófusok, amikor szembesülnek azzal a ténnyel, hogy egyes kollégáik az övékével ellentétes alapvető elmefilozófiai intuíció hatása alatt állnak, a következő álláspontokat dolgozhatják ki.

(A1) Az antirevizionista intuícióból (AI) kiindulva, tekintettel az alternatívája, a revizionista intuíció (RI) természetellenességére és abszurdítására, az antirevizionista tézisek (AT) nyilvánvalóan helyesnek foghatók fel, a velük ellentétes revizionista tézisek (RT) pedig nyilvánvalóan helytelenek, komolytalannak, képtelenek. Valójában nincs olyan filozófus, aki a revizionista intuíció hatása alatt áll és revizionista téziseket képvisel. Amennyiben mégis úgy tünne, hogy egy filozófus revizionista téziseket védelmez, ezt tehát nem lehet úgy értelmezni, mintha támadást intézne az antirevizionista tézisek ellen. Sokkal inkább arról van szó, hogy – mivel az elmefilozófiát szűkebben értelmezi (RT3 értelmében), ezért kutatásának szűkre szabott területén nem foglalkozik a tudattal, vagy csak olyan értelemben (RT1, RT2 értelmében), amely értelmezés megfelel korlátozott céljainak. A látszólag revizionista téziseket hangoztató filozófusok valójában nem cáfolják az antirevizionistákat, nem is vitatják antirevizionista téziseit, hanem másról beszélnek. A két tábor elbeszél egymás mellett.

(R1) A revizionista intuícióból (RI) kiindulva, tekintettel az alternatívája, az antirevizionista intuíció (AI) természetellenességére és abszurdítására – akik természetesnek tartják (AI)-t, valójában félreértik természetes meggyőződéseiket, amennyiben ezeket egy hamis elmélet alapján interpretálják – az antirevizionista tézisek (AT) nyilvánvalóan helytelenek: elavult, hibás tudományos elméletek filozófiai interpretációinak tekinthetők, vagy pedig motiválatlanok, vagyis jó ok nélkül feltételezték őket. Valójában azonban nincs olyan filozófus, aki az antirevizionista intuíció hatása alatt áll és ezért antirevizionista téziseket képvisel. Amennyiben mégis úgy tünne, hogy egy filozófus antirevizionista téziseket védelmez, azt tehát nem lehet úgy értelmezni, mintha valódi vitája lenne a revizionista tézisekkel. Inkább arról van szó, hogy – mivel elavultak a tudományos

ismeretei – ezért azt hiszi, hogy a pszichológia és a természettudomány nem képes megválaszolni bizonyos kérdéseket, és az így keletkező űrt a filozófiai spekulációnak kell betöltenie (AT3). Az antirevizionista szerint a tudomány nem képes megmagyarázni olyan jelenségeket, mint a tudat vagy a tapasztalat milyensége (AT1, AT2), mivel ezek különleges, „nehéz” problémák, melyeket a filozófia feladata megoldani. Valójában azonban az antirevizionista téziseket hangoztató filozófusnak nincs érdemi vitája a revizionistával, mert a tudományról alkotott más (téves, elavult) elképzelésből indul ki. A két tábor elbeszél egymás mellett, mert valójában nem arról vitatkoznak, hogy „a tudomány” (általában) képtelen-e megválaszolni bizonyos kérdéseket, szükségessé téve a filozófiai spekulációkat, hanem csupán szembeállítják egy olyan tudomány képét, amely az említett kérdéseket meg tudja válaszolni, egy olyannal, amely nem.

(A2) Az antirevizionista intuícióból (AI) kiindulva az (A2) állásponton lévő filozófusok számot vetnek az intuíciók sokféleségének lehetőségével. Nem gondolják azt, hogy (RI) természetellenes abban a szigorú értelemben, hogy (RI) valójában nem tekinthető senki természetes meggyőződésének. Ebből következően még ha maguk (AI) hatása alatt is állnak, és nehezen tudják beleélni magukat azoknak a helyzetébe, akik számára (RI) a természetes meggyőződés, elfogadják, hogy nem utasíthatják el az (RT) téziseket mint „nyilvánvalóan abszurd” vagy természetellenes téziseket. Elfogadják, hogy előfordulhat filozófus, aki a revizionista intuíció hatása alatt áll, és revizionista téziseket képvisel. Ebből az következik, hogy azt is elfogadják, hogy vannak filozófusok, akik a tudat vagy az élmény milyenségének realitásának kérdésében radikálisan más állásponton állnak, mint ők. Azonban éppen a radikálisan eltérő intuíciók fennállása egyben az érdemi vita gátja is: aligha tudjuk meggyőzni azt, aki a mieinkkel ellentétes alapvető meggyőződésekkel rendelkezik, így számára egészen más állítások „nyilvánvalóan igazak” vagy „nyilvánvaló képtelenségek”, mint számunkra. Amikor a pluralista antirevizionista filozófus szembesül a revizionista állásponttal, mindössze megállapítja az intuitív meggyőződések radikális különbségét, de érdemi vitát nem tud folytatni.

(R2) A revizionista intuícióból (RI) kiindulva az (R2) állásponton lévő filozófusok számot vetnek az intuíciók sokféleségének lehetőségével. Nem gondolják azt, hogy (AI) természetellenes lenne, vagyis (AI) csupán természetes meggyőződéseinknek hibás elméletek alapján történő félreértelmezésének eredménye lenne. Ebből következően még ha maguk (RI) hatása alatt állnak is, és nehezen tudják beleélni magukat azoknak a helyzetébe, akik számára (AI) a természetes

meggyőződés, elfogadják, hogy nem utasíthatják el az (AT) téziseket mint „nyilvánvalóan abszurd” vagy természetellenes téziseket. Elfogadják, hogy előfordulhat filozófus, aki az antirevizionista intuíció hatása alatt áll, és antirevizionista téziseket képvisel. Ebből az következik, hogy azt is elfogadják, hogy vannak filozófusok, akik a tudat vagy az élmény milyenségének realitásának kérdésében radikálisan más állásponton állnak, mint ők. Azonban éppen a radikálisan eltérő intuíciók fennállása egyben az érdemi vita gátja is: aligha tudjuk meggyőzni azt, aki a mieinkkel ellentétes alapvető meggyőzésekkel rendelkezik, így számára egészen más állítások „nyilvánvalóan igazak” vagy „nyilvánvaló képtelenségek”, mint számunkra. Amikor a pluralista revizionista filozófus szembesül a revizionista állásponttal, mindössze megállapítja az intuitív meggyőzések radikális különbségét, de érdemi vitát nem tud folytatni. A1 és R1 tehát intuíciónomista, A2 és R2 intuíció-pluralista álláspontok.

3. További metafizikai megfontolások

3.1 Chalmers a tudat metaproblémájáról

Chalmers nemcsak a „nehéz probléma” megnevezésével, valamint az arra adott tipikus válaszok (és az azok mögött munkáló intuitív meggyőzések) azonosításával járult hozzá a tudatfilozófiai viták nagy léptékű metafizikai elmezéséhez, hanem a tudat „metaproblémájának” felvetésével is.

A tudat metaproblémája az a kérdés, hogy egyáltalán miért gondoljuk, hogy létezik a tudat problémája (Chalmers 2018: 6). A problémát olyan módon is meg lehet fogalmazni, hogy magyarázatra szorul: miért mutatnak az emberek bizonyos viselkedéseket – miért tesznek problémanyilatkozatokat (*problem reports*), mint például „a viselkedés magyarázása még nem magyarázza meg a tudatot”. Chalmers szerint számos adat gyűlt össze, mely az emberek „intuitív dualista” meggyőzéseiről tanúskodik (Chalmers 2018: 7). A metaprobléma egészen biztosan a „könnyű problémák” közé tartozik, hiszen csakis fizikai-funkcionális magyarázatot tesz szükségessé. Az sem kizárt, hogy a metaprobléma megoldása szertefosztatja a nehéz problémát (Chalmers 2018: 8). Ez például az illuzionisták álláspontja, akik szerint nem a tudatprobléma, hanem az illúzióprobléma vár megoldásra. Ha sikerül magyarázni a tudattal kapcsolatos

hiedelmeinket a tudatra való hivatkozás nélkül, az erős érv az illuzionizmus mellett (Chalmers 2018: 9). A nem-illuzionista elmefilozófia számára pedig felvetődik a metaprobléma-kihívás: ha egy elmélet szerint M mechanizmus a tudatosság alapja, akkor az elméletnek meg kell magyaráznia, hogy az M mechanizmus hogyan játszik központi szerepet a tudatosságról való ítéletek előidézésében (Chalmers 2018: 36). Így például a funkcionalizmust elutasító materialista vagy „biológiai” elméletek számára problémát jelenthet, hogy ezen elméletek szerint egy tudatos rendszer *ex hypothesi* nem tudatos, például szilíciumchipben implementált funkcionális duplikátuma osztani fogja a tudatos rendszer problémaintuícióit.

A probléma-nyilatkozatok tételére irányuló diszpozíciók a *probléma-intuíciók* (Chalmers 2018: 11). Ilyenek például a rés-intuíciók, melyek szerint magyarázati rések vannak a fizikai és tudati folyamatok között, és az antifunkcionalista intuíciók, melyek szerint a viselkedési funkciók megmagyarázása nem elégséges a tudat megmagyarázásához; metafizikai intuíciók, mint például a dualista intuíciók; tudásintuíciók a csak első személyi nézőpontból elnyerhető és az egyéb módon is hozzáférhető tudás viszonyára vonatkozóan; modális intuíciók azzal kapcsolatban, hogy mi elképzelhető vagy lehetséges. Chalmers főleg az első két csoportot (a magyarázati és metafizikai intuíciókat) tartja fontosnak. (Chalmers 2018: 12). Más elmefilozófiai is előfordulnak: Chalmers értékintuíciónak nevezi például azt az elképzelést, hogy a tudatosság az, amiért érdemes élni (utalva Sellarsnak Dennett elbeszélése szerint tett híres megjegyzésére).

Chalmers elismeri, hogy az általa felsorolt intuíciók elterjedtségének kérdése olyan empirikus kérdés, amelyről nem rendelkezünk sok adattal (Chalmers 2018: 14) Ez persze gyanúba keveri azt az elképzelést, hogy a problémaintuíciók keletkezésének mechanizmusának és a tudat keletkezésének mechanizmusának egy valamire való elmefilozófiai elméletben szorosan össze kell kapcsolódnia. Ha a tudatosság előállításának mechanizmusa csak minden második emberben állítja elő a problémaintuíciókat, akkor egy olyan elmélet szerint, amelyik azonosítja a problémaintuíciók és a tudat előállításának mechanizmusát, kétségbe vonható, hogy tudatos-e, akinek nincsen problémaintuíciója. Másfelől ha kitartunk amellett, hogy minden ember tudatos, miközben sokuknak nincsenek problémaintuíciói, akkor a problémaintuíciók elmefilozófiai jelentősége kérdőjeleződik meg.

Csak egyetérthetünk Chalmerssel abban, hogy a problémaintuíciók vizsgálata érdekes kutatási irány, és még sok empirikus kutatásra van szükség ezen a téren. Így például érdekes lehet a különféle problémaintuíciók függőségi viszonyainak tisztázása. Ha például A intuíció sohasem fordul elő B intuíció nélkül, de B előfordul A nélkül, akkor B intuíció tűnik alapvetőbbnek. Lehetséges, hogy a felszínen, a jelenlegi pszichológiai tesztekkel vizsgálható, vagy a filozófiai

álláspontokból leszűrhető intuitív meggyőződések alapvetőbb, de nehezen megfogalmazható meggyőződések kifejezésére tett nem mindig sikeres erőfeszítések eredményei.

Chalmers problémaintuíciónak felveti azt a kérdést, hogy amikor észlelünk egy problémaintuíciónak, azt mennyiben vagyunk jogosultak filozófiai álláspontként interpretálni. Chalmers többek között dualista és antifunkcionalista intuíciónakról beszél, vagy arról, hogy sok ember számára nyilvánvalóan problematikus az elme azonosítása „valami fizikaival”, mint amilyen az agy. Ha az utca embere azt állítja, hogy az elme „biztosan nem fizikai”, akkor ezt nem feltétlenül van jogunk úgy értelmezni, hogy a megszólaltatott riportalany antifizikalista, hiszen egyelőre az sem világos, hogy mit ért az illető „fizikai”-n. Ha elfogadjuk azt a meglehetősen kézenfekvő gondolatot, hogy az „a rózsaszín vagy a kávé íze nem lehet azonos egy másfél kilós szürke fehérjetömeggel” állítás egyszerű érvelési hibán (a szemléletes és nem szemléletes összekeverésén, introjekción) alapul (vö. Feigl 1958: 455), akkor mi értelme van ezt úgy értelmezni, hogy aki ehhez hasonló megjegyzést tesz, az antifizikalista?

Indokolatlan arrogancia lenne vitatni, hogy a laikusok rendelkezhetnek filozófiai álláspontokkal, de még inkább az lenne könnyen cáfolható álláspontokat tulajdonítani nekik egy-két megjegyzés alapján.

3.2 Antirevizionizmus és intrinzikusság

Az a meggyőződés, hogy van tudat, a legalapvetőbb problémaintuíciónak tűnik, és ehhez képest másodlagos, hogy milyen további problémaintuíciónak vannak. Aki úgy gondolja, hogy értelmes kérdés azon töprengeni, hogy a horogra húzott giliszta szenved-e, és ez nem pusztán a giliszta viselkedésére irányuló kérdés, az ettől még nem biztos, hogy nem tartja pusztán időpocsékolásnak a zombikkal kapcsolatos gondolatkísérleteken való rágódást. Előfordulhat, hogy valaki egyáltalán nem antifunkcionalista, miközben értelmesnek tartja a mások színérzetének milyenségére irányuló kérdést. Az antirevizionizmus nem kész filozófiai elmélet, hanem egy olyan intuitív meggyőződés, amely a filozófia során sokféle probléma és megoldási kísérlet komolyan vételében nyilvánult meg. Ezzel szemben meglehetősen nehéz elképzelni, hogy egy revizionista szerint, aki szerint nincs tudat, vagy ha van is, akkor tisztán viselkedési vagy kognitív jelenség, a giliszta-probléma értelmetlen, míg a zombi-hipotézis értelmes.

A különféle problémaintuíciók mögött tehát felsejlik egy hierarchia: a Chalmers által katalogizált problémaintuíciók egy részével, de nem feltétlenül mindegyikével rendelkeznek az antirevizionista meggyőződésűek, míg a revizionisták egyikkel sem. Lehetséges, hogy a különféle problémaintuíciók nem egy tőről fakadnak, vagy nem egyenesen következnek az alapvető antirevizionista intuícióból. Ezeknek a viszonyoknak a feltárása eljövendő kutatások és elemzések tárgya lehet.

Nyilvánvaló, hogy aki szerint az antirevizionista intuíció „dualista intuíció”, annak számára az elmefilozófiai diskurzus alapvetően a dualista és nem-dualista álláspontok közötti konfliktus alapján strukturálódik, míg aki számára az antirevizionista intuíció nem más, mint antifunkcionalista intuíció, az úgy gondolja, hogy az elmefilozófiai viták „frontvonalának” egyik oldalán a funkcionalisták, a másik oldalon pedig a materialisták és a dualisták helyezkednek el.

Ezért meglehetősen nagy a kísértés, amikor valamilyen elmefilozófiai álláspontot védünk, hogy – amennyiben magunk is antirevizionista meggyőződésűek vagyunk – az antirevizionista intuíció helyes értelmezésének azt a filozófiai álláspontot tartsuk, amely megkülönbözteti elméletünket az általunk leginkább elutasított elméletektől, vagy ha revizionista meggyőződésűek vagyunk, akkor álláspontunkat vagy valamilyen azt alátámasztó tézist a revizionista intuíció helyes értelmezésének tekintjük.

Megkockáztatom azt a feltételezést, többek között a Dennett és Block közti vitákra, valamint Rorty elmefilozófia-kritikájára és Livingston elmefilozófia-történeti elemzéseire (ez utóbbiakat lásd később) támaszkodva, hogy az antirevizionista intuitív meggyőződésűek azért nem tudják elfogadni, hogy az olyasmik, mint a tudatosság, a szubjektív élmény, az élvezet és fájdalom pusztán funkcionális terminusokban magyarázhatók, mert a kizárólag funkcionális terminusokban definiált állapotok és folyamatok absztraktnak és relacionálisnak tűnnek, miközben az antirevizionisták szerint csak az igazán valóságos, ami intrinzikus és konkrét, és a szubjektív élmény fenomenális minősége éppen ilyen.

Így például Heil (2013: 61) szerint természetes reakciónk azt gondolni, hogy az olyan érzéseknek, mint a fejfájás, intrinzikus kvalitatív karaktere van, és ennek tagadása teszi hihetetlenné a behaviorizmust. Hill (2009: 34) is úgy véli, hogy introspekciónk ellentmond a kválékkal kapcsolatos funkcionális terminusoknak. Pereboom (2016) szerint a fenomenális tulajdonságok a tapasztalat intrinzikus tulajdonságainak tűnnek, míg a funkcionális terminusok szerint ezek csupán relacionális tulajdonságok. Ezért a funkcionális terminusok ellentmond intuícióinknak, és az illuzionisták arra törekedve, hogy teljes elméletet nyújtsanak elmefilozófiai intuícióinkról, jobban tennék, ha

elutasítanak a funkcionalizmust. Levine (1983: 356) is úgy látta, hogy Block példái azt mutatták meg, hogy a funkcionális leírások túl absztraktak ahhoz, hogy megragadják a kvalitatív érzéki tapasztalat lényeges vonásait.

Egy „intuitív” antifunkcionalista például a következőt mondhatja (a funkcionalizmusnak egy kissé karikatúraszerű változatáról, nem a Dennett (1991: 31) által elképzelt borkóstoló gépről): „Nem hiszem el, hogy a borkóstoló robot csakugyan szeretni kezdi B bort és utálni A bort, míg eddig szerette A bort és utálta B bort csak azért, mert kicserélték az „ízlik” és „nem ízlik” feliratokat a gép állapotjelző lámpái előtt. Az, hogy most hogyan érzem magam, nem értelmezés kérdése”. Az antifunkcionalista egyetért azzal, hogy egy zivatar komputerszimulációja nem okoz esőt (Dennett 1978: 416), és nehezen tudja elfogadni, hogy egy egymilliárd, egymással rádióan kommunikáló kínai által implementált agyszimulációnak (Block 1978) „valamilyen” lenni. Ezen attitűdök mögött az az alapvető gondolat, hogy a tudat valóságos jelenség, de nem pusztán absztrakt mintázatként, amit külsődleges kritériumrendszerek alapján lehet odaítélni egy lénynek vagy elvitatni tőle (lásd Putnam 1980 „Leírás”-fogalmát), hanem konkrét és intrinzikus tulajdonságként.

Valószínű, hogy az antirevizionista meggyőződés kialakulásában fontos szerepet játszik az a komoly filozófiai előképzettséget nem igénylő, és sokak számára magától értetődőnek tűnő meggyőződés, hogy bizonyos esetekben különbség tehető a szimuláció és a szimulált dolog között, meg lehet különböztetni a konkrét, intrinzikus valóságot az absztrakciótól. Az utca embere, akár csak Block vagy Searle (1980), ellenállhat annak a gondolatnak, hogy az emberi agyszimulációja fájdalmat vagy élvezetet él át („az csak szimuláció!”), akkor is, ha sohasem hallott Descartes-ról.

Az utca embere szerint – ha antirevizionista meggyőződésű – értelmes kérdés, hogy egy sérülésre elkerülő magatartással reagáló állat vagy robot „ténylegesen szenved-e”, és erre a kérdésre nem válasz, hogy az állapota bizonyos elméleti keretben „tekinthető fájdalomnak”, mert az antirevizionista intuíció szerint a fájdalom és a gyönyör nem interpretáció kérdése, hanem élményünk konkrét, valóságos jellemzője.

Mint az előző példákban láttuk, világos, hogy az antirevizionizmus hogyan motiválhat antifunkcionalista álláspontokat, de ez nem jelenti, hogy az antirevizionista intuíció „antifunkcionalista intuícióként” írható le. Hiszen aki antirevizionista, az minden további nélkül elfogadhat egy olyan funkcionalizmust, amely választ ad kétségeire. Ilyen például egy olyan funkcionalizmus, mely szerint tudatos elme minden olyan rendszer, amely kielégít bizonyos funkcionális leírást, és az elmélet a funkcionális leírásoknak a fizikai rendszerekre való alkalmazási

lehetőségeinek megfelelő megszorításával gondoskodik róla, hogy az elmélet elkerülje a „liberalizmus” (Block 1978) szélsőségét – ha egyáltalán szükséges antirevizionista gondolkodónk szerint az említett szélsőség elkerülése. Egy ilyen elmélet non-reduktivista, el fogja ismerni, hogy a léteznek olyan konkrét *funkcionális típusok* (esetleg többszörös, de nem korlátlan fizikai megvalósíthatósággal), melyek szerepet játszanak a tudat megmagyarázásában, sőt akár a pszichológiában is.

Az antirevizionista intuíció nem is feltétlenül antimaterialista vagy dualista intuíció. Az antirevizionista abban hisz, hogy a tudatosság „valóságos”, és pedig egy kellően „szigorú értelemben”, vagyis a tudatosság jelenléte vagy távolléte nem interpretáció kérdése, nem valamilyen absztrakciós szinthez vagy absztrakt leíráshoz kapcsolódik, hanem konkrét valóság.

Levine az antirevizionizmus egyik jelentős képviselője a kortárs elmefilozófiában. Ez az álláspontja egyértelműen kiderül abból, hogy szerinte „a materializmus kihagyja a tudatos élményt” (Levine 2001: 10), és mindig „magyarázati rés” marad a kválék és azon fizikai tulajdonságok között, amelyekre a materialista elméletek redukálják őket (Levine 2001: 8), mivel szemben a többi fogalommal, a fenomenális tapasztalatra vonatkozó fogalmaink oly módon „szubsztantívák és meghatározottak”, ahogy más fogalmaink nem. Amikor a pirosságra gondolunk, akkor maga a pirosság ebben a gondolatban mintegy „benne foglaltatik”, míg ha egy asztalra vagy székre gondolunk, fogalmunk nem meghatározott ilyen módon (ezek a fogalmak „prezentációsan vékonyak”). (Levine-nek a magyarázati résre vonatkozó érveit bővebben nem tárgyaljuk, csak jelezzük, hogy Levine konklúziója nem feltétlenül a fizikalizmus elutasítása, csupán az, hogy ha igaz is a fizikalizmus, egyelőre nem teljesen *érthető* elmélet a számunkra.).

Ezek után nem véletlen, hogy Levine elutasítja a fenomenális tudatosság olyan elemzéseit, amelyek a kválékat nem valamiféle nagyon is konkrét és intrinzikus tulajdonságokként értelmezik. Levine szerint a kválékról szóló irodalom a kválék intrinzikusnak vagy relacionálisnak tekintése között ingadozik. A dilemma szerinte az, hogy az intrinzikus kválékat nem lehet beilleszteni egy naturalista képbe, míg relacionálisnak tekintésük nem meggyőző. Valami „pirossága” első ránézésre az intrinzikus tulajdonság paradigmatis esetének tűnik. A materialisták szerint konkrétan valamilyen fizikai tulajdonság, de pont ez van kitéve a magyarázati rés ellenérvének. (Levine 2001: 93). Ugyanakkor szerinte a kválék minden olyan „relacionális” felfogása implauzibilis és kontraintuitív (Levine 2001: 96-100), mint amilyen például egy tisztán funkcionalista elemzés. Ha viszont a kválékat abban az értelemben elemezzük relacionálisan, hogy például a színélményt egy bizonyos felszíni tulajdonságokkal rendelkező tárgy és megfigyelő viszonyára vezetjük vissza,

akkor az ilyen esetekben mindig marad valamilyen „intrinzikus maradék” az elemzés végére. Ha nem is élményünk intrinzikus minősége, de a megpillantott tárgy felszínének intrinzikus minősége (lásd még Armstrong felvetését a másodlagos minőségekkel kapcsolatos azonosságelméletéről: Armstrong 1999: 75) fogja meghatározni azt, hogy „milyen” számunkra az adott felületet látni (Levine 2001: 101). A kválékról szóló elméletek egyik családja a reprezentacionalista álláspontoké, mely szerint a kvalitatív tartalom intencionális tartalom (Levine 2001: 111). Ezek az elméletek az intrinzikus tulajdonságokat, amelyek mindenképpen szerepet játsznak a kválék bármely hihető elméletében, a külvilág tárgyaiba „számúzik”. Ezeknek az elméleteknek a működőképessége többek között azon is múlik, hogy mennyiben vetnek számot a Dretske által azonosított „internalista intuícióval” (Levine 2001: 112). Ugyanis ha kvalitatív állapotainkat intencionális állapotaink határozzák meg (tehát például színélményünk „milyenségét” az, hogy milyenek azok a tárgyak, amelyeket látunk), akkor az elmélet bizonyos változatai szerint kvalitatív állapotaink nem lesznek szűk állapotok (röviden: megváltozhatnak a környezet változásaival akkor is, ha a fejünkben semmiféle változás nem történik). Az internalista intuíció ezt a lehetőséget tartja elfogadhatatlannak. Nem nehéz belátni, hogy ez is egy „konkrétsági” vagy „intrinzikussági” intuíció: élményünk fenomenális minősége nem függhet környezetünk változásaitól akkor, ha ez bennünk nem idéz elő fizikai változást.

A konkrét és az intrinzikus mellett felsorolhatjuk az okkurrenst is, mint további jelzót, mely az antirevizionisták szerint a fenomenális tudatosságot megilleti: Dennett is szembeállította a kválékra vonatkozó diszpozicionális elemzéseket a „kvalofileknek” azzal a meggyőződéssel, hogy a tapasztalat kvalitatív jellemzői okkurrens állapotok; de tulajdonképpen Place (1954) is azon az alapon bírálta Ryle elméletét a tudatról, hogy a Ryle-i diszpozicionális elemzés nem ad meggyőzően számot azokról a „belső folyamatainkról”, amelyekről „élő kommentárt” tudunk nyújtani. Tudatos tapasztalatunk okkurrens jellegére vonatkozó intuíciókat aknáznak ki Antony (1994) gondolatkísérletei is.

Sok filozófus intrinzikus tulajdonságoknak tekinti a fizikai tulajdonságokat, ennek megfelelően egy reduktív materialista álláspont, mely a tudatos mentális állapotokat fizikai állapotokkal azonosítja, megfelel az antirevizionista ama meggyőződésének, hogy a tudatos élmény konkrét valóság, illetve mentális állapotunk intrinzikus jellemzője. Természetesen, egyes szerzők, mint Rorty (1998) számára az is kétséges, hogy akár a fizikai tulajdonságok intrinzikus tulajdonságnak tekinthetők-e.

Felvethető a kérdés, hogy ha az antirevizionista intuíciót egy „intrinzikussági” vagy „konkrétsági” intuíciónak tekintjük, akkor nem olyan túlinterpretálással van-e dolgunk, mint

amelyek szerint az utca embere ösztönös dualista, antifunkcionalista vagy karteziánus lenne. Nyilvánvalóan elsietett lenne az intuitív alapmeggyőződések olyan interpretációja, ami ahhoz az elképzeléshez vezetne, hogy „az utca emberének” olyan kész elmélete van, mely által Dennett álláspontjára helyezkedik a Rorty és Dennett közti vitában (Rorty 1998, Dennett 1991b), vagy Block (2015: 128-129) álláspontjára a Block és Dennett közti vitákban. Interpretálhatjuk azonban az antirevizionista intuíciót egy gyengébb intrinzikussági tézisként. E szerint a tudatos élmény „valóságos” és „konkrét”, azaz interpretációfüggetlen, tehát tapasztalatunk intrinzikus jellemzője. Ez az intrinzikusság azonban elegendő, ha csupán *relatív intrinzikusság*, (lásd Block, Atkinson és Davies, Farah és mások ide vonatkozó gondolatait) például olyan, amilyennel a funkcionális állapothoz képest az őt realizáló fizikai állapot rendelkezik.

Amikor az „utca embere” elfogadja, hogy a számítógépes zivatarszimuláció hatására nem lesz csuromvizes, és amikor ennek analógiájára kételkedik abban, hogy az emberi elme számítógépes szimulációja valódi fájdalmakat vagy gyönyöröket élhet át, akkor ebből nem következik, hogy az illető egyúttal abban a kérdésben is állást foglal, hogy a fizikai tulajdonságok abszolúte intrinzikusak-e, vagy a mikrofizikában minden tulajdonság relacionális.

Az antirevizionista intuíció itt javasolt értelmezése tehát az, hogy az antirevizionisták úgy gondolják: a tudatos élmény valóságos, konkrét, nem mások interpretációjától függő jelenség, hanem tapasztalatunk (legalább relatíve) intrinzikus jellemzője, illetve okkurrens jelenség vagy állapot, továbbá preverbális és prekulturális adottság. Más szóval, a fájdalom vagy a rózsailat pontosan olyan értelemben valóságos, ahogy az arany is valóságos: a fájdalom szimulációja nem fájdalom, ahogy a hamis arany pedig nem arany. Ez a meggyőződés semmit sem mond arról, hogy a mikrofizikai tulajdonságok intrinzikusak-e, és összeegyeztethető számos elmefilozófiai állásponttal, sőt mindegyikkel abban az értelemben, hogy az antirevizionista számára azok az elméletek, melyek teljesen elhanyagolják a tapasztalat intrinzikus minőségének kérdését, olyan elméleteknek tűnnek, amelyek egyszerűen kikerülnek a tudat problémáját.

Ha ez így van, akkor a revizionista intuíció nem egy, az antirevizionizmussal ellentétes tartalmú intuíciónak tűnik, hanem egyszerűen az antirevizionista intuíció hiányának, amely párosulva a tudományos módszer magyarázati sikereinek tudomásulvételével, oda vezet, hogy az emberek egy része számára az antirevizionista beszéd „rejtélyekről” vagy a tudomány számára megoldhatatlan kérdésekről üres okvetetlenkedésnek tűnik. Lehetséges, hogy pusztán személyiségpszichológiai eltérések állnak e különbségek között.

Revizionista szerzők (mint Dennett és Rorty) esetében gyakran megfigyelhetjük, hogy úgy vélik: akik velük nem értenek egyet, azok féltenek valamit (a világ „rejtélyességét” vagy az ember „különleges státuszát”) a mindent absztraháló, matematizáló, relációkká oldó tudományos haladástól (Rorty 1998).

Rorty (1998) például úgy értelmezi az egyfelől Dennett, másfelől Nagel, Searle, Kripke és mások között a tudattal kapcsolatos véleménykülönbséget, hogy a vita az intrinzikusság körül folyik: az utóbbiak (akiket inkább az antirevizionisták közé sorolhatunk, míg Dennett revizionista) ragaszkodnak hozzá, hogy a tudatosságnak intrinzikus, nonrelacionális jellemzői vannak, melyeket nem lehet Dennett módjára „kimagyarázni” (Rorty 1998: 98). Rorty ebben a tanulmányában igen világosan képviseli a revizionista álláspontot. Rorty szerint Nagel törekvése, hogy megakadályozza azt, hogy az olyan szerzők, mint Dennett, az „intrinzikusan ineffábilist az effábilissal helyettesítsék” (Rorty 1998: 102). Ezt a törekvést a magát holistának és verifikacionistának valló Rorty elutasítja. A következő módon fogalmazza meg a revizionista elmefilozófia törekvését: „Mit várunk egy tudatelmélettől? Nem azt, hogy feltárja a tudat intrinzikus természetét[...] hanem azt, hogy olyan beszédmódra tegyünk szert, ami, ha elterjed, megváltoztatja intuícióinkat. Miért akarnánk megváltoztatni intuícióinkat? A szokásos kuhni okokból: régi intuícióink túl sok terméketlen vitához vezetnek [...] túl sok defetista locsogáshoz a ‘tudomány korlátjairól’.” (Rorty 1998: 120).

A pragmatista Rorty idegenkedik minden olyan filozófiai álláspontról, amely arra törekszik, hogy „meghaladja az emberi nézőpontot” (Rorty 1998: 108) és amely olyan „rejtélyek” elismerésével, mint a tudat, akadályozza a tudományos fejlődést. Valóban, ha senki sem gondolná, hogy létezik tudat, akkor egy gonddal kevesebb lenne, és teljesebb lenne a világról alkotott tudományos képünk. Talán a revizionista intuíció pozitív tartalma – ha a revizionisták esetében nem pusztán az antirevizionista intuíció hiányáról van szó – nem más, mint a türelmetlenség a „kimondhatatlannal” vagy a „rejtélyekkel” szemben, és a ragaszkodás ahhoz a gondolathoz, hogy csak azt érthetjük vagy ismerhetjük meg, amiről beszélni tudunk, vagyis Wittgensteinnel szólva gondolkodásunk határai nyelvünk határai (Rorty 1998: 101).

A „tudat tagadása” mindazonáltal egy antirevizionista számára azért abszurd, mert csak akkor lenne igazán meggyőző, ha meggyőzné arról, hogy nem lehetséges különbséget tenni absztrakt és konkrét, vagy intrinzikus és relacionális között, még relatív értelemben sem. Az antirevizionisták szerint ez az a heroikus álláspont lenne, mely szerint például nem lehet különbséget tenni a zivatar és annak szimulációja között, vagy az arany és a hamis arany között, ezért elmaradott tudományellenes okvetetlenkedés lenne az a felvetés, hogy a szimulált fájdalom vagy élvezet megkülönböztethető a valóságos fájdalomtól és élvezettől, sőt, egy élmény attól

függően lenne élvezetes vagy fájdalmas, hogy mások minek minősítik. Az antirevizionisták azonban kitartanak amellett, hogy a zivatar, az arany vagy a fájdalom valamilyen értelemben „konkrét” jelenségek, létezők vagy folyamatok. Ez önmagában nem egy jól meghatározott filozófiai álláspont, és legalább ugyanannyi joggal lehet „materialista intuíciónak” nevezni, amennyi joggal „antimaterialistának”.

Az itt elhangzott megfontolások alapján valamivel tovább juthatunk az antirevizionista intuíció megértésében, mint annak első, ideiglenes interpretációjakor („létezik tudatosság”). Az antirevizionisták szerint a tudatosság létezik, és pedig valamiféle konkrétumként, tapasztalataink intrinzikus jellemzőjeként, vagy legalább az általunk megtapasztalt világ intrinzikus jellemzőiként, illetve okkurrens folyamatként, szemben az absztrakt, extrinzikus és diszpozicionális jellemzőkkel és állapotokkal. Továbbá az a tény, hogy a szubjektív élményünk minőségét az antirevizionisták szerint alkotó intrinzikus, konkrét tulajdonságok az absztrakció más szintjén helyezkednek el, mint azok az állapotípusok, melyek (a népi vagy tudományos) pszichológia magyarázataiban szerepet játszanak, azt eredményezik, hogy a szubjektív élmény minősége (annak konkrétsága) és az emberi viselkedés minden leírása (az előbbinek legfeljebb absztrakt, fogalmi megragadása) között valamiféle *kontraszt* tapasztalható, amit aztán a filozófusok olykor antifunkcionalista, antibehaviorista vagy éppen antifizikalista intuitív meggyőződésként interpretáltak. (Emlékeztetőül: már Lashley is úgy magyarázta a tudattartalmak „érzéki minőségének” látszólagos nehezem megmagyarázható voltát, hogy az „feloszthatatlan” tartalomként áll szemben az introspekcióval. Más szóval, az introspekció számára átláthatatlanok, analizálhatatlanok az érzékelési állapotaink, és ezt a kontrasztot értelmezték Lashley szerint egyes filozófusok antifizikalista módon).

Ha az antirevizionista intuíció voltaképpeni tartalma valóban az, hogy a tudatosság intrinzikus, vagy legalábbis a (népi vagy tudományos) pszichológiai tulajdonságainkhoz képest „relatív intrinzikus” jellemzőnk, akkor lehetséges olyan elméletet alkotni a tudatosságról, amely semmi olyan feltételezést nem tartalmaz, ami a revizionista kép szerint felesleges entitásokkal vagy tulajdonságokkal egészíti ki a valóság képét, miközben megfelel annak az antirevizionista elképzelésnek, hogy a fenomenális tudatosság valamilyen intrinzikus tulajdonság vagy okkurrens folyamat.

3.3 Rorty és Antipodea lakói

Rorty híres gondolat kísérletében (Rorty 1979: 70-128) egy távoli bolygó (Antipodea) lakóit képzelettel el, akik sok szempontból emberszerűek, ám egy lényeges tekintetben különböznek tőlünk. Rendelkeznek ugyan népi pszichológiánk megszokott fogalmaival, azonban az akarásokat, hiedelmeket, vágyakat nem tekintik *mentális* állapotoknak. Különbséget tesznek személyek és nem-személyek között, a személyek közé csak fajtársaikat sorolva, az állatokat és robotokat nem, de ezt a különbségtételt nem az „elme” vagy a „tudat” terminusokat felhasználva indokolják. Röviden, nem rendelkeznek az elme fogalmával, és mentális állapotok helyett agyállapotokat tulajdonítanak maguknak. Ahelyett, hogy „ez fáj”, azt mondják, hogy „ez ingerelte a C-rostjaimat”.

Amikor a földiek felfedezik a bolygót, kultúrájuk e sajátossága miatt lakóit „antipodeaiaknak” nevezik el, utalva az ausztrál materializmusra (Rorty 1979: 72). Az antipodeaiakkal való kommunikáció és kulturális érintkezés során a földlakók a következő rejtvényrel szembesülnek: rendelkeznek-e az antipodeaiak érzetekkel? (Rorty 1979: 74). Az antipodeaiak „egyszerűen tudják”, ha például a C-692 agyállapotba kerülnek (Rorty 1979: 75), ahogy mi is „egyszerűen tudjuk”, hogy milyen színélményünk van. De vajon színélményről számolnak-e be, amikor olyan agyállapotról számolnak be, amelyet mi földiek bizonyos színélményként szoktunk megnevezni?

A problémát Rorty a következőképpen veti fel. Ha hiszünk annak a Kripke (1972/2007: 120) által is kifejtett dualista intuíciónak, amelynek a történetében szereplő földi űrutazók hitelt adnak, mely szerint a fájdalmat a fenomenális minősége alapján azonosítjuk, akkor – mivel az antipodeaiak a fájdalmat másként, jelesül bizonyos sorszámú agyállapotként azonosítják – éreznek-e fájdalmat az antipodeaiak (Rorty 1979: 79) az olyan helyzetekben, melyek relevánsan hasonlóak azokhoz, amelyekben mi fájdalmat érzünk? Fájdalomról számolnak-e be a C-rost-ingerlésről beszámoló antipodeaiak, és megfordítva – még zavarbaejtőbben – *mi* érzésről vagy neurális tevékenységről beszélünk-e, amikor a „fájdalom” szót használjuk? Minthogy az antipodeaiak a miénkhez hasonlóan gazdag kultúrával rendelkeznek az elmefilozófiával kapcsolatos sajátosságukon kívül, ezért „filozófusainknak az azzal kapcsolatos érdeklődését, hogy van-e elméjük, kissé provinciálisnak tarthatják. Mi fontos, töprengenek, múlhat ezen?” (Rorty 1979: 81) Az antipodeaiak fejlett agytudományuk révén nyomban meg tudják mutatni, hogy a mi érzet-szavaink mind agyállapotokra vonatkoznak. Azonban ez az antipodeaiak számára a „test és elme viszonyának” földiek által felvetett kérdését érthetlenné teszi (Rorty 1979: 82-83).

Az antipodeaiak egy része materialista módon értelmezi a földiek beszédét az érzetekről: a földiek érzet-szavai csupán tudományelőtti, ideiglenes szavak (*placeholders*), amelyeket a tudomány fejlődése tudományos agyállapot-szavakkal helyettesít. Használatuk a földiek

agytudományának fejletlenségét tükrözi. Az antipodeaiak más része epifenomenalista, és el tudja képzelni, hogy a földiek esetében – azok állításának megfelelően – van valami, az érzet, ami az agyállapotokat kíséri, de nem azonos velük. Ha a földiek nem tudják megmagyarázni az antipodeaiaknak, mi a test-lélek probléma (Rorty 1979: 83), akkor szembe kell néznünk azzal lehetőséggel, hogy a materialista antipodeaiaknak volt igazuk, és amikor érzetéről beszélünk, akkor igazából agyállapotokról beszélünk. Az, hogy még mindig ezeket a tudományelőtti kifejezéseket használjuk (szemben az antipodeaiakkal), csak kulturális fejlődésünk esetleges adottsága. Rorty a következő módon foglalja össze a földiek és antipodeaiak közti kölcsönös értetlenséget:

„Mi a különbség a fájdalomérzet és aközött, ha a C-rost-ingerlésre egyszerűen a „fájdalom” szóval, elkerülő magatartással stb. reagálunk? És itt hajlamosak vagyunk azt mondani: egyáltalán semmi különbség kívülről nézve, de a lehető legnagyobb különbség belülről tekintve. A nehézség az, hogy soha nem fogjuk tudni ezt a különbséget semmiképpen elmagyarázni az antipodeaiaknak. A materialista antipodeaiak azt gondolják, hogy nincsenek érzeteink, mert azt gondolják, hogy nincs olyan dolog, hogy "érzet". Az epifenomenalista antipodeaiak azt gondolják, hogy talán vannak efféle dolgok, de el sem tudják képzelni, miért csapunk akkora hűhót akörül, hogy rendelkezünk velük. A földi filozófusok, akik szerint az antipodeaiaknak vannak érzeteik, csak nem tudnak róluk, elérték a filozofálás Wittgenstein által említett végső fázisát... még azt sem mondhatják az antipodeaiaknak, hogy „ez másmilyen számunkra belülről”, mert az antipodeaiaknak ... a „belül” azt jelenti, hogy „a koponyában”... azok a földi filozófusok, akik szerint az antipodeaiaknak nincsenek érzeteik, csak azért könnyebb a dolguk, mert rangon alulinak is érzik, hogy elme nélküli lényekkel vitatkozzanak...” (Rorty 1979: 85-86)

Ebből a szakaszból a következő négy álláspontot ismerhetjük meg (az elnevezések forrása részben: Tartaglia 2007:82). A zárójelekben az A/R sémabeli álláspontok körülbelüli megfelelőit tüntettem fel.

antipodeai materialista (R1) : a földiek elme/anyag különbségtétele sajnálatos nyelvi szokás

antipodeai epifenomenalista (R2) : a földiek a c-rost-ingerlésről szóló beszámolókon kívül még fájdalomról szóló beszámolókat is képesek adni

földi egyéb (A1) : az antipodeaiak nyelve „inadekvát”: ha vannak is érzeteik, ezeket nem tudják kifejezni

földi szkeptikus (A2) : mivel az antipodeaiaknak nincsenek érzet-szavaik, érzeteik sincsenek

Körülbelüli megfelelőkről azért indokolt beszélni ezen a ponton, mert Rorty történetében az antipodeaiak materialisták, míg a földiek szemlátomást elkötelezettek a fenomenális tulajdonságok mellett, amelyek problematikusan kapcsolódnak agyállapotokhoz (Rorty 1979: 74), és ez alapján leginkább dualistának vagy tulajdonságdualistának tűnnek.

Rorty rövid úton elintézte a 4 álláspontból kettőt (R2-t és A2-t). Ez a döntés nem meglepő. A filozófiai irodalomban rendszerint kevesebb figyelmet kaptak a 2-es számmal jelölt álláspontok: gyakran nem is tekintik őket komolynak.

Az antipodeai epifenomenalista (R2) álláspont párhuzamba állítható az illuzionizmussal. Az illuzionista filozófus elismeri, hogy miközben „valójában nincs tudat”, egyesek számára úgy tűnhet, hogy mégis van. Ez nem ugyanaz az álláspont, mint az R1-es, mely szerint valójában még a tudatról folyó(nak szánt) diskurzus sem a tudatról szól, hanem tudománytalan vagy a tudományosság kezdetleges fokán álló pszichológiai diskurzus.

A „földi szkeptikus” álláspontnak valóban aligha találjuk párhuzamot a komoly filozófiai vitákban, némely epés megjegyzéseken kívül. Ettől még azonban ez kellőképpen világos, egyben az érdemi vitának alighanem elejét vevő álláspont: az, hogy vitapartnerünk eltérő pszichológiája miatt nem képes hozzászólni a témánkhöz.

Rorty szerint két álláspont vehető komolyan: az antipodeai materialisták nézete, mely szerint az emberek azt hiszik, vannak nyers érzeteik, de ebben tévednek, és a földi filozófusoké, mely szerint az antipodeaiaknak vannak nyers érzeteik, de erről nem tudnak. Rorty kérdése (Rorty 1979: 88): van kiút ebből a patthelyzetből?

A két álláspont szembenállásából Rorty egy másik, 3+1 elemű felosztáshoz jut el. Ebből három álláspont három hagyományos megoldási javaslat az antipodeai kihívás által megvilágított problémákra, egy pedig a Rorty által ajánlott alternatíva: „materializmus azonosság nélkül”.

Rorty szerint mivel a Ryle utáni modern elmefilozófia középpontjában a nyers érzetek problémája állt, azoknak pedig legfontosabb ismertetőjegye, hogy noninferenciális – és a karteziánus hagyomány szerint inkorrigibilis és biztos – tudásunk van róluk, ezért végső soron az inkorrigibilis beszámolók előfordulása tényének megmagyarázása az elme-test probléma kulcsa.

Ha a nyers érzetek lényegi sajátossága, hogy inkorrigibilisen ismerhetők (Rorty 1979: 91), ám az antipodeaiak nem gondolják, hogy bármiről inkorrigibilis tudásuk lenne (Rorty 1979: 76-77), akkor vagy nincsenek az antipodeaiaknak nyers érzeteik, vagy nem tudnak arról, hogy inkorrigibilis tudással rendelkeznek. Mivel az antipodeaiak megtaníthatók a földi „nyers érzet”-diskurzusra, ez vagy azt jelenti, hogy ekkor a nyers érzeteik felismerését tanulják meg, vagy a nyers érzetekről való beszámolók szimulálását. A dilemma tehát úgy is megfogalmazható, hogy vagy igaz, hogy a „valami valamilyennek tűnik” típusú beszámolók némelyike nyers érzetekről számol be, vagy nem igaz.

A pozitív válasz a (P) elv („Amikor inkorrigibilis beszámolót adunk egy állapotunkról, akkor kell lennie egy tulajdonságnak, ami prezentálva van számunkra, amely minket e beszámoló megtételére készít”) elfogadásán alapulhat, amely a fenomenális tulajdonságokon alapuló karteziánus érveket megalapozó elv (Rorty 1979: 84).

Rorty elemzése szerint (P) elfogadása esetén kénytelenek vagyunk az antipodeaiaknak nyers érzeteket tulajdonítani, legalábbis mihelyt elsajátították a földi beszédmódokat. Vagyis a karteziánus elvből paradox módon „egyfajta behaviorizmus következik” (Rorty 1979: 94), ami a tárgyalt kontextusban azért problematikus, mert kizárja az antipodeaiak nyers érzeteivel szembeni szkepszist. A szkepszis csak akkor tartható, ha feltételezzük, hogy a nyersérezet-diskurzust elsajátító antipodeaiak csak szimulálni tanulták meg a nyers érzetekről való beszámolást. Ez a szkepszis azonban alaptalan, „pürrhóni”. Ha viszont az említett dilemmára a válaszuk negatív, akkor szembe kell néznünk azzal a lehetőséggel, hogy nekünk sincsenek nyers érzeteink (Rorty 1979: 95).

A dilemma egyik ágán tehát két álláspont van:

- (1) minden lény, amelynek a nyelve tartalmaz érzet-szavakat, tudatos
- (2) pürrhóni szkepszis az antipodeaiak tudatával kapcsolatban

A másik ágán pedig a következő:

- (3) az antipodeaiak nem tudatos lények, és mi sem vagyunk azok.

Ezek pedig három hagyományos elmefilozófiai álláspontnak felelnek meg Rorty szerint: a behaviorizmusnak, a szkepticizmusnak (ami történetileg leginkább a (tulajdonság-)dualizmusnak feleltethető meg) és a materializmusnak (Rorty 1979: 95).

Mindhárom elmélet háttérében az az elv áll, hogy „ami inkorrigibilisen tudható, az nyers érzet” (Rorty 1979: 93). Ez az elv olyan elméletet alapoz meg, mely szerint az „elme” nem más jelent, mint inkorrigibilisen ismerhető nyers érzetek halmazát, az elmét elmévé pedig különleges episztemikus státusza (inkorrigibilitás) teszi. Rorty javaslata ennek az elvnek a feladása az inkorrigibilitás „társadalmi gyakorlat” (vagy „nyelvi szokás”) felfogása javára (Tartaglia 2007: 86).

3.3.1 Behaviorizmus

Ryle és Wittgenstein behaviorizmusát lehet ontologizáló módon („nincsenek mentális események”) is olvasni, és ennek az értelmezésnek – legalábbis annak lehetőségének – kétségkívül megvolt a maga hatástörténete (Smart 1959: 143).

A másik értelmezési (vagy továbbfejlesztési) lehetőséget a Rorty által idézett Donagan-féle (Rorty 1979: 104) Wittgenstein-interpretáció villantja fel. Wittgenstein (természetesen Donagan és nem Smart Wittgensteinje) elkerüli a karteziánusok és (filozófiai, „metafizikus”) behavioristák közötti patthelyzetet, melynek oka végső soron a „természetes módon adott doktrínája” Rorty szerint, vagyis az az elv, mely szerint bizonyos dolgok közvetlenül adóttak megismerő tudatunknak, míg másokra csak következtethetünk az előbbiekről szerzett tudomásunk alapján. Ez az elv, kibővítve a „platonikus elvvel”, mely szerint a közvetlenül megismerhető dolgok az igazán valóságosak, vagyis „a legmegismerhetőbb a legvalóságosabb” (Rorty 1979: 105) a fizikai mentálisra vagy a mentális fizikaira redukálását eredményezte, attól függően, hogy melyik filozófus milyen fajta entitásokat vagy tudást választott rendszere kiindulópontjául.

Antipodeai nézőpontból azonban nincs helye a „természetes módon adott”-nak. Az antipodeaiak rendelkeznek a közvetlen vagy a noninferenciális tudás fogalmával, azonban az ő esetükben ez nem jelent semmi, a közvetett tudástól radikálisan eltérő: „hogy mit tudunk noninferenciálisan, az azon múlik, hogy történetesen mit ismerünk jól” (Rorty 1979: 106). Ezért antipodeai nézőpontból a behaviorista-szkeptikus vita meglehetősen fölösleges.

3.3.2 Szkepticizmus

A nyers érzetek „ugyanolyan jó partikulárek, mint az asztalok, arkangyalok vagy elektronok ... jó pályázók az ontológiai státuszra” (Rorty 1979: 107), ha elvetjük azt az elképzelést – mint ahogy természetesen Rorty elveti – hogy vannak inkább és kevésbé valódi entitások. Ha azonban azt az észrevételt, hogy a nyers érzetek különösen jól ismertek, párosítjuk a „természetes módon adott” már említett elvével, akkor a nyers érzetek valamiféle speciális fajta entitásokká – „talán egy privát ‘nem-anyagi közegben’ zajló folyamatokká” válnak (Rorty 1979: 107). A szkepticizmus úgy áll elő, hogy a saját elménkhez való privilegizált hozzáférés olyan felfogása, amely szerint mentális állapotainkat tisztán azok „különleges érzett minőségének” (Rorty 1979: 109) köszönhetően ismerjük, együtt azzal a követelménnyel, hogy mások elméje akkor ismerhető meg, ha úgy ismerjük, ahogy az önmagát ismeri (Rorty 1979: 107), természetesen eleve kizárja, hogy mások elméjéről ismeretünk legyen. Ha azonban a privilegizált hozzáférésnek egy deflacionárius, a „fenomenális tulajdonságokkal” való közvetlen ismeretség nélküli képéből indulunk ki, akkor az érzet – „ugyanolyan valami lehet, mint egy asztal” (Rorty 1979: 109). Antipodeai nézőpontról a „privát entitások” fogalma érthető, noha egyben érdektelen (úgy értik, hogy olyan entitások, amelyekről csak egyvalakinek van inkorrigibilis tudása). Ellenben nem lenne számukra értelmes Rorty szerint az olyan entitások fogalma, amelyek annyira privátak, hogy kommunikálhatatlan a róluk szóló tudás: az ilyesmi számukra „nem is entitás” (Rorty 1979: 111). Az antipodeaiak e szkepticizmusa azonban egészen más, mint a más elmékkel szembeni karteziánus szkepszis, amelynek követői egyszerűen kételkednek abban, hogy másoknak vannak-e érzeteik és kész.

3.3.3 Materializmus és Rorty alternatívája

A behaviorizmushoz és a szkepticizmushoz hasonlóan a materializmus is egy problémamentes intuícióból indul ki, azonban azt Rorty szerint paradox állásponttá fejleszti. A kiinduló gondolat az, hogy az emberi viselkedés magyarázatában szerepet játszó „rövidtávú mentális állapotok”-ra történő hivatkozások (Rorty 1979: 114) egyszer lecserélhetők lesznek majd neurológiai struktúrákra és folyamatokra való hivatkozásokkal. A materialista azonban ezen túl még „valami metafizikusát is mondani akar” (Rorty 1979: 115), tudniillik azt, hogy „a mentális állapotok nem mások, mint neurális állapotok”. Ez azonban Rorty szerint „paradoxul hangzik”. A „paradoxon” enyhítésének egyik bevett módja a Place, Smart és Armstrong által képviselt

tárgyközömbös elemzés. Ez Rorty szerint problematikus. Fenntartásait többféleképpen megkísérli megfogalmazni (Rorty 1979: 116):

(1) intuitív különbség van az „az a mentális állapot, amely ...-t okozza...” és az „az a fizikai állapot, amely ...-t okozza...” között;

(2) intuíciónk szerint különbség van a materializmus és parallelizmus között, ezért valami félrevezető van annak tárgyközömbös elemzésében, hogy mi teszi a mentálist mentálissá;

(3) ha elmefogalmunk csakugyan olyan, amilyenek a tárgyközömbös elemzés lefesti, akkor nehéz megmagyarázni, miért létezik egyáltalán az elme-test probléma.

Egy lehetséges válasz a (3)-ban felvetett kérdésre Rorty szerint az, hogy míg a hétköznapi életben van nagyléptékű elméletünk a viselkedésről (a népi pszichológia fogalmi apparátusa), addig hagyományosan hiányzott a finomabb léptékű, neurológiai szintű leírás. Ezért az elme bizonyos aspektusai rejtélyesek maradtak. Ez az elképzelés az elme „hagyományos felfogását” két részre osztja: a kauzális szerepekre és arra, ami a szerepeket betölti. A tárgyközömbös elemzés az előbbire vonatkozik és nem az utóbbira, ezért lehet hiányérzetünk a mentálisra vonatkozó tárgyközömbös elemzésekkel kapcsolatban. Rorty ezt úgy jellemzi, hogy ez a taktika a „mentális állapot” fogalmának felosztása egy, „a materializmussal kompatibilis” részre, és egy másikra, ami nem kompatibilis a materializmussal, majd pedig annak önkényes állítása, hogy az előbbi a „lényeges” a fogalom szempontjából. Ez az önkényes eljárás csupán fordítottja annak a karteziánus állításnak, hogy a fájdalmat lényegileg a „közvetlen fenomenális minősége” határozza meg (Rorty 1979: 116).

A tárgyközömbös elemzésekre támaszkodó („fordítási típusú” vagy redukzív) materializmus alternatívája az „elmék és agyak azonosságának olcsó verziója”, az „eltűnési típusú” (Rorty 1965) vagy más néven eliminatív materializmus lenne. Ám végső soron mindkét fajta materializmus egyformán súlyos nehézségekkel küzd, mindössze más téren: a redukzív materializmus gyenge pontja, hogy a jelentésazonosság-kritériumok körüli problémákba bonyolódhat, míg az eliminatív materializmusé a referenciaazonosság-kritériumok körüliekbe (Rorty 1979: 118).

Mivel azonban az azonosságelmélet mindkét változata csupán „kísérlet arra, hogy kortárs filozófiai zsargonba öntsük az antipodeaiakkal való találkozásra adott természetes reakciónkat”, Rorty végső soron a köztük lévő különbségeket nem tartja lényegesnek: mindkettőt fel kell adni, és vele az elme-agy azonosság elképzelését is. „A megfelelő reakció az antipodeaiak történetére egy

olyan materializmus elfogadása, amely nem azonosságelmélet semmilyen értelemben” (Rorty 1979: 119).

Először is a Rortyt követő materialista ne tegyen metafizikai állításokat. Szorítkozzon arra az állításra, hogy nem veszítenénk semmilyen magyarázó vagy előrejelző erőt, ha antipodeaiul beszélünk (Rorty 1979: 120). Értelmetlen megkérdezni, hogy az antipodeai történet tanulságának „nincsenek mentális állapotok” vagy az „a mentális állapotok igazából neurális állapotok” értelmezése-e a helyes, mert ezen „semmi sem múlik” (Rorty 1979: 120). Abba kell hagyni annak firtatását, hogy mi számít „mentális”-nak és mi nem (Rorty 1979: 121).

3.3.4 Rorty értékelése

Rorty szemlátomást tisztában volt a „tudat tagadása” és a „tudat igenlése” álláspontja közti patthelyzet jelentőségével az elmefilozófiában, sőt célul tűzte ki ennek a problémának a megoldását, hiszen a földi kutatók számára az a kérdés vetődik fel, hogy az antipodeaiaknak vannak-e érzeteik, és az ennek nyomán előálló dilemmára Rorty végül is valamiféle megoldást kínál. Az antipodeaiakkal folytatott kommunikáció során 2x2 álláspont (két földi és két antipodeai) kristályosodik ki attól függően, hogy a földiek, illetve antipodeaiak tulajdonítanak-e egymásnak érzeteket vagy sem. Rorty példázatának egyik lehetséges értelmezése tehát az, hogy a történetben az A1, A2, R1, R2 álláspontok jelennek meg.

Az álláspontok közötti patthelyzetekre adott sikertelen megoldási kísérletek fényében Rorty konklúziója az, hogy ideje elismernünk a földi elmefilozófia csődjét, és túllépnünk ezen a vallásos, metafizikai, karteziánus stb. eredetű, haszontalan diskurzuson. Ha azonban nem osztozunk az elmefilozófiának ebben az értékelésében, valamint Rorty általánosan metafizikaellenes, az objektív valóság megismerhetőségét és minden intrinzikus természet elfogadhatóságát elutasító filozófiai programjában, akkor nem biztos, hogy kielégítőnek tarthatjuk Rorty megoldását.

Az antipodeaiakról szóló példázat másik lehetséges értelmezése, hogy a Rorty által vizsgált 2x2 álláspont strukturális hasonlóságot mutat ugyan az A- és R-álláspontokkal, de tartalmilag nem azonos ezekkel. Figyeljük meg ugyanis, hogy Rorty példájában a földiek és antipodeaiak közötti vita az inkorrigibilitás körül forog, ami Rorty antikarteziánus filozófiai törekvései szempontjából fontos ugyan, de a karteziánus szigorú vagy „logikai” inkorrigibilitást a materialista elmefilozófia (például Armstrong 1968: 100-113) elutasította anélkül, hogy ezzel irrealistává vált volna az

elmével vagy tudattal kapcsolatban. Lehet, hogy a földiek és antipodeaiak vitája nem az antirevizionisták és revizionisták vitája, hanem egy karteziánus-materialista vita (hogy így van-e, azt nehéz megállapítani a szövegből: rortyi fanyar humorral „a Földön feledésbe merült régi iskoláról” (ausztrál materializmus) nevezik el Antipodeát; másrészt vannak földi materialisták a vitában), ami egy revizionista szemszögéből egy provinciális episztemológiai és pszichológiai vita, semmi több.

Az előző két értelmezés szintézise egy olyan olvasat, mely szerint Rorty valóban az A- és R-álláspontok közti vitát mutatja be, de a megoldást nem ennek a dilemmának a meghaladásában vagy egy, a kétféle intuíció között egyensúlyt vagy szintézist teremtő elméletben látja, hanem a vita eldöntésében az egyik (konkrétan a revizionista) oldal javára. Ezt az olvasatot támogatja az, hogy Rorty hangsúlyozza, hogy semmi sem múlik a klasszikus elmefilozófiai álláspontokon, ami a revizionisták szerint így van, de aki felfüggeszti a döntést az (anti)revizionizmussal kapcsolatban, az nem mondhat ilyet. Ezen értelmezés szerint Rorty a revizionista állásponton belül mozog, és az antirevizionizmus és revizionizmus közti dilemmát úgy oldja meg, hogy a revizionista álláspont mellett tör lándzsát. Rorty időnként explicit módon képvisel bizonyos R-téziseket, így például úgy gondolja, hogy az elmefilozófiai problémák megfelelő kulturális változásokkal kiküszöbölhetők. Egy antirevizionista nem fogadná el azt, hogy nyelvi szokásaink változtatásával megszabadulhatunk a tudatosság jelenségeitől, melyek az antirevizionista számára konkrét, intrinzikus valóságnak tűnnek, és fennmaradnak függetlenül az elméleteink változásaitól.

Rorty gondolatmenetének értékelésében további nehézséget okoz, hogy míg az antirevizionisták és revizionisták köztünk járnak, és a földi filozófusok mind az A1, A2, R1 és R2 álláspontokat is képviselték, addig Rorty egy sci-fi történetet mesél egy elképzelt társadalomról, melynek tagjai fejlett tudománnyal és kultúrával rendelkeznek, mégis értelmezhetetlennek tartják a földi elmefilozófiát, elvileg egyszerűen azért, mert materialisták (és sohasem voltak karteziánusok). Tartaglia szerint azonban az antipodeaiaknak könnyűszerrel el lehet magyarázni a földi elmefilozófia problémáit, ha viszont ez csakugyan nem lehetséges, mint ahogy Rorty állítja, akkor az antipodeaiak pszichológiailag annyira különböznek a földlakóktól, hogy kérdéses, hogy lehet-e a segítségükkel a földiek számára tanulságos állításokat illusztrálni (Tartaglia 2007: 84–85).

Rorty filozófiai tevékenységének fő célja az volt, hogy „rábeszéljen minket arra, hogy mondjunk le az esszencializmusról, intrinzikusságról, és így az ahistorikus, objektív igazságról” (Tartaglia 2020: 5). Mivel az elmében az „intrinzikusság utolsó menedékét” látta, meg akarta mutatni, hogy az elme is csupán egy „relacionális csomópont hiedelmeink kulturálisan konstruált hálójában” (Tartaglia 2020: 5). Ezért harcolnia kellett az olyan beszédmódok ellen, amelyek szerint

a tudatosságnak intrinzikus természete van, ami ellenáll a fizikai leírásoknak. Rorty célja tehát annak bebizonyítása volt, hogy a tudatosság nem rendelkezik olyan intrinzikus természettel, ami gátjává válhatna a tudomány haladásának, és metafizikát motiválhat (Tartaglia 2020: 7). Rorty szerint a tudat problémája az újkorban keletkezett, amikor úgy tűnt, hogy az új, matematizált természettudomány a valóságot „relációk hálójává” változtatja úgy, hogy ezekben nincs helye az olyan, intrinzikusnak tűnő manifeszt minőségeknek, mint amilyenek a színek vagy az ízek. Descartes ezeket a másodlagos minőségeket az elmébe menekítette a tudomány elől. Rorty szerint jobban tette volna, ha hagyja, hogy a tudomány elsöpörje az intrinzikusság minden maradványát. Ám Descartes erre a lépésre nem volt képes, mert vallásos hajlamai nem engedték, hogy lemondjon az objektív, „független természetű” valóságról (Tartaglia 2020: 7).

A vallás visszaszorulása után is még sokan kitartottak amellett, hogy az ember „különlegességébe” vetett hitüket a tudatosságra hivatkozva fenntartsák. Ryle azonban felismerte, hogy „nem kell komolyan vennünk a tudatot” (Tartaglia 2020: 9). E tradíció a tudat „lingvisztifikációjában” érte el végcélját, vagyis abban az álláspontban, hogy az elméről vagy tudatról beszélni (vagy nem beszélni) csupán kényelmi kérdés. Azok a filozófusok, akik ennek ellenálltak arra hivatkozva, hogy a tapasztalatnak intrinzikus, szubjektív „minőségi” természete van, a vallási jellegű „transzcendencia-ambíció” rabjai, vagyis arra vágynak, hogy a megismerhető határán túl, felfedezzék a valóság „valódi természetét”. A valóságról való ismereteink relacionálisak, ezért Rorty szerint ez az intrinzikus természet iránti vágyódás nem érhet célt.

Rorty igencsak pontosan és tömören fogalmazza meg a revizionisták központi hiedelmét, amikor a zombikkal kapcsolatban kijelenti: „az eléggé széles körben elfogadott szimuláció maga a valódi dolog” (Tartaglia 2020: 11). Ezzel szemben csak úgy lehetne érvelni, ha valami ineffábilisra hivatkoznánk, aminek könnyűszerrel ellent lehet állni verifikacionista érvekkel (Tartaglia 2020: 11). Gondoljunk csak arra, hogy míg Block nyilvánvalónak tartja, hogy a tudatosság nem lehet tisztán funkcionális jelenség (vagy éppen kulturális konstrukció), milyen nehéz számára pozitívan érvelni a tudat létezésével. *LABM* példa lehet az ineffábilisra való hivatkozásra, ám egy verifikacionistát ez aligha fog meggyőzni bármiről.

Tartaglia szerint egyébként egyszerűen hamis Rortynak az a központi állítása, hogy az (elmefilozófiai kérdések felvetésére alkalmas adó értelemben vett) elme az újkori filozófusok találmánya. Amikor Kirké Odüsszeusz társait disznóvá változtatta, a testük megváltozott, de elméjük megmaradt. Mint Tartaglia megállapítja, a test változásaitól független elme ötletét Homérosz feltehetően nem Descartes-tól vette. (Meg kell azonban jegyeznünk, hogy Hacker (2012)

az idézett Homérosz-verssorokat éppen ellentétes módon értelmezi. Úgy tűnik, az antirevizionista és revizionista meggyőződések a versek értelmezésére is rányomják a bélyegüket).

A tudatos tapasztalatok intrinzikus minőségének elutasítása és a tudat „lingvisztifikálása” Tartaglia szerint oda vezet, hogy Rortynak és a többi „ryleánusnak” azt kell gondolnia, hogy prelingvisztikus tudatosság nincs, így a beszélni még nem tudó csecsemők, vagy a kutyák, amikor fájdalom hatására felsírnak, akkor egyszerűen „úgy nyikorognak, mint egy ajtózsalu” (Tartaglia 2020: 13). Ám a prelingvisztikus tudatossága lehetetlensége egyszerűen abból következik, hogy Rorty számára nincs tudat (a szó szokásos értelmében), csak a tudatról való beszéd; az pedig természetes, hogy amíg nincs beszéd, addig a tudatról folyó beszéd sincs, viszont mihelyt van beszéd, a tudatról folyó beszéd is lehetséges. A tudat tagadásával és bizonyos beszédmódokkal való helyettesítésével azonban „elérkeztünk a filozófiai implauzibilitás legvégső határaihoz” (Tartaglia 2020: 15). Tartaglia szerint Rorty végeredményben azért jutott el ilyen extrém álláspontra, mert attól tartott, hogy „mindennek, ami úgy jelenik meg, mint ami akadályt gördíthet a világ kimerítő fizikai leírásának útjába, köze van a valláshoz”. Rorty számára a vallás elleni hadjárata olyan fontos volt, hogy szemben materialista kortársaival, ő még a tudományban sem hitt: még az objektív tudományos igazságba vetett hit is túlságosan a vallásos hitre, az abszolút igazságra irányuló „metafizikai sóvárgásokra” emlékeztette őt (Tartaglia 2020: 23).

Így aztán annak ellenére, hogy Rortynak az antipodeaiakról szóló történetét az A/R séma és a négyféle olvasat (A1, A2, R1, R2) fontos előfutárának vagy bizonyos lehetséges értelmezés szerint első felvázolójának tekinthetjük, és megállapíthatjuk, hogy Rorty igen fontos megállapítást tett, amikor rámutatott az *intrinzikussággal* kapcsolatos intuíciók centrális jelentőségére az alapvető elmefilozófiai meggyőződések szempontjából, nem érthetünk mindenben egyet Rortyval. Vitatható, hogy az elmefilozófia feladása az egyedüli adekvát válasz az alapvető elmefilozófiai intuíciók ellentéte által eredményezett patthelyzetekre. Talán ki lehet dolgozni olyan elmefilozófiai álláspontot, amely másféle kiutat kínál a dilemmákból. Szintén vitatható, hogy az elmefilozófiai kérdéseknek ne lenne gyakorlati jelentőségük. Nem kötelező követnünk Rortyt mindenféle intrinzikus és objektív valóság elutasításában sem, sőt abban sem kell egyetértünk vele, hogy minden jelentős elmefilozófiai nézetkülönbség az inkorrigibilitás kérdésével lenne kapcsolatos, vagy hogy az elmefilozófia funkciója kimerülne az ember „különleges státuszának” elméleti megalapozásában.

3.4 Az egyensúlyi elmélet keresése

Értekezésem eddigi részében azt igyekeztem megmutatni, hogy az elmefilozófiai viták résztvevői között alapvető nézetkülönbségek vannak még abban a kérdésben is, hogy egyáltalán létezik-e tudatosság. Mind a laikusok, mind a filozófusok egy része számára a legnyilvánvalóbb igazságnak tűnik, hogy létezik tudatos tapasztalat. Más filozófusok és nem-filozófusok szerint azonban ez vagy nem is tűnik így, vagy ha így tűnik, az tévedés vagy illúzió. A laikusok egy része számára könnyen érthetőek azok a problémafelvetések, melyekkel a tudatfilozófiai érveléseket motiválni szokás, míg mások számára teljesen értelmetlennek tűnnek.

Kétféle intuitív meggyőződést különböztettem meg, az R (revizionista) intuíció és az A (antirevizionista) intuíció nyomában támadt meggyőződéseket. Ezt követően megvizsgáltam, hogy az elmefilozófiai viták értelmezhetők-e olyan módon, mint az egymással ellentétes intuitív meggyőzésekkel rendelkező filozófusok közti viták, azaz értelmezhetők-e az elmefilozófiai elméletek úgy, mint az R- vagy A-intuíció filozófiai explikációjára irányuló törekvések eredményei.

Megvizsgáltunk néhány olyan filozófiai tanulmányt, melyeket példaként szoktak felhozni, amikor a tudatprobléma jogosultsága melletti kiállásról vagy ellenkezőleg, a „tudat tagadásáról” van szó. Noha a filozófiai tanulmányok bevett recepciója, illetve a tanulmányok szerzőinek saját szándéknyilatkozatai azt az első benyomást keltik, mintha ezek a szerzők csakugyan kiállnának a revizionista álláspont („igazából nincs tudatosság”) vagy az antirevizionista álláspont („van tudatosság”) mellett, a szövegek alaposabb tanulmányozása ezt a benyomást kétségessé teszi. Ez azért van így, mert minden elmefilozófiai elméletnek elő lehet állítani egy revizionista (R-) vagy egy antirevizionista (A-) olvasatát.

Ennek oka pedig az, hogy amikor a revizionista szerzők a saját szándékaik (és revizionista olvasói interpretációja) szerint a tudatosság létezése ellen érvelnek, akkor az antirevizionista olvasó számára úgy tűnik, hogy a revizionista szerző vagy nyilvánvalóan hamis állításokat tesz, ami felveti azt a lehetőséget, hogy az ő pszichológiája egészen más, mint az az antirevizionista olvasóé – például ő zombi, szemben az olvasóval – (A2 olvasat), vagy pedig egyszerűen témát vált, valójában nem tagadja a tudatosság létezését, hanem másról beszél, miközben vélhetően ugyanannyira szilárdan hisz a tudatosság létezésében, mint az antirevizionista olvasó (A1 olvasat).

Amikor azonban egy antirevizionista szerző megkísérli alátámasztani a tudatprobléma létjogosultságát, és ennek érdekében, vagy más megfontolásból a tudat, a tudatos tapasztalat, az élmény „milyenségének” létezése mellett foglal állást, akkor ugyan antirevizionista olvasói úgy

vélik, hogy pontosan értik, miről van szó, és képesek követni a gondolatmenetet, a revizionista olvasók számára azonban az antirevizionista szövegek kevésbé meggyőzőek. Nekik ugyanis mindig rendelkezésükre áll az az olvasat, mely szerint az antirevizionista szerzőnek a tudattal kapcsolatos alapvető meggyőződései azonosak az övékével, csak hogy – a viselkedési vagy pszichológiai mechanizmusokkal kapcsolatos tudatlanság, tudományos tájékozatlanság miatt – álproblémákba bonyolódnak (R1 olvasat), és elavult megoldásokat adnak rájuk. R1 olvasatban az antirevizionista szerző az *élan vital* vagy a flogisztionelmélet hívójéhez hasonló episztemikus pozícióban van. A másik lehetőség azt állítani, hogy az antirevizionista szerző pszichológiája meglehetősen más, mint a revizionista olvasóé, például azért, mert az antirevizionista szerző nem zombi, míg az olvasó igen, vagy azért, mert az antirevizionista szerző gondolkodását illúziók vezették tévútra, míg az olvasó nem szenved ezekben az illúziókban (R2 olvasat).

Az antirevizionista szerző nem fogja tudni filozófiai érvekkel rávenni revizionista olvasóját, hogy gondolja át a pozícióját és csatlakozzon az antirevizionisták táborához, két okból. Egyrészt az A- és R-intuíciók a filozófiai álláspontoknál alapvetőbb álláspontoknak tűnnek, és egészen pontosan még nem is tudjuk, hogy mi a filozófiai interpretációkat megelőző, lényegi tartalmuk. Ezért közvetlenül aligha támadhatók filozófiai érvekkel. Az antirevizionizmus motiválhat dualizmust, de motiválhat akár materializmust is; ha valakinek megingatjuk a bizalmát a dualizmusban, ezzel nem tesszük feltétlenül revizionistává. Másrészt az antirevizionista szerző nem tud rámutatni arra (a tudatra), amit hiányol a revizionista filozófiából. Amikor a rózsa illatáról vagy a kávé ízéről beszél, akkor a revizionista ezt például úgy értelmezi, hogy tárgyak fizikai tulajdonságairól beszél. Az antirevizionisták olyan módon próbálják elkerülni ezt a – szerintük – félreértést, hogy példákat sorolnak (rózsa illata, kávé illata, tej színe, gól belövése miatt érzett öröm, féltékenység, sárgás-narancsszínű utókép...) és reménykednek, hogy ezek hatására végbemegy a revizionista olvasó vagy vitapartner elméjében a szükséges *gestalt switch*, ezt azonban hiába várják, ha hallgatóságuk nem eleve meggyőződéses antirevizionista.

Amikor az antirevizionista úgy próbálja érzékeltetni a revizionista számára, hogy amiről beszélni óhajt, az nem a rózsa vagy a kávé valamilyen érzéki tulajdonsága, hanem valami *nagyon rejtélyes* dolog, az, ami *kimarad* a tudományos beszámolókból, és ami még az élenjáró tudomány számára is megismerhetetlen, akkor a revizionista végre megérti, honnan fúj a szél: az antirevizionista szemlátomást tudományos felvilágosításra szorul, mert nem érti, hogy a szagérzékelés milyen kémiai reakciókon alapul, vagy a szín hogyan feleltethető meg a fény hullámhosszának.

A revizionisták sincsenek éppen könnyű helyzetben, ha antirevizionista hallgatóságukat arról kívánják meggyőzni, hogy nincsen tudatprobléma, mert nincsen tudatosság bizonyos könnyen megmagyarázható kognitív funkciók ellátásán túl, és ennek megfelelően vagy értelmetlen az a kérdés, hogy egy termosztátnak lehet-e melege vagy hogy egy emberi agy számítógépes szimulációja lehet-e szerelmes, vagy pedig könnyen megválaszolható egy olyan kétértelmű válasszal, ami az antirevizionistákat csak nagyobb tanácstalanságban hagyja, mint a válaszadás előtt.

Ugyanis az antirevizionisták a revizionista érvelését vagy úgy fogják értelmezni, hogy egyszerűen nem arról a témáról beszél, amiről szó van: a tudományos anekdoták végeláthatatlan sora, a különféle kognitív mechanizmusok működésének elmagyarázása nem is érinti a tudat kérdését – így semmilyen vita nincs köztük és a revizionista között, vagy pedig úgy, hogy a revizionista – zombiként vagy doktrinér ideológusként – annyira más pszichológiai tulajdonságokkal rendelkezik, mint hallgatói, hogy belátásai az utóbbiak számára nem hasznosíthatók.

Miközben úgy tűnik, hogy az elmefilozófiai viták egyik oka az A- és R-intuíciók különbsége, aközben az is világos, hogy az elmefilozófiai elméletek, noha szerzőik gyakran világossá teszik, hogy preferálják műveik R- vagy A-olvasatát, nem tudnak kikényszeríteni egy ilyen olvasatot. Ezért aztán az R-elméletnek és A-elméletnek szánt elmefilozófiai elméletek nem is ütköznek össze azon a szinten, hogy az A-tézisek képviselői valódi érveket tudnának felhozni az antirevizionizmus mellett, és erre válaszul az R-tézisek képviselői valódi érveket hoznának fel a revizionizmus mellett. Ehelyett az történik, hogy az antirevizionista vagy feltartott kézzel elismeri, hogy vitapartnere zombi (A2), vagy azt állítja, hogy vitapartnere témát vált, nem arról beszél vagy ír, amiről voltaképpen kellene (A1), míg a revizionista vagy elismeri, hogy vitapartnere tudatos lény (szemben vele) (R2), vagy pedig azt állítja, hogy vitapartnere a saját tudományos tudatlanságát keveri össze azzal a ténnyel, hogy van, amit a tudomány „kihagy” és nem tud megmagyarázni (R1).

Egy antirevizionista indokoltnak tarthatja a fizikaitól különböző mentális szubsztanciák vagy tulajdonságok feltételezését, ha sikerül meggyőznie magát arról, hogy ez valahogyan „megoldja a tudatproblémát”. Ezzel szemben egy revizionista számára a szubsztancia- vagy tulajdonságdualizmus, amennyiben a nem fizikai szubsztanciák és tulajdonságok létezésének feltételezésének célja az, hogy a revizionista számára nem létező tudatproblémát oldják meg vele, fölösleges feltételezés (R2). A revizionista továbbá azt is gondolhatja, hogy a dualista igazából semmiféle olyan tudatproblémában nem hisz, amiben ő sem, csupán azért feltételezi például a testtől különböző elmét, mert nem tudja elképzelni, hogy a koponyában hogyan férhet el egy, az

elme képességeivel rendelkező szerv: tehát a dualizmus igazából egy téves, elavult tudományos magyarázatkísérlet, amit az agykutatás eredményei fölöslegessé tesznek (R1).

Amit a dualizmusról mondtunk, elmondható a funkcionalizmus antirevizionista értelmezéséről is, amely nonreduktív materializmusba (vagy más nonreduktív elméletbe) torkollik. Egy revizionista számára egy ilyen elmélet felesleges feltételezéseket tartalmaz, hiszen nem létező problémát próbál megoldani.

A dualizmus és a nonreduktív elméletek, amennyiben forrásuk nem tudományos tájékozatlanság és fogalomzavar (mint némely revizionisták gyanítják), akkor kizárólag egy antirevizionistának lehetnek vonzók, hiszen revizionista szemszögből fölösleges elemeket tartalmaznak. Problémájuk az, hogy egyrészt az antirevizionisták valójában nincsenek rászorulva, hogy ilyen elméletek mellett álljanak ki, másrészt pedig ezeknek az elméleteknek is van revizionista olvasatuk, még ha „ellentétesnek” is tűnik azokkal a szándékokkal, amely az efféle elméletek gyártása mögött munkál.

Legújabbban az illuzionisták kísérelték meg, hogy komolyan vegyék az R2 alternatívát, és pedig nem abban a komolytalan formában, hogy „míg mi zombik vagyunk, az antirevizionisták nem”, hanem abban a formában, hogy feltételezik: az antirevizionisták *illúziókban* szenvednek, ezek hatására foglalnak el hamis filozófiai álláspontot. (Természetesen az R1 illuzionizmus is létezik, melynek képviselői elismerik, hogy ők maguk is szenvednek a tudat illúziójában. Így például Dennett egy helyütt azt állítja, hogy ő is rendelkezik azokkal az intuíciókkal, amelyek Chalmerst félrevezetik, csak éppen nem ad nekik hitelt (Dennett 2012: 88, Dennett 2016: 67).

A mai illuzionizmust egy olyan ígéretes irányzatnak tekinthetjük, amelynek képviselői alaposan számot vetettek a kétféle olvasat problémájával. Kérdéses, hogy egy antirevizionistát ugyanakkor meg tudnak-e győzni arról, hogy nem tudatos lény, csupán illúziókban szenved, ám még nagyobb probléma, hogy az illuzionizmusnak is van antirevizionista olvasata: az tudniillik, ahol úgy értelmezik, hogy „a tudatosság illúziója”, amiről az illuzionista beszél, valójában ugyanaz, mint a tudat, amiről az antirevizionista beszél. Hogy elhárítsa a A-olvasat lehetőségét, az illuzionistának világossá kell tennie, hogy amikor illúzióról beszél, akkor tisztán kognitív illúzióról beszél, mely maradéktalanul „funkcionalizálható”, ám ez a manőver természetesen azt az A-olvasatot is kiprovokálhatja, mely szerint az illuzionista témát vált, mert csak a tudatosságról való gondolkodásról beszél, de nem beszél „magáról az élményről”. Kérdéses tehát egyelőre, hogy az illuzionizmusnak sikerül-e megoldania a kétértelműség problémáját.

E rövid összefoglaló után egy követelményt állíthatunk fel az elmefilozófiai elméletekkel szemben. Az ennek a követelményeknek megfelelő elméletek kevésbé vannak kitéve azoknak a kétértelműségeknek, mint azok, amelyek a követelménynek nem felelnek meg, továbbá az ilyen elméletek nem elfogadhatatlanok intuitív alapon sem a revizionisták, sem az antirevizionisták számára.

(EE) *Egyensúlyi elmélet*: egy elmefilozófiai elméletnek egyensúlyi helyzetet kell elfoglalnia az antirevizionista és revizionista intuíciók vonatkozásában. Ez azonban nem egyszerűen azt jelenti, hogy az elméletnek legyen A- és R-olvasata is, hiszen minden elmefilozófiai elméletnek van A- és R-olvasata. Az egyensúlyi elméletet az jellemzi, hogy teljesül

(MK) *A megfelelőség kritériuma*: az elmélet határozott válaszokat ad az antirevizionista tradícióban felvetődő kérdésekre, vagy legalábbis világosan látható, hogy az elmélet keretein belül ezekre elvileg hogyan lehetne egyértelműen válaszolni. Az elméletben a fenomenális tudatosságnak valamilyen intrinzikus vagy legalább relatíve intrinzikus tulajdonságokon kell alapulnia.

(RK) *A redundancia kritériuma*: az elmélet semmilyen olyan entitást vagy tulajdonságot nem feltételez, semmi olyasmivel nem bővíti ontológiánkat, aminek a bevezetése revizionista nézőpontból fölösleges, és csak azért feltételezték, hogy antirevizionista igényeket elégítsenek ki.

Nyilvánvaló, hogy a dualizmus és a non-reduktívizmus vét a redundancia kritériuma ellen, hiszen revizionista szemszögből nem létező problémák megoldása érdekében bocsátkozik kalandos feltevésekbe, miközben a revizionista funkcionalizmus (például Dennett 1991-es álláspontja) vét a megfelelőség kritériuma ellen. Antirevizionista nézőpontból nem csupán az a probléma Dennett felfogásával, hogy Dennett olyan könyvet írt a tudatról, mely a tudat témájához egyáltalán nem szól hozzá (mint Lanier (1995) panaszolta), hanem az, hogy a revizionista szerzők álláspontja alapján nem világos, mi a tradicionális antirevizionista problémák (mások színlátásának problémája, gilisztaprobléma) megoldása.

Az az elmélet, mely megfelel mind a megfelelőség, mind a redundancia kritériumának, a reduktív materializmus, mely szerint a fenomenális tulajdonságok fizikai tulajdonságokkal azonosak. Ezért amennyiben a filozófiai vitákat sújtó kétértelműségeket és félreértéseket az

elmefilozófia haladása szempontjából szerencsétlennek tartjuk, akkor jó okunk van preferálni a reduktív materializmust a nonreduktív fizikalista és a nem fizikalista álláspontokkal szemben.

Mielőtt rátérek annak kifejtésére, hogy a reduktív materializmus miért felel meg a fent megfogalmazott követelményeknek, még szükséges annak diagnosztizálása, hogy a filozófusok sok esetben miért nem tudnak megbirkózni elméleteik kétértelműségével, vagyis a revizionisták miért nem tudják meggyőzni az antirevizionistákat arról, hogy nem térnek el a tárgytól, amikor a tudatosságról beszélnek, és az antirevizionisták miért nem tudják a számukra oly „nyilvánvalóan létező”-nek tűnő kválékat vagy tudatos tapasztalatokat olyan módon megnevezni vagy felmutatni, hogy a revizionisták is elfogadják azok létezését.

A diagnózis pedig már több mint fél évszázada rendelkezésre áll, aminek tükrében meglehetősen különös, hogy a 70-es évektől napjainkig az elmefilozófiában számos olyan vita folyt, amelyekre nem került volna sor, ha bizonyos korábbi eredményeket komolyan vettek volna (Tartaglia 2013: 5-6).

4. A fenomenális hozzáférhetetlensége

Az elmefilozófiai vitákat megnehezítő kétértelműségek és félreértések forrása *a fenomenális kognitív és szemantikai hozzáférhetetlensége*. Az egyensúlyi elméletnek ezért figyelembe kell vennie megismerési képességeinknek ezt a jellegzetességét.

Már szó volt róla, hogy annak a kérdésnek, hogy két különböző személy ugyanolyannak látja-e a piros vagy zöld színt, van egy sajátosan filozófiai változata vagy értelmezése. A kérdés feltevése ebben a sajátos értelemben még akkor is jogosultnak tűnik, ha a két személy minden körülmények között ugyanolyan módon reagál a szóban forgó színekre, tehát úgy tűnik, nem a színekkel kapcsolatos ízlésükben vagy a színlátásukban (a színek megkülönböztetésére való képességükben) rejlő különbségeket firtatjuk. Noha ugyanúgy reagálnak a színekre, ugyanúgy nevezik őket, ugyanolyan attitűdökkel rendelkeznek velük kapcsolatban, honnan tudjuk, hogy „kvalitatíve” is azonos a színlátási tapasztalatuk (Farrell 1950: 177)? Ha ezt a kérdést tesszük fel, akkor például a zöld szín látásának élményének *fenomenális minőségére* kérdezzünk rá.

Ugyancsak ezzel a minőséggel kapcsolatos a „milyen ...-nek lenni” kifejezésnek az a „sajátosan filozófiai” változata, amely elkülöníthető a kérdés filozófiailag nem problematikus, hétköznapi változatától. Sohasem tudhatjuk meg, milyen érzés olimpiai aranyérmét nyerni vagy

háborúban harcolni (kivéve persze, ha magunk is olimpiai bajnokok vagy háborús veteránok vagyunk), és egy másik értelemben sohasem tudhatjuk meg, hogy milyen érzés egy másik ember számára egy piros foltot nézni. Az első esetben arról a gyakorlati nehézségről van szó, hogy a nyelv és a képzelőerő nem elég gazdag ahhoz, hogy tökéletesen ábrázolja, illetve teljes részletességgel el tudjuk képzelni ezeket a helyzeteket. A második esetben arról az elvi akadályról, hogy saját nézőpontunkkal rendelkezünk csupán, másokét nem vehetjük fel. A problémának az utóbbi a sajátosan filozófiai változata, hiszen sokan úgy vélik (filozófusok és nem-filozófusok egyaránt), hogy egy bizonyos értelemben sohasem tudjuk meg, másvalaki milyennek lát egy piros foltot, annak ellenére, hogy ezzel kapcsolatban biztosan nem merül fel olyan jellegű gyakorlati akadály, mint az olimpikon vagy a háborús veterán esetével kapcsolatban.

A revizionisták azonban a kérdés e „sajátosan filozófiai” változatának értelmességét kétségbe vonják. Ha egy revizionista intuíció-pluralista (R2 állásponton van), akkor vélheti úgy, hogy aki ezt a kérdést felteszi, az illúzió áldozata (érez valamit, vagy legalábbis azt hiszi, hogy érez valamit a szóban forgó elképzelt szituációkkal kapcsolatban, amit azonban nem tud úgy kifejezni, hogy a revizionistát is meg tudja győzni arról, hogy a kérdés „sajátosan filozófiai változata” értelmes vagy érdekes). Ha intuíció-monista (R1 állásponton van), akkor pedig a kérdést olyan kérdésként értelmezi, amelyre valójában sohasem hiányzott a válasz, mihelyt megmagyaráztuk például a színérzékelés fiziológiai folyamatait. Csak annak tűnik úgy, hogy van még nyitott kérdés, aki nem értette ezeket a magyarázatokat, és fogalomzavarokkal küszködik. Más szóval, a kérdést értelmesnek tartja, csak éppen nem annak „filozófiai változatát”.

A revizionisták számára tehát a színérzékelés során semmiféle olyan „fenomenális minőség” nem jut szerephez, ami valamiféle többlet lenne a látott tárgy színén (mely fizikai tulajdonság) vagy a színt látó személy agyának állapotán (neurofiziológiai tulajdonság) felül. Filozófiailag ez az állítás például egy *transzparenciatézis*ként juthat kifejeződésre, melynek értelmében amikor például egy piros almát látunk, és az alma pirosságára irányítjuk figyelmünket, akkor figyelmünk tárgya az alma (felszínének) valamilyen tulajdonsága, nem pedig „tapasztalatunk intrinzikus tulajdonsága” (Harman 1990: 41).

Ha egy materialista elmefilozófus azt állítja, hogy a színérzékelés tapasztalatának fenomenális minősége nem más, mint valamilyen fizikai tulajdonsága annak a neurofiziológiai állapotnak, amibe akkor kerülünk, amikor meghatározott fizikai tulajdonságú tárgyat látunk, akkor ennek az állításnak van revizionista és antirevizionista olvasata is. Antirevizionista olvasatban ugyanis a materializmus annyit állít, hogy létezik tudatos tapasztalat, „valamilyen” a látás folyamatát átélni, és ez a sajátos szubjektív minőség nem más, mint a látás neurofiziológiai

folyamatának valamilyen fizikai jellemzője. Revizionista olvasatban a materialista álláspontot úgy interpretálhatjuk, hogy nincs semmiféle „szubjektív érzetminőség”, nincs értelme arról beszélni, hogy „valamilyen” átélni a látás folyamatát, csupán fizikai és fiziológiai folyamatok vannak.

Ha egy filozófus annak érdekében, hogy kikényszerítse az érzéki észlelésről adott elméleti beszámolójának antirevizionista olvasatát, feltételezi, hogy az észlelésben az észlelt tárgy fizikai tulajdonságai és az elme perceptuális állapotai közé közvetítőként még valamilyen speciális „fenomenális tulajdonság” is betolakodik, amely önmagában felelős azért, hogy „valamilyen” átélni a kérdéses perceptuális állapotot, akkor ezzel a tudat filozófiai problémája szempontjából felesleges munkát végez. Ennek az elméletnek ugyanis van revizionista olvasata, mely szerint az nem más, mint egyszerűen az érzetadatok elmélete, egy meglehetősen sokat bírált (Ryle 1949/1999: 259-273) észleléselemélet, amely ráadásul pontosan olyan jól (vagy rosszul) lenne alkalmazható a filozófiai zombik észlelésének leírására, mint a tudatos lényekére.

Livingston (2004) szerint az elmefilozófia történetét végigkíséri az a folyamat, hogy a szerzők újra és újra megkísérelték elméleti kereteikben elhelyezni a tudatos tapasztalat fenomenális minőségét, de ebben minduntalan kudarcot vallottak, mert mindig strukturális magyarázatokra törekedtek, a fenomenális minőség azonban a *strukturával* szembeállítható (Schlick kifejezésével) *tartalom*, a tapasztalat olyan intrinzikus minősége, mely a relacionális, strukturális tudományos elméletekben nem játszik szerepet. Ennek megfordításaként pedig bármilyen entitás vagy tulajdonság, ami a tudományos elméletekben magyarázó szereppel bírhat, már nem a fenomenális minőség, és így egy fogalom, amely szerepet játszik tudományos magyarázatokban, nem referálhat a fenomenális minőségre.

A következőképpen értelmezhetjük azt az ebből következő, az elmefilozófia történetében többször is visszatérő jelenséget, melyet „fogalmi inflációnak” is nevezhetünk. Olykor a filozófusok új terminus technicust definiálnak, melynek (legalábbis a szándéknyilatkozatok szintjén így tűnik) a tapasztalat fenomenális minőségére kellene valamilyen módon referálnia. Amíg csak egy szó kiválasztása történt meg erre a feladatra, továbbá a releváns szándéknyilatkozat hangzott el, addig úgy tűnik, hogy az új fogalom alkalmas lehet erre. Elnevezhetjük például a tapasztalat fenomenális minőségét *nyers érzetnek*, megkülönböztetve azt az *érzetektől*, melyek már meglévő észleléseleméletekben vagy ismeretelméletekben szerepelnek mint a tudás konstituálásában valamilyen szerepet játszó tényezők. Ily módon hangsúlyozhatjuk, hogy amikor nyers érzetéről beszélünk, akkor az ismeretszerzésben vagy észlelésben betöltött funkciótól, pszichológiai, biológiai vagy fizikai vonatkozásuktól elvonatkoztatva szeretnénk valami olyasmiről beszélni, ami megfelel annak, hogy „milyen” egy bizonyos színt látni vagy hangot hallani.

Amikor azonban pontosabban meghatározzuk az új terminus *technicus* („nyers érzet”) jelentését, meg kell határoznunk helyes használatát különféle kontextusokban, és azt, hogy a nyers érzetekről szóló kijelentések milyen logikai viszonyban állnak más tárgyú kijelentésekkel. Ezek a kontextusok észleléseleméletiek, ismeretelméletiek vagy éppen fizikaiak is lehetnek. Mihelyt elvégeztük terminusunk jelentésének e pontosítását, a terminus már „elvesztette tisztán fenomenális jellegét”: ugyan már használható egy észleléseleméletben vagy episztemológiai elméletben, de arra már éppen nem alkalmas, hogy a segítségével valami olyasmit fejezzünk ki, aminek biztosan nincs revizionista olvasata. Az érzet fogalmának triviális módon adódik revizionista használata: ha az érzet az érzékszervek által a tapasztalóval a megismerés valamilyen korai fázisában közölt információt jelenti, vagy a külvilágról szerzett ismereteinknek azt a részét, amely a legközvetlenebbül függ a tárgyak valamilyen fizikai jellezőjétől (szín, szag stb.), akkor érzet tulajdonítható filozófiai zombinak vagy akár termosztátnak is. Ezen próbál segíteni az antirevizionista filozófus, amikor bevezeti a nyers érzet fogalmát. Ám ahhoz, hogy egyáltalán értsük, mit ért nyers érzeten, azt valamiképpen definiálnia kellene, vagy egyéb módon be kellene vezetnie. Ha azonban a *nyers érzet*et a kávé illatára, az ég színére, vagy egy Bach-fúga hallgatásának élményére hivatkozva definiáljuk, akkor ez a definíció fizikai, illetve pszichológiai tulajdonságokat fog felsorolni. Mi kapcsolja össze ezeket? Mi a közös a Bach-fúgákban, a rózsákban és a kávéban? Ezt az antirevizionista filozófus csak úgy tudja megfogalmazni, ha említést tesz arról, hogy az említett objektumokat érzékeljük, illetve bizonyos hatással vannak ránk. Revizionista olvasatban ekkor a nyers érzet valami olyasmit fog jelenteni, mint egy perceptuális állapot, amit például a kávé illata vagy az ég színe, esetleg érzelmi állapot, amit például a zene alakít ki egy tapasztalóban. Ha a *nyers érzet*et annak a percepcióelméletnek vagy ismeretelméletnek a terminusaiban definiáljuk, amelyben a *nyers érzet* fogalmának szerepet szánunk, akkor a *nyers érzet* percepcióelméleti vagy ismeretelméleti fogalom lesz: a *nyers érzet* lehet például „az észlelés első fázisában szereppel bíró tulajdonság” vagy „az ismeret analízisének egyik végpontja”. Nyilvánvaló, hogy egy így definiált fogalom revizionista olvasatban is tökéletesen értelmezhető és használható, vagy más szóval, ha egy zombiról állíthatjuk, hogy észlel vagy megismer, akkor nyers érzetei is vannak. Ez viszont azt jelenti, hogy a nyers érzet fogalmának megalkotásával egyetlen lépéssel sem kerülünk közelebb egy egyértelműen antirevizionista elmélet megalkotásához. Csak addig tűnik úgy, hogy lehetséges csak antirevizionista módon értelmezhető fogalmat definiálni, amíg semmit sem mondtunk erről az új fogalomról, vagyis igazából még nem definiáltuk. Mihelyt definiáltuk, a kétértelműség előállt. Revizionista szemszögből viszont az efféle új fogalom bevezetése csak feleslegesen túlbonyolítja elméleteinket.

Ennek következtében mihelyt számottevő mennyiségű filozófiai irodalom keletkezik, amelyik használja a nyers érzet fogalmát, azoknak a filozófusoknak, akik szeretnék egyértelművé tenni, hogy a tudatról alkotnak elméletet, és nem pedig valamilyen ismeretelméleti vagy észleléseleméleti kérdésről, vagy akár a minket körülvevő tárgyak fizikai tulajdonságairól, előbb-utóbb új fogalommal kell előállniuk, aminek ezúttal ténylegesen a fenomenális minőségre kell referálnia. Ez az új fogalom, például a „kválé” fogalma, egészen addig nem lesz kitéve a kétértelműségnek, amíg nem kap ismeretelméleti, észleléseleméleti, pszichológiai vagy egyéb publikus használatot, vagyis mihelyt többet mondunk róla, mint azt, hogy a kválé az a bizonyos valami, amire utalva meg tudjuk fogalmazni az elmefilozófia illékony, egyesek számára szembetűnő, mások számára nem is érthető alapproblémáit. Mihelyt aztán sor kerül arra, hogy a kváléről többet mondunk ennél, a kválékra hivatkozó elmefilozófiai írások kézenfekvő revizionista olvasatokat tesznek lehetővé, és nemsokára – ha eltérő terminológiával is – de újra megfogalmazódnak a tudattal kapcsolatos standard filozófiai rejtvények (mint a mások színlátás-élményére vonatkozó kérdés).

Míg az 50-es és 60-as évek elmefilozófiai irodalmában általában *érzetekről* (Smart 1959) vagy az egyértelműség érdekében *nyers érzetekről* (Sellars 1965) van szó, az 1970-es évektől mindinkább divatba jön a *kválé* fogalma (Shoemaker 1975, Levine 1983), noha az utóbbi terminust korábban is alkalmazták, olykor párhuzamosan a többivel. Az illuzionizmus körül kirobbant vitákban ugyanakkor már a kválé fogalma, mely az 1990-es évek óta, többek között Chalmers munkásságának eredményeként, nagy népszerűségnek örvendett (annak dacára, hogy több szerző, mint például a már idézett Dennett, élesen bírálta), is gyanúba keveredett. Pete Mandik például úgy véli, hogy „az olyan terminusok, mint a ‘kválé’, ‘fenomenális minőség’ stb. híján vannak elégséges tartalomnak ahhoz, hogy bármi informatívát állíthasson az, aki megerősíti vagy tagadja a létezésüket” (Mandik 2016: 148). Ezzel szinte szó szerint ugyanazt mondja a kválékról, mint Farrell a „tapasztalat”-ról (Farrell 1950: 180). Mandik „kválé-kvietista”: a kváléfogalmat nem tartja alkalmasnak arra, hogy értelmesen állítsuk a kválék létezését vagy nemlétezését; sőt „meta-illuzionista” is, aki kételkedik abban, hogy valóban rendelkezünk azokkal az illúziókkal, melyekkel Frankish szerint rendelkezünk. (Tekintettel az elmefilozófiai elméleteket sújtó, eltérő intuitív meggyőződésekkel fakadó kétértelműségekre, ezen aligha csodálkozhatunk.) Mit tehet az a filozófus, aki a tapasztalat fenomenális minőségeiről szeretne beszélni, de elégedetlen a kválé (vagy éppen a nyers érzet vagy az érzet) fogalmával? Definiálhat új fogalmakat, vagy a meglévőket minimalista, minél kevesebb problematikus feltevés mellett elkötelező módon is felfoghatja.

Így például Schwitzgebel (2016) megkísérelte a fenomenális tudatosságot olyan módon definiálni, hogy a definíció ne legyen kitéve a dennetti és illuzionista kritikáknak. Nem szükséges tehát a kváléknak egyszerűnek, ineffábilisnak, privátnak stb. lennie. Schwitzgebel egészen minimális („a lehető legártatlanabb”) definícióra törekszik, amelynek nagy része pozitív és negatív példák felsorolásából áll (pozitív példa: hallani a ventilátor zúgását, negatív példa: diszponálva lenni arra, hogy egy matematikai kérdésre megfelelő választ adjunk). Ezen kívül megkísérli elhatárolni a fenomenálist a kvázifenomenálistól, az illuzionisták szerint ugyanis a fenomenális azonos a kvázifenomenálissal (bármivel, ami arra készlet bennünket, hogy fenomenális állításokat tegyünk). Ettől Schwitzgebel szerint a valódi fenomenális a következő („csodálatossági”) feltétel alapján különíthető el: a fenomenális tudatosság népi pszichológiai fogalma nyitva hagy bizonyos kérdéseket, például azt, hogy a csigák rendelkeznek-e fenomenális tudatossággal. A fenomenális tudatosság javasolt meghatározása tehát: a fenomenális tudatosság az a népi pszichológiai fogalom, amely megfelel a felsorolt pozitív és negatív példáknak, és lehetővé teszi, hogy problematikusnak tartsuk az említett nyitott kérdéseket.

Minden ilyen javaslat nehézsége, hogy a példák felsorolása csak akkor tenné egyértelművé, hogy a példák felsorolója mit szeretne közvetetten definiálni, ha csakugyan mindenki ugyanazokkal az intuitív meggyőződésekkel rendelkezne, ez azonban a kísérleti filozófiai felmérések eredményei tükrében kétséges. Mihelyt azonban egyértelműen meg akarnánk határozni, hogy mi a szisztematikus különbség a pozitív és negatív példák között, referálni kellene valamilyen módon a fenomenálisra. Természetesen erre történt kísérlet, amikor például a tapasztalat olyan intrinzikus, ineffábilis és abszolúte privát minőségeként fogták fel a filozófusok a kválékat, amikről tévedhetetlen tudásunk van. Az efféle elméletekkel szemben felhozott súlyos ellenvetések vezettek aztán ahhoz az igényhez, hogy „a lehető legártatlanabb” módon, vagyis a lehető legkevesebb problematikus feltevéssel határozzák meg a fenomenális tudatosságot azok, akik nem akartak a „tudat tagadóinak” táborába lépni. A „lehető legártatlanabb” definíció az lenne, ami semmit sem állítana a fenomenális tudatosságról, azon kívül, hogy az létezik, az ilyesmi azonban nem alkalmas definíciónak, mert nem tudja kiválasztani a definiálni kívánt jelenséget. Schwitzgebel definíciója ennél valamivel többet, tulajdonképpen annyit állít, hogy a fenomenális tudatosság az, ami népi pszichológiánk szerint közös a látás- és hallásélményekben (de nem merül fel az emésztésben), és amivel kapcsolatban értelmesen felvethető a kérdés, hogy a csigák rendelkeznek-e vele. Ez a definíció azonban csak antirevizionista olvasatban tudja kiválasztani a fenomenális tudatosságot, hiszen revizionista olvasatban úgy is értelmezhetjük, hogy a definíció, mondjuk, az „észlelés” vagy az „élet” népi pszichológiai fogalmát próbálja megragadni.

A következőkben bemutatom, hogyan fogalmazták meg a fenomenális hozzáférhetetlenségének problémáját azok a szerzők, akik jelentős hatással voltak a reduktív materialista elmefilozófia klasszikus megfogalmazóira.

4.1 Schlick, Wittgenstein, Farrell, Ryle

4. 1. 1 Schlick

Amikor élményekről beszélünk, például arról, hogy milyenek látunk egy adott színfoltot, akkor valójában sohasem közvetlenül erre a fenomenális minőségre referálunk, és az általunk használt nyelvi kifejezések sem kifejezetten ezt a minőséget írják le, továbbá sohasem képezik tudásunk részét kizárólagosan és közvetlenül a tapasztalat fenomenális minőségére vonatkozó hiedelmek. Schlick így ír erről:

Egy zöld levél fekszik az asztalomon... Lehetséges közölni *zöltségét* valakivel, aki nem látja a saját szemével? ... Bizonyos értelemben tökéletesen lehetséges beszélni a színekről és arról, hogy miféle zöldet látok, de... egy másik értelemben egyáltalán nem lehetséges közölnöm semmiféle fogalmat erről a bizonyos fajta zöld színről annak, aki maga nem látja közvetlenül (Friedl 2013: 178-179).

Milyen értelemben közölhetetlen egy zöld levél zöld volta? Schlick szerint minden kommunikáció „csakis struktúrát fejez ki és közöl” (Friedl 2013: 182). Ami strukturálódik, az mindig valamilyen *tartalom*, azonban maga a tartalom nem tárgya a kommunikációnak. Noha beszélhetünk a tartalomról abban a minimális értelemben, hogy éppenséggel kimondhatjuk, hogy szemben a formával, áll valamiféle tartalom, nem tudjuk kimondani, micsoda a tartalom:

Nem tudod leírni, mit értesz a „zöld” szón, ha azt a sajátos tartalmat vagy minőséget akarod jelölni vele, aminek ... „éppen tudatában vagy”, miközben erre a zöld levélre

nézel... a „zöld” *tartalom* kifejezhetetlen. Ez nem valami különleges felfedezés vagy új észrevétel, ezt mindenki tudja. (Friedl 2013: 184-185)

Természetesen értelmesen beszélhetünk arról, és ennek során jól megérthetjük egymást, hogy egy bizonyos zöld levél milyen színű, például más levelekkel, ismert zöldárnyalatokkal összehasonlítva, ezek azonban strukturális információk közlései. Nem magát a levél színének látásának intrinzikus milyenségét közöljük, hanem a különféle zöld objektumok színeinek egymással fennálló hasonlósági viszonyaira vonatkozó tényekről számolunk be. Egy másik lehetőség egyszerűen megmutatni a zöld színt, amiről beszélni akarunk, ám ez sem jelenti, hogy a *tartalom* kommunikálható lenne. Felvetheti ugyanis valaki, hogy egy bizonyos levél zöld színe „közlhető” lenne úgy, ha egy papírlapra megfelelő színű foltot festenénk, és ezt a lapot eljuttatnánk a címzetthez. Csakhogy ez a festékfolt „nem pusztán kifejezné a zöldet, hanem ő maga lenne zöld” (Freidl 2013: 200). Ez nem a zöld szín közlése, hanem csupán szállítása lenne. A *tartalom* „nem lép be a nyelvbe, és mindig mintha elsiklana előlünk, amikor már azt hisszük, elkaptuk” (Friedl 2013: 202).

Mindamellettt értelmesen lehet vitatkozni például arról a kérdéstről, hogy két ember ugyanolyannak látja-e ugyanazt a zöld foltot. Ekkor azonban nem a zöld folt vagy a két megfigyelő színlátási élményének intrinzikus minőségéről beszélünk, arról ugyanis nem tudunk beszélni, hanem mindig a szituáció valamilyen strukturális jellemzőjéről. Így például arról, ha az egyik személy színlátása ép, míg a másik színvak, ami végső soron viselkedésükből (vagy fiziológiai tulajdonságaikból) is megállapítható. Olyan „sajátosan filozófiai” kérdést Schlick szerint valójában meg sem tudjuk értelmesen fogalmazni, ami nem erre irányulna, hanem arra, hogy például két teljesen ép színlátású, neurofiziológiailag normális ember ugyanazt a színfoltot nézve láthatja-e azt egészen más színűnek, de úgy, hogy a helyzet minden más vonatkozásban megegyezik egy normális szituációval. Mindezzel együtt Schlick egyáltalán nem állítja azt, hogy *nincs* közvetlen, intrinzikus fenomenális minősége a zöld színfolt látásának. Csak ennek közölhetetlenségét állítja.

Mindez pedig a legkevésbé sem jelenti, hogy a színek bármilyen gyakorlati értelemben rejtélyesek lennének, vagy azt, hogy nem tudnánk bánni a színekkel (vagy más *tartalommal*) vagy sikeresen kommunikálni róluk. Erre a legjobb példát a lírai költészet jelenti, amelynek célja, hogy szubjektív érzelmi, hangulati tartalmakat közöljön olvasójával, és messzemenőig sikeres lehet abban, hogy az olvasóban akár olyan hangulati állapotokat keltsen fel, amire nincsenek hétköznapi szavak. Ez azonban szigorúan véve nem *tartalmak* közlése. A lírai költészet az olvasóval nem

érzelmeket közöl, hanem az olvasó a költemény segítségével reprodukál, előállít magában bizonyos érzelmi állapotokat, azokat közvetlenül megtapasztalva (Freidl 2013: 248).

Schlick felhívja a figyelmet a „tudni, milyen...” kifejezés bizonyos használatának félrevezető jellegére is. Ez akkor következik be, amikor azt gondolják, hogy tudást eredményezhet a „közvetlen ismeretség”. Ha életemben először piros színt látok, akkor ettől kezdve „tudom, milyen” piros színt látni, mert „közvetlen ismeretségbe” kerültem a piros színnel, és „ismeretségből szerzett tudásra” (Russell 1912, V. fejezet) tettem szert arra nézve, hogy „milyen” piros színt látni. Ez Schlick szerint hibás megfogalmazás. Nevezhetjük ugyan a piros szín látását a piros színnel való „közvetlen ismeretségbe kerülésnek”, sőt, azt is elfogadhatjuk, hogy valaki onnantól kezdve, hogy egyszer már közvetlen ismeretségbe került a piros színnel, fel tudja idézni vagy fel tudja ismerni a piros színt, és ezt kifejezhetjük azzal a szófordulattal is, hogy „most már tudja, milyen piros színt látni”. Szükségtelen azonban „közvetlen ismeretségből vett tudásról” beszélni, amikor bőven elég közvetlen ismeretségről. Tudás tartalma csak strukturális jellegű lehet, tartalom nem. Egy színnel kerülhetünk ugyan ismeretségbe, de ezzel nem teszünk szert tudásra. „A tartalom tartalom; semmit sem lehet vele csinálni, egyszerűen ott van” (Freidl 2013: 218).

Ha nem szerezhetünk tudást a tartalomról, akkor egyáltalán mit tehetünk vele, hogyan viszonyulhatunk hozzá? A tartalmat megtapasztaljuk, illetve *átéljük*. Londoni előadásaiban Schlick az „élvezni” (enjoy) szót használta, mivel az angolban nincs pontos megfelelője az *erleben*-nek. Az Erfahrung (tapasztalat) szó nem alkalmas, mert tudás megszerzését konnotálja, az angol experience pedig inkább az Erfahrunghoz áll közel. Az élvezet (átélés) pedig „nem a tudás legmagasabb vagy akár legalacsonyabb foka, hanem egyszerűen a leírhatatlan, ami minden mást megelőz” (Freidl 2013: 221). Az „élvezni” szót ebben az értelemben Samuel Alexander kezdte használni, szembeállítva a „szemlélettel”, az elmének tárgyakra irányuló, tudást eredményező aktusaival. Az elme „*élvezi* magát és *szemléli* tárgyait” (Alexander 1920: 12). Schlick szerint „ha vízzé válnánk, víz lennénk, de... ez nem jelenti, hogy tudnánk, mi a víz. Ismeri az arany az arany természetét?... a közvetlen szemlélet élvezet [értsd: élmény], az élvezet pedig élet, nem tudás.” (Freidl 2013: 225). Schlick tehát hasonlóan Farrellhez és Hackerhez, úgy vélte, hogy a „tudni, milyen ...-nek lenni” szófordulat semmi többet nem jelent, mint a „...-nek lenni” jelenlegi vagy befejezett tényét.

Amikor valaki megtudja, „milyen piros színt látni”, akkor bekövetkezik az a tény, hogy az illető átélte, a piros szín látását. Ez pedig képessé teszi olyan tettekre, amelyekre eddig nem volt képes. Így például felidézhető epizodikus emlékévé válik a piros szín látásának tapasztalata. Mivel az emberek képesek különféle emlékeiket egymással, illetve jelenlegi érzeteiket emlékeikkel összehasonlítani, az újszerű emlékekkel rendelkező személy új képességekre is szert tesz, így

például ránézésre azonosítani tudja a továbbiakban a piros dolgokat (ha a piros szín első látásakor tudta, hogy piros színt lát). Lewis és Nemirow (Lewis 1988/2008: 346-351) ilyen módon válaszolta meg *képességhipotézisükben* azt a kérdést, hogy milyen új tudásra tesz szert Jackson gondolatkísérletének hőse, a szintudós Mary (Jackson 1982: 130), amikor fekete-fehér börtönéből kiszabadulva először lát piros színt. A képességhipotézis kompatibilis Schlicknek azzal az álláspontjával, hogy a tudás mindig strukturális természetű, illetve Farrellnek azzal az állításával, hogy a fenomenális érzetminőségek nem vetnek fel olyan filozófiai problémát, ami szükségessé teszi a viselkedés és élmény (experience) közötti különbségtételt.

Egyik esetben sem arról van szó, hogy az említett filozófusok tagadták volna, hogy „valamilyen” megtapasztalni a piros szín látását, vagy „van olyan, mint” a piros szín megtapasztalásának élménye. Csupán arról van szó, hogy az átélés *ténye*, noha végső soron hatással lehet a képességeinkre is (és így közvetetten arra is, hogy milyen tudásra tehetünk szert), Schlick kifejezésével, „csak” élet, nem pedig tudás.

4. 1. 2 Wittgenstein

Egy másik klasszikus gondolatmenet, amely a fenomenális hozzáférhetőségének vagy hozzáférhetetlenségének kérdésével kapcsolatban klasszikus hivatkozási pont, Wittgenstein érvelése (Wittgenstein 1953/1998. 244-315. §.).

Wittgenstein itt fogalmazza meg azt a tézist, amit Hill (1991: 161) nyomán expresszizmusnak nevezünk. Az expresszizmus szerint azok a kijelentéseink, melyek látszólag a tapasztalataink fenomenális minőségének leírásai (annak leírásai tehát, hogy milyen nekünk lenni egy bizonyos szituációban, illetve milyen számunkra valamilyen inger érzékelését megtapasztalni), valójában nem leírások, hanem bizonyos állapotaink *kifejezései*. Az expresszizmusához talán már az „élvezet” és „szemlélődés” megkülönböztetését kidolgozó Alexander is közel került, amikor a következőt írta:

nincs fajtabeli különbség az undorodás révén kifejezett érzés és az ezen érzés reflektív pszichológiai elemzési között. Helyettesítsük a ‘pfuj!’ kifejezést a jól kidolgozott beszéd egész apparátusával; a határozatlan undorézés helyett élvezzük ki az érzelem elemeit külön-külön, és megkapjuk az undor teljeskörű introspekcióját. (Alexander 1920: 18).

Ryle a következőképpen fejtette ki az expresszvizmust bizonyos hangulati állapotokkal kapcsolatban:

A „Levertnek érzem magam” beismerésekor nem teszünk mást, mint hogy végrehajtjuk az egyik olyan dolgot, mégpedig a beszélve csinált dolgok egyikét, amelynek elvégzésére a levertség hangol minket. (Ryle 1949/1999: 127)

Wittgenstein közismert expresszivista tézise szerint pedig „a fájdalom szóbeli kifejezése helyettesíti az üvöltést, s nem leírja” (Wittgenstein 1953/1998. 244. §.) Wittgenstein továbbá hasonló szellemben bírálja a „tudni, milyen” (megtapasztalni valamit) kifejezés „sajátosan filozófiai” használatát, mint Farrell vagy Schlick (Wittgenstein 1953/1998. 246 §.).

Wittgenstein elutasítja a privát, „belső élményeinket leíró” nyelv lehetőségességét (Wittgenstein 1953/1998. 256-261. §). Egy nyelv szavainak jelentése nem irányulhat privát élményeink minőségére: amennyiben értelmesek ezek a szavak, annyiban használatuknak publikus tényekre vonatkoztatott szabályai vannak, ha viszont ilyen szabályok nincsenek (a szavak ténylegesen privátok, és semmi mással nem állnak vonatkozásban, mint privát élményünkkel), akkor nem érthetőek. Szó sincs arról, hogy a hétköznapi életben ne tudnánk értelmesen beszélni a színekről vagy fájdalmainkról, azonban minden ilyen esetben úgy tűnik, mintha maradna egy rés az élmény szubjektív minősége és aközött a szerep között, amelyet az élményt magában foglaló fajtájú esemény a gyakorlatban betölt:

A privát élményben valójában nem az a lényeges, hogy mindenki saját példánnyal rendelkezik, hanem hogy senki sem tudja, vajon a másik is *ezt* birtokolja-e, vagy pedig valami mást. Fel lehetne tehát tételezni... hogy az emberiség egyik részének *egyfajta* pirosérzete van, a másik részének pedig egy másfajta. (Wittgenstein 1953/1998. 256-261. 272. §).

Wittgenstein nem bagatellizálja a filozófus (vagy a nem-filozófus) nehézségeit, amelyek akkor kínozzák, ha a fenomenális megnevezésének, megismerésének problémáival szembesül.

„Annyira szívesen mondanánk: ‘A fontos *ez*’, és közben önmagunknak rámutatunk az érzetre” (Wittgenstein 1953/1998. 298. §.). Sőt: „Ha az ember filozofál, akkor végül eljut oda, hogy már csak egy artikulálatlan hangot szeretne kiadni.” (Wittgenstein 1953/1998. 261. §).

Wittgenstein nem azt a következtetést vonja le, hogy nincsenek érzetek (vagy nem „valamilyen” átélni tudatos tapasztalatainkat), sőt, még azt sem, hogy értelmetlen erről a kérdéstről beszélni. A Vizsgálódások következő passzusait tekinthetjük az expresszizmus tömör megfogalmazásának:

Tételezzük fel, hogy mindenkinek volna egy skatulyája, benne pedig valami, amit „bogár”-nak nevezünk[...] mindenki azt mondja, hogy csak a *saját* bogarának látványából tudja, mi egy bogár. [...] lehetséges volna, hogy mindenkinek valami más tárgy van a dobozában. Sőt, az is elképzelhető volna, hogy egy ilyen dolog szüntelenül változzék. – De mi lenne, ha ezeknél az embereknél a „bogár” szónak mégiscsak volna használata? – Úgy ez nem abban állna, hogy egy dolgot jelölnek vele. [...] ha az érzet kifejezésének grammatikáját a ‘tárgy és megjelölés’ mintájára konstruáljuk, akkor a tárgy mint irreleváns kihullik a vizsgálódásból. (Wittgenstein 1953/1998. 293. §)

Továbbá:

„De azt csak el fogod ismerni, hogy különbséget jelent, hogy a fájdalomviselkedést kíséri-e fájdalom, vagy sem.” – Hogy elismerem-e? Mely különbség lehetne nagyobb! [...] Az érzet nem valami, de azért nem is semmi! Az eredmény csak az volt, hogy a semmi megtenné ugyanazt, mint egy olyan valami, amiről semmit sem lehet állítani. [...] A paradoxon csak akkor tűnik el, ha radikálisan szakítunk azzal az elképzeléssel, hogy a nyelv mindig *egyetlen* módon funkcionál, s hogy mindig ugyanarra a célra szolgál: gondolatokat közvetít” (Wittgenstein 1953/1998. 304. §).

Így például a verbális fájdalomviselkedés a fájdalom kifejezése, nem pedig leírása. Az érzet „nem valami”, amennyiben félrevezető lenne olyan *dologként* kezelni, amelyről leírásaink vannak. Ám ettől még nem is *semmi*. Mivel „a semmi megtenné ugyanazt, mint egy olyan valami, amiről

semmit sem lehet állítani”, ezért akár azt a következtetést is levonhatnánk, hogy érzetek nincsenek – a dobozok üresek, nincs bennük bogár. Másrészt Wittgenstein számos ponton jelzi, hogy nem utasítja el azt a kézenfekvőnek tűnő feltételezést, hogy mégiscsak vannak érzetek, ha nem is lehet róluk úgy és abban az értelemben beszélni, ahogy a publikus tárgyról. Mintha a revizionizmus és antirevizionizmus ellentétét látnánk itt, vagy legalábbis egy példát arra, hogy miért zavarbaejtően plauzibilis mind a revizionizmus (nincsenek érzetek; a dobozok üresek) és az antirevizionizmus (vannak érzetek; a dobozok nem üresek; mégiscsak nem mindegy, hogy érzünk-e fájdalmat vagy sem).

4. 1. 3 Farrell

Farrell Tolmant, a korabeli pszichológust idézi, aki szerint az érzékelés pszichológiája „kihagy valamit” magyarázataiból. Ugyan magyarázatokat nyújt az ingerek megkülönböztetésének képességeire, de a „megkülönböztetések pszichológiájából” kimarad valami: a „nyers érzetek”. A nyers érzetek szerinte „nem kezelhetők tudományosan” (Farrell 1950: 174). Az efféle fogalmi újításra azért van szükség, mert a korabeli elterjedt nézet szerint

Ha leírjuk, hogy X alany hogyan reagál a különféle ingerekre, akkor világos, hogy egyáltalán nem X érzeteivel foglalkozunk, hanem a viselkedésével; és a pszichológiának azok a felfedezései, melyek állítólagosan az „érzetekkel” kapcsolatosak, egyáltalán nem az „érzetekről” szólnak, hanem az organizmusoknak a fizikai ingerekkel szembeni érzékenységről. (Farrell 1950: 173)

Még ennél is zavarbaejtőbb Edgar Adrian következő (1946-ra datálható) megjegyzése:

Mégis megmarad a közbeeső események nyomasztó problémája. Az emberi elme valahol közbejön az érzékelési és mozgató idegpályák fizikai eseményei közti oksági láncban. (Farrell 1950: 170)

Az utóbbi idézet revizionista olvasatban annyit jelent, hogy nem tudjuk, hogyan működik az emberi elme, pontosan mi közvetít érzékelés és mozgás között. Ez kétségtelenül nem könnyű tudományos kérdés, és 1946-ban sem számított annak. Másfelől értelmetlen és szükségtelen misztifikáció lenne azt gondolni, hogy az idézet valami mást is jelent ezen felül. Antirevizionista olvasatban viszont az az értelmezés is plauzibilisnek tűnik, hogy a fenti idézet mintha annak az intuitív meggyőződésnek próbálna igazságot szolgáltatni, hogy az „érzékelést a mozgással összekötő oksági lánc” nem elégséges kép az emberről, hiszen ez az embert valamiféle automatának látta, holott az emberi állapotnak van egy olyan jellemzője is, amit úgy fogalmazhatunk meg, hogy *valamilyen* embernek lenni, vagyis rendelkezünk tudatos élménnyel.

Ha ez önkényes belemagyarázásnak is tűnik, az eggyel korábbi idézet esetében már ezt aligha állíthatjuk. Amikor valaki azt állítja, hogy az érzetekről szóló elméletek igazából nem az érzetekről szólnak, akkor ezt revizionista olvasatban nehéz másként értelmezni, mint úgy, hogy aki ezt a kritikát fogalmazza meg, az valamilyen egyéb elméleti elköteleződése miatt elégtelennek tartja az elméletet. Például vitalista, ezért számára egy olyan érzetelmélet, amely az *élan vital* szerepére nem tér ki, nem teljes. Vagy egyszerűen mást gondol az érzékelés fiziológiájáról, és az érzékelés fiziológiájával kapcsolatos tudományos viták keretein belül elutasítja az érzetekről szóló eddigi elméleteket, esetleg egy általa preferált elmélet javára. Antirevizionista olvasatban azonban meglehetősen kézenfekvő az az értelmezés, hogy aki arról beszél, hogy az érzetelméletek nem szólnak az érzetekről, egyszerűen csak nem tudja elég világosan kifejezni azt, hogy mit hiányol a meglévő érzetelméletekből, megfelelő terminológia híján. Megkockáztathatjuk, hogy aki ezt mondja, az igazából azt szeretné állítani, hogy az érzetelméletek nem szólnak a *nyers érzetekről*, vagy éppen a *nyers érzetek fenomenális karakteréről*, a *kválékról*.

Ha angol nyelvterületen valaki azt állítja, hogy a tapasztalat (experience) elméletei valójában nem magyarázzák meg a tapasztalatot (experience), akkor antirevizionista olvasatban megkockáztathatjuk, hogy igazából azt szeretné kifejezni: a tapasztalat (Erfahrung, experience) elméletei nem szólnak az élményről (Erlebnis, ‘enjoyment’). Mihelyt azonban pszichológiai művek jelennek meg az *élményről*, biztosak lehetünk abban, hogy az „élmény” fogalma, mely kezdetben alkalmas lehetett arra, hogy szembeállítva a tapasztalat fogalmával kifejezze azt a hiányzó összetevőt, ami az antirevizionisták szerint a tapasztalat *minden* elméletéből hiányzik, „kontaminálódik”, például túl sok korrelációt találnak a szerzők az „élmény” és a testi érzetek, érzelmi állapotok, vagy akár érzéki észlelési állapotok között. Innentől kezdve felmerülhet az a kritika, hogy az élmény elméletei már nem is az élménnyel foglalkoznak.

Visszatérve a Farrell által idézett véleményekre, az első idézet mintha a másodikban felvetett problémára reagálna: a „nyers érzetek” fogalmának bevezetésével lehet meghaladni annak a megfogalmazásnak az esetlenségét, hogy az érzet-elméletek nem szólnak az érzetéről.

Tanulmányában Farrell a „tapasztalat” fogalmát vizsgálja, bár az „élmény”-hez közelebb eső értelmében. Célpontja az a sokak által megfogalmazott elképzelés, hogy a tudományokból, melyek csak az élőlények viselkedését írják le, „kimarad valami”, éspedig a „tapasztalat”. Ha kimarad a tapasztalat a viselkedéssel foglalkozó elméletekből, akkor a tapasztalatot a viselkedéssel szembeállítva, attól élesen megkülönböztetve lehet értelmezni (Farrell 1950: 177), ahogyan például ez abban a meghatározásban is megnyilvánul, hogy a pszichológia a viselkedésről és a tapasztalatról, sőt, a viselkedésről és a tudatról szóló tudomány (Farrell 1950: 171).

Farrell példájában egy kísérleti alanyt megkérnek a laboratóriumban, miközben piros foltot mutatnak neki, hogy írja le, mint lát. Ekkor elmondhatjuk, hogy „Ha csak az X által a piros alakzatra adott megkülönböztető válaszokat és erre való készségeket és más ilyesmiket vesszük figyelembe, akkor kihagyjuk azt a tapasztalatot, amit átél, miközben rá néz.” (Farrell 1950: 173). Ez a mondat azonban Farrell szerint egy kicsit más, mint a hétköznapi életből hozható hasonló példamondatok, így például az, hogy „ha csak azt veszed figyelembe, hogy Y mit mond vagy tesz, akkor kihagyod, mit érez valójában” (pókerarca mögött). A különbség ugyanis az, hogy amit a hétköznapi példában „kimarad”, az olyan információ, amiről Y akár be is számolhatna, ha megkérdeznék.

Annak érdekében, hogy megmutassuk, hogy az, ami állítólag kimarad, valami nagyon különös, a „viselkedés” szó jelentéstartományát addig terjesztjük ki, amíg le nem fedti legalábbis a személy rejtett verbális és egyéb válaszait, válaszokra való készenléteit, a releváns testi állapotait [...] de e kiterjesztés dacára még mindig elégedetlenek vagyunk. Még mindig határvonalat akarunk húzni X ez értelemben vett viselkedése és tapasztalata közé [...] mert ez a distinkció még mindig értelmesnek látszik. Például még mindig fel akarjuk tenni a kisiskolás kérdést: „Amikor X és Y egyaránt piros foltot néz [...] kvalitatíve azonos a tapasztalatuk? Még mindig azt akarjuk mondani, hogy amikor egy született vak ember látását helyreállítják, akkor nemcsak az a helyzet, hogy ingerek egy új tartományára képes differenciáltan reagálni. [...] Még mindig különbséget akarunk tenni a robotok és az emberek között azt mondva, hogy az előbbiek nem rendelkeznek érzetekkel. [...] Így aztán még ha igaz is [...] hogy „az elméről és tudatról szóló

állítások [végső soron csak a] viselkedésről szóló állítások, még mindig úgy érezzük, azt akarjuk erre felelni, hogy ez legfeljebb csak pszichológiai szempontból igaz... (Farrell 1950: 177)

Miért „valami nagyon különös” Farrell szerint az, ami „kimarad” X viselkedésének leírásából? Azért, mert a kihagyott „tapasztalat” (a „nyers érzet”, vö. Farrell 1950: 188) tulajdonság- vagy részletnélküli (*featureless*). Ennek megfelelően X semmiféle megkülönböztető viselkedést nem tud vele kapcsolatban tanúsítani. Farrell szerint még azt is félrevezető lenne emiatt állítani, hogy „de legalább X tapasztalata megtörténik” (Farrell 1950: 178), másfelől az az állítás is problematikus, hogy „nem létezik olyan, hogy X tapasztalata”. Minden más „Q nem létezik” állításnál ugyanis tudjuk, miféle Q létezését tagadjuk, itt viszont ezt nem mondhatjuk el (Farrell 1950: 180).

Farrell tulajdonképpen Wittgenstein „nem valami, de nem semmi”-jéhez hasonló felfogáshoz jut el: „X tapasztalatáról mint folyamatról vagy eseményről stb. beszélni csak zavarokat okoz” (Farrell 1950: 180). Wittgenstein szavaival: „Az első lépés teljesen észrevétlen. Folyamatokról és állapotokról beszélünk, ám eldöntetlenül hagyjuk természetüket! [...] De éppen ezáltal horgonyoztunk le egy meghatározott szemléletmódnál. Nekünk ugyanis meghatározott fogalmunk van arról, hogy mit jelent egy folyamatot közelebbről megismerni [...] Tagadnunk kell tehát a folyamatot [...] És ezért látszik úgy, mintha a szellemi folyamatokat tagadtuk volna.” (Wittgenstein 1953/1998. 308. §.)

Farrell számos példán keresztül érvel amellett, hogy nem szükséges a viselkedéssel szembeállított „tapasztalatról” beszélni ⁴ például ahhoz, hogy pontosan megértsük az olyan kérdéseket, mint azt, hogy „milyen lehet denevérnek lenni és úgy hallani, mint egy denevér”. Gondolatmeneteinek lényegét úgy foglalhatjuk össze, hogy tagadja, hogy az élmény („enjoyment”) a megismerés egy különleges (vagy bármilyen) fajtája lenne, és tagadja, hogy az „első kézből való” tudás (mint például „első kézből tudni”, milyen ópiumot szívni vagy denevérnek lenni) másféle tudás lenne, mint a szokványos („másodkézből való”) tudás. Természetesen csak akkor tudjuk „első

4...ami persze Chalmers számára az egész elmefilozófiai vizsgálódásának kiindulópontja: a pszichológiai és fenomenális tudatosság szembeállítása, vö. Kuusela 2016: 201-202, Tanney 2016: 217. Kuusela és Tanney számára úgy tűnik, hogy bizonyos elmefilozófiai kérdésfeltevések chalmersi „újrafelfedezése” Wittgenstein és Ryle bizonyos eredményeinek elfelejtésén alapul.

kézből”, hogy milyen ópiumot szívni, ha mi magunk ópiumot szívunk, azonban a *tudás*, amire ezáltal szert teszünk, nem tér el attól, mint amire elvileg más módon is szert tehetnénk.

Farrellt Strawson „a tudat tagadói” közé sorolta, idézve Farrellnek azt a lábjegyzetét, mely szerint Boringnak az az állítása, hogy „a tudat az, amit *közvetlenül* tapasztalunk”, „komikus és patogén megjegyzés” (Farrell 1950: 188-189). Farrell azonban aligha sorolható egyértelműen a tudat tagadói közé annak ellenére, hogy bizonyos értelemben eliminativista a tudattal szemben, hiszen számára a tudatos tapasztalat mint olyan nem dolog, nem esemény, nem folyamat, sőt, még azt sem lehet róla állítani, hogy nincs.

Farrell pozitív álláspontja nem könnyen megragadható, hiszen szerinte elfogadhatatlan a (tudatos élmény értelmében vett) „látás” nemlétezésének állítása éppúgy, mint azonosítása valamilyen neurális eseménnyel. Ez az elfogadhatatlanság azonban azért (illetve abban az értelemben) áll fenn, mert a mindennapi és tudományos nyelvhasználat gyakorlatában „nyilvánvalóan hamis” az az állítás, hogy „amikor látok valamit, akkor mindössze csak bizonyos neurális események (vagy folyamatok) történnek”. Másfelől a látás és a neurális események azonosítása ellen érvelni (szerinte) csak úgy lehet, ha kétféle eseménysorról beszélünk. Ebből az egyik a „látás” lesz mint „nyers érzetként” felfogott tapasztalás (Farrell 1950: 193). A „nyers érzetként” értett tapasztalatfogalomra csak akkor van szükség, ha azt szembe akarjuk állítani a viselkedési vagy természettudományos fogalmakkal. Ha a mentális eseményeket „nyers érzetként” fogjuk fel, akkor „láthatjuk, hogy abszurd még megkísérelni is ezeket azonosítani neurális [eseményekkel], mert ezek nem olyan fajta dolgok, amiket bármivel is azonosítani lehetne” (Farrell 1950: 194).

Farrell behaviorista álláspontja szerint a tapasztalat és viselkedés közötti különbségtétel csak akkor tartható fenn, ha a (viselkedéssel szembeállított) tapasztalat a „nyers érzetekre” korlátozódik (Farrell 1950: 194), amely manővert Farrell elutasítja. Így, mivel a tudománynak szempontjából semmi haszna nincsen a „tapasztalat” fogalmának, a jövőben a tudományos fejlődés hatására ez a fogalom eltűnhet, hasonlóan olyan más elavult fogalmakhoz, mint a „boszorkányság” (Farrell 1950: 195). Tanulmányának e szakaszaiban mintha határozottan a „tudat tagadóinak” táborához csatlakozna: a pszichológusoknak nem kell attól tartaniuk, hogy „kihagynak valamit”, elég „a viselkedés fiziológiai korrelátumait” megadniuk, és más dolguk nincs is (Farrell 1950: 197). Nemcsak az elme és anyag viszonyára vonatkozó kérdések értelmetlenségek, hanem az összes válasz is: a materializmus és annak szokásos cáfolatai egyaránt. Ezeket a konklúziókat azonban úgy is értelmezhetjük, mint ajánlásokat a jövő tudománya számára. Farrell sehol sem tesz olyan metafizikai kijelentéseket, melyek szerint nincs tudat, élmény vagy érzet, csak viselkedés vagy

anyag. Amikor azt az állítást elemzi, hogy egy emberszerűen viselkedő robotnak nincsenek érzetei, akkor arra mutat rá, hogy az ember/robot különbség kifejezéséhez nem szükséges a „nyers érzetekre” hivatkozni:

[Broad szerint] „Bármilyen tökéletesen átmegy egy test az intelligencia összes viselkedési tesztjén, továbbra is teljesen értelmesen felvethető kérdés marad: Valóban van elméje, vagy pusztán automata?... Mivel ez a kérdés felvethető [...], világos, hogy amikor elmét vagy mentális folyamatot tulajdonítunk egy testnek, akkor nem azt akarjuk állítani, hogy az bizonyos módon viselkedik.” (Farrell 1950: 191)

Farrell szerint Broad érve látszólag azért meggyőző, mert egyrészt kihasználja, hogy a „test” (*external body*) szót hallva általában nem gondolunk arra, hogy az ilyesminek elméje van, másrészt a „viselkedés” szóról a külsődleges, felszíni viselkedésre asszociálunk. Ha azonban egy robot kellő alapossággal reprodukálja „belső viselkedésünket” (*covert behaviour*) is, akkor az már talán nem is robot, hanem egy új típusú organizmus. Másfelől az „érzetek” tulajdonítását fenntarthatjuk kizárólag az organizmusoknak, vagy megengedhetjük akár robotok számára is. Végző soron az intelligens robotnak vagy elmét fogunk tulajdonítani, feltételezve, hogy organizmus (annak pedig, hogy mi organizmus, finom őrlésű behaviorisztikus kritériumai is lehetnek), vagy pedig nem, amennyiben az elmén nem azt értjük, amivel „bizonyos módon viselkedő testek” rendelkeznek. Broadnak azonban nem sikerül amellet érvelnie, hogy az elmén nem azt értjük, hogy amivel bizonyos módon viselkedő *organizmusok* rendelkeznek. Antirevizionista olvasatban viszont mindez azt jelenti, hogy Farrell nem feltétlenül állítja, hogy „nincs elme” vagy „nincs tudat”, vagy bármi hasonlót, hanem elfogadható számára egy olyan elme vagy tudat fogalma, amely bizonyos módon viselkedő organizmusoknak tulajdonítható.

4. 1. 4 Feyerabend

Feyerabend (1970) Farrellhez hasonlóan diagnosztizálta azt, hogy a „nyers érzeteket” jelölő fogalmak tartalomszegénységük révén tesznek szert arra a különleges státuszra, aminek hatására az a látszat keletkezik, hogy a tudatos tapasztalat kezelhetetlen a tudományos kutatás, illetve a materialista filozófia számára.

A „közvetlen ismeretségből vett” materializmus-ellenes érv szerint a mentális folyamatok mások, mint a fizikaiak, ugyanis az utóbbiakkal szemben „közvetlen ismeretségbe” kerülhetünk velük. Ez azt jelenti, hogy teljesen biztos tudásunk lehet saját mentális folyamatainkról, míg a fizikai objektumokkal kapcsolatos tudásunk mindig közvetett és sohasem teljesen biztos. Szemben a fizikai tárgyakkal, ami a mentális eseményeket illeti, „lényeg és megjelenés” ott egybeesik. Ha fájdalmat érzek, akkor tudom, hogy fájdalmat érzek, és ezzel kapcsolatban nincs helye kétségnek, nincs lehetőség tévedésre. Ez nem lenne így, ha a fájdalom esetében is különbséget lehetne tenni lényeg és megjelenés között éppúgy, mint a fizikai tárgyaknál.

Feyerabend kérdése: miből fakad a mentális folyamatainkkal kapcsolatos e bizonyosságunk? Válasza az, hogy ezen állapotok *tartalomnélkülisége* a kérdéses bizonyosság forrása. A fizikai objektumokra vonatkozó állítások „gazdag tartalommal rendelkeznek” (Feyerabend 1970: 148). Predikciókat fogalmazhatunk meg rájuk nézve, és ha ezek kudarcot vallanak, akkor az állítások visszavonandók. Nem ez a helyzet a mentális állapotainkra vonatkozó állításokkal. Tartalmuk olyan szegényes, hogy nem szűrhető le belőlük semmilyen predikció, így visszavonásuk sem válhat szükségessé. Ezt jelenti a velük kapcsolatos bizonyosság.

Feyerabend szerint a kortárs mindennapi nyelvben a mentális állapotra vonatkozó terminusok tartalomszegények, mert semmi olyasmire nem referálhatnak a mentális terminusok, amiknek köze van a testünkhöz vagy más fizikai entitáshoz (például a „fájdalom” mint hétköznapi fogalom csupán egy érzést jelent, és semmiféle fiziológiai jelentése nincs). Ha tartalmukat gazdagítanánk, és így hasonlóvá tennénk őket a fizikai objektumokat jelölő terminusokhoz, akkor „megszűnnének azon a sajátos módon működni, ahogy a mentális terminusok szoktak” (Feyerabend 1970: 149).

Az azonban, hogy nyelvünkben így működnek a mentális terminusok, tehát mentális terminusaink „viszonylagos szegényessége”, Feyerabend szerint nem szükségszerű, hanem „filozófiai újítás” következménye (Feyerabend 1970: 151). A régebbi („primitív”) kultúrákban például gyakori volt a mentális fogalmak objektivizálása.

A mentális állapotok csak azért tudnak „adottnak”, „közvetlenül ismertnek” tűnni, mert szegényes a tartalmuk. Mihelyt ez a tartalom gazdagabb lenne, nem lehetne a mentális eseményeket pusztán „ismeretség révén” ismerni, hanem ugyanúgy, mint a fizikai objektumok esetében, lehetne velük kapcsolatban természetükről és megjelenésükről beszélni, illetve tévedni velük kapcsolatban (Feyerabend 1970: 152).

Hogyan alkothatók ilyen fogalmak? Feyerabend szerint ehhez vennünk kell egy megfigyelésre vonatkozó kijelentést, majd meg kell tisztítanunk minden prediktív tartalmától és minden, publikus tárgyakra vonatkozó következményétől. Az így nyert „jelsorozat” objektíve ugyanolyan alkalmakkor fogják kiejteni, mint a kiindulását, azonban ez már egy csak „ismeretség” révén ismert objektumot fog jelölni. Kezdjük például a „Fájdalmat érzek” állítással (Feyerabend 1970: 155). Értsük úgy, mint egy ismeretség révén tudható tényekre vonatkozó kijelentést. Ez persze megváltoztatja az állítás eredeti tartalmát, de tegyünk úgy, mintha még mindig egy „értelmes állítással” lenne dolgunk:

Ha fájdalmat érzek, azt mondom, „fájdalmat érzek”, és természetesen van valami önálló elképzelésem arról, mik a fájdalmak. Nem lakoznak asztalokban és székekben; megszüntethetők gyógyszerek szedésével; [...] nem fertőzők [...] De ezen elképzelések egyike sem járulhat hozzá a közvetlenül adotról szóló új állítás jelentéséhez [...] Ha mindezeket az elemeket eltávolítjuk, mit fejezünk ki az új állítással? [...] Kimondhatom fájdalomérzés alkalmával [...] használhatom metaforikusan [...] vagy talán megtanítottak arra, hogy [...] akkor mondjam ki, amikor kellemes érzéseket érzek [...] Világos, hogy ezek a használati módok most már mind legitimek, és mind leírásai a „közvetlenül adott fájdalom”-nak. (Feyerabend 1970: 155)

Feyerabend ebben a példájában azt is szemlélteti, hogy miért lenne nyilvánvaló hiba a fájdalom fogalmát csupán fenomenális minőségek terminusaiban meghatározni, hiszen a fájdalomnak lényeges további vonása hedonikus karaktere (az tehát, hogy kellemetlennek tartjuk, szenvedést okoz), ráadásul még a fájdalom „milyensége” is alig elválasztható kauzális körülményektől (Ryle 1949/1999: 131-132). Csakhogy a funkcionális, kauzális terminusokban definiált fájdalomfogalom már más, mint a közvetlen ismeretség által megismerhető, tisztán „nyers érzetként” értelmezett fájdalom teljesen tartalom nélküli fogalma.

4. 1. 5 Ryle

Érdeemes Ryle-nak az érzetekről alkotott felfogásával is megismerkednünk, már csak azért is, mert Ryle elképzeléseinek kritikája, illetve továbbfejlesztésének szándéka fontos kiindulópontja volt a modern reduktív materialista tudatfilozófiának annak 1950-es évekbeli kezdeteinél. Ryle-t gyakran nevezik behavioristának, sőt, az 1950-es évek materialista elmefilozófusai kifejezetten felrötták neki, hogy elmélete kiegészítésre szorul, mert kimarad belőle az érzetek, illetve a tapasztalat „milyen”-ségének kielégítő tárgyalása.

Ahogy Wittgenstein is igen nehéz feladatnak találta az érzetekkel kapcsolatos megfelelő álláspont megfogalmazását, úgy Ryle esetében is hasonlót tapasztalhatunk. Ryle a fejezet elő- és és zárszavában elismerte, hogy az *A szellem fogalma* érzetekről szóló fejezetével elégedetlen volt, és meglehetősen sok kérdést nyitva hagyott ezzel a tárgykörrel kapcsolatban.

Az említett fejezetet tartalmilag úgy lehetne felosztani, hogy az érzetadatok elméletének, az érzéki benyomások elméletének és az érzetek „szellemi” entitásként való felfogásának kritikáját igyekszik nyújtani, összhangban Ryle-nak azzal az alapvető célkitűzésével, hogy bizonyítsa: „a ‘szellemi’ nem jelöli a dolgok vagy események olyan státusát, aminek következtében értelmesen megkérdezhetnénk róluk, szellemiek-e vagy fizikaiak” (Ryle 1949/1999: 246). Más szóval, Ryle célpontja a „hivatalos doktrína”, mely szerint az embernek különálló teste és szelleme van. Ez persze a karteziánus dualizmus (Ryle 1949/1999: 17.). A „hivatalos doktrína” szerint az érzetek szellemi események, vagyis nem-fizikaiak és nem nyilvánosak („magántermészetűek”, csak az őket átélő szubjektum számára ismerhetők), ugyanakkor aki rendelkezik velük, az mindenképpen tud róluk és tévedhetetlen velük kapcsolatban.

Az érzetadat-elmélet kritikája előtt azonban Ryle azt is világossá teszi, hogy „szellemi magaviseletet jelölő fogalmaink logikájának kiigazítását” (Ryle 1949/1999: 23) célul kitűző művéből az érzetek elemzése akár ki is maradhatna: „semmi ‘szellemi’ nincs az érzetekben [...] Az, hogy érzeteink vannak, nem jelenti értelmi vagy jellembeli képesség kifejtését. Ezért a csúszómászók nem lesznek nagyobbak a szemünkben, ha elismerjük, hogy vannak érzeteik.” (Ryle 1949/1999: 252) Csakhogy mégis szükséges az érzetek elemzése, mivel vannak, akik éppen az „érzetek, érzések és képzetek folyamát akarják az emberi szellem alapanyagává megtenni, és [...] így akarják alátámasztani azt a dogmát, hogy a szellem sajátos alapanyagból álló sajátos státusú dolog” (Ryle 1949/1999: 253).

Ryle szerint érzeteink privát volta az, ami miatt valami egész másnak tűnnek az érzetek, mint a külvilágbeli tárgyak, és így az érzetek folyamaiból álló szellemi életünk valami egészen más jellegű folyamatnak tűnik, mint a külsődleges, fizikai világban zajló életünk. Így lehet az érzetek

sajátos privátsága a dualista elképzelések támasza. Ryle válasza erre az, hogy igyekszik lebontani azt a feltételezett különbséget, ami az érzetek állítólagosan radikális, különleges privátsága és más, könnyen érthető, hétköznapi privátságok között áll fenn a népszerű elképzelések szerint. „Igaz, hogy a cipész nem figyelheti meg a kínokat, amiket én érzek, ha szorít a cipőm. De hamis az a kijelentés, hogy én viszont megfigyelem őket.” (Ryle 1949/1999: 253) Érzeteink sohasem lehetnek *megfigyeléseink* tárgyai; nem azért nem érzi másik személy a mi fájdalmunkat, mert a fájdalmak a „szellemi világ”-ban helyezkednek el, ahol csak egyvalaki számára *megfigyelhetők*, szemben a fizikai világban található publikus tárgyakkal, hanem egyszerűen csak logikai képtelenség, hogy más érezze, amit mi érzünk, ahogy „lehetetlen, hogy te lödd be az én gólokat [...] Lengyelország sem rendelkezhet Bulgária történetével” (Ryle 1949/1999: 258).

Ryle számára tehát az a fontos, hogy megdöntse azt az elképzelést, hogy az érzetek valahogyan, például introspekció révén megfigyelhető „dolgok” valamilyen szellemi közegben. Amikor ezt az elképzelést támadja, azzal nem állít semmi olyasmit, hogy nincsenek érzetek. Azt ugyan voltaképpen állítja, hogy saját projektje szempontjából az érzetek jelentősége csekély, de azt sem állíthatjuk, hogy az érzetek problémájának vizsgálata kimaradt volna művéből.

Ryle az érzetadatok elméletének támadásával szintén azt igyekszik megmutatni, hogy nem szükséges feltételeznünk különleges ontológiai státuszú, nem fizikai entitások („érzéki benyomások”, „érzetadatok”) létezését csak azért, hogy megmagyarázzunk olyan eseteket, mint amikor egy tárgyat másnak látunk, mint amilyen: „Amikor valaki azt mondja, hogy a tányér elliptikusnak néz ki, akkor nem valamilyen külön objektumot, nevezetesen ‘egy kinézetet’ jellemez elliptikusnak, hanem azt, ahogyan a ferdére állított kerek tányér valójában kinéz, ahhoz hasonlítja, ahogyan a nem ferdére állított elliptikus tányérok valóban kinéznek” (Ryle 1949/1999: 268). Amikor pedig másodlagos minőségekre vonatkozó kijelentéseket teszünk, akkor Ryle szerint a közönséges tárgyak publikus tulajdonságairól beszélünk. Abban a példamondatban, hogy „a gyümölcs valójában keserű, bár nekem egészen íztelennek tűnik” (Ryle 1949/1999: 271) nem arról számolunk be, hogy rendelkezünk egy keserűség-érzettel, hanem arról, hogy a gyümölcsöt íztelennek érezzük, vagyis olyan állapotban vagyunk, mint általában mások akkor, amikor íztelen gyümölcsöt esznek, noha ez a gyümölcs keserű. A keserűség pedig bizonyos gyümölcsök tulajdonsága, nem pedig érzeteké.

A mindennapi nyelvhasználatban (tehát filozófiai érveléseken kívül) az érzetekről mindig mint valaminek az érzetéről beszélünk, nem pedig pusztán érzetekről, így végső soron fizikai tárgyak fizikai tulajdonságairól, nem pedig valamilyen különleges státuszú szubjektív entitásokról és az ő speciális szubjektív tulajdonságaikról. Amikor mégsem erről van szó, az olyan származtatott

szóhasználat, amely a szokványosabb példákon alapul. Például ha szemgyulladás esetén valaki „homokot érez a szemhéja alatt”, akkor ezzel azt fejezi ki, hogy olyan állapotban van, mintha homok lenne a szemhéja alatt. A „mintha” kötőszó után mindig valami publikus, objektív világbeli tényállás leírása következik (Ryle 1949/1999: 296).

Ryle azonban nem érzi az elméletét teljesnek, mert az érzetekről szóló összes kifejezéssel nem lehet ilyen módon elszámolni. Van ugyanis az érzet-szavaknak egy olyan csoportja, amely máshogyan viselkedik, mint azok, amelyek publikus tárgyak tulajdonságaira utalva vagy „észlelési szavakként” értelmezhetők. Ilyen például az, hogy „fáj” vagy „viszket”. *Látni* valamit egy észlelés eredményét fejezi ki, vagy az, hogy „zöld”, egy tárgy tulajdonságát. A „szörnyen viszketek” kifejezés viszont Ryle szerint nem sorolható egyik csoportba sem (Ryle 1949/1999: 300). Az ilyen kifejezések megfelelő elemzését Ryle még megoldatlannak látta.

Ryle-nak az érzetekről adott elemzéseit tehát nem értelmezhetjük úgy, hogy Ryle teljes mértékben kihagyta volna elméletéből vagy tagadta volna az elmének ezt az aspektusát. Igaz, az érzetekre, érzéki benyomásokra, érzetadatokra és más hasonló entitásokra hivatkozó állításokat legnagyobb részt visszavezette a publikus objektumokra és azok megfigyelhető tulajdonságaira vonatkozó állításokra, illetve ezekkel és a szubjektum állapotaival kapcsolatos hipotetikus állításokra. Wittgensteinhez hasonlóan elutasította azt, hogy az érzetek esetében valamiféle megfigyelhető (de nem az anyagi világban, hanem egy másik, szellemi világban tartózkodó) objektumokról lenne szó. Azonban Ryle nem volt elégedett azzal, ahogy elmélete a fájdalommal és viszketéssel bánik, és ezzel összefüggésben elvégezetlen feladatokról beszélt. A fájdalom vagy a viszketés Ryle számára sem „valami” (dolog), de nem éppen „semmi”.

4.2 Livingston a strukturális magyarázatokról

A fenomenális megragadására irányuló kísérletek visszatérő kudarcait Livingston (2004) azzal magyarázza, hogy a XX. századi analitikus filozófia a természettudományos módszereket tartva eszményének, az elméről is *strukturális* magyarázatot kívánt adni. Livingston meghatározása szerint a *strukturális magyarázatok* olyan magyarázatok, melyek egyes elemekről olyan módon adnak beszámolót, hogy relációk egy struktúrájában helyezik el őket. A *fizikalizmus* vagy materializmus az a doktrína, mely szerint minden valós jelenséget meg lehet magyarázni alapvető fizikai terminusokban. A jelenségeket *kauzális* relációk hálózatában elhelyezve magyarázza meg. A

funkcionalizmus pedig az a doktrína, mely szerint a mentális állapotokat, beleértve a tudatos vagy tudati állapotokat, meg lehet magyarázni a többi mentális állapothoz és fizikai állapotokhoz való *funkcionális* viszonyok terminusaiban. (Livingston 2004: 4) Livingston szerint az egész analitikus tradícióban rányomja a bélyegét a tudattal kapcsolatos gondolkodásra a *struktúra* és a tudat közötti ellentét. (Livingston 2004: 14). A fizikalista és funkcionalista magyarázatok, mivel strukturális jellegűek, szükségképpen kihagyják a tudatot, chalmersi értelemben vett „nehéz problémát” vetve fel (Livingston 2004: 17).

A strukturális magyarázatok és a tudat közti feszültség természetének fel nem ismerése vezet Livingston szerint ahhoz a „frusztráló dialektikus szituációhoz, amelyben a vita... maga a tudat jelenségének pusztá létezése vagy nemlétezése körül forog” (Livingston 2004: 19). Az egyik tábor tagadja a tudat létezését, mert véleményük szerint az objektív világban nem fordulhat elő semmi, ami azokkal a különleges tulajdonságokkal rendelkezne, mint állítólag a tudatosság. A másik tábor egyszerűen magától értetődőnek tekinti, hogy a tudat létezik. „Ezen a ponton a vita intuíciók pusztá összecsapásává válik” (Livingston 2004: 19).

Livingston rámutat, hogy számos filozófus érzékelte a feszültséget az intrinzikusként felfogott kválék és a különféle relacionális, vagyis strukturális (behaviorista, fizikalista, funkcionalista) magyarázatok között. A közvetlen, nonrelacionális tartalomként felfogott kválék és a relacionális magyarázatok közötti feszültségek újabb és újabb innovációkhoz vezettek a mai napig (Livingston 2004: 19).

Livingston a tudat és a strukturális magyarázat közötti szembenállást fogalmilag stabil és történetileg instabil fogalmi konfigurációként jellemzi. Egyfelől minden strukturalista magyarázati programban szükség van valamilyen alapvető elemekre, amelyek a struktúrák relátumai. Ebben az értelemben az említett szembenállás stabil. Másfelől instabil, mert az alapvető elemek megmagyarázhatatlanok lesznek (hiszen a magyarázatok strukturálisak) és így a tudat intrinzikus karaktere *kimondhatatlan* (Livingston 2004: 21–22).

Livingston szerint ez az instabilitás meghatározó jelentőséggel bír a XX. századi elmefilozófia fejlődésében. Újra és újra megismétlődött az a jelenség, hogy egy-egy sikeresnek tűnő strukturalista magyarázati program megfeneklett a tudattal kapcsolatos hétköznapi nyelv valamely aspektusának megmagyarázásának problémáján. Azonban mivel ezekben a helyzetekben a filozófusok nem ismerték fel a strukturális magyarázat és a tudatosság intrinzikus jellege közötti alapvető ellentét jelentőségét, azt hitték, hogy egy másik strukturalista program megoldhatja azokat a problémákat, amelyeket az előző nem tudott. Így például Ryle elmefilozófiája nem tudott

megbirkózni az érzetek problémájával, ami az azonosságelmélet megalkotásához vezetett, az azonosságelmélet azonban nem tudott beszámolót nyújtani a pszichológiai nyelv jelentéséről, ami a funkcionalizmus kifejlesztéséhez vezetett, a funkcionalizmus azonban zátonyra futott a kválék problémáján (Livingston 2004: 22-23).

Livingston szerint Ryle megközelítésének elterjedt értelmezése, hogy zsákutca volt, egy olyan behaviorista álláspont, mely különösebb érvelés nélkül tagadta mind a „belső” mentális élet jelenségeinek létezését, mind az olyan „belső” fizikai vagy funkcionális állapotokét, amelyek megfeleltethetők lennének ezeknek vagy előidézhetnék ezeket (Livingston 2004: 114). Ez azonban Ryle projektjének félreértése, és Livingston szerint ebben a félreértésben osztoztak az azonosságelmélet korai képviselői is (Livingston 2004: 115). Livingston szerint Place és Smart úgy értelmezték Ryle elméletét, hogy Ryle az érzeteket és más tudati állapotokat viselkedési diszpozíciókkal azonosította, és ennek korrigálásaként javasolták saját elméleteiket. Ryle-nak azonban valójában nem állt szándékában az érzetokról vagy a tapasztalatról behaviorista vagy diszpozicionalista elméletet adni, hanem éppenséggel nyelvi analízis révén meg akarta világítani és el akarta oszlatni azokat a tévedéseket, amelyek miatt a filozófusok azt hitték, hogy az elméről olyan szubsztanciális elméletet tudnak alkotni, mint amilyen a karteziánus.

A ryle-i analízis semmi olyan információt nem tár fel, ami nem foglaltatik már benne a közönséges nyelvhasználatban (míg ezzel szemben az azonosságelméletek elméletüket empirikus hipotézisnek tartották, amit jövőbeli tudományos felfedezések alátámaszthatnak). Livingston tehát úgy véli, hogy Place és Smart félreértette Ryle céljait, mert félreértette a módszereit. Csakhogy ez a félreértés nem véletlen, mert Ryle-nak az érzetokról szóló elemzésének problémáira vezethető vissza (Livingston 2004: 116).

Livingston szerint az *A szellem fogalma* érvelésének gyenge pontja, amelynek problémái láttán Place és Smart úgy érezték, hogy ki kell javítaniuk Ryle álláspontját, az érzetokról szóló fejezet volt (melynek előszavában Ryle megállapította, hogy nincs megelégedve ezzel a fejezettel). Ryle célja az, hogy megmutassa, nincs okunk feltételezni, hogy az érzetokról beszámoló kifejezéseink irreducibilisen „szellemi” eseményekről számolnak be. (Livingston 2004: 126)

Livingston szerint Smart és Place úgy értelmezte Ryle álláspontját, hogy szerinte a mentális állapotokról szóló beszámolók (beleértve az érzetokról és mentális képekről szólókat) pusztán viselkedésként vagy viselkedési diszpozícióként elemzendők; és bár az efféle elemzés megfelelő lehet a tudás, akarás vagy érzelmek esetében, az érzetek és a mentális képek esetében elégtelen. Fizikalistákként azonban Place és Smart nem térhettek vissza a karteziánus dualizmushoz, így azt a

nézetet képviselték, hogy az érzetek és mentális képek ténylegesen agyi folyamatok (Livingston 2004: 131). Livingston szerint Smart és Place azt az empirikusan alátámasztható hipotézist fogalmazzák meg, hogy az érzetek agyi eseményekkel vagy folyamatokkal azonosak, míg Ryle semmi szubsztantívat nem akart állítani arról, hogy mi jellemző a mentális fogalmak tárgyaira azon túl, ami a természetes nyelvhasználat alapján kikövetkeztethető. Place és Smart azonosságtézise nem konceptuális, hanem empirikus: nem azt állítják, hogy amikor mentális állapotokról beszélünk, akkor agyi állapotokról beszélünk (a szavaink jelentése szerint), hanem azt, hogy amikor mentális állapotokról beszélünk, akkor történetesen agyi állapotokról beszélünk, mert a mentális állapotok igazából agyi állapotok. Smart azt az állítást, hogy az érzetek agyi folyamatok, egy olyan (elképzelt, vagy bizonyos Wittgenstein-olvasattal azonosított) állásponttal állítja szembe, amely behaviorista módon redukálná az érzeteket: a fájdalomról való beszámoló e szerint „szofisztikált felszisszenés” lenne, az érzet pedig egyszerűen kísértés arra, hogy érzetről számoljunk be (Smart 1959: 141). Livingston úgy véli, hogy Place és Smart azt hitték, hogy Ryle ténylegesen a viselkedésre (vagy viselkedési diszpozíciókra) akarta redukálni az érzeteket, és erre Ryle-nak az érzetekről szóló elemzése voltaképpen okot is adott (Livingston 2004: 139-141).

Livingston szerint Place és Smart álláspontja előrelépést jelentett a Ryle-éhoz képest annyiban, hogy komolyan vették a természetes nyelv érzetekről szóló beszámolóinak referencialitását. Azonban ennél tovább is mentek, és elméletet alkottak arról, hogy pontosan mi is a természete azoknak a dolgoknak, amelyekre az érzetekről szóló beszámolók referálnak. Az a belátás tehát, hogy az érzetekről szóló beszámolók referálnak valamilyen objektumokra, egy olyan elmélet alkotásához vezetett az ő esetükben, amely a mentális állapotokat fizikai (agy)állapotokkal azonosította a világ teljes fizikalista leírásának részeként. Ez azonban – tekintettel arra, hogy a tudatosság ellenáll a strukturalista magyarázatoknak – nem lehetett teljesen sikeres vállalkozás (Livingston 2004: 142). Miért nem?

Smart abból indult ki, hogy az érzet-nyelv valódi beszámoló szereppel bír. Önmagában ez csupán egy bizonyos kifejezések használatáról szóló állítás, mely emlékeztet minket arra, hogy a hétköznapi nyelvben olykor fájdalmakról vagy utóképekről számolunk be, és az olyan elméletek abszurditására, amelyek ezt vitatják (ezek között van maga Ryle elmélete is, hiszen Ryle abbeli igyekezetében, hogy az érzet-nyelvet a publikus tárgyokról való beszámolók nyelvére vezesse vissza, az érzetekről szóló fejezetben – szemben a könyvének többi részében követett módszerrel – szembement a természetes nyelvhasználati szokásokkal). (Livingston 2004: 146). Csakhogy Place és Smart strukturalista fizikalizmusának keretei között azt a megállapítást, hogy az érzet-terminusok

igenis *referálnak* valamire, annak kell követnie, hogy meg is határozzuk, mi ezeknek a valamiknek a természete. A referencia pedig egy fizikalista elméletben nem lehet más, mint valami fizikai.

Place és Smart fizikalizmusa (és az analitikus elme filozófia későbbi fejleményei) a tudatosság számára a világ *struktúráján* belül akart helyet találni. Azt remélték, hogy a természetes nyelv elköteleződése és a tudatosságnak a természet rendjén belüli helye közti viszonyt majd a jövő kifejlett tudománya magyarázhatja meg. (Livingston 2004: 150). Livingston szerint azonban a tudat ellenáll a strukturalizálásnak; csak az alapvető elemek relációinak mintázatain kívüli pozícióból lehet megérteni.

Az elme strukturalista felfogásának következő változata a funkcionalizmus volt. (Livingston 2004: 151). Mivel a tudat jelensége ellenáll a strukturalista magyarázatoknak, az elterjedt vélemény szerint a funkcionalizmus nem tud magyarázatot adni a kválékra vagy nyers érzetekre, és pedig azért, mert a funkcionalista magyarázat strukturális jellege folytán nem magyarázhatja a tudat intrinzikus karakterét. A funkcionalizmus elleni Block-féle gondolat kísérletek mind abból az alapvető problémából merítik erejüket, hogy egy strukturalista elmélet nem tudja elégséges módon megragadni a tudatos tapasztalatot (Livingston 2004: 193).

4.3 Összefoglalás

Az e fejezetben elemzett szerzők arra mutattak rá, hogy élményeink fenomenális minőségéhez önmagában és közvetlenül, vagyis minden mástól elvonatkoztatva, minden nem-fenomenális tartalomtól függetlenül – nem férünk hozzá: valójában nem vagyunk képesek megnevezni, közölni, és nem rendelkezünk róla tudással. Amikor arról beszélünk, hogy „milyen” átélni egy helyzetet, látni egy képet, vagy hallani egy zenét, akkor vagy publikus tárgyak megfigyelhető tulajdonságairól számolunk be, vagy pedig hipotetikus helyzetek felvázolásával tesszük lehetővé hallgatóságunk számára, hogy a megfelelő helyzetbe képzelve magát (képzeletben felvéve a releváns nézőpontot) virtuálisan átélje a kérdéses szituációt. Amikor kommunikáció révén megkíséreljük valakivel közölni, hogy *milyen* átélni valamilyen helyzetet, akkor kommunikációnk tartalma sok mindenre kiterjedhet, egyedül csak magát ezt a milyenséget nem fogja tartalmazni. Minthogy hallgatóságunk, ha előadásunk sikeres, az elmondásunk alapján képes a maga számára „szimulálni” (elképzelni) a leírt helyzetet, ezért normális körülmények között bízhatunk abban, hogy *közvetetten* sikerült közölni vele a kérdéses élményminőséget, különösen akkor, amikor olyan

egyszerű dolgokról van szó, mint a paradicsom pirossága. Amikor azt mondjuk, „piros”, akkor publikus tárgyak publikus tulajdonságára referálunk, de ennek a szónak a használata potenciálisan felhívja hallgatóságunkat arra, hogy elképzelje, milyen az ő számára abban a helyzetben lenni, amikor piros tárgy van előtte, és ilyen értelemben a „piros” szó azt is jelöli közvetett módon, hogy a szó mindenkor hallgatója vagy olvasója számára *milyen* a piros szín.

Mindez azt jelenti, hogy a fenomenális „hozzáférhetetlensége” nem abszolút abban az értelemben, hogy bár sohasem közvetlenek vagy kizárólag fenomenálisak a közléseink, a fenomenálisról folytatott kommunikáció az említett kerülő úton sikeres lehet – normális körülmények között. Mi a helyzet a nem normális körülményekkel? Azért kell kikötnünk, hogy a fenomenális közvetett közlése sikeres lehet „normális körülmények között”, mert lehetséges, hogy beszélgetőpartnerünknek például más a színlátása (piros-zöld színtévesztő), amikor a piros vagy zöld színről beszélünk neki. Ezen kívül felvethetjük azt a szkeptikus lehetőséget is, hogy két személy más fenomenális minőségű színérzékeléssel bírnak még akkor is, ha fiziológiailag és viselkedés szempontjából mindketten normálisak. Ez azonban csak akkor lehetséges, ha elfogadjuk a fizikai tulajdonságoktól független kválék létezését, vagyis valamilyen nem fizikalista álláspontot.

A felcserélt színek esete nem jelent problémát azoknak, akik azt állítják, hogy a fenomenálisról közvetlenül nem tudunk kommunikálni vagy ismeretet szerezni, közvetetten viszont (viselkedésbeli különbségek vagy akár az agy vizsgálata révén) igen, hiszen a felcserélt színlátású egyének a normál látású személyhez képesti „abnormalitása” objektív fiziológiai különbségekkel magyarázható (Armstrong 1968: 257).

A fizikalizmussal összeegyeztethetetlen szkeptikus lehetőség azt jelenti, hogy két személy, bár fiziológiailag és pszichológiailag minden releváns tekintetben hasonló (ugyanúgy vélekednek a színekről, érzelmileg ugyanúgy viszonyulnak hozzájuk, ugyanolyan az ízlésük, ami a színeket illeti), mégis eltérő színkváléval rendelkezik, amikor ugyanazt a színt látják.

Ez a felvetés azonban elveszti látszólagos filozófiai fontosságát, ha pontosan megértettük az eddig elhangzottakat. Miért tűnhet egyesek számára úgy, hogy a kválékkal kapcsolatos szkeptikus feltételezést értelmes dolog felvetni, míg nem értelmes dolog hasonló szkeptikus feltételezések megfogalmazása más esetekben? Így például felvethetnénk, hogy olyan tárgyak, például (Smart 1959: 153 kifejezését kölcsönvéve) „kísértet-anyag” vesz körül minket, amely semmiféle kölcsönhatásba nem lép a fizikai tárgyakkal, vagy a tárgyakkal vannak „kísérteties tulajdonságaik”, amelyek semmilyen hatással nincsenek a tárgyak fizikai tulajdonságaira. Mégsem tartja az ilyesmit senki komoly filozófiai felvetésnek. A kváléinverzióval kapcsolatos antifizikalista spekulációk

azonban pontosan olyanok, mint a kísérteties anyaggal vagy tulajdonságokkal kapcsolatosak. Aki szerint ez nem így van, az nyilván arra gondol, hogy a kválékról tudomásunk van (vagy legalábbis „ismeretségben” vagyunk velük), és ezért minőségük valamilyen értelemben hatással vannak ránk, szemben a kísérteties tulajdonságokkal és entitásokkal. A fenomenális hozzáférhetetlenségének doktrínája szerint azonban ez nincs így. A kválékkal „ismeretségbe kerülünk”, de nem szerzünk tudást a természetükről.

Nem *tudjuk*, milyen nekünk lenni, amennyiben tudáson proposicionális tudást értünk. Az, hogy milyen nekünk lenni, nem valami olyasmi, amit megfigyelhetünk, leolvashatunk, majd megfigyelésünk eredményét közölhetjük. Mivel a „tudjuk, milyen ...-nak lenni” sajátos, „filozófiai” értelemben – az „első kézből való” tudás értelmében – nem mást jelent, mint azt, hogy képesek vagyunk felvenni (vagy felvettük) a szóban forgó nézőpontot (egyszerűen fogalmazva: például emberek vagyunk, viszont nem vagyunk denevérek), ezért ebben a sajátos értelemben viszont tudjuk, milyen embernek lenni (és nem tudjuk, milyen denevérnek lenni) (Hill 1991: 89-90).

Akik komolyan veszik a kváléinverzióval kapcsolatos szkeptikus felvetést (tehát nem azt a legitimen felvethető, valószínű lehetőséget, hogy lehetséges, hogy fiziológiai vagy pszichológiai eltérések miatt az emberek színérzékelése bizonyos mértékig eltér), miközben elutasítják a kísérteties tulajdonságokra vonatkozó szkeptikus felvetést, a kettő közötti különbséget motiválniuk kell valamivel, és ez pedig az, hogy a kváléinkról állítólag „közvetlen tudomásunk” van. Ha azonban ez nem igaz, akkor nincs különbség a kválékra vonatkozó és az egyéb, komolytalan szkeptikus felvetések között. Hasonlóan érvel a zombikkal kapcsolatban Tye (2008: 192-193).

Mivel pedig ténylegesen nem *tudunk* (a szó szokásos értelmében) semmit „közvetlen ismeretség révén” a kváléinkról (Tye (2008: 145) szerint „nincs tudomásunk a fenomenális tudatosságról”), még referálni sem tudunk rájuk közvetlenül, ezért nem fordulhat elő, hogy bármiféle „kváléinverzió” meglepetéssel szolgáljon számunkra, hiszen a „kválé” abban az értelemben, ahogy a szkeptikus felvetésben szerepel, kísérteties tulajdonság, aminek semmire sincs hatása.

Valójában azoknak a helyzeteknek az elképzelése, amelyeket ténylegesen el tudunk képzelni, amikor azt hisszük, hogy a kváléinverziót képzeljük el, a szituáció pszichológiai, fiziológiai és fizikai paramétereinek megváltoztatásán alapul. Ha elképzeljük, hogy felcserélődnek a kváléink, akkor igazából azt képzeljük el, milyen lenne számunkra az, ha egy olyan világba kerülnénk, ahol például az ég piros, vagy milyen lenne számunkra az, ha olyan módon változnánk meg (akár színes szemüveg viselése, akár a színlátás valamilyen betegség miatti zavara miatt), hogy

az eget pirosnak látnánk. Amiatt, hogy az ilyen esetek nagyon is *elképzeltetők*, úgy tűnhet, hogy a zombik és a kváléinverzió *elgondolhatók*.

Mihelyt elfogadjuk tehát, hogy az, hogy *milyen* nekünk lenni, végső soron attól függ, hogy kik (mik) vagyunk és milyen helyzetbe (környezetbe) kerültünk, és azt, hogy tapasztalataink fenomenális minőségei annak aspektusai, hogy milyen nekünk lenni⁵, valamint lemondunk a fenomenális közvetlen megismerhetőségének elképzeléséről, a kváléinverzió ugyanolyan motiválatlan felvetésként lepleződik le előttünk, mint a „kísérteties tulajdonságokra” vonatkozó egyéb szkeptikus feltételezések.

Van azonban még egy probléma, ami felmerülhet akkor, ha nem hiszünk a fenomenális közvetlen megismerhetőségében. Honnan tudjuk, hogy egyáltalán „valamilyen” nekünk lenni, honnan tudjuk, hogy „valamilyen” tapasztalni? Honnan tudjuk, hogy nem vagyunk zombik? Ez utóbbi kérdést eredetileg Dretske (2003) vetette fel. Ugyan hangsúlyozta, hogy sem ő maga, sem olvasója nem zombi, és erről mindketten tudnak, nem sikerült megállapítania, *honnan* tudják: Dretske szerint „semmi nincs az észlelésünkben, ami azt mondaná nekünk, hogy tudatosak vagyunk”, és az sem segít, ha az észlelési állapotainkra irányuló másodrendű állapotokra vagy az introspekcióra hárítjuk tudatosságunk észlelésének feladatát. Mondhatjuk például, érvel Dretske, hogy a zombik csak protofájdalmakkal és protoviszketéssel rendelkeznek, azonban ezekről nincs „tudomásuk”, emiatt ezek nem valódi fájdalmak vagy viszketések. A probléma persze az, hogy nem világos, honnan tudjuk, hogy mi viszont fájdalmakat érzünk és nem protofájdalmakat. Chalmers a „fenomenális ítélet paradoxonának” (Chalmers 1996: 164) nevezi azt, hogy a tudatosság magyarázatilag irreleváns a tudatról szóló állításaink és ítéleteink szempontjából: „Zombi ikrem... nevetséges mennyiségű időt tölt egy számítógép fölé görnyedve, fejezet után fejezetet írva a tudat rejtélyeiről” (Chalmers 1996: 167). Chalmers e megjegyzése szerint egy zombit még az alapján sem lehet megkülönböztetni egy tudatos lénytől, hogy más módon filozofál a zombikról vagy a fenomenális tulajdonságokról, mint egy tudatos lény.⁶

5 Peressini (2017) amellet érvel, hogy ahelyett, hogy a „milyen ...-nek lenni” fogalmát próbálnánk a tapasztalat fenomenális tulajdonságai alapján meghatározni, az átélő szubjektum „milyen neki lenni”-ségét kell elsődlegesnek tekinteni. Ez a „fenomenális holizmus” lehetővé teszi azt, hogy a kválékkal kapcsolatos elméleti bonyodalmaktól függetlenül el lehessen ismerni a tapasztalat „milyen valaminek lenni” aspektusát.

6 Így aztán a Dennettel fennálló nézetkülönbségnek sem lehet az az oka, hogy Dennett talán zombi (Chalmers 1996. 176.).

Tye-hoz (2008: 192-193) csatlakozva a dretskei kérdésre azt felelhetjük, hogy nincs róla közvetlen tudomásunk, hogy nem zombik vagyunk. Ez azonban nem jelenti, hogy komolyan számolnunk kellene azzal az eshetőséggel, hogy esetleg zombik vagyunk, csak nem tudunk róla olyan módon, ahogyan például (tehát megfigyelés révén) tudunk arról, hogy milyen színű az előttünk lévő asztal. Zombikról ugyanis csak akkor van értelme beszélni, ha szembeállítjuk őket az emberekkel. A különbséget úgy jellemezhetjük, hogy az ember fenomenálisan normális, míg a zombi fenomenálisan abnormális (noha fizikailag vagy funkcionálisan, a konkrét elmélettől függően normális).

Arról természetesen senki sem képes „közvetlen tudomást” szerezni, hogy ő maga fenomenálisan normális-e, így ha a zombik elgondolhatók (ami nem áll fenn, ha elfogadjuk azt a materialista elvet, mely szerint fizikai duplikátumunk egyúttal fenomenális duplikátumunk is), akkor nem tudhatjuk, nem vagyunk-e zombik.

A gyakorlatban nem „közvetlenül észleljük”, hogy nem vagyunk zombik, hanem következtetünk erre azáltal, hogy elutasítjuk azt a lehetőséget, hogy zombik vagyunk, amire az jogosít fel, hogy nem számolunk azzal a teljesen motiválatlan feltevésével, hogy olyan nem fizikai, viselkedésünkre vagy tapasztalatunkra hatással nem lévő tulajdonságok hiányoznak belőlünk, amelyek másokban megvannak.

Mit jelentene fontolóra venni azt a lehetőséget, hogy zombi vagyok? Ha elfogadjuk Farrell, Ryle és Wittgenstein álláspontját, akkor ez vagy azt jelentené, hogy fontolóra veszem azt a lehetőséget, hogy valójában nem tudom felvenni vagy nem vettem fel a saját első személyű nézőpontomat, vagyis nem vagyok én, vagy azt, hogy esetleg sohasem láttam, hallottam vagy éreztem semmit (vagy számomra minden színtelen, hangtalan, szagtalan stb.). Mivel az első értelmezés logikai képtelenségnek tűnik, a második pedig hamisnak, ezért csak akkor tudnám komolyan venni azt a lehetőséget, hogy zombi vagyok, ha megengedném, hogy ez annyit jelent, hogy hiányoznak olyan tulajdonságaim, amik másoknak megvannak, de ez tapasztalataimat és tudásomat semmilyen számomra megismerhető módon nem befolyásolja. Ezt fel tudom tételezni, csak a feltételezés fontossága, tétje marad homályos.

A tudatosság állítólagos ténye, vagyis az, hogy „valamilyen” nekünk lenni, az elmondottak alapján semmiféle magyarázó, előrejelző erővel nem bír, hiszen tudás sem szól róla, még leíró módon beszélni sem tudunk róla. Hiszen ez utóbbit megtenni olyasfajta bravúr lenne, mint becsúszni a nyelvvel a fájdalom és annak kinyilvánítása közé (Wittgenstein 1953/1998. 245. §.). A revizionisták szerint ezért a tudat létezésének feltételezése teljesen fölösleges. Ugyanakkor gyanús

lehet számukra, hogy a tudat filozófiai problémája, a zombikkal, kválékkal stb. kapcsolatos gondolatmenetek csupán a lélekkel vagy a túlvilággal kapcsolatos vallásos elképzelések maradványai, és az ezekhez fűződő remények – a vallásos és mitikus elképzelések modern filozófiai elméletekként való rehabilitálásának szándéka – motiválják azokat, akik a tudatot a tudomány és a filozófia előtt álló nagy kihívásként, megoldhatatlan rejtélyként vagy az eddigi eredményeket megkérdőjelező problémaként állítják be. Így látta ezt például Rorty.

Másfelől úgy tűnik, hogy nem mindenki revizionista még a tudományosan vagy filozófiailag képzettek körében sem. Revizionisták számára kézenfekvő lenne erre az a magyarázat, hogy az antirevizionisták azok, akik utóvédharcokat folytatnak a tudományos haladással vagy a materialista és pozitivista filozófiai irányzatokkal szemben, hogy megmentsék a tudat „rejtélyét” vagy a tudatosság „nyilvánvaló tényét”. Vannak revizionisták, akik egyenesen úgy gondolják, hogy az antirevizionisták meghatározott filozófiai doktrínákhoz (például dualizmus) ragaszkodnak. Erre azonban rengeteg ellenpélda van. Egyáltalán nem kell dualistának vagy vallásosnak lenni ahhoz, hogy a tudat problémáját komolyan vegyük, és ne tartsuk elégséges megoldásnak a téma ignorálását.

Mivel a tudat „nem valami, de mégsem semmi”, és bizonyos értelemben az, amit „közvetlenül ismerünk”, ám egy másik értelemben mégsem tudunk róla semmit, ezért igen nehéz fogalmat alkotni róla. Ebben (a fenomenális kognitív és szemantikai hozzáférhetetlenségében) látom az okát annak, hogy a tudatosságról merőben ellenkező elképzelések is hihetőnek tűnnek egyesek számára, míg mások számára hihetetlenek, a tudatról szóló elméletek pedig rendszeres módon kétértelműnek bizonyulnak, a tudattal kapcsolatos filozófiai vitákat a félreértések, félremagyarázások, illetve az ellentétes intuitív meggyőződések összecsapásának terepévé téve.

5. A reduktív materializmus mint egyensúlyi elmélet

A tudatosság egyensúlyi elméleteinek válaszolnia kell az antirevizionisták által a tudatosság természetére vonatkozóan feltett kérdéseire anélkül, hogy olyan entitás vagy tulajdonság létezését feltételeznék, amelyek feltételezése revizionista nézőpontból teljesen fölösleges.

Ebben a fejezetben megvizsgálom a reduktív materialista elmefilozófiát (az elme–agy azonosságelméletet) abban a formájában, ahogyan az 1950-es években először megfogalmazták, némi kitekintéssel későbbi formáira is. Az elméletet azóta sokféleképpen újrafogalmazták, illetve

továbbfejlesztették. Bőven meghaladná ennek az értekezésnek a terjedelmét, ha ezeket a fejleményeket megkísérelnénk akár csak áttekinteni. Ezért ez a fejezet csupán arra szorítkozik, hogy az elmélet legkorábbi megfogalmazásait vizsgálja meg. Úgyszintén nem célom alaposan vizsgálni a materializmus ellen felhozott érveket és az ezekre adott materialista válaszokat, általában a materializmus védelmezésének problémáit. Ez a fejezet csupán azt hivatott megmutatni, hogy a reduktív materialista elmefilozófia klasszikusai tudatában voltak a fenomenális hozzáférhetetlenségének problémájával, és igyekeztek elkerülni az ebből adódó filozófiai hibákat. Ezért, ha maguk nem is éltek ezzel a terminológiával, közel kerültek ahhoz, hogy egyensúlyi elméletet próbáljanak alkotni, ami nem hanyagolja el a tudat problémáját, illetve hozzájárul annak körvonalazásához, hogy a tudatossággal kapcsolatban az antirevizionisták által megfogalmazott „rejtvények” (mint a színcsere problémája vagy a giliszta fájdalomérzetének problémája) hogyan lennének megoldhatók, mindezt azonban a materialista világgépen belül éri el.

A három klasszikus álláspontjának alapos vizsgálata arra is fényt derít, hogy ezek a szerzők jóval árnyaltabban álltak hozzá a tudatossággal kapcsolatos különféle intuitív meggyőződések által támasztott nehézségekhez, mint azt általában gondolni szokás, ezért túlzottan leegyszerűsített az elmefilozófia történetének az a felfogása, hogy ezek a szerzők nem vették elég komolyan a fenomenális tudatosságot, vagy abszurd módon figyelmen kívül hagytak nyilvánvaló tényeket, ezért a materialista elmefilozófia későbbi változatai – így például a fenomenális fogalmakat elfogadó elméletek – feltétlenül előrelépést jelentenek az elmélet eredeti változataihoz képest, vagy pedig igazából nem foglalkoztak a tudat mai értelemben vett problémájával, hanem azt újra fel kellett a későbbiekben fedezni (Nagel, Jackson vagy Chalmers színre lépésével).

Place, Smart és Feigl, az elme-agy azonosságelmélet első klasszikus kifejtői kétféle szempontnak igyekeztek megfelelni, amikor elméletüket megfogalmazták. Egyrészt igyekeztek a tudatról (Place), az érzetéről (Smart), illetve a nyers érzetéről és a tapasztalat kváléiról (Feigl) mondani valami szubsztantívat, elkerülve azt, hogy ezeket „kihagyják” elméleteikből úgy, mint azt véleményük szerint a behavioristák tették. Másrészt el akarták kerülni azt, hogy elköteleződjenek a fenomenális tárgyak, fenomenális mező vagy fenomenális tulajdonságok létezésével, illetve egy különleges fajta tudás, a „közvetlen ismeretségből vett” tudás mellett. Az efféle speciális státuszú entitások vagy tulajdonságok, illetve a hozzájuk fűződő e különös episztemikus viszony feltételezése ugyanis nem tűnt számukra összeegyeztethetőnek egy fizikalista világgéppel.

5.1 Place

Place egyik kiindulópontja Ryle-nak a figyelemről alkotott ama diszpozicionális elméletének (Ryle 1949/1999: 196, 169-185) bírálata volt, amely segítségével Ryle a tudatosság fogalmát próbálta értelmezni. Ryle szerint ugyanis amikor tudatosságról beszélünk, akkor (legalábbis az esetek egy fontos részében) figyelmesen végrehajtott cselekményekről, vagy figyelmünk középpontjába került észlelésekről számolunk be, vagy arról a képességről, hogy be tudunk számolni nemrég átélt észleléseinkről vagy cselekedeteinkről. Továbbá tudatosnak nevezünk az olyan észleléseinket vagy tetteinket, amelyekről utóbb vagy közben be tudunk számolni, illetve amelyeket figyelmesen hajtunk végre. A tudat nem valami, a többi cselekedetünktől vagy az érzeteinktől elválasztható, különálló tevékenység vagy folyamat, hanem a tudatosság jellemzi bizonyos esetekben a tevékenységünket vagy tapasztalatainkat.

Place azonban elégtelennek tartotta a figyelemmel kapcsolatos fogalmak pusztán diszpozicionális értelmezését. Szerinte amikor „figyelemfogalmakat” (mint például „figyel”) alkalmazunk, akkor nemcsak valamilyen, egy megfigyelt vagy észrevett dologra adott reakcióra vonatkozó diszpozícióról (készenlétre, diszkriminációs képességre stb.) beszélünk, hanem egyszersmind az egyén egy belső állapotáról is, mely az efféle diszpozíció fennállásának feltétele (Place 1954/2004: 40).

Túl azon a meggyőződésen, hogy minden diszpozíciónak van valamilyen kategoriális alapja is, Place szerint a tudat diszpozicionális elméleteinek az a fő nehézsége, hogy „mindig értelmes megkérni egy egyént, hogy írja le, milyen nézni, hallgatni, megfigyelni valamit, vagy tudatában lenni annak, miközben nincs értelme megkérdezni, milyen egy bizonyos képességgel vagy tendenciával rendelkezni. [...] Ténylegesen [...] egyfolytában leírjuk, milyen figyelni, megnézni, meghallgatni vagy érezni dolgokat.” (Place 1954/2004: 41). Továbbá: „Amikor egy tudatállapotot írunk le, akkor általában az egyik dolog tudatában levést más dolgok tudatában levéséhez hasonlítjuk. Mindazonáltal van egy-két kifejezés, mint ‘kellemes’, ‘kellemetlen’, ‘élénk’, ‘homályos’, ‘éles’, ‘tompá’, melyeket magukra a tudatállapotokra alkalmazunk. Kissé szokatlan jelzők lennének készenléti állapotok leírására” (Place 1954/2004: 41).

Ezek a legfeltűnőbb pontok, ahol úgy tűnik, Place (Levine 1993/2008 kifejezését kölcsönvéve) a „milyenség kihagyását” rója fel Ryle-nak. Place visszaállítja a tudat Ryle által elutasított „kontemplatív felfogásának” egy változatát, mely szerint „figyelni” lehet egyfajta

tevékenység, nem pedig pusztán arról van szó, hogy tevékenységeinket több vagy kevesebb figyelemmel végezzük.

Csak hogy Ryle azért igyekezte a figyelmet és a tudatosságot a viselkedési diszpozíciók terminusaiban elemezni, hogy ne kelljen elismernie semmiféle „belső folyamatot”, például a „figyelés” privát tevékenységét, a tudatosság folyamatát, amely privát, belső mentális eseményeinket kíséri. Attól tartott ugyanis, hogy éppen az efféle privát, titokzatos entitások és folyamatok létezésének elismerése motiválta a „gépben lakozó kísértet” dogmáját, tehát az általa bírált dualizmust. Amikor azonban Place elismeri egy „belső folyamat” létezését, akkor felvethető, hogy megteszi a wittgensteini első, ártalmatlannak tűnő lépést (Wittgenstein 1953/1998: 308. §) afelé, hogy a tudatosságot valami dolognak tekintse, ami aztán tévútra vezeti a filozófusokat.

Place azonban nem gondolja, hogy (utalva arra, hogy Ryle a karteziánus dualizmus a gépben lakozó kísértet dogmájának nevezi) „visszahelyezi a kísértetet a gépbe” (Place 1954/2004: 43), és pedig azért nem, mert szerinte Ryle elméletének nagy része tartható, csupán a tudatról adott elemzése és a figyelem diszpozicionális elmélete szorult kiigazításra. Ráadásul Place szerint maga Ryle is elismerte, hogy az olyan szavak, mint a „figyel” vagy „hallgat”, végső soron valamilyen rejtett belső folyamatra utalnak. A különbség csak annyi Place szerint, hogy Ryle ezeknek a folyamatoknak nagyon csekély jelentőséget tulajdonított, és ezért „kitartóan kerülte” a tudat mint „belső folyamat” megemlítését.

Antirevizionista olvasatban Place Ryle elméletének voltaképpen az élmény milyenségének kihagyását, a tudat folyamatának figyelmen kívül hagyását róttta fel. Másrészt Place csak futólag említi meg az élményeink milyenségéről való beszámolás problémáját. Arra, hogy Ryle elméletét elégtelennek tartja, sok egyéb oka is van, és az ezekkel kapcsolatos érveknek tisztán kognitív olvasata is lehetséges. Ezen kívül ha elméletét úgy olvassuk, mint amely szerint a tudat nem más, mint az a folyamat, amely lehetővé teszi, hogy „élő kommentárt adjunk” érzeteinkről vagy más mentális állapotainkról, akkor a tudatnak tisztán funkcionális értelmezését adtuk, megfelelően egy revizionista olvasatnak.

Híresebb tanulmányában (Place 1956/2004), melynek *A figyelem fogalma* közvetlen előzménye volt, Place folytatja egy olyan elmélet kidolgozását, amely elkerüli a behaviorizmus hibáját (a tudat kihagyását), de anélkül, hogy ez dualizmushoz vezetne. Place ebben a cikkében kora fizikalizmusát behaviorisztikusként jellemzi; e szerint a tudat, illetve a tudatos mentális események valamilyen sajátos fajta viselkedésekkel vagy viselkedési diszpozíciókkal azonosak. A kognitív és

akarati fogalmakkal kapcsolatban alapvetően helyesnek tartja azoknak a viselkedési diszpozíciók terminusaiban történő (Wittgenstein és Ryle által kifejtett) elemzéseit. Ezt az álláspontot azonban elégtelennek tartotta, és továbbfejlesztette 1954-es gondolatmenetét. Place szerint ugyanis a behaviorisztikus elemzés után is fennmarad fogalmak egy „kezelhetetlen maradéka”, melyek a „tudat, tapasztalat, érzet és mentális képek fogalmai körül csoportosulnak”. Ezekkel kapcsolatban elkerülhetetlen feltételeznünk egy „belső folyamatot”:

[kijelentések] fájdmakról és nyilallásokról, vagy arról, hogy a dolgok hogyan néznek ki, hangzanak és érződnek, valamint a képzeletünkben és álmainkban felmerülő képekről szólnak, olyan eseményekkel és folyamatokkal kapcsolatosak, amelyek bizonyos értelemben privátak vagy belsőek annak számára, akiről állítják őket. (Place 1956/2004: 45)

Place fő célja annak bizonyítása, hogy a „belső folyamat” létezésének elfogadása nem teszi elkerülhetlenné egy olyan dualista álláspont elfogadását, így tehát összeegyeztethető a materializmussal, ezért nem lehet a priori kizárni azt a tézist, hogy a tudat egy agyi folyamat. Más szóval ismét nekigyürkőzik annak, hogy bebizonyítsa, a „belső folyamat” feltételezése nem helyezi vissza a kísértetet a gépbe. A tudatnak egy agyi folyamattal való azonosítását „*értelmes tudományos hipotézisnek*” tekinti, vagyis olyan állításnak, ami nem önellentmondó és nem is magától értetődő, és az olyasfajta tudományos azonosításokkal állítható párhuzamba, mint az, hogy „a villám elektromos töltések mozgása”.

Az ilyen állítás *nem magától értetődő*, mivel az azonosság-tézis nem analitikus, fogalmi állítás: nem azt fejezi ki, hogy amikor érzetekről számolunk be, akkor agyi eseményekről számolunk be abban az értelemben, hogy az agyi eseményekkel kapcsolatos fogalmaink és az érzetekkel kapcsolatos fogalmaink szinonimák, vagy az érzetekkel kapcsolatos fogalmaink agyi eseményekkel kapcsolatos fogalmak terminusaiban elemezhetők úgy, ahogyan például az akarattal vagy vélekedésekkel kapcsolatos kijelentések viselkedési diszpozíciók terminusaiban elemezhetők a behavioristák szerint. És *nem is ellentmondásos*, mert az azonosság-elmélettel szemben felhozott szokásos ellenérvek hamisak. Place tanulmánya nagy részét az ilyen ellenérvek cáfolatának szenteli.

Place szerint azok tartják azt az állítást, hogy „a tudat: agyi folyamat”, ellentmondásosnak, akik azt hiszik, hogy ha két kifejezés (mint a „tudat” és az „agyai folyamat”) jelentését egymástól

független módon definiáltuk, akkor a két kifejezés nem jelölheti ugyanazt a tárgyat vagy tényállást. A tudományos azonosítások azonban erre ellenpéldát szolgáltatnak. A felhők lebegő vízcseppek tömegei. Ez közlelről nézve nyilvánvaló, hiszen a felhőbe belépve ködszitalást tapasztalhatunk, távolról nézve azonban nem nyilvánvaló, hogy a felhők vízcseppekből állnak. Ha csak távolról látnánk felhőket, akkor hihetnének azt, hogy lehetnének éppen vattaszerű képződmények is. Sőt, nyelvünkben nincs „logikai kapcsolat” a felhők és a parányi vízcseppek tömegének fogalma között. Beszélhetünk olyan felhőkről is, amelyek nem vízcseppekből állnak. A „felhő” és az „apró lebegő vízcseppek tömege” kifejezések tehát eltérő jelentéssel bírnak. Mégsem gondoljuk, hogy egy adott felhő és a lebegő vízcseppek tömege két különböző dolog. Annak, hogy a fogalmaink jelentése közötti viszony így alakult, Place szerint az az oka, hogy a felhőket sohasem látjuk egyszerre közlelről és messziről, vagyis sohasem egyszerre verifikáljuk az „ez egy felhő” és „ez itt lebegő vízcseppek tömege” kijelentéseket. A kétféle megfigyelés sohasem esik egybe, annyira, hogy kétféle szóval írjuk le azt, amit e megfigyelések során tapasztalunk.

Place szerint az „a villám elektromos töltések mozgása” állítás a villám és az elektromos töltések áramlása között hasonló *kompozíciós* viszonyt példáz, mint amilyen a vízcseppek és a felhő közötti viszony. Bár a felhő és a vízcseppek példája közérthető, egyben tökéletlen is, mert a felhő és a vízcseppek megfigyelése esetében két olyan megfigyelésről van szó, amelyek között folyamatos átmenet lehetséges, így könnyű megérteni, hogyan alkothatnak a vízcseppek felhőt. Nincs szó hasonló szemléletes viszonyról a villám és az elektromos töltések áramlása között, valamint az agyi folyamatok és a tudat között. Ezekben az esetekben ugyanis a kétféleképpen leírt eseményekre vonatkozó kijelentések verifikációja egészen eltérő módon történik (míg mind a felhőé, mind a vízcseppeké vizuálisan). Bármilyen közlelről vizsgáljuk meg (vizuálisan) a villámot, nem fogjuk látni az áramló elektromos töltéseket, és bármennyire is alaposan „magunkba nézünk”, introspekciónk nem tárja fel az idegi impulzusok mozgását a szinapszisok között (Place 1956/2004: 49).

Ennek ellenére sikeres volt a villám tudományos azonosítása az elektromos töltések áramlásával. Nem pusztán azért, mert a villámokra és elektromos töltések áramlásaira vonatkozó megfigyelések között szisztematikus korrelációkat figyelhetünk meg. Place szerint akkor gondoljuk, hogy megfigyelések két sorozata ugyanazokról az eseményekről szól, ha a tudományos megfigyelések (a megfigyelések egyik sorozata) magyarázatot kínálnak az „utca emberének” megfigyeléseire (ez a megfigyelések másik sorozata).

A tudat és az agybeli folyamatok azonosításához tehát az szükséges, hogy legalábbis lehetséges legyen (és egyszer megvalósuljon) az introspektív megfigyeléseknek az agybeli

folyamatokra hivatkozó megmagyarázása. Place azonban arra figyelt fel, hogy kortársai (így például a jeles agykutató, Sherrington) ezt a lehetőséget meglehetősen nehezen tudták elképzelni. Úgy látták ugyanis, hogy miközben érzéki tapasztalatokat élünk át, két olyan eseménysor történik, melyek „nyilvánvalóan” különböznek egymástól: egy fizikai és egy pszichikai. Egyfelől elektromos impulzusok mozognak az agyunkban, másfelől viszont látjuk a Napot. Hogyan azonosítható a kétféle, szemléletmást különböző esemény?

Place „fenomenológiai tévkövetkeztetésnek” nevezi azt a gondolatmenetet, ami szerinte az efféle ellenvetések mögött munkál. Ez az a téves elképzelés, mely szerint amikor egy alany leírja tudatos tapasztalatait, akkor egy „belső moziban vagy televíziós képernyőn” létező tárgyak tulajdonságait írja le, vagyis a „fenomenális mező” tárgyainak fenomenális tulajdonságait. Ha ez így van, akkor amikor valaki egy zöld utóképről számol be, akkor valamilyen, az ő elméjében létező zöld tárgyról számol be, amilyen tárgynak „nincs helye a fizikai világban” (Place 1956/2004: 51). Valójában azonban semmiféle zöld tárgy nincs a zöld utóképet látó személy környezetében, sem az agyában. Egy agyi folyamat nem olyasmi, ami zöld vagy más színű lehet.

A fenomenális tévkövetkeztetés Place szerint azon a tévedésen nyugszik, hogy amikor a környezetünkben lévő tárgyakat leírjuk, akkor elsődlegesen „érzékelési tapasztalatainkat” írjuk le, és csak ezek révén, indirekt módon, a tárgyakat. Valójában Place szerint ez éppen fordítva van: csak azután tudunk tapasztalataink minőségéről beszélni, ha már megtanultuk leírni a környezetünkben lévő tárgyakat. Azért tudunk beszámolni arról, hogy zöld utóképet látunk, mert már korábban láttunk zöld tárgyakat és megtanultuk a zöld szín fogalmának használatát:

Tudatos tapasztalatainkat nem a mitologikus ‘fenomenális tulajdonságok’ terminusaiban írjuk le, melyek állítólagosan a mitologikus ‘fenomenális mező’ mitologikus ‘objektumaira’ jellemzők, hanem azon konkrét fizikai tárgyak, események és folyamatok valóságos fizikai tulajdonságaira referálva, melyek általában [...] előidézik azt a fajta tudatos élményt, amit le akarunk írni. Más szavakkal, amikor egy utóképet zöld színűként írunk le, akkor nem azt mondjuk, hogy van valami, az utókép, ami zöld, hanem azt mondjuk, hogy az a fajta tapasztalatunk van, amilyen szokványosan akkor szokott lenni, amikor [...] egy zöld fényfoltot nézünk (Place 1956/2004: 51).

Place tanulmányának két szembetűnő jellemzője van. Az egyik, hogy határozottan állítja, hogy a tudat, a tapasztalatok és érzetek „kezelhetetlen maradékot” alkotnak az elme behaviorista elméletei szempontjából, és elkerülhetetlennek tűnik ezzel kapcsolatban a tudat „belső folyamatának” elismerése. Place-t nehéz lenne azzal vádolni, hogy figyelmen kívül hagyta a tudat problémáját, hiszen akkor akár elégedett is lehetne a behaviorista elméletekkel. Éppen ezért elsietett lenne Place-t a „tudat tagadói” közé sorolni pusztán azért, mert „mitologikusként” utasítja el a fenomenális mezőt, a fenomenális objektumokat, vagy a fenomenális tulajdonságokat, és azt állítja, hogy amikor tapasztalatainkról beszélünk, akkor előfordulásuk körülményeire utalva, közvetett módon referálunk rájuk, nem pedig közvetlenül referálva fenomenális tulajdonságaikra.

Place tanulmányának figyelemreméltó vonása, hogy fő állítása az: *a tudat egy agyi folyamat*, annak mintájára, ahogy *a villám elektromos töltések áramlása*. „Azonosság-tézise” valójában egyfajta kompozíciós viszonyt feltételez, és az azonosítás két oldalán szereplő tényezőkről csak a legnagyobb általánosságok szintjén beszélt.

5.2 Smart

Smart klasszikus tanulmánya (1959) szorosan kapcsolódik a Place-éhez, részben annak érveit fejlesztve tovább. Smart Place-hez hasonló célokat tűz ki: egyrészt az érzetek valóságosságát kívánja állítani egy elképzelt szélsőséges behaviorista állásponttal szemben, másrészt viszont igyekszik kimutatni, hogy ez nem teszi szükségessé a fizikalizmus elutasítását.

Az említett behaviorista álláspont az, hogy amikor egy sárgás-narancs színű utóképről számolunk be, akkor „valójában nem számolunk be semmiről”, csupán kifejezzük azt a kísértést, amit arra érzünk, hogy azt mondjuk, hogy van a falon egy sárgás-narancsszín folt. Ezt a nézetet Smart egy lehetséges (de nem feltétlenül helyes) Wittgenstein-interpretációként jellemzi, párhuzamot vonva a fájdalom kifejezésével kapcsolatos wittgensteini expresszivitással (mely szerint amikor fájdalmat fejezünk ki, akkor nem beszámolunk a fájdalomról, hanem egyfajta szofisztikált fájdalomviselkedést tanúsítunk). Ezzel a nézettel szemben Smart azt szeretné védelmezni, hogy amikor utóképről (vagy éppen fájdalomról) számolunk be, akkor ezek valódi beszámolóok, nem pusztán állapotunk kifejezései – ugyanakkor sem a fájdalom, sem az utóképek nem valamilyen irreducibilisen pszichikus jelenség, hanem magyarázat adható róla a materializmus keretein belül, és nincs a dualizmus mellett szóló kényszerítő erejű filozófiai érv (Smart 1959: 143).

Az említett szélsőséges Wittgenstein-olvasat előnye Smart szerint, hogy e nézet szerint „bizonyos értelemben nincsenek érzetek. Az ember fizikai részecskék hatalmas együttese, de nincsenek ezen kívül érzetek vagy tudatállapotok” (Smart 1959: 143). Egy ilyen álláspont (Smart kifejezésével az érzet-kijelentések „expresszív” értelmezése) tagadhatatlan előnye, hogy ennek keretei közt a dualizmus végképp motiválatlan marad.

Smart azonban nem éri be ezzel az állásponttal: „úgy tűnik számára” (Smart 1959: 144), hogy aki utóképről számol be, az valóságos beszámolót nyújt valamiről, aki pedig fájdalomra panaszkodik, az valami többet is tesz annál, mintha pusztán fájdalomviselkedést folytatna. A fájdalomról adott beszámoló beszámoló, nem pedig szofisztikált felszisszenés. Nehéz másként értelmezni Smart „úgy tűnik”-jét, mint annak elismerését, hogy antirevizionista intuíciói miatt valószínűtlennek érzi azt a revizionista elméletet, mely szerint „nincsenek is érzetek”. Smart tehát azt a feladatot szabja magának, hogy az érzetekkel, fájdalmakkal stb. kapcsolatos beszámolók vonatkozásában elfogadja a deskriptivizmus valamilyen változatát, de anélkül, hogy ez valamilyen dualista álláspont elfogadásához vezetne.

Ez a célkitűzés hasonlítani látszik Place törekvéseire. Place elégtelennek tartja a behaviorizmust, és úgy igyekszik kiegészíteni, pozitív beszámolót adva a tudatról, hogy ez ne vezessen dualizmushoz. Ennek során azonban Place szükségesnek látja hangsúlyozni, hogy a fenomenális objektumok és tulajdonságok „mitologikusak”, így sohasem referálunk rájuk közvetlenül. Smart viszont mintha deskriptivista akarna lenni az érzetekkel kapcsolatban, ha figyelembe vesszük, hogy az expresszivisták értelmezéssel szemben határozta meg pozícióját. Ráadásul szemben Place-szel, akinek állítása az volt, hogy „a tudat: agyi folyamat”, Smart az *érzetek* agyi eseményekkel való azonosítására törekszik.

Hasonlóan Place-hez, Smart is kifejti, hogy tézise nem az, hogy az „utókép” vagy a „fájdalom” szinonim lenne egy olyan kifejezéssel, mint az „X típusú agyi folyamat”, hanem az, hogy amennyiben az „utókép” vagy a „fájdalom” beszámoló valamilyen folyamatról, akkor ez a folyamat történetesen egy agyi folyamat (Smart 1959: 144)

A későbbiekben kommentátorok, illetve Smart elméletének kritikusai nagy jelentőséget tulajdonítottak annak, hogy Smart az érzetek és agyi folyamatok „kontingens azonosságát” állította. Smart kétségtelenül állít olyat, hogy „lehetségesek ‘A azonos B-vel’ alakú kontingens állítások, és egy személy nagyon jól tudhatja, hogy valami A anélkül, hogy tudná, hogy B”.

A „kontingens azonosítással” kapcsolatos problémák tárgyalása messzire vezetne. Csupán egy megjegyzést teszünk. Armstrong (1999: 70-72) szerint Place és Smart számára (mint általában a

Kripke előtti analitikus filozófusok számára) nem volt teljesen világos a kontingens és empirikus azonosítás közötti különbségtétel, de Smart számára az utóbbi lehetett a fontos.

Nagy jelentőségű Smart cikkében az azonosságelmélettel szemben megfogalmazott 3. ellenvetés, mely *Max Black ellenvetése*ként közismert, és Smartnak erre adott válasza. Az ellenvetés szerint még ha nem is bizonyítható, hogy az érzetek nem azonosak agyi folyamatokkal, az érzetek tulajdonságai mindenesetre különböznek az agyi folyamatok tulajdonságaitól. Ha nincsenek is irreducibilisen mentális folyamatok, nem sikerül elkerülni az ilyen tulajdonságok létezésének elfogadását. Ilyen lehet például az érzetek valamilyen tisztán fenomenális vagy kizárólag introspekcióban feltáruló minősége. (Egy érzet másmilyennek mutatkozik az introspekcióban, mint a neurofiziológiai megfigyelések során.) Az érzet valamilyen „fenomenális tulajdonsággal” rendelkezik, míg egy agyállapot nem. Smart erre válaszul vezeti be az érzetekkel kapcsolatos kijelentések „tárgyközömbös” (Ambrus 2007: 21) értelmezését:

Amikor egy személy azt mondja: ‘Sárgás narancsszín utóképet látok’, akkor valami olyasmit mond, hogy ‘*Valami olyasmi történik, mint ami akkor történik, amikor a szemem nyitva van, ébren vagyok, és egy jól megvilágított narancs van előttem...*’ (Smart 1959. 149)

A dőlt betűkkel kiemelt kifejezéseket Smart „kvázilogikus vagy tárgyközömbös” kifejezéseknek nevezi. Segítségükkel anélkül tudunk beszámolni érzeteinkről, hogy bármi szubsztantívát állítanánk az érzetek természetéről. Ehelyett csak a hasonló érzetek előfordulásainak körülményeit nevezzük meg a fenti példában. Ennek jelentősége az, hogy ha ilyen módon hivatkozunk az érzetekre, akkor semmi olyasmit nem mondunk róluk, amely összeférhetetlen lenne azzal a materialista tézissel, hogy az érzetek agyi folyamatok vagy események.

Érzeteink akár agyi folyamatok is lehetnek, mégis beszámolhatunk róluk akkor is, ha semmit sem tudunk az agyi folyamatokról, hiszen csupán „nagyon absztrakt módon” (Smart 1959. 150) számolunk be róluk. Ez pedig megmagyarázza a fenomenális hozzáférhetetlenségét, vagy Smart kifejezésével a „a ‘nyers érzetek’ egyedülálló illékonyágát – miért tűnik úgy, hogy senki sem képes tulajdonságokkal felruházni őket.” (Smart itt a már sokszor idézett Farrellre hivatkozik.) A nyers érzetek Smart szerint „ugyanamiatt színtelenek, amiért a *valami* is színtelen. Ez nem jelenti azt, hogy az érzeteknek nincsenek tulajdonságaik, hiszen ha agyi folyamatok, akkor biztosan

vannak tulajdonságaik. Csak azt jelenti, hogy amikor úgy beszélünk róluk, mint egymáshoz hasonlóakról vagy egymástól különbözőktől, akkor nem kell ismernünk vagy említenünk ezeket a tulajdonságokat” (Smart 1959: 150).

Smart lényegében azt fejt ki részletesebben, amit már Place is felvetett. Place szerint „mitologikusak” a fenomenális mező tárgyai és tulajdonságaik, Smart szerint pedig „bizonyos értelemben nincs is olyan dolog, hogy utóképp vagy érzetadat, bár van olyan, hogy az utóképpel rendelkezés élménye, és ezt az élményt indirekt módon lehet leírni materiálisobjektum-nyelven, nem pedig fenomenális nyelven, mivel olyan nincs.” (Smart 1959: 151) Nem létezik a „privát minőségek nyelve”, érzeteink minőségét azonban le tudjuk írni közvetett módon, a környezetünkben lévő tárgyak fizikai tulajdonságainak leírására szolgáló nyelven (Smart 1959: 154).

Így aztán, miközben Smart elutasította azt a szélsőséges nézetet, hogy érzetek „nincsenek is”, és azt állította, hogy „beszámolunk valamiről”, amikor érzetekről számolunk be, álláspontja nem sokban tér el a Place-étől. Az érzetekről nyújtott beszámolóink Smart szerint is közvetettek, illetve a megnevezett objektumok és tulajdonságok mindig külvilágbeli, publikus objektumok és ezek publikus tulajdonságai. Az érzetek, amikről ilyen közvetett módon, de ténylegesen beszámolunk, pedig agyi események.

5.3 Feigl

Az elme-agy azonosságelmélet első generációjához tartozó három klasszikus tanulmány közül a legrészletesebb a Feiglé (Feigl 1958).

Feigl szerint az elme-test problémát kora tudósainak jelentős része (és bizonyos filozófiai irányzatok) látszatproblémának tekintette (Feigl 1958: 370-371, 476). Feigl azonban ezt „elfojtásnak” tekinti, ami csak ideig-óraig fenntartható, és „tünetei” vannak – paradoxonok és inkonzisztenciák.

Feigl szerint a „durva materialisták” vagy „kemény behavioristák” (egy helyütt a logikai behavioristákat „illogikus behavioristáknak” is nevezi (Feigl 1958: 423)) számára a mentális és fizikai viszonyának egész problémája nem létezik; az ilyen olvasókhöz olyan retorikai kérdésekkel fordul, mint az, hogy „Te nem kérnél érzéstelenítést műtét előtt?” (Feigl 1958: 390). Feigl szerint a kegyetlenség morális elítélésének is előfeltétele, hogy másoknak „közvetlen tapasztalatot” (vagy

közvetlen élményt) tulajdonítunk. Feigl egyértelművé teszi, hogy itt a tapasztalat fenomenális minőségére gondol (Feigl 1958: 391).⁷

E megjegyzésekből is sejthető, hogy Feigl meglehetősen biztonsággal az antirevizionista szerzők közé sorolható. Feigl számára is az az alapprobléma, hogy hogyan lehet azt az intuitív meggyőződést, hogy ténylegesen átélünk élményeket, „valamilyen” számunkra fájdalmat érezni vagy színeket látni, összeegyeztetni a fenomenális hozzáférhetetlenségével kapcsolatban elért filozófiai belátásokkal, éspedig a materializmus keretein belül.

Feigl azonosságtézise az, hogy „[a] közvetlen tapasztalat nyers érzetei, ahogyan „rendelkezünk velük”, empirikusan azonosíthatók [...] bizonyos neurofiziológiai fogalmak referenciáival” (Feigl 1958: 445). Feigl egyértelművé teszi, hogy azonosságállításának egyik oldalán a nyers érzetek, a „közvetlen tapasztalatnak” azok az állapotai szerepelnek, amelyeket „átélünk” (Feigl 1958: 446). A *mentális* lényege (a test-elme probléma szempontjából) Feigl számára a *fenomenális* (Feigl 1958: 405), nem pedig például az intelligencia kérdése (Feigl 1958: 419). Ennél egyértelműbben aligha fogalmazhatná meg antirevizionista célkitűzését.

Feigl többek között Lashley-t, Ryle-t, Carnapot és Skinnert sorolja fel azon szerzők között, akik kételkedtek a tiszta fenomenális nyelv lehetőségességében (Feigl 1958: 372). E szerzők szerint a mentális életünket leíró terminusok használatát is interszubjektív használat során tanuljuk meg. Ha kivonnánk belőlük ezt az interszubjektív tartalmat, akkor pusztán „ineffábilis nyers érzetek” maradnának vissza, vagy kognitív jelentés nélküli verbális gesztusok. Feigl azonban ragaszkodik ahhoz, hogy az introspekció lehetővé teszi számunkra, hogy leírjunk „elemeket... a közvetlen tapasztalat fenomenális mezején” (Feigl 1958: 372): a fájdalmat érző beteget vizsgáló orvos „megengedheti magának, hogy behaviorista legyen”, a páciens viszont nemcsak gondolkodik vagy beszél a fájdalomról, hanem *átéli* a fájdalmat, amiről azután a közvetlen tapasztalat és introspekció révén beszámol.

⁷ Későbbi szerzők rámutattak, hogy korántsem nyilvánvaló az, hogy a fenomenális minőségnek kizárólagos vagy döntő szerepe, vagy akár bármiféle szerepe lenne annak meghatározásában, hogy egy szubjektum szenved-e vagy sem (Levy 2014). Feiglnek ez a megállapítása azért is problematikus, mert úgy tűnik, Feigl elfogadja, vagy legalábbis közel kerül annak a később Kripke (1972/2007: 120) által is felhasznált tézisnek az elfogadásához, hogy a fájdalmat fenomenális minősége teszi fájdalommá, ráadásul erről a minőségről valamilyen módon közvetlen tudomásunk van. Ha ez így van, akkor Feigl a materializmus fenomenális fogalmakkal operáló változatának előfutáraként olvasható, és elmélete ki van téve a kripkei típusú ellenvetéseknek. Mint azonban látni fogjuk, Feigl tanulmányának legtöbb állítása nem támogatja ezt az olvasatot.

Ezen a ponton úgy tűnik, hogy Feigl nem „eliminativista” a fenomenális mezővel és annak tárgyaival, illetve azok tulajdonságaival szemben úgy, mint Place és Smart. Mintha Feigl számára intuitíve nyilvánvaló tény lenne, hogy létezik fenomenális mező fenomenális tárgyakkal, mivel ténylegesen átéljük fájdalmainkat és más érzeteinket. Így például kijelenti, hogy a „fizikalisták” (akik ebben a kontextusban azok a szerzők, akiket a „tudat tagadói” közé sorolhatunk: szerintük nincsenek érzetek, csupán fizikai folyamatok) minden érve ellenére „még mindig meg [van] győződve arról, hogy a tisztán fenomenális kijelentések értelmeseek” (Feigl 1958: 392), azonban példája nem meggyőző (jóindulatúbb olvasatban: ártalmatlan a fenomenális hozzáférhetetlenségének elvére nézve): „amikor úgy ítélem meg, hogy egy bizonyos fájdalom erősödik [...] akkor az a bizonyos *az*, ami később megtalálhatja helyét a világ kauzális struktúrájában, az mindenekelőtt és önmagában véve, a *közvetlen tapasztalat adata*” (Feigl 1958: 392).

A gond az, hogy amikor valaki arról számol be, hogy a fájdalma erősödik, akkor először is, akármilyen legyen a fájdalomnak természete, a fájdalommal kapcsolatban azt a *strukturális* tényt közli, hogy egy bizonyos időpontban a fájdalom nagyobb mértékű, mint egy másik időpontban. A növekedés strukturális, illetve matematikai (az idővel és mennyiségekkel kapcsolatos) fogalom. Elmondhatjuk, hogy a fájdalom egy bizonyos pillanatban erősebb, mint egy másikban, anélkül, hogy ehhez ki kellene fejtenünk, mi a fájdalom önmagában. Továbbá a fájdalmat (Place és Smart nyomán) tárgyközömbös fordítások módszerével, illetve funkcionálisan (vö. Lewis 1972) is definiálhatjuk.

Ha persze a fájdalmat tisztán funkcionálisan definiáljuk, akkor olyan elméletet kapunk, amely nem elégíti ki azokat, akik antirevizionista intuíciókkal rendelkeznek. Az antirevizionista intuíciónak csak a fájdalom olyan elmélete felelhet meg, amiben valamilyen módon szerepel a fájdalom intrinzikus természete, fenomenális minősége.

A probléma az, hogy ha azt, hogy az élmény fenomenális aspektusa a „közvetlen tapasztalat adata”, úgy értelmezzük, hogy a közvetlen tapasztalatban (egy különleges módon: „ismeretség révén”) tudásra teszünk szert a szubjektív élményünk vagy érzeteink sajátos, csak ilyen módon hozzáférhető fenomenális tulajdonságairól, vagy a fenomenális mező fenomenális tárgyairól, akkor ez azt jelenti, hogy létezik a tulajdonságok vagy tárgyak egy olyan osztálya, amelyekről más módon (és csakis ezen a módon) teszünk szert tudásra, mint a fizikai tárgyairól. Ezzel kitarjuk a kaput Max Black ellenvetése előtt. Mint Feigl is elismeri, a dualizmus mellett szóló egyik legfontosabb érv, hogy léteznek a „közvetlen tapasztalat” nyers érzetei (Feigl 1958: 375). Place és Smart megoldása az volt, hogy elutasították a fenomenális tárgyakat és fenomenális mezőt, illetve a fenomenális

tulajdonságokat, a látszólag fenomenális kijelentéseket pedig a tárgyközömbös fordítás módszerével visszavezették a fizikaiakra.

Feigl egyfelől azt akarja állítani, hogy (Schlick terminusát használva) a *tartalom*, ama bizonyos „az”, ami részt vesz a kauzális viszonyokban, nem elhanyagolható, vagyis a tudatosságot nem lehet pusztán viselkedési terminusokban kimerítően megmagyarázni. Ugyanakkor el kell utasítania azt, hogy a *tartalomról* olyan típusú *tudásunk* lehet, amilyen fajta *tudásunk* nem lehet a fizikai objektumokról. Úgy tűnik, hogy Feigl végül is megérti ezt a problémát, amikor nagy jelentőséget tulajdonít a közvetlen ismeretségnek, de úgy, hogy egyúttal elhatárolja a közvetlen ismeretségből vett tudástól. Terminológiája azonban ingadozik.

Többször megkísérli tisztázni, pontosan mit gondol az ismeretség és az ismeretségből vett tudás viszonyáról (Feigl 1958: 391), illetve a fenomenális nyelv lehetőségéről, a fenomenális terminusok jelentéséről. A kirajzolódó összkép alapján nehéz megállapítani Feigl pontos álláspontját már csak azért is, mert ő maga is elismeri bizonytalanságát, és bizonyos kérdéseket nyitva hagy. Tanulmánya azonban a kérdés érdekes megközelítéseinek tárháza, és jól szemlélteti a filozófusok nehézségeit, amelyekkel akkor szembesülnek, ha a tudatos tapasztalatot nem kívánják figyelmen kívül hagyni vagy a létezését tagadni, ám különleges, a fizikai világon kívüli entitássá vagy tulajdonsággá sem kívánják változtatni.

Feigl nézetét a következőképpen rekonstruálhatjuk. Megkülönböztethetjük a „közvetlen ismeretséget” a „közvetlen ismeretségből vett tudás”-tól. Amikor azt mondjuk, hogy (sajátosan filozófiai értelemben) nem tudjuk, milyen denevérnek lenni, vagy a fekete-fehér szobába zárt Mary nem tudja, milyen zöld színt látni, akkor a „közvetlen ismeretség” hiányáról számolunk be. Amikor Mary először lát színeket, vagy valaki – például miután egy boszorkány denevérré változtatta – sikeresen felveszi egy denevér első személyű nézőpontját, akkor közvetlen ismeretségbe kerülnek a színlátás, illetve a hanglokátorral való tájékozódás élményével. Ez a „megtudni, milyen” (a szó sajátosan filozófiai értelmében) nem más, mint a „közvetlen ismeretségbe kerülés” valamilyen tulajdonsággal: Az „ismeretség mint olyan [...] egyszerűen a közvetlen tapasztalatot magát jelenti, ahogyan azt átéljük, élvezzük vagy elszenvedjük; az ismeretségből vett tudás viszont propozicionális.” (Feigl 1958: 404). Az „ismeretségből vett tudás” egyszerűen a noninferenciális tudást jelenti (Feigl 1958: 403).

Feigl arra a kérdésre, hogy mi hiányzik a vak ember tapasztalatai közül összehasonlítva a látóéval, azt feleli, hogy a vak esetében hiányzik a látási kválékkal kapcsolatban mind a közvetlen ismeretség, mind a közvetlen ismeretség általi tudás. Ez azonban nem jelenti, hogy Feigl ellent

mondana a közvetlen ismeretség és a közvetlen ismeretség általi tudás közti különbségtételnek, vagy annak a tézisnek, hogy a tudás tartalma csakis strukturális lehet. Az ismeretség ugyanis ebben az esetben a látási kválékkal való puszta rendelkezést (azok átélését) jelenti, míg az ismeretségből vett tudás képességeket (mint a színekre való ráismerés vagy színek elképzelésének képességét). Feigl a képességhipotézis előfutárának tekinthető (Feigl 1958: 416). Más esetekben a „közvetlen ismeretségből vett tudás” egyszerűen a közvetlen érzéki észlelésből eredő, noninferenciális tudást jelenti. (Feigl 1958: 435). Ilyen például látásból ismerni Churchillt (Feigl 1958: 403), nem csupán újságból.

Az utóbbi példa arra vonatkozik, hogy Feigl szerint a „közvetlen ismeretségnek” kétféle fogalma különböztethető meg: egy „szűkebb, filozófiai” fogalom, mely a nyers érzetek kváléival való közvetlen ismeretségre korlátozódik, és a szélesebb, hétköznapi vagy fizikalisztikus fogalom, amely magában foglalja a fizikai tárgyak megfigyelhető tulajdonságainak észlelését is (Feigl 1958: 432). Amikor olyan fizikai kutatásokat végzünk, melyeknek eredménye nem szemléletes, például az atom szerkezetével vagy az elektromossággal kapcsolatban, akkor sohasem kerülünk közvetlen ismeretségbe ezek „természetével”. Ez csak annyit jelent, hogy ezek számunkra nem közvetlenül szemlélhető dolgok. Ugyanilyen viszonyban van egy vak ember a tárgyak vizuális tulajdonságaival. Ugyanúgy, ahogy a normál látású emberek, akik számára nincs szemléletes közvetlen észlelés az elektromosságról vagy a radioaktivitásról, egy vak tudós is kutathatja akár a látást is, és minden törvényszerűséget megállapíthat róla. Amiben hátrányt szenved egy látó tudóssal szemben, az nem az, hogy végül kevesebbet fog tudni. Ami „ismeretség” révén hozzáadható tapasztalataihoz, az kognitíve irreleváns (Feigl 1958: 433). Feigl azt a radikális állítást is megteszi, mely szerint a kválékkal való ismeretség nem szükséges feltétele a kválékról való tudásnak (Feigl 1958: 431). *Pace Jackson, Mary*, a színtudós sokkal többet *tudhat* fekete-fehér szobájában a színekről, mint az odakint. a színes világban élők, akik nála kevesebbet tanultak a színek fizikájáról és a színlátás fiziológiájáról.

Feigl így Schlickhez hasonlóan elutasítja Bergsonnak azt az elképzelését is, hogy a közvetlen ismeretség vagy az intuíció lenne a legtökéletesebb tudás forrása (vagy egyáltalán tudást eredményezne) (Feigl 1958: 432). A nyers érzetekkel való puszta *rendelkezés* ugyanis nem tudás (Feigl 1958: 477), ami tudás – akkor is, ha ismeretségből vett tudás – mindig ellenőrizhető indirekt módszerekkel (Feigl 1958: 477).

Feigl szerint ugyanakkor csak azért tudjuk megismerni a világot, mert bizonyos mértékig közvetlen ismeretségbe kerülünk vele. Ez azonban magától értetődő, hiszen valamiféle interakciónak fenn kell állnia köztünk és a világ között, ha meg akarjuk ismerni. Éppen ezért az

viszont messzemenőig mindegy, hogy közvetlen ismeretségünk „hídfőállása” a világban mekkora, és milyen érzéki modalitásban áll fenn. Egy születésétől fogva vak és süket személy is képes lehetne elvileg csillagászati kutatásokat folytatni (Feigl 1958: 433). Feigl mondanivalója azonban nem az, hogy először is „közvetlen ismeretség általi tudásra” teszünk szert, amely a többi („leírás alapuló”) tudás alapját képezi, hanem arról, hogy bármiféle kapcsolat a külvilággal, aminek kognitív jelentősége lehet, magában foglal érzéki észlelést (tehát közvetlen ismeretséget a szó hétköznapi, egyszersmind sajátosan filozófiai értelmében is).

Feigl Chalmers „táncoló kválék” gondolat kísérletének egy előfutárát is megalkotta. Ebben a gondolat kísérletben valakivel kválé-inverzió történik, aminek tényét igazolja az illető agyának cerebroszkópos megfigyelése. A műszer például jelzi, hogy piros színt kellene látnia, holott ő zöld szín látásának élményét éli át. Feigl ezt „logikailag elképzelhető” helyzetnek tekinti, de hozzáteszi, hogy a „fogoly elme” ismeretségből vett tudása saját kváléinak invertált voltáról privát és kommunikálhatatlan marad (Feigl 1958: 476). Sőt, amikor valakinek megváltoznak a kváléi, de a világ kauzális berendezkedése nem sérül, ez az illető észre sem veszi saját kválé-inverzióját sem, és nem tud róla kommunikálni. (Feigl 1958: 477)

Feigl tehát szem előtt tartotta a fenomenális hozzáférhetetlenségének elvét, mindamellett többször is megemlíti tanulmányában, hogy a „fenomenális nyelvet” értelmesnek tartja. Az eddig elhangzottak fényében valószínűleg arra gondolhat, hogy képesek vagyunk beszélni élményeink minőségéről. Ugyanis nincsenek abszolúte privát mentális állapotaink (Feigl 1958: 448-49): állapotainkat le tudjuk írni publikus nyelven.

Az elhangzottak alapján pontosítható a feigli azonosságtézis. Emlékezhetünk rá, hogy Place csupán annyit állított, hogy a tudat egy agybeli folyamattal azonos, Smart tézise pedig az volt, hogy az érzetek agyi események. A feigli azonosságtézis az, hogy „az, amit átélünk a tapasztalatban és ismeretség révén ismerünk, leírás alapján tárgya a viselkedés nagy léptékű elméletének”, továbbá azonos a neurofiziológia által leírt bizonyos folyamatokkal. Másként fogalmazva, a neurofiziológiai nyelv némely terminusainak jelölete egybeesik némely fenomenális terminusok jelöleteivel, és egyben a pszichológiai nyelv némely terminusainak jelöletével. Így például ha valakinek utóképet tulajdonítunk, akkor először is egy olyan állapotot tulajdonítunk neki, amelyet megfigyelhető ingerek és rájuk adott válaszok terminusaiban lehet definiálni (ahogy azt a logikai behavioristák kidolgozták). Az ilyen módon definiált állapotot, melyről introspekció révén lehet tudomása a átélőjének, és amelyre ezáltal a „fenomenális nyelven” is tud hivatkozni, a későbbiekben tudományos felfedezésként empirikusan azonosíthatni lehet egy neurofiziológiai állapottal. (Feigl 1958: 448).

Természetesen Feigl nem ragaszkodott ahhoz a naiv elképzeléshez, hogy a viselkedés elméletében (a behaviorisztikus pszichológiában) szerepet játszó típusok sora 1:1 megfelelésbe hozható a hétköznapi „fenomenális nyelv” terminusaival, majd pedig a fenomenális terminusok jelöletei a jövő agyutatásai eredményeként hasonlóan szabályos és egyértelmű megfelelésbe állíthatók neurofiziológiai állapot típusokkal. Feigl nyitott kérdésnek tartotta, hogy amikor a tudomány majd megállapítja azokat a tényleges azonosságokat, amelyeknek ő csak legelvontabb sémáját vázolta fel, akkor a neurofiziológiai eseményekkel közvetlenül azonosított fenomenológiai és pszichológiai típusok „az elemzés melyik szintjén” lesznek találhatóak (Feigl 1958: 468).

Ezt azért érdemes kiemelni, mert egy igen elterjedt felfogás szerint a Place, Smart és Feigl által képviselt elmélet gyorsan és látványosan megbukott (Jackson 2012: 150), többek között a többszörös realizálhatóság (Tózsér 2008: 39) problémáján. A „tankönyvi nézet” szerint Place, Feigl és Smart mellett érveltek, hogy a mentális típusok azonosak neurofiziológiai típusokkal (például M típusú érzet P típusú agybeli eseménnyel). (Shapiro, Polger 2012: 265), amelyek aztán azonosak bizonyos biokémiai típusokkal és így tovább egészen a legalsó (fizikai) szintig. Egy ilyen elméletet, mely szigorú 1:1:1... megfeleléseket feltételez a mentális típusok szintjétől a fizikai típusok szintjéig, a típusoknak az azonosítások bármelyik szintjén bekövetkező heterogenitása falszifikálná. Egy efféle elmélet valóban nem lenne plauzibilis, csak éppen téves, vagy legalábbis fölöttébb rosszindulatú lenne ezt Place-nek, Smartnak vagy Feiglnek tulajdonítani. (Shapiro, Polger 2012: 268). Place eleve kompozíciós viszonyról beszélt tudat és agyi események között, ami nyitva tartja a mentális-fizikai megfelelések közötti kompozíciós változatosságot. Feigl számolt a „sok:egy” típusú megfelelésekkel, melyekben a mentális-fizikai azonosítás formája az, hogy a mentális oldalon egyetlen predikátum található, míg a fizikai oldalon predikátumok diszjunkciója. Smart számára pedig irreleváns volt a típusazonosság szempontjából, hogy egy típus tagjai között van-e valamilyen belső vagy összetételi variabilitás (Shapiro, Polger 2012: 269). Az azonosságelmélet első képviselői nem állították, hogy a mentális és agyállapotok közti azonosításoknak az absztrakció bármely tetszőleges szintjén igaznak kell lenniük (Shapiro, Polger 2012: 270).

Szokás a három szerzőt olyan módon is csoportosítani, hogy míg Place egy kompozíciós viszonyt feltételezett a tudat és az agybeli események között, addig Smart és Feigl mentális és fizikai (neurofiziológiai) típusok közti típusazonosság. Ez pedig azt jelenti, hogy a mentális és fizikai típusok fogalmai ugyanarra referálnak. (Shapiro, Polger 2012: 267) Mivel minden szerzőnk kiemelte, hogy a fogalmak e koreferenciája nem analitikus vagy pusztán a nyelv elemzése révén megállapítható tény, hanem empirikus, tudományos kutatások eredményeként megállapítható tény,

ezért a posteriori, empirikus azonosságállításokról van szó. Ezeket „kontingens azonosságállításoknak” is nevezték, amivel szemben felhozható az a jogos kritika, hogy az azonosság sohasem áll fenn kontingensen. Valójában Place, Smart és Feigl lényegi mondanivalója az volt, hogy az azonosítás *a posteriori* (Armstrong 1999: 71), viszont Kripke munkássága előtt minden a posteriori állítást kontingensnek tartottak. A későbbiekben Lewis volt az a szerző, aki kifejezetten ragaszkodott a „kontingens azonosság” fogalmához, ő azonban a kontingens azonosságon egy valóban kontingens relációt, a szerepbetöltést értette.

5. 4 Armstrong és Lewis: érett reduktív fizikalizmus

Place, Smart és Feigl megítélésem szerint olyan szerzők, akik, ha ezt nem is így fogalmazták meg, egyensúlyi elmélet megalkotására törekedtek. Egyfelől megkíséreltek igazságot szolgáltatni annak a sokak számára intuitívnak tűnő antirevizionista kritikának, mely szerint az elme behaviorista megközelítéseiből „kimarad” valami fontos (és mint láttuk, sem Wittgenstein, sem Ryle nem vádolható azzal, hogy egyszerűen elsiklottak volna e probléma felett). Másfelől úgy próbálták megvalósítani az elméletek szükséges kiigazítását, hogy közben a fizikalizmus keretei között maradjanak. Ennek következtében elméleteik revizionista szempontból nem tartalmaztak fölösleges feltételezéseket, míg az antirevizionisták méltányolhatták az elméleteikben azt, hogy nem ignorálták, hanem megpróbálták megoldani a tudat problémáját, illetve megkísérelték megmutatni, hogy az antirevizionista intuíciókat, ellentétben az elterjedt nézetekkel, nem feltétlenül kell antifizikalista intuícióként értelmezni.

Mindhárman szem előtt tartották a fenomenális hozzáférhetetlenségéről a logikai pozitivisták és behavioristák által kidolgozott érveket. Bár Feigl kevésbé tűnik radikálisnak ebben a kérdésben, valójában minden terminológiai ingadozása ellenére úgy tűnik, nem mond ellent a másik két szerzőnek (mivel világos különbséget tesz a közvetlen ismeretség és a közvetlen ismerettség között). Egyik szerző esetében sem tűnik plauzibilis értelmezésnek, hogy „a tudat tagadói” közé tartoztak volna.

Mint már szó volt róla, egy egyensúlyi elméletnek fizikalistának kell lennie, ráadásul redukcionista, hiszen az ezektől az álláspontoktól való ellépést csak antirevizionista aggodalmak motiválhatják. Mit várnak az antirevizionisták a tudat filozófiai elméleteitől? A fenomenális tudatosságnak ezekben az elméletekben valamilyen intrinzikus jellemzőnek kell lennie (a tapasztaló

szubjektum valamilyen intrinzikus tulajdonságának, vagy legalábbis a megtapasztalt külvilágbeli tárgyak valamilyen intrinzikus tulajdonságának); a tapasztalat „milyenségét” meghatározó folyamatoknak okkurrens folyamatoknak kell lenniük.

Az egyensúlyi elméletnek számot kell adnia arról is, hogy miért tűnik némelyek számára rejtélyesnek a tudat, és miért lehet vele kapcsolatban rejtvényeket megfogalmazni. Chalmers kifejezésével a tudat metaproblémájára is válaszolnia kell.

Az egyensúlyi elméletnek válaszolnia kell az antirevizionisták kérdéseire (lehetséges-e a zombik, két személy ugyanolyannak látja-e a színeket, szenvednek-e a giliszták, valamilyen-e csigának lenni), vagy legalábbis a keretein belül felvázolhatónak kell lennie, hogyan lehet ezeket a kérdéseket egyszer majd megválaszolni.

A reduktív fizikalizmus egyik legalaposabban kidolgozott változata David Lewis elmélete, mely a típus-azonosságelméletet a funkcionalizmussal ötvözi. Mivel Lewis pályafutása során számos tanulmányt publikált, és ezek között lehettek kisebb-nagyobb eltérések, hangsúlyeltolódások, ezért nem törekszem arra, hogy nagy pontossággal rekonstruáljam Lewis álláspontját. Inkább általánosságban egy „lewisi elmélet” összefoglalását kísérem meg, mely rendelkezik az egyensúlyi elmélet előnyeivel.

Lewis (1966) az azonosságelméletét először úgy fogalmazza meg, mint azt a tézist, mely szerint minden tapasztalat-típus azonos valamilyen fizikaiállapot-típussal (valamilyen „neurokémiai állapottal”). Ez az azonosság nem szükségszerűen, hanem történetesen, pusztán tényként áll fenn.

Ez a híres kontingens azonosság (kevésbé félreérthető lenne „szerepbetöltési relációnak” nevezni), mely a későbbiekben sok vitát váltott ki. Mit értett Lewis kontingens azonosságon? Lewis szerint egy tapasztalat-típus meghatározó jellemzője kauzális szerepe, „tipikus okainak és hatásainak szindrómája” (Lewis 1966. 17), és ezeket a jellegzetes kauzális szerepeket, melyek analitikus szükségszerűséggel (definíció szerint) jellemzőek egy-egy tapasztalat-típusra, történetesen (kontingens tényként) valamilyen fizikai állapotok töltik be a materializmus szerint. Népi pszichológiánk a fájdalmat, dühöt vagy a piros szín látásának élményét ezen állapotok kauzális, funkcionális szerepeinek terminusaiban definiálja. A materialista filozófus pedig azt állítja, hogy azok az állapotok, amelyek betöltik az így definiált szerepeket, éppenséggel neurofiziológiai állapotok (és nem például valamilyen szellemi szubsztancia állapotai).

A materializmus melletti (Lewis és a hasonló elméletet kidolgozó Armstrong által alkalmazott) érvelésnek két fázisa van. Az első, bonyolultabb lépés megmutatni kognitív, akarati, perceptuális vagy éppen érzelmi állapotokat jelölő fogalmainkról, hogy ezek mind definiálhatók

kauzális-funkcionális terminusokban, sőt, természetes definíciójuk ilyen jellegű. Ez tulajdonképpen a Place és Smart által kezdeményezett „tárgyközömbös fordítás” módszerének általánosítása és következetes végigvitele. Ezt követően már nem tűnik paradoxnak a mentális állapotok fizikaiakkal való azonosítása az érvelés második fázisában.

Armstrong hasonlóan fogalmazza meg célkitűzéseit, mint Place és Smart: célja a behaviorizmus „természellenességének” elkerülése (Armstrong 1970: 72) a materializmus védelmezése közben (Armstrong 1970: 67). Hasonlóan Place behaviorizmus-ellenes kritikájához, Armstrong szerint is igazságot kell szolgáltatni annak a nyilvánvalónak tűnő megállapításnak, hogy amikor gondolkodunk, akkor „valami ‘zajlik’ bennünk, a ‘zajlik’ legerősebb és legszószerintibb értelmében” (Armstrong 1970: 72), tundiillik a gondolkodás, akkor is, ha a gondolataink nem vezetnek cselekvéshez. Ezért nem elég a viselkedési diszpozíciókról beszélni, hanem a viselkedési diszpozíciók alapjául szolgáló aktuálisan fennálló állapotokkal kell azonosítani bizonyos mentális állapotokat. Armstrong tehát olyan módon módosította a Ryle nyomán kialakult behaviorista elemzési stratégiát, hogy a viselkedési diszpozíciókat azok kategoriális alapjaira vezette vissza (Armstrong 1968: 86).

A behavioristák beérték azzal, hogy viselkedési diszpozíciókról beszéljenek, Armstrong szerint azonban a diszpozíciók alapja valamilyen fizikai állapot. A tudományos magyarázatok nem állnak meg a diszpozíciók pusztá nyilvántartásba vételénél. Például az üveg törékenységét az üveg mikroszerkezete alapján magyarázzák. Hasonlóan, a mentális állapotok a viselkedésnek a személyben rejlő okai. Ezeknek az okoknak a természete tudományosan kutatható, és a modern tudomány fényt derített arra, hogy ezek az okok a központi idegrendszer működései (Armstrong 1970: 73).

Ily módon Armstrong az elme „centrális állapot” elméletéhez jutott, melyben mentális állapotaink nem mások, mint a központi idegrendszernek azok az állapotai, amelyek tipikusan bizonyos viselkedéseket okoznak (Armstrong 1968: 89 skk.) vagy bizonyos ingerek által váltódnak ki.

Armstrong elmélete abban is különbözik Place és Smart elméletétől, hogy Smart elmélete kifejezetten az érzetokról szól, míg Place-é a behaviorista elemzések számára „kezelhetetlen maradékról” kíván számot adni. Armstrong viszont a mentális-fizikai azonosságelméletet kiterjesztette (Armstrong 1999: 81) azokra az (intencionális, akarati, érzelmi stb.) állapotokra is, amelyekről Place még úgy gondolta, hogy behaviorisztikus elemzésük megfelelő. Armstrong sem vitatja természetesen a behavioristák legtöbb elemzését a mentális terminusokról, csak éppen,

rámutatva arra, hogy a viselkedési diszpozícióknak kategoriális alapja van, mely az emberi idegrendszer állapotaival azonosítható, átfogó művében (Armstrong 1968) az elme minden aspektusáról materialista beszámolót ad.

Lewis ezt a fajta elemzést a későbbiekben olyan módon fejlesztette tovább, hogy a mentális állapotok kauzális-funkcionális definíciójában lehetővé tette, hogy a mentális állapotok tipikus okai és következményei között más mentális állapotok is szerepeljenek. Lewis szerint tehát a mentális állapotokat kauzális szerepük definiálja, vagyis a többi mentális állapotot, tipikus viselkedéseket és ingereket tartalmazó kauzális hálózatban elfoglalt helyük.

Lewis szerint az a premisszája, hogy a tapasztalatok meghatározó jellemzője azok kauzális szindrómája, Smart tárgyközömbös fordítási módszerének általánosítása (Lewis 1966. 20). Ez önmagában nem materialista elv, hanem olyan elv, amit mindenki elfogad, aki szerint „a tapasztalatok valóságosak, és hatásuk van a saját tartományukon kívül”, vagyis a viselkedés magyarázatában szerepük lehet. Ez az elv behaviorista eredetű, de a behaviorizmust több ponton továbbfejleszti. Először is megengedi, hogy a tapasztalat, akár csak kiváltó okai és kiváltott következményei, „valami valóságosak legyenek” (Lewis 1966. 21). Másodszor, lehetővé válik, hogy a tapasztalatok tipikus okai és következményei között további tapasztalatokat soroljunk fel, ami többek között lehetővé teszi a tapasztalati állapotok introspektív hozzáférhetőségének megmagyarázását is, hiszen egy tapasztalat introspektív hozzáférhetőség azt jelenti, hogy egy rá irányuló másik tapasztalatot tud előidézni.

Lewis harmadik javítása a korábbi elméletekhez képest az, hogy a tapasztalatok definiálásakor beérhetjük a *tipikus* okok és hatások felsorolásával, nem kell az adott tapasztalathoz minden esetben járuló okokat és hatásokat felsorolni. A mentális állapotoknak tehát lehetségesek olyan kivételes esetei, amelyek nem elégítik ki a tipikus kauzális szindrómát. (Lewis 1966. 22). A behaviorizmusban erre nem nyílik lehetőség, hiszen (Lewis értelmezésében) a behaviorizmus a tapasztalatot azonosítja okaival és hatásaival (kauzális szindrómájával), ezért súlyos problémát jelent számára az olyan esetek értelmezése, mint a teljesen béna emberé vagy a tökéletes színészé.

Amikor Lewis a mentális állapotokat funkcionális szerepeik alapján definiálja, akkor nem azt állítja, hogy az olyan szavak jelentésével kapcsolatban, mint a piros szín látásának élménye vagy a fájdalom, valami tudományos felfedezések által ránk kényszerített fogalmi innovációt javasol. Ellenkezőleg: a mentális állapotokra vonatkozó mindennapi, természetes fogalmaink kauzális-funkcionális jelegűek, még ha a mindennapi életben nem is használunk olyan alapos definíciókat, mint egy pszichológus. Ha felsorolnánk az összes olyan közhelyet, amit a mentális állapotok, érzelmi

ingerek és viselkedések közötti összefüggésekről ki tudunk mondani, akkor ezekből (legalábbis egy részükből) egy elméletet lehetne konstruálni. Mentális terminusaink ennek az elméletnek a teoretikus terminusai lennének. Történetileg ugyan nem igaz, hogy mindennapi pszichológiai fogalmaink egy ilyen elméletalkotás eredményei, de ha így történt volna, akkor olyasmi fogalmakat alkotott volna az elméletalkotó, mint amilyenekkel jelenleg rendelkezünk (Lewis 1972: 256-257).

A materialista elmefilozófus tehát Lewis megfogalmazásában következőképpen azonosítja a mentális állapotokat fizikai állapotokkal:

M mentális állapot = az M szerep betöltője (analízis révén),
P fizikai állapot = az M szerep betöltője (a tudomány szerint),
tehát $M = P$. (Lewis 1994/2003: 205)

$M=P$ *kontingens azonosság*, amennyiben kontingens tény, hogy milyen fizikai állapot tölt be milyen kauzális szerepet. Lewis funkcionalizmusa realizáló-funkcionalizmus, nem pedig szerepfunkcionalizmus (Tózsér 2008: 62).

Kripkének azt a felvetését, hogy bizonyos mentális terminusaink, mint a „fájdalom”, merev jelölők, Lewis azzal a megjegyzéssel intézi el, hogy ugyan ez Kripkének „élénk intuíciója”, ő maga nem rendelkezik efféle intuícióval (Lewis 1994/2003: 205).

Lewis a következő érvet szegezi szembe Kripke állításával. Gondoljunk egy olyan esetre, amikor erős fájdalmat éreztünk. Ekkor valamilyen állapot betöltötte a fájdalom-szerepet. Ezután képzeljük el egy olyan helyzetet, amikor valami más tölti be ugyanezt a szerepet. Ekkor is be van töltve a fájdalom-szerep, tehát fájdalmat érzünk, és nem könnyű (ha egyáltalán lehetséges) megkülönböztetnünk az utóbbi esetet az előbbitől. De ha a „fájdalom” mereven jelöli az első helyzetet, akkor a másodikban nem érzünk fájdalmat. Tehát a „fájdalom” nem merev jelölő (Lewis 1994/2003: 206). Lewis még hozzáteszi:

Nem látom, hogy a szerep tényleges betöltőjének fizikai vagy nem-fizikai természetének bármi köze lenne ahhoz, hogy a kétféle szituáció megkülönböztethető-e. A „fenomenális karakterről” és hasonlókról való beszéd ezen nem segít. (Lewis 1994/2003: 206)

Míg Kripke úgy vélte, hogy a fájdalom lényege valamilyen fenomenális minőség, amiről közvetlen tudásunk és fenomenális fogalmunk van, addig Lewis szerint fájdalomfogalmunk kauzális-funkcionális, és nem rendelkezünk közvetlen tudással a mentális állapotainkat realizáló neurofiziológiai állapotaink természetéről – szellemes megállapítása szerint ha arra képesek lennénk, valamivel könnyebb lenne neurofiziológiai felfedezéseket tenni, mint most (Lewis 1995: 142).

Mindazonáltal mégiscsak úgy tűnik, mintha ismerhetnénk tapasztalataink kváléit: valamilyen kváléfogalom Lewis szerint bele van épülve népi pszichológiánkba, amit az is mutat, hogy amikor egy filozófus megpróbálja bevezetni a kválék témáját, hallgatósága elég gyorsan megérti, miről van szó. Lewis szerint végső soron egy materialista is hihet a kválékban. Nem adhat hitelt annak a szerinte a népi pszichológiánkban feltételezett elméletnek, mely szerint a tapasztalat kváléit képesek vagyunk megbízhatóan azonosítani, ami előfeltételezi, hogy lényegi természetül feltárul előttünk, amikor ismeretségbe kerülünk velük. De ha ennél gyengébb episztemikus viszonyokat feltételezünk a tapasztalatunk kváléival, illetve a népi pszichológiabeli kválé-szerepnek csak „nem tökéletes, de elég jó betöltőit” feltételezzük, akkor ebben a mérsékelt értelemben egy materialista is hihet a kválékban (Lewis 1995: 143).

Mint láttuk, Lewis elképzelhetőnek tartja, hogy fizikailag egymástól eltérő állapotok is betölthessék a fájdalom szerepét. Ez a többszörös realizáció kérdésköréhez vezet. A mentális állapot típusokat fizikai állapot típusokkal azonosító elméletekkel szemben ugyanis felhozták (Putnam 1980), hogy amennyiben egy mentális állapot típus többféle fizikai állapot által is realizálható, akkor végső soron nem sikeres a mentális állapot típusok azonosítása fizikai *típusokkal*. Ha a fájdalom más fizikai állapotokkal azonos az emberekben, mint a polipokban, akkor a fájdalom végső soron nem azonosítható egy fizikai típusal. Egyes filozófusok ennek hatására a típusazonosság elmélettől visszavonultak a példány-azonosság elméletéhez (Lewis 1994/2003: 2007). Ez nem kívánatos fejlemény, hiszen a példány-azonosság elmélet nem reduktív materialista elmélet (Tózsér 2008: 41), hanem a fizikaiakra nem redukálható mentális tulajdonságokkal is számol. Ezzel vét a redundancia kritériuma ellen.

Lewis egyébként nem tartotta végzetesnek az általa képviselt elméletre nézve a többszörös realizáció érvét. A mentális-fizikai azonosság elmélet híve ugyanis például fajokra korlátozhatja az általa definiált mentális állapot típusok érvényét: „a fájdalom egyfajta agyállapot az emberben, és valamilyen más agy (vagy nem-agy) állapot a puhatestűek esetében” (Lewis 1980b: 233).

Nem mindenki elégedett meg Lewisnak ezzel a válaszával, és a reduktív materializmus és a többszörös realizáció összeegyeztethetőségének kérdése a későbbiekben számos vitát váltott ki, amelyeknek áttekintésére itt nincs mód. Egy biztos: önmagában arra a lehetőségre hivatkozni, hogy fizikailag többszörösen megvalósíthatók a mentális állapot típusok, vagyis ugyanazt a mentális szerepet többféle fizikai állapot is betöltheti, nem perdöntő érv a Lewis által megfogalmazott típusazonosságelmélettel szemben.

Mint már szó volt róla, Lewis a mentális állapotokat kauzális-funkcionális terminusokban definiálja. A fájdalom tehát olyan állapot, amibe bizonyos ingerek, fizikai behatások (például szúrás, égés stb.) hatására kerülünk, és amelyet bizonyos érzelmi és motivációs állapotok jellemeznek (például azt akarjuk, hogy szűnjön meg a fájdalom; fájdalomcsillapítót veszünk be vagy orvoshoz fordulunk; nem tudjuk a figyelmünket összpontosítani; aggodalmat érzünk stb.). A fájdalom ilyen definíciója nem a fenomenális karaktere, hanem a kauzális szindrómája segítségével definiálja a fájdalmat. Lewis álláspontját úgy értelmezem, hogy a fájdalom „fenomenális karaktere”, amennyiben azt el tudnánk különíteni minden egyéb aspektusától, nem más lenne, mint a fájdalom-állapotot realizáló állapot fizikai természete. Amikor látszólag a fájdalom „minőségéről” beszélünk, például arról, hogy milyen (szúrós, égető stb.) fájdalmat éltünk át a fogorvosnál, akkor valójában nem a fenomenális karakterről (végső soron a fájdalom szerepet betöltő fizikai állapot fizikai természetéről) beszélünk, hanem csak arról, amiről tudomásunk lehet, vagyis az említett szituáció pszichológiai jellemzőiről. Egy „szúró fájdalom” például egy olyan fájdalom, ami olyan, mint azok a fájdalmak, amelyek szúrás hatására következnek be. Ilyen körülírásokkal, közvetetten meg tudjuk ragadni a kválékát: „tökéletlenül, de elég jól”, mondaná Lewis. Természetesen egy ilyen közvetett leírás olyan pontos is lehet, hogy egyértelműen kijelölheti azt a fizikai állapotot, ami akkor áll elő, ha szúrási inger éri a releváns idegvégződéseket.

Miközben tehát közvetlenül nem férünk hozzá a mentális állapotaink fenomenális karakteréhez, azaz nem látunk bele introspekció révén vagy más közvetlen módon azok implementációjának fizikai részleteibe, aközben normál esetben (például amikor fájdalomról számolunk be) ez semmiféle nehézséget vagy rejtélyt nem okoz, hiszen mentális állapotaink pszichológiai aspektusa (kauzális szerepe) és fenomenális aspektusa (a realizáció fizikai módja) a gyakorlatban elválaszthatatlan.

Ennek ellenére a tudatfilozófia rejtélyei előállnak, mihelyt megértjük a „fenomenális hozzáférhetetlenségét”, ami abból ered, hogy introspekcióval nem látunk bele mentális állapotaink implementációs részleteibe.

Ha Lewis elméletét úgy értelmezzük, hogy az elme pszichológiai aspektusáról funkcionalista beszámolót ad, míg a fenomenális aspektusa vonatkozásában fizikai-fenomenális típus-azonosságelméletet fogalmaz meg, akkor az elmélet alapján az antirevizionista rejtvények – mint amilyen a mások színlátásának kérdése – értelmesek lesznek, sőt, meg is lehet válaszolni legalább egy részüket. Arra is választ kapunk, hogy a tudat miért tűnhet egyszerre „rejtélyesnek” és a „legnyilvánvalóbb adottságnak” is.

Revizionista szempontból Lewis semmilyen felesleges feltételezéssel nem él. Azt, hogy pszichológiai állapotaink valahogyan realizálva vannak, nem teszi szükségessé, hogy a fizikai objektumokon és tulajdonságokon kívül mást is feltételezzünk. Lewis elutasítja, hogy léteznének különleges státuszú „fenomenális tulajdonságok”, amelyekről tévedhetetlen, közvetlen, lényegfeltáró tudásunk van. A tapasztalat fenomenális tulajdonságai fizikai tulajdonságok.

Antirevizionista szempontból Lewis elméletének erénye, hogy nem redukálja az elme fenomenális aspektusát a pszichológiai aspektusra. A tudattal kapcsolatos „rejtélyeket” a fenomenális és a pszichológiai közötti kontraszt generálja: az elme fenomenális aspektusa, vagyis az, hogy „milyen nekünk lenni” a szó fenomenális értelmében, a fizikai implementáció részleteivel függ szorosan össze, míg az elme pszichológiai aspektusa kauzális-funkcionális terminusokban megmagyarázható struktúra. Miközben sokak erős intuíciója, hogy „valamilyen” nekünk lenni, aközben ez – legalábbis „sajátosan filozófiai” értelemben – ineffábilis marad, pontosan úgy, ahogy Block példájában egy számítógép számára hozzáférhetetlen az, hogy milyen színűek benne a vezetékek.

Miközben agyunk fizikai tulajdonságai rejtve maradnak introspekciónk előtt, és nem képezik közvetlen, noninferenciális tudás tárgyát, aközben ez a rejtélyességük és ineffabilitásuk nem jelenti, hogy az agykutató vagy a fizikus szemszögéből is rejtélyes tulajdonságok lennének.

Tehát az az elmélet, amely a tudat fenomenális aspektusát a fizikai tulajdonságokkal azonosítja, míg funkcionalista az elme pszichológiai aspektusa vonatkozásában, anélkül tud megfelelni az antifunkcionalista („intrinzikussági”) intuícióknak, és anélkül tudja megmagyarázni, hogy a tudatosság miért tűnhet rejtélyesnek és megfoghatatlannak, hogy ehhez feltételeznie kellene, hogy a tudat valamilyen nem-fizikai tulajdonság, entitás vagy folyamat.

Lewis elmélete alapján megválaszolhatunk néhány jellegzetes antirevizionista kérdést, de korántsem mindegyiket. Láthatom-e pirosnak, amit pontos fizikai másolatom zöldnek lát? Nem, hiszen két fizikai másolat mind pszichológiai, mind fenomenális másolatok. Lewis elmélete ugyanakkor mintha nyitva hagyná azt a kérdést, hogy láthatja-e (kizárólag a fenomenális aspektus

értelmében) „pirosnak” egy lény azt, amit a másik „zöldnek” lát úgy érve, hogy viselkedésük – beleértve verbális és kognitív viselkedésüket a legapróbb részletekig – egyébként azonos? Ehhez csupán az kell, hogy a két lény eltérő fizikai felépítésű legyen.

Ilyen példák jelennek meg Lewis *Mad Pain and Martian Pain* című tanulmányában (Lewis 1980a). Az „őrült” olyan lény, aki fizikailag hozzánk hasonló felépítésű, ám a fájdalom az életében nem játssza a fájdalom szerepét: ami másnak fájdalmas inger, az számára nem az. A „marslakó” más felépítésű, mint mi: agyában nincsenek neuronok, az agya hidraulikus elven működik. Ugyanakkor bizonyos ingerek fájdalom szerepét játszanak az életében. Lewis szerint egy hihető elmefilozófiai elméletnek értelmeznie kell tudnia mind az őrült-, mind a marslakó-fájdalmat. A két elképzelt lény példájának tanulsága Lewis szerint az, hogy egyfelől a fájdalom csak kontingensen kapcsolódik egy kauzális szerephez, és csak kontingensen kapcsolódik egy bizonyos fizikai realizációhoz. Egy funkcionista elmélet képes értelmezni a marslakó-fájdalmat, egy egyszerű azonosságelmélet képes értelmezni az őrült-fájdalmat. Lewis szerint a megoldás a kevert elmélet, ami szerint a marslakó azért érez fájdalmat, mert a megfelelő funkcionális állapotban van, míg az őrült azért, mert a megfelelő fizikai állapotban. A fájdalom fogalma egy olyan állapot fogalma, amely bizonyos kauzális szerepet tölt be, ám a fájdalom fogalma nem a kérdéses szerepet betöltő neurális állapot fogalma (Lewis 1980a: 218).

Lewis az ember és a marslakó esetére azt a megoldást javasolja, hogy „ember-fájdalomról” illetve „marslakó-fájdalomról” beszélhetünk, amikor azokról az állapotokról beszélünk, melyek ezekben a populációkban betöltik a fájdalom szerepét. Hasonló szűkítésekkel, relativizálásokkal lehet kezelni az őrült (sőt, az őrült marslakó) esetét is. Más esetek azonban kevésbé világosak Lewis (1980a: 220) szerint: ide tartozik a spektrum inverzió esete is. Amikor a spektrum inverzió áldozata „pirosat lát, amikor a fűre néz”, akkor a fűre nézve abban az állapotban van, ami a többi embernél a piros látás állapotának funkcionális szerepét szokta betölteni. Másrészt viszont azt is lehet mondani, hogy zöldet lát ilyenkor, mert abban az állapotban van, ami számára a zöldet látás szerepét szokta betölteni. (Hozzátehetjük, hogy ez a dilemma valószínűleg eldönthető lenne, ha kiderülne, hogy az illető fiziológiailag abnormális, például a szemében másfajta színérzékelő sejtek vannak, mint a normál populációnak).

Lewis tanulmányának azonban utolsó fejezete az egyik legérdekesebb, ahol igyekszik elébe menni annak az ellenvetésnek, hogy tanulmányában nem foglalkozott azzal, milyen marslakónak vagy őrültnek lenni: „Milyen az állapotának fenomenális karaktere? Ha fájdalomnak érződik számára, akkor fájdalom is, akármilyen kauzális szerepe vagy fizikai természete legyen. Ha nem, akkor meg nem.” (Lewis 1980a: 221).

Lewis válasza az, hogy nem is lehet azon gondolkodni, hogy egy állapot fájdalom-e, ha közben nem gondolunk arra is, hogy milyen ebben az állapotban lenni. Az egyikről gondolkodni a másik nélkül olyan lenne, mint egy szám párosságán gondolkodni anélkül, hogy az osztóira gondolnánk (Lewis 1980a: 222).

A fájdalom egy érzés. [...] Egy állapot fájdalom-volta és fájdalomnak érződése [...] egy és ugyanaz. Egy elmélet arról, hogy egy állapot milyen értelemben fájdalom, elkerülhetetlenül egy elmélet arról is, hogy milyen abban az állapotban lenni, hogy milyennek érződik az az állapot, hogy milyen az állapot fenomenális karaktere. Távollattól, hogy figyelmen kívül hagytam volna annak kérdését, hogy milyennek érződnek az állapotok [...] semmi másról nem beszéltem! Csak ha abban hiszel [...] hogy a kauzális szereppel és fizikai realizációval kapcsolatos megfontolások nem játszanak szerepet abban, hogy egy állapot fájdalom-e, mondhatod, hogy nem játszanak szerepet abban, milyennek érződik egy állapot. (Lewis 1980a: 222)

Ez az idézet nem látszik egyértelműen támogatni azt a Lewis-olvasatot, amely szerint Lewis számára elfogadható az elme pszichológiai és fenomenális aspektusának szétválasztása, és az előbbi funkcionista, az utóbbi materialista elemzése. Lehet azonban, hogy a probléma az, hogy tanulmánya a fájdalomról szól, amely egy meglehetősen összetett, sajátos funkcionális, érzelmi, motivációs (lásd: Klein 2015, Dennett 1978) és fenomenális karakterrel is rendelkező állapot. Eltérő „milyenségű” fájdalmak lehetnek egyformán *kellemetlenek*, de nem minden kellemetlen érzés egyben *fájdalom*. Vagyis a fájdalom esetében igaza lehet Lewisnak abban, hogy mind a fizikai realizációval, mind a kauzális szerepekkel kapcsolatos megfontolások szerepet játszanak annak eldöntésében, hogy egy állapot fájdalom-e. De valóban különleges példa a fájdalom ebből a szempontból? Lehet, hogy nem is lehet igazi példát hozni egy olyan mentális állapotra, amelyet csak fenomenális vagy csak funkcionális vonatkozásai alapján is lehet definiálni.

6. Konklúzió és kitekintés

Dolgozatom kiindulópontja az az észrevétel volt, hogy amikor a tudatosság kérdéséről van szó, akkor mind a filozófusok, mind a nem-filozófusok körében homlokegyenest ellenkező

elképzelések jelentkeznek „nyilvánvalóan helyes” álláspontként. Ez a nézetkülönbség olyan jelentősnek tűnik, hogy sok esetben érdemi viták helyett patthelyzeteket, félreértéseket, vagy az egymás melletti elbeszélés, illetve a témaváltás vádjainak hangoztatását figyelhetjük meg.

Úgy tűnik ugyanis, hogy ami a tudatosság körüli filozófiai kérdéseket illeti, olykor a legalapvetőbb kérdésben sincs egyetértés: abban, hogy van-e egyáltalán tudat.

Revizionista (R-) intuíciónak neveztem azt az intuíciót, amely olyan (revizionista) meggyőződések kialakulásához vezet, melyek szerint nincs tudatosság mint sajátos filozófiai problémákat felvető, vagy a tudományos magyarázatok számára különös nehézséget okozó jelenség. Filozófiai elméletekben a revizionista meggyőződés revizionista tézisekként jutott kifejezésre.

Antirevizionista (A-) intuíciónak neveztem azt az intuíciót, amely olyan (antirevizionista) meggyőződések kialakulásához vezet, melyek szerint a tudat létező jelenség, mely a filozófia és a tudományok számára fontos kutatási terület, mely sok antirevizionista szerint sajátosan „nehéz problémákat” vet fel. Filozófiai elméletekben az antirevizionista meggyőződés antirevizionista tézisekben jutott kifejezésre.

Felmerül az a lehetőség, hogy az elmefilozófiai elméletek osztályozhatók úgy, mint amelyek szerzői az R- vagy az A-intuíciót igyekeztek kifejezésre juttatni. Amikor azonban megvizsgálunk revizionistaként számon tartott szövegeket, rendszerint könnyen előállíthatónak bizonyul az elmélet antirevizionista olvasata, az antirevizionista szövegeknek pedig revizionista olvasata. Az elmefilozófiai elméletek állításai gyakran kétértelműnek bizonyulnak, ezért a revizionista szerzők nem tudják meggyőzni antirevizionista vitapartnereiket és viszont. Ehelyett rendszerint két malomban örölnek, vagy álláspontjuk félreértését róják fel egymásnak.

Rorty (1998), Levine (2001), Block (1978, 1995) és más szerzők munkássága alapján az antirevizionista intuíció tartalma pontosítható – az antirevizionisták szerint a tudat létezik, és a tudatosság valamilyen intrinzikus tulajdonság, illetve okkurrens állapot vagy folyamat.

Az a tény, hogy az elmefilozófiában két, gyökeresen eltérő intuitív meggyőződésekkel rendelkező tábor van, felveti a kérdést, hogy mely szerzők intuíciói állnak közelebb természetes meggyőződéseinkhez. Ennek megválaszolásához mindenekelőtt Sytsma munkásságát hívtam segítségül, aki empirikus felméréseket végzett a nem-filozófusoknak a tudatossággal kapcsolatos intuícióval kapcsolatban. Az eredmények nemcsak azért voltak zavarba ejtők, mert a legtöbb kérdésben nem volt konszenzus, így nem lehet egyértelműen kijelenteni, hogy a revizionizmus vagy az antirevizionizmus-e a természetesebb meggyőződés, hanem azért is, mert a kutatások során fény

derült arra, hogy a kutatók által feltett kérdések gyakran éppen olyan kétértelműek, mint amilyenek az elmefilozófiai elméletek tünnek az elmefilozófiai viták során, így a válaszokat is nehéz értelmezni.

Az elmefilozófiai viták vizsgálata arra enged következtetni, hogy minden elmefilozófiai elméletnek lehetséges revizionista és antirevizionista olvasata. Továbbá amikor a filozófusok egymás műveit olvassák, négyféle olvasat (A1, A2, R1, R2) állhat rendelkezésükre abban a tekintetben, hogy az értelmező revizionista vagy antirevizionista, illetve intuíció-monista vagy intuíció-pluralista.

Ez a négyféle olvasat talán megegyezik azzal a négyféle értelmezéssel, melyet Rorty példájában a földi tudósok és Antipodea lakói alkalmaznak egymásra. Ha helytálló ez a Rorty-olvasat, akkor Rortyt az ebben az értekezésben ismertetett „négyféle olvasat” első felfedezőjeként ünnepelhetjük. Ha ez így van, akkor Rorty volt az első szerző, aki diagnosztizálta az elmefilozófiai elméletek rendszeresen jelentkező kétértelműségét, és megpróbálta meghaladni az antirevizionista és revizionista álláspont konfliktusát. Egy másik értelmezés szerint viszont Rorty nem annyira meghaladni kívánta az antirevizionizmus és revizionizmus ellentétét, mint eldönteni a kettő közötti vitát a revizionizmus javára.

Szembesülve a négyféle olvasat és az intuíciók összecsapásának problémáival, azt a követelményt fogalmaztam meg a tudat filozófiai elméleteivel szemben, hogy az ideális elméletnek „egyensúlyi elméletnek” kell lennie, vagyis az antirevizionisták és a revizionisták számára egyaránt plauzibilisnek. Az egyensúlyi elmélet „nem érzékeny” a kétértelműségre, a következő értelemben: az egyensúlyi elmélet semmilyen feltételezéssel nem él, amely revizionista szempontból fölösleges, ugyanakkor antirevizionista nézőpontból nem vádolható azzal, hogy „kihagy valamit”.

Állításom az, hogy a reduktív materializmus egyensúlyi elmélet, ezért preferálandó mindazokkal az elméletekkel szemben, melyeket azzal a szándékkal alkottak, hogy az antirevizionista (vagy a revizionista) intuitív meggyőződéseknek szerezzenek érvényt, de közben revizionista szemszögből felesleges feltételezésekbe bocsátkoztak, vagy antirevizionista szemszögből „kihagytak valami fontosat” (például az élmény „milyenségét”).

Ha ez a gondolatmenet sikeres, akkor új érvet jelent a reduktív materializmus elfogadása mellett.

Értekezésem negyedik fejezetében megkíséreltem feltárni az elmefilozófiát sújtó kétértelműségek okát. Véleményem szerint ez az ok a fenomenális kognitív és szemantikai hozzáférhetetlensége, amit azonban nagyon könnyű félreérteni vagy rosszul értelmezett intuitív

meggyőződések, illetve bizonyos, nem jó okkal elfogadott filozófiai előfeltevések alapján elutasítani.

A fenomenális hozzáférhetetlenségének doktrínáját sokféleképpen megfogalmazták olyan szerzők, mint Schlick, Wittgenstein, Ryle, Farrell vagy Feyerabend. Beszéltek a (struktúrával szembeállított) tartalom kommunikálhatatlanságáról, különbséget tettek ismeretség és ismeretségből vett tudás között, vagy éppen a nyers érzetek „tulajdonságnélküliségét”, „ürességét” hangoztatták. A legfontosabb tanulság az, hogy nem rendelkezünk közvetlen ismeretségből vett tudással tapasztalataink kvalitatív jellemzőiről, miközben bizonyos értelemben viszont „közvetlen ismeretségben vagyunk” tapasztalataink fenomenális minőségével.

Akik összetévesztik azt az állítást, hogy tapasztalataink fenomenális karakteréről nem rendelkezünk közvetlen „ismeretségből vett tudással”, azzal az állítással, hogy fenomenális karakter nincs is, azok amennyiben igazságot akarnak szolgáltatni antirevizionista meggyőződéseiknek, gyakran a „fenomenális tudás” vagy a „közvetlen ismeretség” olyan elméletét fogadják el, amely nem egyeztethető össze a fizikalizmussal. A filozófia történetében visszatérő motívum, hogy az antirevizionista intuíciót antifizikalistaként értelmezték. Ez az értelmezés azonban nem szükségszerű. Úgy tűnik, már Place, Smart és Feigl is úgy gondolták, hogy ezt csak a fenomenális hozzáférhetetlenségének tudomásul vételével lehet elkerülni. Egyikük sem tagadta, hogy tapasztalataink fenomenális minősége valóságos, intrinzikus tulajdonság lett volna. Éppenséggel azonosították is, milyen intrinzikus tulajdonságokról van szó: neurofiziológiai, illetve fizikai tulajdonságokról. Másrészt hozzátették, hogy ezekhez a tulajdonságokhoz az introspekcióban közvetlenül nem férünk hozzá.

Mivel azt a meglehetősen árnyalt pozíciót, amit a fenomenális tudatossággal kapcsolatban a korai reduktív materialisták és „logikai behaviorista” előfutárai elfoglaltak, igen könnyű félreérteni, az antirevizionista intuíciókat könnyű antifizikalistaként félreértelmezni.

Értekezésem utolsó fejezetében röviden bemutatom a reduktív materialista elmefilozófiát abban a formában, ahogyan azt legkorábban kifejtették, valamint egy érettebb, Lewis munkásságára visszavezethető formában. Amellett érvelek, hogy ezek egyensúlyi elméletek. Semmi sem marad ki belőlük, ami egy antirevizionista számára fontos, ugyanakkor semmilyen olyan entitás vagy tulajdonság létezését nem feltételezik, amely egy antirevizionista szemszögéből felesleges lenne.

Ebben az értekezésben a reduktív materialista elmefilozófiának csak néhány, ma már klasszikusnak számító változatát vizsgáltam meg. Ezek az elemzések csak az első lépések voltak egy lehetséges kutatási irányban, hiszen Place, Smart, Feigl, Armstrong és Lewis óta rengeteg

materialista tudatfilozófiai elmélet született. Érdeemes lenne megvizsgálni, vajon ezek is egyensúlyi elméleteknek tekinthetők-e. Némelyik elméleti újítással kapcsolatban az lehet a gyanúnk, hogy vétettek a redundancia kritériuma ellen, túlinterpretálva az antirevizionista intuíciót.

Így például a 90-es évektől többek között Brian Loar hatására az a felfogás került előtérbe, hogy tudati állapotainkról alkotott fogalmaink nem tárgyközömbös elemzés által leírt módon referálnak, ahogy azt a reduktív materializmus 50-es és 60-as évekbeli képviselői állították, hanem közvetlenül (Tartaglia 2013: 3). Ezzel kibontakozott az elmefilozófiában a fenomenális fogalom stratégia (FFS). Tartaglia szerint a FFS hívei Nagel hatására gondolták úgy, hogy csak így lehet megérteni azt a tényt, hogy „valamilyen” átélni a bennünk zajló (fizikai) folyamatokat, és erről fogalmat tudunk alkotni. Ezeknek a fogalmaknak az élmény valamilyen lényegi, intrinzikus tulajdonságát kell megragadniuk. Míg Place, Smart, sőt Lewis is elutasították a fenomenális fogalmakat, az elmefilozófusok új generációja úgy gondolta, szemben az 50-es és 60-as évek szerzőivel, hogy megengedhető ilyen fogalmak létezése anélkül, hogy ebből dualizmus következne (egy ártalmatlan „fogalmi dualizmuson” túl, vö. Papineau 2002: 96). A FFS hívei nem utasítják el a fenomenális tulajdonságokat (melyekre a fenomenális fogalmak referálnak), mint Place és Smart, hanem úgy gondolják, ezek azonosíthatók fizikai tulajdonságokkal. Azt pedig, hogy vannak fenomenális tulajdonságok és róluk alkotott fenomenális fogalmak, a FFS képviselői kétségbevonhatatlan adottságnak tartják. Loar szerint a fenomenális fogalmak elfogadása következik abból, ha tapasztalatunk fenomenológiáját „névértéken vesszük” (Tartaglia 2013: 5). Tartaglia azonban úgy látja, hogy Loar elmélete azon a címen, hogy „névértéken veszi fenomenológiánkat”, egészen annak feltételezéséig jut el, hogy elkerülhetetlenül dualista módon gondolkodunk tapasztalatainkról, a dualizmus mintegy „bele van drótozva” elménkbe, és nem pedig pusztán egy XVII. századi filozófiai elmélet (Tartaglia 2013: 17).

Vajon valódi előrelépést jelent a FFS fentebb jellemzett változata a fenomenális hozzáférhetetlenségét elfogadó elméletekhez képest? A FFS körüli élénk viták azt tanúsítják, hogy ez kétségbe vonható. Kétséges, hogy olyan elméletekre van szükség, melyek az antirevizionista intuíciókat dualista intuícióként értelmezik, majd ennek próbálják valamilyen módon, például „fogalmi dualizmus” bevezetésével, elvenni az élet. Így például Goff (2011) szerint még az sem igaz, hogy a stratégia jobban megfelel a fenomenális tulajdonságokkal kapcsolatos intuícióinknak (melyek akár hamisak is lehetnek), mint a FFS előtti fizikalizmus.

Természetesen nem lehet ennyivel elintézni a FFS és ellenzőinek munkásságát. Mindössze csak rámutattam arra, felidézve Goff és Tartaglia kritikáját, hogy attól, hogy egy elmélet későbbi, mint Place, Smart, Armstrong vagy Lewis elmélete, még nem feltétlenül képvisel egyértelmű

előrelépést hozzá képest. Egy biztos: a filozófusok újra és újra megkísérlik artikulálni az antirevizionista meggyőződéseket a filozófiában, és persze, mint az illuzionisták esetében láttuk, megérkezik erre a revizionista reakció is. Place, Smart és Feigl korában az olyan antirevizionisták, akik nem fogadták el a fenomenális hozzáférhetetlenségét, rendszerint valamilyen dualista álláspontonál lyukadtak ki, mert a „közvetlen ismeretség” révén megismert „fenomenális tulajdonságokat” kellett feltételezniük. A FFS hívei ezt a dualizmust Papineau kifejezésével „fogalmi dualizmussá” kívánják szelídíteni azáltal, hogy a „fenomenális fogalmakról” materializmussal kompatibilis elméletet alkotnak. Ha ez sikerül is a FFS követőinek, még nem biztos, hogy az elméletük jobb a fenomenális fogalmak nélküli materializmusnál: például azért, mert lehet, hogy ez az elmélet már nem egyensúlyi. Olyan „dualista intuíciókat” próbál összeegyeztetni a materializmussal, amelyekkel a legtöbb ember talán nem is rendelkezik. Különösen revizionista nézőpontból dolgozik igen sokat feleslegesen.

Egy biztos: még nagyon hosszú út áll előttünk az elmefilozófiai álláspontok mögött álló intuitív meggyőzések feltárásában. Mint láttuk, olyan filozófusok, mint Chalmers vagy Rorty, és még sokan mások az ebben az értekezésben elemzett szerzők közül, már megkezdték ezeknek az intuícióknak a feltárását.

Értekezésemben ehhez a tevékenységhez igyekeztem hozzájárulni. Igyekeztem megmutatni, hogy azok a gyakran meglehetősen frusztráló félreértések, kétértelműségek és patthelyzetek, melyek az elmefilozófiai vitákban gyakran visszaköszönnek, nem véletlenül fordulnak elő, hanem az A- és R-intuíció közti ellentétre és a négyféle olvasat lehetőségére vezethetők vissza, amelynek oka pedig az, hogy rendkívül sajátos episztemikus viszonyban vagyunk tapasztalatunk fenomenális jellemzőivel. Javasoltam, hogy tekintettel ezekre a problémákra, részesítsük előnyben az egyensúlyi elméletet, mely a legkevésbé érzékeny a kétértelműségre, ugyanis a fenomenális hozzáférhetetlenségének doktrínáját tudomásul veszi, és nem interpretálja túl az antirevizionista intuíciót. Még nem világos, hogy az a keskeny út, amit a megfelelés és a redundancia kritériuma jelöl ki, elvezet-e további érdekes elméletekhez vagy hozzásegít-e tudományos felfedezésekhez, de talán az ebben az értekezésben felvetett szempontok egyébként is hozzájárulnak, hogy valamivel jobban megértsük az elmefilozófiai vitákat.

Bibliográfia

- Alexander, S. (1920). *Space, Time and Deity: The Gifford Lectures at Glasgow, 1916-1918*. Volume I. London: Macmillan and Co. Ltd.
- Ambrus G. (2007). *A tudat metafizikája*. Budapest: Gondolat.
- Antony, M. V. (1992). The where and when of what? *Behavioral and Brain Sciences*, 15/2, 201-202.
- Antony, M. V. (1994). Against Functionalist Theories of Consciousness. *Mind and Language*, 9, 105-123.
- Armstrong, D. M. (1968) *A Materialist Theory of the Mind*. London: Routledge & Kegan Paul.
- Armstrong, D. M. (1970). The nature of mind. In: *The Mind-Brain Identity Theory. Controversies in Philosophy*. London: Palgrave. 67-79.
- Armstrong, D. M. (1995). Perception-consciousness and action-consciousness? *Behavioral and Brain Sciences*, 18/2, 247-248.
- Armstrong, D. M. (1999). *The Mind-Body Problem: An Opinionated Introduction*. Boulder, Co.: Westview Press.
- Aronson, J., Dietrich, E., Way, E. (1992). Throwing the conscious baby out with the Cartesian bath water. *Behavioral and Brain Sciences*, 15/2, 202-203.
- Atkinson, A. P., Davies, M. (1995). Consciousness without Conflation. *Behavioral and Brain Sciences*, 18/2, 248-249.
- Block, N. (1978). Troubles with Functionalism. *Minnesota Studies in the Philosophy of Science*, 9, 261-326.
- Block, N. (1992). Begging the question against phenomenal consciousness. *Behavioral and Brain Sciences*, 15/2, 205-206.
- Block, N. (1995). On a Confusion about a Function of Consciousness. *Behavioral and Brain Sciences*, 18/2, 227-247.

Block, N. (2015). The Canberra Plan Neglects Ground. In Terence Horgan, Marcelo Sabates, David Sosa (eds.), *Qualia and Mental Causation in a Physical World: Themes from the Philosophy of Jaegwon Kim*. Cambridge University Press. 105-133.

Chalmers, D. J. (1995/2004). Szemközt a tudat problémájával. Fordította Sutyák Tibor. *Vulgo*, 2004/2, 14-35.

Chalmers, D. J. (1996). *The Conscious Mind: In Search of a Fundamental Theory*. Oxford: Oxford University Press.

Chalmers, D. J. (1997). Moving Forward on the Problem of Consciousness. *Journal of Consciousness Studies*, 4(1), 3-46.

Chalmers, D. J. (2018). The Meta-Problem of Consciousness. *Journal of Consciousness Studies*, 25(9-10), 6-61.

Church, J. (1995). Fallacies or analyses? *Behavioral and Brain Sciences* 18/2, 251-252.

Churchland, P. M. (1981). Eliminative Materialism and Propositional Attitudes. *Journal of Philosophy*, 78/2, 67-90.

Clark, A. (1992). Experiential facts? *Behavioral and Brain Sciences* 15/2. 207-208.

Cottingham, J. (2016). Descartes and the Problem of Consciousness. In Leach, S., Tartaglia, J. (eds.), *Consciousness and the Great Philosophers: what would they have said about our mind-body problem?* London: Routledge. 63-72.

Demeter T. (2008). Népi pszichológia (Bevezető tanulmány). In Ambrus G., Demeter T., Forrai G., Tózsér J. (szerk). *Elmefilozófia*. Budapest: L'Harmattan. 411-425.

Dennett, D. C. (1971). Intentional systems. *Journal of Philosophy*, 68/feb, 87-106.

Dennett, D. C. (1978). Why You Can't Make a Computer that Feels Pain. *Synthese*, 38/3, 415-449.

Dennett, D. C. (1988). Quining Qualia. In A. Marcel and E. Bisiach (eds.), *Consciousness in Modern Science*, Oxford: Oxford University Press.

Dennett, D. C. (1991). *Consciousness Explained*. Boston, MA: Little, Brown and Co.

Dennett, D. C. (1991b): Real patterns. *Journal of Philosophy*, 88/1, 27-51.

Dennett, D. C., Kinsbourne, M. (1992): Time and the Observer: The Where and When of Consciousness in the Brain. *Behavioral and Brain Sciences*, 15/2, 183-247.

- Dennett, D. C. (1995). The path not taken. *Behavioral and Brain Sciences*, 18/2, 252-253.
- Dennett, D. C. (2012). The Mystery of David Chalmers. *Journal of Consciousness Studies*, 19/1-2, 86-95.
- Dennett, D. C. (2016). Illusionism as the Obvious Default Theory of Consciousness. *Journal of Consciousness Studies*, 23/11-12, 65-72.
- Dennett, D. C. (2019). Welcome to Strong Illusionism. *Journal of Consciousness Studies*, 26/9-10, 48-58.
- Dretske, F. (2003). How do you know you are not a zombie. In Brie Gertler (ed.), *Privileged Access: Philosophical Accounts of Self-Knowledge*. Ashgate. 1-14.
- Farah, M. J. (1992). The distributed pineal gland. *Behavioral and Brain Sciences*, 15/2, 209-209.
- Farah, M. J. (1995). Is consciousness of perception really separable from perception? *Behavioral and Brain Sciences*, 18/2, 254-255.
- Farrell, B. A. (1950). Experience. *Mind*, 59/april, 170-198.
- Feigl, H. (1958). The 'mental' and the 'physical'. *Minnesota Studies in the Philosophy of Science*, 2, 370-497.
- Feyerabend, P. (1970). Materialism and the mind-body problem. In: *The Mind-Brain Identity Theory. Controversies in Philosophy*. London: Palgrave. 142-156.
- Flanagan, O., Wallace, H. (2016). William James and the problem of consciousness. In Leach, S., Tartaglia, J. (eds.), *Consciousness and the Great Philosophers: what would they have said about our mind-body problem?* London: Routledge. 152-161.
- Frankish, K. (2016). Illusionism as a Theory of Consciousness. *Journal of Consciousness Studies*, 23/11-12, 11-39. Az oldalszámok a következő online változatra vonatkoznak: https://nbviewer.org/github/k0711/kf_articles/blob/master/Frankish_Illusionism%20as%20a%20theory%20of%20consciousness_eprint.pdf
- Friedl, J., Rutte, H. (2013). Form and Content. In: Friedl, J., Rutte, H. (eds) *Moritz Schlick. Moritz Schlick Gesamtausgabe*, vol 1.2. Springer: Vienna.
- Glymour, B., Grush, R., Hardcastle, V. G., Keeley, Ramsey, B., Shagrir, O., Watson, E. (1992). The Cartesian Theater stance. *Behavioral and Brain Sciences*, 15/2, 209-210.

Goff, Ph. (2011). A posteriori physicalists get our phenomenal concepts wrong. *Australasian Journal of Philosophy*, 89/2, 191-209.

Goff, Ph. (2016). Bertrand Russell and the problem of consciousness. In Leach, S., Tartaglia, J., (eds.), *Consciousness and the Great Philosophers: what would they have said about our mind-body problem?* London: Routledge. 185-191.

Hacker, P. M. S. (2002). Is there anything it is like to be a bat? *Philosophy*, 77/300, 157-174.

Hacker, P. M. S. (2012). The Sad and Sorry History of Consciousness: being, among other things, a Challenge to the 'Consciousness-studies Community'. *Royal Institute of Philosophy Supplement*, 70, 149-168.

Harman, G. (1990). The intrinsic quality of experience. *Philosophical Perspectives*, 4, 31-52

Harman, G. (1995). Phenomenal fallacies and conflation. *Behavioral and Brain Sciences*, 18/2, 256-257.

Heil, J. (2013). *Philosophy of Mind: A Contemporary Introduction*. New York: Taylor and Francis.

Hellie, Benj. (2013) Against Egalitarianism. *Analysis* 73 (2):304-320

Hill, Ch. S. (1991). *Sensations: a Defense of Type Materialism*. Cambridge: Cambridge University Press.

Hill, Ch. S. (2009). *Consciousness*. New York: Cambridge University Press.

Irvine, L., Sprevak, M. (2020): Eliminativism about Consciousness. In U. Kriegel (ed.), *Oxford Handbook of the Philosophy of Consciousness (2020)*. Oxford: Oxford University Press, pp. 348–370. A hivatkozások a következő online változat oldalszámaira vonatkoznak:

<https://marksprevak.com/pdf/paper/IrvineSprevak---Eliminativism%20about%20consciousness.pdf>

Jackson, F. (1982). Epiphenomenal qualia. *Philosophical Quarterly* 32/apr, 127-136.

Jackson, F. (2012). In defence of the identity theory mark I. In Hill, Ch., Simone, G. (eds.), *New Perspectives on Type Identity: The Mental and the Physical*. New York: Cambridge University Press. 150-165.

James, W. (1879). Are we automata? *Mind*, 4/13, 1-22.

James, W. (1904). Does "consciousness" exist? *Journal of Philosophy, Psychology, and Scientific Methods*, 1/18, 477-491.

Klein, A. (2020). The Death of Consciousness? James's Case against Psychological Unobservables. *Journal of the History of Philosophy*, 58/2, 293-323.

Klein, C. (2015). What Pain Asymbolia Really Shows. *Mind* 124, 493-516.

Kripke, S. (1972/2007). *Megnevezés és szükségszerűség*. Fordította Bárány Tibor. Budapest: Akadémiai Kiadó.

Kuusela, O. (2016). Wittgenstein and Consciousness. In Leach, S., Tartaglia, J., (eds.), *Consciousness and the Great Philosophers: what would they have said about our mind-body problem?* London: Routledge. 199–208.

Lanier, J. (1995). You can't argue with a zombie. *Journal of Consciousness Studies*, 2/4, 333-345.

Lashley, K. S. (1923). The behavioristic interpretation of consciousness. *Psychological Review* 30/4, 237-272., 30/5, 329-353.

Levine, J. (1983). Materialism and Qualia: the Explanatory Gap. *Pacific Philosophical Quarterly*, 64 , 354-361.

Levine, J. (1993/2008). A milyenség kihagyásáról. In Ambrus G., Demeter T., Forrai G., Tözsér J. (szerk). *Elmefilozófia*. Budapest: L'Harmattan. 354-372.

Levine, J. (1995). Phenomenal access: A moving target. *Behavioral and Brain Sciences*, 18/2, 261-261.

Levine, J. (2001). *Purple Haze: The Puzzle of Consciousness*. New York: Oxford University Press USA.

Levy, N. (2014). The Value of Consciousness. *Journal of Consciousness Studies*, 21/1-2, 127-138.

Lewis, D. (1966). An Argument for the Identity Theory. *Journal of Philosophy*, 63, 17-25.

Lewis, D. (1972). Psychophysical and theoretical identifications. *Australasian Journal of Philosophy*, 50/3, 249-258.

Lewis, D. (1980a). Mad Pain and Martian Pain. In Ned Block (ed.), *Readings in the Philosophy of Psychology*. Cambridge: Harvard University Press. 216-222.

Lewis, D. (1980b). Review of Putnam. In Ned Block (ed.), *Readings in Philosophy of Psychology*. Cambridge: Harvard University Press. 232-233.

Lewis, D. (1988/2008). Amire a tapasztalat tanít. In Ambrus G., Demeter T., Forrai G., Tózsér J. (szerk). *Elmefilozófia*. Budapest: L'Harmattan. 325-352.

Lewis, D. (1994/2003). Reduction of Mind. In O'Connor, T., Robb, D. (eds.) *Philosophy of mind. Contemporary readings*. London: Routledge. 197-209.

Lewis, D. (1995). Should a Materialist Believe in Qualia? *Australasian Journal of Philosophy*, 73/1, 140-144.

Livingston, P. M. (2004). *Philosophical History and the Problem of Consciousness*. Cambridge: Cambridge University Press.

Lloyd, D. (1992). Toward an identity theory of consciousness. *Behavioral and Brain Sciences*, 15/2, 215-216.

Lloyd, D. (1995). Access denied. *Behavioral and Brain Sciences*, 18/2, 261-262.

Mandik, P. (2016). Meta-Illusionism and Qualia Quietism. *Journal of Consciousness Studies*, 23, 140-148.

Nagel, T. (1974/2004). Milyen lehet denevérek lenni? Fordította Sutyák Tibor. *Vulgo*, 2004/2, 3-12.

Papineau, D. (2002). *Thinking About Consciousness*. Oxford: Oxford University Press.

Pereboom, D. (2016). Illusionism and anti-functionalism about phenomenal consciousness. *Journal of Consciousness Studies*, 23, 172-185.

Peressini, A. F. (2017). There is nothing it is like to see red: holism and subjective experience. *Synthese*, 195, 4637-4666.

Place, U. T. (1954/2004). The Concept of Heed. In Graham, G., Valentine, E. R. (eds.) *Identifying the Mind. Selected Papers of U. T. Place*. Oxford University Press. 30-44.

Place, U. T. (1956/2004). Is Consciousness a Brain Process? In Graham, G., Valentine, E. R. (eds.) *Identifying the Mind. Selected Papers of U. T. Place*. Oxford University Press. 45-52.

Pratt, J. B. (1922). Behaviorism and Consciousness. *Journal of Philosophy*, 19, 596-604.

Putnam, H. (1980). The Nature of Mental States. In Block, N. (ed.), *Readings in Philosophy of Psychology*, Volume I. Harvard University Press. 223-231.

Revonsuo, A. (1995). Conscious and nonconscious control of action. *Behavioral and Brain Sciences*, 18/2, 265.

Rey, G. (1988). A Question about Consciousness. In H. R. Otto, J. A. Tuedio (eds.), *Perspectives on Mind*. Dordrecht: D. Reidel. 5-24.

Rey, G. (1995). Block's philosophical anosognosia. *Behavioral and Brain Sciences*, 18/2, 266-267.

Rorty, R. (1965). Mind-body identity, privacy, and categories. *Review of Metaphysics*, 19/1, 24-54.

Rorty, R. (1979). *Philosophy and the Mirror of Nature*. Princeton, NJ: Princeton University Press.

Rorty, R. (1998). Daniel Dennett on Intrinsicity. In Rorty, R.: *Truth and Progress: Philosophical Papers, Volume 3*. Cambridge: Cambridge University Press. 98-121.

Russell, B. (1912). *The Problems of Philosophy*.
<https://gutenberg.org/cache/epub/5827/pg5827.txt>

Ryle, G. (1949/1999). *A szellem fogalma*. Fordította: Altrichter Ferenc. Budapest: Osiris Kiadó.

Schwitzgebel, E. (2016). Phenomenal Consciousness, Defined and Defended as Innocently as I Can Manage. *Journal of Consciousness Studies*, 23/11-12, 224-235.

Searle, J. (1980). Minds, brains, and programs. *Behavioral and Brain Sciences*, 3/3, 417-457.

Seager, W. (1993). The elimination of experience. *Philosophy and Phenomenological Research*, 53/2, 345-65.

Sellars, W. (1956/1963). Empiricism and the Philosophy of Mind. In Sellars, W. (ed.): *Science, Perception and Reality*. London: Routledge. 127-196.

Sellars, W. (1962/1963). Philosophy and the Scientific Image of Man. In Sellars, W. (ed.): *Science, Perception and Reality*. London: Routledge. 1-40.

Sellars, W. (1965). The Identity Approach to the Mind-Body Problem. *Review of Metaphysics*, 18, 430-451.

Shapiro, L. A., Polger, Th. W. (2012). Identity, variability, and multiple realization in the special sciences. In Hill, Ch., Simone, G. (eds.), *New Perspectives on Type Identity: The Mental and the Physical*. New York: Cambridge University Press. 264-285.

Shepard, R. N. (1995). What is an agent that it experiences P-consciousness? And what is P-consciousness that it moves an agent? *Behavioral and Brain Sciences*, 18/2, 267-268.

Shoemaker, S. (1975). Functionalism and Qualia. *Philosophical Studies*, 27, 291-315.

Singer, E. A. (1911). Mind as an Observable Object. *Journal of Philosophy, Psychology and Scientific Methods* 8/7, 180-186.

Smart, J. J. C. (1959). Sensations and Brain Processes. *Philosophical Review*, 68, 141-156.

Smith, D. W. (1988). *Rey Cogitans: The Unquestionability of Consciousness*. In H. R. Otto, J. A. Tuedio (eds.), *Perspectives on Mind*. Dordrecht: D. Reidel. 25-34.

Strawson, G. (2018). The Consciousness Deniers. *NYR Daily*,
<https://www.nybooks.com/daily/2018/03/13/the-consciousness-deniers/>

Sytsma, J. (2010). Phenomenal consciousness as scientific phenomenon? A critical investigation of the new science of consciousness. Doktori értekezés, University of Pittsburgh.
<http://d-scholarship.pitt.edu/8358/>

Sytsma, J. (2014). Attributions of Consciousness. *WIREs Cognitive Science*, 5, 635-648. A hivatkozások oldalszámai a következő online változatra vonatkoznak: http://philsci-archive.pitt.edu/11467/1/Attributions_of_Consciousness_PREPRINT.pdf

Tanney, J. (2016). Ryle and the problem of consciousness. In Leach, S., Tartaglia, J. (eds.), *Consciousness and the Great Philosophers: what would they have said about our mind-body problem?* London: Routledge. 217-226.

Tartaglia, J. (2007). *Routledge Philosophy Guidebook to Rorty and the Mirror of Nature*. London: Routledge.

Tartaglia, J. (2013). Conceptualizing Physical Consciousness. *Philosophical Psychology*, 26/6, 817-838. A hivatkozások oldalszámai a következő online változatra vonatkoznak: <https://jamestartaglia.com/doc/A10.pdf>

Tartaglia, J. (2020). Rorty's Philosophy of Consciousness. in A. Malachowski (ed.) *A Companion to Rorty*. Oxford: Wiley-Blackwell. 43-58. A hivatkozások oldalszámai a következő online változatra vonatkoznak: <https://jamestartaglia.com/doc/A2.pdf>

Tőzsér J. (2008). Általános bevezetés: a test-lélek probléma. In Ambrus G., Demeter T., Forrai G., Tőzsér J. (szerk). *Elmefilozófia*. Budapest: L'Harmattan. 9-86.

Tye, M. (1995). Blindsight, orgasm, and representational overlap. *Behavioral and Brain Sciences* 18/2. 268-269

Tye, M. (2008). *Consciousness Revisited: Materialism Without Phenomenal Concepts*. Cambridge, MA: MIT Press.

Van Gulick, R. (1992). Time for more alternatives. *Behavioral and Brain Sciences* 15/2, 228-229.

Wittgenstein, L. (1953/1998). *Filozófiai vizsgálódások*. Fordította Neumer Katalin. Budapest: Atlantisz Kiadó.