

Az elme pragmatistája

Dennett és a filozófia

Bevezetés

Jelen dolgozat tárgya Daniel C. Dennett, amerikai filozófus rendkívül szerteágazó munkásságának vizsgálata filozófiai szempontból. A vizsgálódás során tárgyalni szeretném Dennett filozófiatörténeti elődeit, hatásait, és el szeretném őt helyezni hozzávetőlegesen a kortársai által kirajzolt filozófiai térképen. Ezen kívül munkásságának két aspektusára kívánok különös tekintettel lenni: egyrészt arra, ahogy a neo-pragmatistának nevezhető filozófiai hagyományba illeszkedik gondolkodása, másrészt arra a döntő hatásra, amelyet Wittgenstein kései korszaka gyakorolt rá. Ezen két témakör vizsgálata azért is ígérkezik gyümölcsöző területnek, mert így a neo-pragmatizmus és Wittgenstein viszonyáról is messzebbmenő következtetések vonhatók le. Dennett gondolkodásának vizsgálata tehát azzal kecsegtet, hogy miközben áttekinthetők rajta keresztül az elmúlt fél évszázad analitikus filozófiájának legfontosabb belátásai, körvonalazható az a hatás is, melyet a múlt század talán legnagyobb hatású filozófusa gyakorol még mindig a kortárs gondolkodásra. Nem titkolt szándékom tehát olyan gondolkodó képét festeni Dennetről, aki elődeitől és tanáraitól a lehető legtöbb hasznosítható megfontolást magába olvasztotta, hogy egyéni karakterét kialakítsa.

Ezen kívül természetesen gondolkodásának azokra az egyedi jegyeire is fel szeretném hívni a figyelmet, melyek segítségével szinte minden filozófus kortársától jól megragadható pontokon eltér. Dennett gondolkodása ugyanis rendkívül polemikus természetű: kritikus meglátásai igen széles rétegét érintik az elmefilozófia, a mesterséges intelligenciakutatás és az evolúcióbiológia legkülönbébb képviselőinek. Ami mégis áttekinthetővé teszi munkásságát, az véleményem szerint az, hogy a sokféle és szerteágazó munkásság alapját egy viszonylag egyszerű alapintenció vezeti, melynek magva az *intencionális alapállás* („*the intentional stance*”) stratégiája. Az intencionális alapállás nagyon röviden összefoglalható lényege az, hogy a világban végbemenő folyamatok értelmezésénél sok esetben egyszerűen hatékonyabbak lehetünk mind a folyamatok megértése, mind előrejelzése során, ha az adott esemény szereplőinek intencionális állapotokat, azaz hiteket, vágyakat és vélekedéseket tulajdonítunk. Dennett stratégiájának kulcsa az, hogy ezt a bizonyos szempontból antropomorf nézőpontot nem kizárólag az emberi tényező teszi indokolttá, hanem például a bonyolultsági szint. Például (ez Dennett kedvenc példája) egy sakkozókompputerrel folytatott

játék során nem a gép fizikai jellemzőire vagyok tekintettel (*fizikai alapállás - physical stance*) és nem is pusztán a gép tervezettségi jellemzőire, és komputációs folyamatainak algoritmusaira (*ez a tervezettségi alapállás-design stance*), hanem a gép bejósolásait próbálom kijátszani: úgy gondolom, hogy a gép nyerni *akar*, és úgy *véli* a következő lépésével előnyt szerezhet, mert nem *tudom*, hogy mi a *szándéka*. Ennek aztán számtalan filozófiai következménye lesz az eredeti intencionalitás doktrínájának (Searle) elvetésétől egészen az evolúció értelmezéséig, ami azonban mindezt Dennett alapintuíciójává teszi, az az a sajátos kép, mely ily módon az emberi racionalitás státuszáról, vagy még inkább az ember helyéről a világban ezáltal kirajzolódik. Nem azt mondja ugyanis Dennett, hogy az ember végső soron gép, hanem fordítva: bizonyos bonyolultsági szint felett gépek is intencionális rendszerekként kezelhetők leghatékonyabban. Azt sem állítja, hogy az emberi racionalitás teljesen azonos szinten lenne az állatvilággal: hisz az intencionalitás szintjei jól meghatározhatóak: az „azt hiszi, hogy” például első szintű intencionalitást, az „azt hiszi, hogy azt hiszi” pedig másodrendű intencionalitást feltételez, a sort tetszőlegesen bővíthetjük, és meghatározhatjuk kísérletekkel, mely állatfajok milyen szintű intencionalitást használnak vagy mennyire lennének képesek. Dennett filozófiájának helyét tehát az intencionális alapálláson keresztül lehet a legjobban megközelíteni. Hisz elmefilozófiai szempontból ez indokolja a népi pszichológia tekintetében fikcionalizmusát és instrumentalizmusát, a kválék létezése felől anti-realizmusát, a Putnam-féle Ikerföld-probléma megoldását „siker-szemantikájának” segítségével, behaviorizmus- és eliminativizmus-krikáját, illetve naturalizmusának jellegét. Tágabb összefüggéseiben pedig az intencionális alapállás státuszának vizsgálata elvezet a Darwinizmus, mint átfogó világkép szerepének filozófiai konzekvenciáihoz is.

Elődök, mesterek és módszerek

(Quine, Ryle, Sellars)

Történetileg szemlélve Dennett gondolkodását, azt mondhatjuk, hogy a Quine utáni amerikai analitikus filozófia pragmatista fordulatának egyenes ági leszármazottja. A klasszikus logikai pozitivizmus Amerikában főként Carnap és Reichenbach nyomán vált népszerűvé, akiknek legfőbb törekvése a filozófia számára olyan szabatos eszköztár biztosítása volt, mely éles határt von tudományos és logikai igazságok, értelmes és értelmetlen mondatok között. A jelentések vizsgálata ebben a felfogásban a verifikációs kritériumok meghatározásával volt

egyenlő, vagyis annak megadásával, mely körülmények közt igaz az adott kijelentés (ha nem adható meg, akkor értelmetlen a kijelentés). Az elemi kijelentések jelentése tehát önmagukban, izoláltan meghatározható Carnap szerint, az igazságfeltételeken keresztül. A logikai igazságok ellenben analitikus igazságok: jelentéseik pusztán az őket alkotó szavak jelentései által igazak.

Ez a két pont, vagyis az analitikus/szintetikus megkülönböztetés, valamint a redukcionizmus Carnap-féle felfogása volt a célpontja Quine 1951-es nagy hatású cikkének, *Az empirizmus két dogmája* című írásának¹. Ha meg akarjuk határozni, mely mondatok analitikusak, definícióink egymásra fognak támaszkodni és végül körbenforgáshoz jutunk. A redukcionizmus esetében pedig be kell látnunk Quine szerint, hogy állításaink nem egyesével, hanem egymáshoz kapcsolódva állnak a tapasztalat ítélőszéke előtt, az empirizmus szempontjából fontos egység a tudomány maga, vagy ahogy későbbi írásaiban finomítja holizmusát: az adott elmélet. A jelentés holisztikus természetének felfogását látjuk tehát a tudományos gyakorlatra alkalmazva. (Ismertebb nevén ez a Duhem-Quine tézis.) A holizmus következménye az elmélet tapasztalat általi aluldetermináltsága: elvben akármilyen elméletet összhangba hozhatunk a tapasztalattal, ha megfelelően módosítunk rajta. A megfigyelési mondatok pusztán abban különböznek a logikai vagy akár metafizikai állításoktól, hogy közelebb vannak a tapasztalathoz, vagyis a tudomány egységes testének a peremét alkotják. Ez azt is jelenti, hogy a logikai állítások sem a priori igazak, például az elméleti fizika, vagy a kvantummechanika fejleményei nem egy esetben változásokat indukálnak a matematikában és a matematikai logikában. Így az analitikus/szintetikus állítások különbségének eltörlése egyben a filozófia és tudomány közti határ eltűnését is jelenti: „*E nézet szerint az ontológiai problémák egy szinten vannak a természettudomány problémáival.*”² Ez Quine pragmatista empirizmusának vázlatos képe, melyet kiválóan szemléltet az a (Neurathtól kölcsönzött) metaforája, melyben a tudós és a filozófus egyazon hajóban eveznek, és a hajót úgy kell újjáépíteniük, hogy az nem süllyedhet el.

Mit jelent ezek alapján quine-i szellemben filozofálni? Vagyis mit köszönhet a Quine-tanítvány Dennett mesterének? Dennett gondolkodása egészen alapvonalaiig tekinthető a fenti értelemben „pragmatistának”: eredendően filozófiai jellegű kérdéseire nem szégyell a tudomány térfelén választ keresni és fordítva: filozófiai belátásait alkalmazhatónak véli a tudományok oldalán. Előbbire példák az elme természetére vonatkozó klasszikus filozófiai

¹ Quine, W. V.: Az empirizmus két dogmája. In: Forrai Gábor – Szegedi Péter (szerk.), *Tudományfilozófia: Szöveggyűjtemény*. Budapest: Áron Kiadó, 1999. 131-153. p.

² id. mű. 152. p..

kérdések: Mi az elme? Mi a tudatosság? Mi az emberi elme sajátossága? Az utóbbira pedig jó példák az etológia adaptacionista fordulatát érintő észrevételei, melyek közvetlenül az intencionális alapállás kiterjesztésének tekinthetők, de ezen kívül a mesterségesintelligencia-kutatás terén tett együttműködései is figyelemreméltók. Dennett gondolkodásának kerete tehát a tudomány, de úgy, hogy eredendően filozófusnak tekinti magát. Tudomány alatt Dennett esetében mindenekelőtt darwinizmust kell értenünk: az evolúció az, amely ezt a sokrétű gondolkodást egységben tartja, ebbe ágyazódik bele minden az elme filozófiájától egészen a keresztény világkép kritikájáig.

Dennett azonban nem csak ebben az általánosabb értelemben követője Quine-nak, nem csak „alapvonalaiban” beteljesítője víziójának, hanem egy jóval technikaibb, filozófiailag pedig szintén rendkívül fontos meglátását is kamatoztatni tudja gondolkodásában. A szintén nagy jelentőségű *Arról, hogy mi van*³ című tanulmányról van, szó, melyben Quine visszautasít néhány olyan érvet, melyek segítségével ontológiai elköteleződések bizonyíthatók rá valakire. Amikor egy olyan elképzelést adunk elő valakinek, melyben egy olyan dolog létét tagadjuk, melyet az ellenfél állít, gyakran adódhatna következő ellentmondásos következtetés: állítást teszünk valamiről, aminek a létét nem ismerjük el. Az eszközt, amivel a problémát kezelhetőnek tekinti Quine, Russelltől veszi. A szinguláris terminust felhasználó kijelentés bármikor helyettesíthető leírással, ekkor pedig a következő helyzet áll elő: ismerjük a kifejezés jelentését és tagadhatjuk létezését. Ugyanis nem kell feltételeznünk azt, hogy jelentésnek megfelel egy tárgy, létezni a szállóigévé vált kifejezés értelmében annyi, mint egy „kötött változó értékének lenni”. Dennett szempontjából ez egy döntő jelentőségű pont, ugyanis az intencionális alapállás egyik legkényesebbnek tűnő kérdése az, hogy vajon léteznek-e intencionális állapotok, el kell-e ismernünk mentális tárgyak létét, elköteleződünk-e ontológiailag nem-fizikai létezők mellett? Dennett többféleképpen megfogalmazta álláspontját, a legtalálóbbs talán az az állítása, miszerint ő realista kis r-rel, de nem Realista nagy R-rel. Ezzel Jerry Fodor definíciókísérletére válaszol⁴: Fodor szerint Realistának azt nevezhetjük a mentális tartalmakkal kapcsolatban, aki elismeri, hogy léteznek mentális állapotok, melyek viselkedést okozhatnak a mindennapi vágy-vélekedés pszichológiának megfelelően, és hogy ezek szemantikai szempontból értékelhetőek. Dennett kis betűs realizmusa nem ilyen: ő nem gondolja, hogy a hétköznapi vágy-vélekedés pszichológia meghatározhatná a mentális tárgyak funkcionális szerepét. Az intencionális alapállást felvéve

³ Quine, W. V.: *Arról, hogy mi van*, in: *A tapasztalattól a tudományig*, Osiris 2002, 115-137. p.

⁴ Ld: Fodor, Jerry: „Fodor kalauza a mentális reprezentációhoz: az intelligens nagynéni segédlete” in, *Kognitív tudomány*, Osiris 1996, Pléh Csaba szerk.

úgy járunk el, hogy az adott ágensnek olyan hiteket-vágyakat tulajdonítunk, melyek leginkább közelítenek egy normatív racionalitásfogalomhoz. Ez tehát egy beszédmód, a valóságnak egy olyan leírása, mely normatív, logikai előrejelzését teszi lehetővé az intencionálisan felfogott rendszernek, anélkül azonban, hogy a rendszer felépítésére, alkotóelemeire vonatkozó akármilyen elméletet igazolna vagy cáfolna. A hitek és vágyak tehát ebből a szempontból létező dolgok, mint absztraktumok, melyek racionalitást rendelnek az adott ágenshez. A Reichenbachtól származó megkülönböztetés értelmében tehát az intencionális alapállás *abstrakta* tételezi a hiteket és vágyakat, mint az elmélet által konstruált entitásokat, én nem *illata*, eleve adott, leírandó tárgyakként. Így fogalmaz: „*Kezdetől fogva szembeállítottam egymással az agyakra s azok különböző neurofiziológiai részeire, állapotaira s folyamataira vonatkozó realizmusomat, valamint a hiedelemállapotokra vonatkozó instrumentalizmusomat, (...)*.”⁵ Ezzel Dennett arra is utal, hogy nem menekül az ontológiai viták elől, melyek a mentális tárgyak létezésével kapcsolatosak, instrumentalizmusa soha nem volt annyira általános, mint azoké, akik például az elektronokat vagy a sejteket is *abstrakta* akarták felfogni. Tagadja a realizmus egyik értelmét, de kidolgozza a sajátját. Úgy gondolom, tekinthetjük Quine említett tanulmányának hozadékának azt a módot, ahogy Dennett ugyanazokat a fogalmakat, melyeket mások használnak, újradefiniálja és teljesen más módon kezeli egy másik elméletben.

Még mindig a Quine-hatásnál maradva, de már attól el is távolodva, nem hagyhatjuk említés nélkül azt a szintén mestere örökségeként felfogható jegyet Dennett gondolkodásában, melyet az előbbieken részben már érintettünk: a harmadik személyű illetékesség szerepét. Az előbb említett Fodor-kritikában is lényegi szerepe van annak, hogy nem fogadhatjuk el a népi-pszichológia „ontológiáját” kutatási keretként : „(...) úgy vélem, hogy a viselkedést okozó *tényleges belső állapotok funkcionálisan nem úgy lesznek egyedíthetők, mint ahogy a vélekedés-vágy pszichológia felszabdalja a világot*”⁶. A bizalmatlanság, ami Dennettet jellemzi a hétköznapi „népi pszichológia” általánosításai iránt, már Quine-nál is megvolt, bár sokkal szélsőségesebb formában. A radikális fordítás gondolat kísérletében⁷ a bennszülött törzset vizsgáló kutató bizonyos esetekben azzal szembesül, hogy egymástól eltérő szótárak is megfeleltethetők a valóságnak: a nyúlra mutató és a „gavagai” szót kiejtő bennszülött kifejezése egyaránt jelenthet „nyulat”, „leválasztatlan nyúlrészt” vagy „a nyulakat”. Ez azt jelenti, hogy a fordító nem csak a szó jelentésének megfelelőjét nem képes megtalálni, de azt

⁵ Dennett, Daniel: Újfent az instrumentalizmusról In: Az intencionalitás filozófiája, Osiris, 1998. 82. p. Pléh Csaba ford.

⁶ Id. mű 81. p.

⁷ Quine, W. V.: Ontológiai relativitás, in: A tapasztalattól a tudományig, Osiris 2002, 137-177.p.

sem tudja eldönteni, mely szónak keresi a megfelelőjét: a referencia kifürkészhetetlenségéről van szó. Quine azt a következtetést vonja le, hogy lehetetlen egy elméleten belül szabatosan meghatározni, mely tárgyra referál az elmélet, nem lehet teljesen interpretált. Minél távolabb kerülünk a megfigyelési mondatoktól, az említett kifürkészhetetlenség annál nagyobb lesz. Az egyetlen megoldás az, ha feltételezünk egy háttérelméletet (vagy a gondolatkísérletnél maradva háttérnyelvet), melynek ontológiáját úgy határozzuk meg, hogy az tartalmazza mindkét elmélet tárgyait (mindkét nyelv referenciáit). Ez nem a két elmélet rejtett ontológiájának felszínre hozása és összebékítése, hanem egy teljesen független ontológia megalkotása. Quine-nál ez a metodológia fizikalizmusát támogatja: úgy gondolja, hogy a mentális kifejezések referenciái a megfigyelési mondatok még viszonylag könnyen kifürkészhető referenciáihoz képest annyira képlékenyek, hogy érdemes őket redukálnunk, amennyiben a tudományos pontosság a cél. A mentális kifejezések redukálásának útja Quine szerint a behaviorizmus, tehát azok megfeleltetése viselkedési diszpozícióknak. A mentális állapotok ugyanis nem magyarázhatók meg viselkedési jellemzőkkel, csak neurológiaikkal. Viszont arra jók a viselkedési jellegzetességek, hogy rögzítsék a referenciát: olyan dolgot neveznek meg, melynek objektíve megadható fizikai sajátosságai vannak, úgymond elhelyezik a fizikalista koordinátarendszerben. Dennett ennek a behaviorisztikus redukciónak nagy ellenzője, pontosabban az intencionális alapállás stratégiájának egyik alapintuíciója, hogy felszámolja a mentalista kifejezések objektíválhatatlanságának doktrínáját, és olyan objektíve megadható keretrendszert ajánl, mellyel racionálisnak tekintett ágensek magatartását hatékonyabban bejósolhatjuk, illetve jobban megérthetjük, mint a másik két alapállást igénybe véve; a fizikai és a tervezettség alapállást. Ennek ellenére úgy gondolom az imént leírt quine-i megoldás, mely azt célozta, hogy a köznyelvet és a mentalista kifejezéseket tudományosan objektív tényekre vezesse vissza, Dennettnél is visszaköszön. Csakhogy Dennett úgy gondolja, teljesen felesleges viszonylag jól működő modelleket „fundamentálisabbakra” visszavezetnünk, amikor a meglévőt is felhasználhatjuk, ha meghatározzuk azokat a minimális keretfeltételeket, melyek mellett az aránylag jól működik. Azt állítom, hogy az a mód, ahogy Dennett a népi pszichológia Fodor-féle felfogását kezeli⁸, nagy vonalaiban azonos azzal, ahogy Quine kezelte a mentális kifejezések nem-behaviorista elemzését, és hasonló indíttatású is. Hiszen mindkét esetben arról a problémáról van szó, hogy a referencia ingatagsága miatt ugyanarról a belső vágy-vélekedés-mechanizmusról tetszőleges számú alternatív leírás lehetséges, ez pedig lehetetlenné teszi az

⁸ Id ehhez: Valódi mintázatok, in IF 233. p, ahol explicit formában is Quine-ra támaszkodik Fodor „Nagyipari Realizmuával” szemben.

objektivitást. Dennett megoldása az, hogy *kívülről* rendel az intencionális rendszerhez vágyakat és vélekedéseket, olyanokat, melyek a feltételezett racionalitáshoz közelítenek. A racionalitás azért nem vezethető le empirikus generalizációkból, mert *normatív* fogalom. Később még látni fogjuk, hogy Dennett normatív racionalitásfogalma, melyet az intencionális rendszerekhez rendel, alapjaiban az evolúciós világnézet talajából sarjad. Most azonban, a történeti hatás, és a quine-i örökség szempontjából fontosabb azt látni, hogy a kisebb, eltérő és bizonytalanabb elméletek lecserélésének igényére mindketten a kívülről meghatározott *háttérelméletben* látják a megoldást. A különbség annyi, hogy míg Quine a viselkedések objektív fizikai létének tulajdonít kizárólag tudományos relevanciát, addig Dennett meghatározhatónak tartja mentális kifejezések használatát egy szigorú, harmadik személyű perspektívából. Dennett tehát Quine-nal tart az objektivitás kritériumának igényét illetően, miközben elveti mentalizmus-kritikáját.

A kérdés, ami ezután felvetődik, hogy mi alapján gondolhatjuk azt, hogy Dennett intencionális alapállása elkerüli azt, ami az említett rivális elképzelés hibája, ti. mi okunk lehet azt gondolni, hogy a vágyak, vélekedések és hitek tulajdonításának dennett-i módja nem egyenrangú a népi pszichológiát Realista módon kezelő fodor-i módszerrel, és miért nem esik az is áldozatul a referencia kifürkészhetetlenségét illető aggályoknak? Többféle válasz adódik, maga Dennett sem ad egyenes választ, érvelései jóformán a működő elmélet előnyeire világítanak rá a riválisokkal szemben, semmint azok hibájára. A kérdésre adódó legkézenfekvőbb válaszként talán azt lehetne kiemelni, hogy mivel az intencionális alapállás úgy működik, hogy van egy előfeltevésünk arról, hogy a leírni kívánt ágens racionalis, a bejósolás során arra vagyunk tekintettel, hogy olyan rendszerrel van dolgunk, mely eleve olyan hitekkel, vágyakkal *kell, hogy rendelkezzen*, melyek számára a környezetét kezelhetővé teszik. Úgymond evolúciós garanciánk van Dennett szerint arra, hogy a szóban forgó ágens mentális reprezentációi és a hozzájuk kapcsolódó intencionális állapotok nagyrészt „működőképesek”. A racionalitás kerete az alapunk akkor is, amikor a leírni kívánt ágensünk működésképtelen, például amikor egy elmebetegről van szó, hisz az irracionalitás is úgymond a racionalitáson élösködik, a diszfunkciót is, mint korlátozott, rosszul működő funkció érthető meg a legegyszerűbben. Ugyanaz működik tehát ilyenkor, mint akkor, amikor egy autó hibásodik meg: arról beszélünk hogyan kéne funkcionálni az adott egységnek és ehhez képest mit tapasztalunk mint diszfunkciót. Nem a gép legelemibb részecskéinek működésével foglalkozunk tehát, mert ebben az esetben egyszerűen nem szolgál olyan hatékony információval. Ez persze nem jelenti azt, hogy az elvi lehetőségét zárjuk ki annak, hogy ilyen leírás segítségével is megoldható az adott helyzet, hisz elképzelhetünk olyan szuper-

fizikusokat (Dennett példája), akik a tőzsdén a tintafoltok hipergyors vizsgálata által képesek előrejelezni az árfolyamokat. A valóságban azonban mindennek az esélye rendkívül kicsi. Látszólag intuitív évről van tehát szó, közelebből azonban látszik, hogy gyakorlatorientált, pragmatikusnak nevezhető attitűdként tekinthetünk leginkább az intencionális alapállásra, semmint minden oldalról megtámogatott, támadhatatlan érvrendszerre. A fenti kérdés elméleti szempontból nagyon kényes pontja az elképzelésnek, ezt maga Dennett is elismeri: „*Az intencionális stratégia által egyformán szavatolt, mégis radikálisan különböző értelmezések mumusa elméletileg fontos – talán inkább úgy mondhatnánk, metafizikailag fontos -, de gyakorlatilag eltekinthetünk tőle, ha figyelmünket a legnagyobb és legkomplexebb intencionális rendszerre korlátozzuk: az emberre.*”⁹ A metafizikai probléma tehát csak addig kell, hogy zavarjon, amíg ezen a szinten mozgunk, a gyakorlat világára tekintve már intuitíve, a valószínűsége és a sikerre alapozva is dönthetünk az intencionális alapállás mellett, pontosabban; már döntünk is eleve, evolúciósan így vagyunk programozva. Filozófiai szempontból talán ez a leggyengébb pontja az egész dennetti gondolkodásnak: igaz, hogy az érvelése nem körbenforgó, hisz a kiindulópont az evolúció által szavatolt adaptív képességünk, a racionalitás feltételezése, az alapállás mellett szóló érvek pedig gyakorlatiak, és az elmélet sikerére támaszkodnak, mégsem szavatol semmi afelől, hogy ez a keret elvileg is a legjobb. Feltételezhetjük elméletileg akár azt is, hogy egy az evolúciós adaptivitásnál jobb keretbe illeszkedve plauzibilisebb racionalitásfogalmat kapunk, melynek a prediktív ereje is jóval nagyobb. Ugyanis Dennett szemében az intencionális alapállás legnagyobb ereje az, hogy a leghatékonyabb előrejelzést teszi lehetővé, amikor racionális ágenseket feltételezünk. Kristin Andrews „*Our understanding of other minds*”¹⁰ című cikkében épp egy ilyen elképzelés körvonalait látjuk Dennett intencionális alapállásával szembeállítva. Andrews véleménye az, hogy amennyiben az intencionális alapállás stratégiáját követjük, akkor adódik egy olyan probléma, hogy például amikor majmoknak vélekedéseket és vágyakat tulajdonítunk, miközben az egymással folytatott kommunikációt írjuk le, akkor a majmoknak másodrendű intencionalitást tulajdonítunk, ami azt vonja maga után, hogy a majmoknak elméletiük van (pl: „az egyik majom azt gondolja, hogy a másik azt hiszi”). Ezt az implauzibilisnek tűnő magyarázatot könnyedén elkerülhetjük, ha nem az intencionális alapállást vesszük igénybe, hanem úgy gondoljuk, hogy amikor az egyik állat meghall egy vészjelzést társa szájából, akkor egyszerűen *induktív* általánosítást végez, arra következtet,

⁹ Dennett, D: Az igazhívők: az intencionális stratégia és sikerének forrásai, in: Az intencionalitás filozófiája, 288.

¹⁰ *Journal of Consciousness Studies*, 7, No. 7, 2000, pp. 12-24
<http://www.yorku.ca/andrewsk/andrews/OTHER%20MINDS%20-%20OCR.pdf>

hogy akárhányszor ezt a hangot hallja, veszély áll fenn, illetve amikor ezt a hangot kiadja, a többiek rendre biztonságba vonulnak. Az *induktív alapállást* felvéve tehát könnyen leírhatunk állatok közti kommunikációs interakciót anélkül, hogy elméletiát tulajdonítanánk nekik. Ezen kívül- és ez már közvetlenül a fenti lehetséges ellenvetés egyik kidolgozott példájává teszi az elméletet- az induktív alapállással az emberi gondolkodást is modellezhetjük. Például vehetünk egy esetet, mely szerint egy illető hozzászokott ahhoz, hogy egy adott országban, mikor kávézóba megy a pultnál kéri az italt, majd azzal nyugodtan helyet foglalhat bárhol az asztaloknál. Egy másik országban viszont emberünk meglepődve tapasztalja majd, hogy fizetés után a pincérnő megkéri, hogy legyen szíves kávéját a pultnál elfogyasztani. Teszi ezt azért, mert az asztaloknál való fogyasztás esetén a kávé drágább, mint a pultnál való fogyasztás esetén. Emberünk azonban ezt nem tudja, és arra következtet, hogy legközelebb, miután bemegy a kávézóba, egyből az asztalokhoz kell leülnie, ha ott akar fogyasztani. Következtetésének igazsága beigazolódik, hisz valóban: a következő alkalommal már asztalhoz viszik neki az italt. Előrejelezte tehát sikeresen az eseményt, melyben intencionális rendszerek szerepelnek, anélkül, hogy bármilyen feltevést megfogalmazott volna a pincérnő hiteiről vagy szándékairól. Egyszerűen azt a következtetést vont le, hogy „Errefelé ez a szokás.”. Semmi nem szól tehát amellet, hogy az intencionális alapállás a leghatékonyabb eljárás vagy az egyetlen járható út, amennyiben racionálisnak tételezett ágensek viselkedésének leírása a cél.

Ha visszakanyarodunk a fejezetünk szűkebb témájához, a történeti hatások és tanárok kérdésköréhez, akkor magyarázatot kaphatunk arra a kérdésre, hogy milyen megfontolások alapján alakíthatta ki Dennett az intencionális alapállás normatív doktrínáját, és miért tulajdonítja neki azt a szerepet, amit tulajdonít. Dennett egy másik nagyon jelentős tanára Gilbert Ryle volt, akinek logikai behaviorizmusa nagyon közeli rokona Quine elképzelésének. Mindkettejükre Skinner pszichológiai behaviorizmusa hatott. Ryle szerint le kell számolnunk azzal a „hivatalos doktrínával”, mely szerint az elme egy különleges hely, úgymond szellem a gépben, melynek működése egészen más szabályok alapján írható le, mint a fizikai világ. Le kell mondanunk Descartes mítoszáról, arról, hogy elménkhez első személyű hozzáférésünk van, saját elménk ismerete tekintetében tévedhetetlenek vagyunk, míg a külvilág tényei mások számára is ugyanúgy érhetők el. A kulcs Ryle szerint abban van, hogy meg kell különböztetnünk a tudás két típusát, a „tudni hogyan” a „tudni, hogy mi-től”, nem pedig abban, hogy más természetű dolgokként tekintünk az elmére és a „kiterjedt valóságra” (a *res cogitans*ra és a *res extensa*sára). A mentális állapotok egytől egyig leírhatók diszpozíciókként, annak a fogalmaival, hogy az adott személy bizonyos körülmények között mire lenne képes.

Így például aki tud sakkozni, az nem egy explicit szabálygyűjteményt tárol valahol az elméjében, hanem ha leül valakivel szemközt egy sakktábla mellé, akkor tudja, milyen lépéseket tehet, ha nyerni akar. Saját mentális képességeink leírása esetén is arról van szó, hogy harmadik személyű, objektív perspektívából tekintve megítéljük, adott helyzetben mire lennénk képesek, nem szükséges feltételeznünk az elsőszemélyű hozzáférés kitüntettségét még ekkor sem. Ezzel elkerüljük, hogy a tudás esetén például valami olyannak a jellegzetességeit írjuk le, ami a „fejben van”: „A „tud” és a „nem tud” pedig nem eseményeket jelölő, hanem modális szavak.”¹¹ Dennett a műhöz írt előszavában explicit formában is megfogalmazza, hogy az intencionális alapállás létrejötte egyfajta reakció volt Ryle koncepciójára¹². Ez a reakció pedig nagyon közeli Wilfrid Sellars, amerikai filozófusnak az *Empiricism and the philosophy of mind* című művében körvonalazotthoz. Sellars művének konklúziója röviden az, hogy a mentális folyamatokról adott leírásaink alkalmával nem egy introspekcio segítségével elérhető tartomány adott rendjéhez férünk hozzá („the myth of the given”), hanem egy elméletet használunk arról, hogy mi történik az elmében, ez az „elmélet-elmélet”. Ennek megvilágítására Sellars előad egy gondolat kísérletet, melynek szereplőit a mi „ryle-i elődeinknek” nevezi¹³. Ezek az elődök egy olyan egyszerűsített nyelvet beszéltek melyben kizárólag a külvilágban szemmel látható objektumokra tudtak referálni, tárgyakra és egymás viselkedésére. A hatalmas változás ezen őseink fejlődésében akkor állt be, amikor egyszer csak valaki előállt egy elképzeléssel, miszerint létezik egy folyamat, mely szintén az időben zajlik, és akkor is működik, amikor éppen nem kimondva, nyilvánosan szavakban gondolkodunk hangosan, hanem akkor is, amikor semmi létező jele nincs annak, hogy ez a folyamat, mely a gondolkodás, végbemegy valakiben. Ráadásul az elmélet része az is, hogy a kimondott szavak csak ezen folyamat betetőzései. Az elképzelés hozadéka, hogy ezen fiktív elődeink immár rendelkeznek egy eszközzel, mellyel modellezhetnek olyan folyamatokat, referálhatnak olyan műveletekre, mint a gondolkodás, illetve egy idő után ezt az eljárást magukra is alkalmazni fogják, és képessé válnak rá, hogy beszámoljanak saját gondolkodásukról. Az első személyű illetékesség ezért nem garantáltan helyes, ahogy a mások mentális állapotaira vonatkozó leírásaink sem mindig működnek. Nem gondolhatjuk tehát azt, hogy amikor valaki az „azt hittem, hogy esik az eső” állítást mondja, akkor saját propozicionális attitűdje felől biztosan igaz állítást tett, mert előfordulhat, hogy rosszul tudja. Amikor gondolatokat tulajdonítunk másoknak és magunknak, akkor ugyanúgy járunk el, mint

¹¹ Ryle, Gilbert: A szellem fogalma, Osiris 1999. 59. p.

¹³ Ehhez ld. Sellars, Wilfrid: Empiricism and the philosophy of mind, XII: Our Rylean Ancestors.
<http://www.ditext.com/sellars/epm.html>

amikor molekulákat tételezünk fel a testben, hogy jobban leírassuk: az azonban nincs garantálva, hogy jól írjuk le, azaz természetesen használhatjuk rosszul az elméletet. Sellars mítosza tehát anti-mítosz: az ellen a karteziánus mítosz ellen szól, mely szerint saját állapotaink felől tévedhetetlenek vagyunk. (Magyarán a „*cogito ergo sum*” következtetés első fele ugyanolyan hipotetikus állítás, mint az, hogy a kandalló mellett ülök, vagy, hogy ruha van rajtam.). Ezt a megállapítást többféleképpen értelmezhetjük, és annak megfelelően, hogy a népi pszichológia státuszát illetően milyen állásponton vagyunk, egészen eltérő következményei lehetnek. A Churchland házaspár eliminatív materializmusa szerint például úgy kell gondolnunk, hogy a mítosz második felében megjelenő elmélet használhatatlan: vagyis, hogy a népi pszichológia alapvetően rossz elmélet, és ugyanúgy el fog tűnni, ahogy a flogiszonelméletet felváltotta az oxigén felfedezése: ahogy ma már senki nem beszél flogiszonról, úgy hitekről és vágyakról sem fog egy idő után, ahogy a neurofiziológia fejlődése kiszorítja majd ezeket a fogalmakat. Jerry Fodor realista nézőpontja szerint a népi pszichológiából kell kiindulnunk, ha olyan elméletre vágyunk, mely a gondolkodás plauzibilis leírását adja. Dennett ebből a szempontból, a mítosz metaforáinál maradva, azt mondhatnánk, hogy bizalmatlan az elmélettel szemben, de meghatározhatónak tart *az eredetiből kiindulva* egy *sajátot*, ami szerinte működőképes. Pontosabban: koránt sem biztos abban, hogy az elméletet abban a formában a leghelyesebb használnunk, ahogy azt a mítosz szereplői teszik, de elismeri, hogy a normatív módon meghatározott elmélet az intencionális állapotokat illetően alapvetően járható út. Leginkább a racionalitás normatív felfogása teszi tehát alapvetően Sellars-követővé Dennettet, de nem ez az egyetlen lényegi pont. A Sellars utáni funkcionista elméletek legnagyobb előnye a behaviorizmussal szemben az volt, hogy a gondolkodás műveleteinek leírását ígérte, nem akadt meg a viselkedések (és diszpozíciók) leírásánál, vagyis a bemenet-kimenet modellt kiegészíteni kívánta a kettő közötti műveletek formális leírásával is. Dennett ebben az értelemben tehát olyan funkcionista, aki a népi pszichológiát nem tekinti kielégítőnek, ellenben az intencionális alapállással kívülről az ágenshez rendelt műveletek posztulálását sikeres eljárásnak tekinti. A kérdésre tehát, hogy Dennett miért az intencionális alapállást tekinti a leghatékonyabb bővítésnek a tervezeti és a fizikai alapállások mellé, azt válaszolhatjuk, hogy azért, mert a behaviorista modellt, melyet tanárai, Quine és Ryle Skinner nyomán használtak, sellars-i funkcionista szellemben tartja kibővíthetőnek, azaz külső, normatív értelemben a(z intencionális) rendszerhez rendelt elméletként tartja csak elképzelhetőnek a racionalitás feltételezését.

Wittgenstein szelleme

1. Behaviorizmus-kritika és a természetes gondolkodás rehabilitálása

A következőkben Dennett gondolkodását egy olyan szempontból szeretném vizsgálni, melyet általában figyelmen kívül hagynak vagy nagyon felszínesen vagy épp tévesen intéznek el, ez pedig annak a hatásnak a kérdése, hogy (főleg a kései) Wittgenstein milyen döntő hatást gyakorolt Dennetre. Például Tadeusz Zawidzki Dennetről szóló könyvében Wittgensteint mindössze a Ryle-féle logikai behaviorizmus képviselőjeként hozza szóba mint tényezőt Dennett gondolkodásában¹⁴. Ennek véleményem szerint épp az ellenkezője igaz, hiszen Wittgenstein a Filozófiai vizsgálódásokban rendkívül módszeres kritikával illeti a behaviorizmust. Ez pedig Dennett gondolkodására ezáltal lényegibb hatást kell, jelentsen, főleg ha tekintetbe vesszük ezen kívül azt is, hogy Dennett saját bevallása szerint alapvetően példaképként tekintett Wittgensteinre gondolkodói pályájának korai szakaszában. Wittgenstein behaviorizmus-kritikája olyan döntő hatással volt Dennetre, hogy *Consciousness explained* című könyvének *Appendix A (For philosophers)*¹⁵ című részében az őt behaviorizmussal vádolók megjegyzéseire válaszolva egyszerűen bemásolja a Vizsgálódások azon két szöveghelyét, ahol Wittgenstein legrészletesebben elemzi a behaviorizmus hibáját, majd utána elmeséli, hogyan kezdett neki oly módon wittgenstein-i szellemben filozófiával foglalkozni, ahogy az akkoriban nem volt jellemző. (Később részletesebben rá fogok térni, hogy Wittgenstein-hatás szempontjából Dennett a mai napig egy teljesen egyedi úton jár.)

Nézzük tehát kezdésként azt a részt, amely elindította Dennettet a behaviorizmus-ellenesség rögzös útján:

¹⁴ Id. Zawedzki, Tadeusz: Dennett, Oneworld Publications, Oxford 2007, 25. p.

¹⁵ Dennett, Daniel: *Consciousness explained*, BackBay Books, 1991, 462. p.

„307. „Nem vagy te a lelked mélyén mégis behaviorista? Nem azt állítod-e, hogy alapjában véve minden fikció, kivéve az emberi viselkedést?” – Ha fikcióról beszélek, akkor grammatikai fikcióról.

308. Csakhogy hogyan jutunk a lelki folyamatok, állapotok és a behaviorizmus problémájához? – Az első lépés teljesen észrevétlen. Folyamatokról és állapotokról beszélünk, ám eldöntetlenül hagyjuk a természetüket ! Talán egyszer majd többet fogunk tudni róluk – véljük. De éppen ezáltal horgonyoztunk le egy meghatározott szemléletmódnál. Nekünk ugyanis meghatározott fogalmunk van arról, hogy mit jelent egy folyamatot közelebbről megismerni. (A döntő lépés a bűvészműtatványban megtörtént, és épp ez a lépés tűnt ártatlannak számunkra.) _ És íme, széthullik a hasonlat, amelynek gondolatainkat megfoghatóvá kellett volna tennie a számunkra. Tagadnunk kell tehát folyamatot, amelyet még nem is értünk, és amely olyan közegben játszódik le, amelyet még nem fürkészünk ki. És ezért látszik úgy, mintha a szellemi folyamatokat tagadtuk volna. Pedig természetesen nem akarjuk tagadni őket! ¹⁶”

Mi történik tehát Wittgenstein szerint? Az, hogy miközben tudatállapotokról gondolkodunk, akaratlanul is megjelenik előttünk valaminek a képe, ami azonban nem az, amit kerestünk. Magyarán a definiendum természetét látszólag meghatározatlanul hagyva kutatunk a definiens után, ami azt eredményezi, hogy a definiensünk nem definiálja a definiendumot, hanem a végén helyettesíti azt. Kicsit hasonló a helyzet tehát a Menón-paradoxonhoz, ahol, ha tudjuk, mit keresünk, akkor már nem kell keresni, ha nem tudjuk, mit keresünk, akkor pedig nem tudunk elindulni a keresésben. A behaviorizmus esetében az történik, hogy a tudatállapotokról gondolkodva azok képeként beugorhatnak arckifejezések, viselkedések, melyek aztán nem elégítik ki azt az intuíciónkat, hogy itt valóban a folyamatról van szó, ezért tagadni akarunk. Amit tagadnunk kéne, az Wittgenstein szerint az, hogy ezek a képek felelnek meg a kezdeti fogalmunknak, melyet definiálni akarunk. Ehelyett a behaviorista gondolkodás úgy következtet, hogy maga a definiendum meghatározhatatlan, jobb, ha felhagyunk vele, kimondjuk, hogy nem létezik az adott dolog. A fikció tehát, melyet Wittgenstein tagad, az az, hogy a mód, ahogy az adott dolgot a gondolkodásunk megjeleníti, valóban a dolog képe lenne. Mivel a gondolkodás nyelvben történik, „grammatikai fikciónak” nevezi. Az egész Vizsgálódásokon ez a módszer húzódik végig: annak megmutatása a cél, hogy nyelvünk hol

¹⁶ Wittgenstein, Ludwig: Filozófiai vizsgálódások, Atlantisz 1992. (a továbbiakban FV) 154. p.

ejti rabul a gondolkodásunkat, a mód, ahogy a gondolkodásunk reprezentálja a valóságot, hol vezet abszurd következtetésekhez. Wittgenstein szerint a filozófiai problémák ilyen tévedések eredményei, melyből a filozófiai terápia általa javasolt módján gyógyulhatunk ki. Így a logikai behaviorizmus is csak egy betegség, olyan, mint egy neurózis, ahol a beteg elfojt valami feszültséget, melyet máshol ad ki. Vagy még inkább olyan, mintha valakivel például filozófiáról szeretnénk beszélgetni, de az illetőnek lelki szemei előtt egy roppant ellenszenves egykori tanárát kíséző benyomás jelenik meg, majd inkább másik témát javasol a társalgás tárgyául. Számára ugyanis a tanárát kíséző kellemetlenség, az adott óra légköre vagy bármi hasonló az, amely a filozófiára, mint olyanra jellemző. Ha felhívnánk a figyelmét rá, hogy feltehetőleg nem arra gondol, amiről mi beszélünk, akkor úgymond egy gondolati betegségből gyógyítanánk ki emberünket.

Az előző fejezet végén láttuk, hogy milyen választ adott Dennett a behaviorizmusra Sellars szellemében. Hogy tudnánk ezen a ponton kapcsolatba hozni ezt Wittgenstein hatásával? Úgy gondolom, az intencionális alapállásnak az a leginkább wittgenstein-i jellemzője, hogy a döntő ponton a hétköznapi nyelvhasználat egyfajta rehabilitálásával éri el a célt. Hisz vágyakról, hitekről, szándékokról beszélni a mindennapos gondolkodásunk része, és a népi pszichológiával, mentalizmussal kapcsolatos kérdés épp abból a kérdésből fakadt Quine-nál főleg, hogy a hétköznapi nyelvhasználat fogalmai ezen a szinten ingatagok. Ha elfogadjuk azt a státuszt, amit Dennett az intencionális alapállásnak tulajdonít, vagyis kutatási stratégiát látunk benne elsősorban, akkor megkockáztathatjuk, hogy Dennett a viselkedés tudományos vizsgálata terén arra tett kísérletet, amivel Wittgenstein a filozófiai problémák területén próbálkozott. Wittgenstein módszere ugyanis abból áll, hogy megnézi, a természetes nyelvhasználat vizsgálata mely pontokon képes feltárni, hogy filozófiai problémák esetén hol „ment szabadságra a nyelv”. Dennett gondolkodása a rivális tudományos (és részint filozófiai) elméletek esetében szintén azt a stratégiát követi, hogy megvizsgálja először is, milyen eredményeket ér el az adott elmélet, majd a hiányosságok kezelése során a hétköznapi intuíciónkból, pontosabban a mindennapi gondolkodásunkból táplálkozó elméletével kívánja orvosolni a problémát. Nézzünk egy konkrét példát! Az *intencionális pszichológia három fajtája*¹⁷ című írása jellemző példája ennek az érvelésmenetnek, ott ugyanis a tanulmány első fejezetében (*A népi pszichológia, mint elméletforrás*) először is ismerteti, mindennapi gondolkodásunk mely értelemben tekinthető tudományosan is releváns eszköznek („*azt javaslom, hogy mint egy racionalisztikus értelmező és előre jelzőkalkulust tekintjük a naiv*

¹⁷ In: Dennett: Az intencionalitás filozófiája, 39.-78. p.

pszichológiát”), majd a tanulmány végén konkrét példáját adja, hogyan képes az a keret, melybe az intencionális alapállás stratégiája illeszkedik, sikeresen megbirkózni azokkal a problémákkal, melyekkel a korábbi elméletek nem tudtak. A személyi szint alatti kognitív pszichológia például nem képes megmagyarázni, hogy a nyelv szintaktikai vizsgálatát hogyan lehetne összekapcsolni az agy neurofiziológiai vizsgálatával úgy, hogy magyarázatot találjon a szemantikára. Az intencionális alapállás azonban Dennett szerint megoldást kínál erre a kérdésre, azáltal, hogy a redukálendő makroszintet másban határozza meg. A makroszint az evolúció és a túlélés szintje, melyet úgy kapcsolhatunk össze a mikroszinttel, hogy például meghatározzuk mely ösztönökkel kell a túlélés érdekében rendelkeznie az ágensnek, majd azt, hogy azok milyen intencionális állapotokat, vagyis vágyakat, vélekedéseket implikálnak. A szintaxisunk vizsgálata tehát úgy vezethet szemantikához, hogy az evolúciós keretnek elvben garantálnia kell azt, hogy intencionális állapotaink olyan fogalmakat használhassanak, melyekkel biztosítható az egyed túlélése. Vagyis úgy vagyunk tervezve, hogy fogalmainkkal képesek vagyunk a minket övező valósághoz alkalmazkodni.

A wittgenstein-i motivációt ott kell látnunk, hogy abban az esetben, amikor az adott tudományos, filozófiai diskurzusok problémákkal szembesülnek, úgy alkotunk valami újat, hogy megnézzük, hétköznapi, naiv felfogásunkat milyen mértékben tudjuk hasznosítani, hogy a problémák okára és a megoldásukra fény derüljön. Ezt természetesen nem úgy értem, mintha Wittgenstein egy új elméletet hozott volna létre a *Filozófiai vizsgálódásokkal*, hanem úgy, hogy ha Dennett kutatási stratégiának nevezi módszerét, akkor Wittgenstein eljárását is tekinthetjük gondolkodási stratégiának. Ha az utca emberét megkérdeznénk, rendelkezik-e vágyakkal, hitekkal, szándékokkal, nagy esélyünk lenne rá, hogy a többségük nem válaszolna nemmel, és nem mondaná például, hogy ő csak viselkedési diszpozíciókkal vagy idegpályák közti kölcsönös gerjesztettséggel rendelkezik. Amennyiben nem minősítjük ezen hétköznapi elképzelést alapvetően hibásnak, hanem a tudományos elméletünk hiányosságainak orvoslására látjuk benne az esélyt, akkor eljárásunk nagyon hasonló Wittgensteinéhez, aki szintén a „*probléma előtti állapotot*” szeretné tanulmányozni, azt, hogy hol romlott el valami a természetes gondolkodásunk menetében, hogyan keletkezett filozófiai probléma: „*E problémákat nem úgy oldjuk meg, hogy új tapasztalatot vonultatunk fel, hanem azt rendezzük, amit régóta ismerünk.*”¹⁸” Dennett esetében láttuk már, hogy mit jelent ez a „rendezés”: azt, hogy nem tartja elfogadhatónak a népi pszichológia általánosításait, naiv, pre-teoretikus beállítódásunknak bizonyos szegmensét mégis rehabilitálhatónak tartja, ha elméletforrásként

¹⁸ FV 109. pont, 79. p.

tekint rá (a normatív racionalitás-felfogás értelmében). A gondolkodás irányának hasonlóságát látjuk tehát, de nem csak ebben láthatjuk az egyezést, hanem abban is, hogy mindkét esetben egyfajta eszközként tekinthetünk a két módszerre, mely ígéretes eredményeket ígér. Dennett esetében az intencionális alapállás nagy prediktív ereje az elérhető cél, a természetes, pre-filozófiai nyelv wittgensteini vizsgálata pedig azt ígéri, hogy kigyógyulhatunk a filozófiai betegségekből, melyekkel a nyelven keresztül fertőződünk meg. Wittgenstein saját módszerének ezt az újítását nagyon találóan fogalmazza meg a következő idézetben: „„*Vívmányom*” nagyon hasonlatos a matematikuséhoz, aki feltalál egy számítást.¹⁹”

2. A metafora „atmoszférája”, mint filozófiai intuíció és az „eredeti intencionalitás” esete

A kései Wittgenstein gondolkodásának van egy további mozzanata, amely szintén lényeginek tűnik Dennett szempontjából. Ez pedig a természetes nyelv használatával kapcsolatos azon észrevétele, hogy hajlamosak vagyunk a nyelvhasználat bevett és átvitt értelmű használatáról beszélni, a filozófiai gondolkodás során pedig gyakran előfordul, hogy ezen *benyomásainknak* úgymond döntő szerep jut. Wittgenstein egyenesen odáig megy, hogy a filozófiai gondolkodás egyik legfőbb jellegzetességeként tekint erre a módra. Így írja ezt le:

„598. *Amikor filozofálunk, érzéseket hiposztazálunk ott, ahol nincsenek. Az érzések arra szolgálnak, hogy gondolatainkat megmagyarázzák nekünk.*

‘Gondolkodásunk magyarázata itt érzést követel!’ Olyan ez, mintha meggyőződésünk e követelés nyomán jutna el a magyarázathoz.²⁰”

Majd így vonja le a konklúziót:

¹⁹ Wittgenstein, Ludwig: Észrevételek, 75. p.

²⁰ If 228. p.

„599. A filozófiában nem következtetnek. A „De hát ilyennek kell lennie” nem filozófiai állítás. A filozófia csak azt rögzíti, amit mindenki el is ismer neki²¹.”

Arról van ilyenkor szó, hogy a (tradicionalis értelemben vett) filozófus, mielőtt következtetni, érvelni kezdene, olyan érzésekre hagyatkozik, melyek nem valódiak, pusztán a nyelv szavaihoz kapcsolódó berögződések, állandósult, a szavakhoz tapadt intuíciók. Ezeket „érzi” az intuícióra hagyatkozó gondolkodó, mielőtt kimondaná a mondatot, melyből majd következtet, érvel mellette stb. Annyit lehetne ezekhez az alap állításokhoz végső soron hozzátenni, ha le akarnánk írni azt a módot, ahogy a filozófus hagyatkozik ezekre az érzésekre, hogy „ilyennek lennie kell, mert érzem, a szükségszerűség benyomását keltve került elém”. Ezt pedig már nem írják le, ezért ez nem filozófiai állítás. A paradoxon az, hogy épp az az állítás már nem kerül kimondásra, amely a végső érv lehetne a filozófiai állítás mellett, vagyis a *filozófiai intuíció fellépését* megállapító mondat már nem egy filozófiai mondat. Pontosabban az az „érv” már nem egy érv.

Ugyanaz történik, folytatja Wittgenstein a példákat, mint amikor belépünk a szobába rég ismert bútordarabjaink közé. Természetesen ismerjük a környezetet, de nem *ismerünk rá*. Ha valaki megkérdezné, megismerjük-e ezt a helységet vagy a bútordarabokat, hogy azok a mieink-e, igennel válaszolnánk, de azt nem állíthatjuk, hogy mindig ez a folyamat játszódik le, akkor is, amikor úgy lépek a szobába, hogy ezt senki nem kérdezi meg tőlem, és nekem sem jut eszembe semmi ezzel kapcsolatos. (600.-607.) Wittgenstein szerint az illúzió abban áll, hogy amikor leírjuk, hogy ráismerünk valamire, úgy fejezzük ezt ki, mintha két esemény azonosságát állítanánk: „ugyanabba a szobába léptem be, mint amelyikből reggel eljöttem”. Pedig inkább arról van szó, hogy a tárgy, vagy egy tárgy képe hozza létre a saját atmoszféráját maga körül, mintha a „*tárgy beburkolózna egy képpel*.”²²

Ennek a nagy horderejű észrevételnek olyan téma kapcsán lesz érezhető a hatása Dennett gondolkodásában, melyek a kortárs elmefilozófiai diskurzusban jelölik ki egyedi pozícióját, ez pedig az eredeti intencionalitás kérdése. Az intencionalitás kérdésének kapcsán John Searle nevéhez fűződik az *eredeti* és a *származtatott* intencionalitás megkülönböztetése. Searle híres kínai-szoba argumentuma arról szól, hogy amennyiben egy embert egy szobába helyezünk majd kínai írásokkal teli lapokat adunk neki, és egy szabálygyűjteményt, amin az áll, milyen angol írással teli lapokat kell a másik ajtón kiadogatnia, akármilyen bonyolultsági szintet is érjen el az utasításgyűjtemény, soha nem fogjuk tudni azt mondani a szobán kívülről, hogy a

²¹ Uo.

²² FV 605. 229. p

szobában levő, kínaiul nem tudó emberünk érti a kínai jeleket. A gondolatkísérlet lényege, hogy Searle szerint akármilyen fejlettségi szintet is érjen el a mesterséges intelligencia kutatása, soha nem tekinthetünk majd egy gépre tudatosként, még akkor sem, ha megközelítőleg nagyon hasonló műveleteket végez is, mint mi, mert azok csak a belétáplált szintaxisunk felett végzett komputációs műveletek eredményei, a gépnek nincs szemantikája. Így csak átvitt értelemben tulajdoníthatunk egy gépnek tudatosságot, ez pedig nem ugyanaz, mint az emberi intelligencia: csak származtatott intencionalitás, szemben az ember eredeti intencionalitásával. Dennett válasza úgy foglalható össze, hogy mivel a normatív racionalitásfogalmat funkcionális terminusokban határozzuk meg, azaz az intencionális fogalmaink közti műveleteket posztulálva tartjuk objektíve előrejelezhetőnek az intencionális rendszerek viselkedését, nem arról van szó, hogy a gépet közelítjük az emberi természethez ebben az elképzelésben, hanem az embert írjuk le, mint olyan evolúciósan fejlett funkcióhalmazt, mely számtalan kis gépből tevődik össze a legmagasabbtól a legalacsonyabb szintekig. A kulcsszó tehát az evolúció: ha úgy gondoljuk, hogy a gép szintaxisát csak mi állítottuk be úgy, hogy látszólag a valóságot szimbolizálva szemantikával rendelkezzen, akkor ennyi erővel azt is mondhatnánk, hogy mi is csak programozva vagyunk, Természetanya táplálta génjeinkbe a túlélés programját. Az eredeti intencionalitás ekkor az lenne, hogy a gén szándéka a továbbörökítődésre az eredendő, a mi hiteink, vágyaink és szándékaink csak származtatottak.

Vagyis el kell ismernünk, hogy ha a gép intencionalitása csak metafora, akkor az emberé is az. Dennett tehát alapvetően wittgenstein-i módon érvel az ellen az álláspont ellen, hogy beszélnünk kéne eredeti vagy intrinzikus intencionalitásról és másodlagos intencionalitásról. Érdekes módon Wittgenstein is egy, a metaforikus nyelvhasználatot érintő megjegyzésének továbbfűzése által jut a fent vázolt kritikájához a filozófiai intuíciónak:

589. „Lélekben elszántam magam rá.” És közben még arra is hajlamos az ember, hogy a mellére mutasson. Az efféle beszédmód pszichológiailag komolyan veendő. Miért kéne kevésbé komolyan venni, mint azt a kijelentést, hogy a hit a lélek egy állapota? (Luther: „A hit a bal mellbimbó alatt van”)²³

Donald Davidson, amerikai filozófus 1978-ban megjelent *What metaphors mean*²⁴ című cikkében is ilyen metafora felfogás körvonalazódik. Davidson szerint sincs semmi alapunk azt

²³ IF. 225. p.

²⁴ Critical Inquiry, Vol. 5, No. 1, Special Issue on Metaphor (Autumn, 1978), pp. 31-47Published

gondolni, hogy a metafora esetében beszélhetünk elsődleges és másodlagos jelentésről. A metafora azzal éri el a kívánt hatást, hogy az újdonság benyomását kelti, szokatlan módon fejezi ki az értelmét. A nyelvbe épített alapvető esztétikai funkció kerül ilyenkor előtérbe, olyan ez, mintha egy szimfóniát hallgatnánk. Dennett és az intencionalitás kérdéskörére alkalmazva ezt, azt mondhatjuk, hogy Searle halott metaforát szeretne élesztgetni az eredeti intencionalitás doktrínájával, Dennett pedig úgy gondolja, semmi akadályja annak, hogy új metaforát hozzon létre, és halottnak nyilvánítsa a korábit. Az 599. alapján Wittgensteinnel szólva érvelhetünk Dennett mellett, mondván, hogy Searle egy olyan érzés fogja gondolkodása során, mely egyfajta atmoszférát kölcsönöz egy szónak, majd mikor az felbukkan gondolkodásában, ezen berögződése az érvelés szerepét veszi át, majd ezután a gondolatkísérletek, a Turing-tesztre hivatkozás és a többi érve a szó újfajta felbukkanása ellenében már csak következménye ennek az alapberögződésnek, melytől képtelen szabadulni. Dennett hasonló szellemben áll szemben az erős MI összes ellenfelével (Putnam, Kripke), akik az emberi értelem kitüntetettségeinek valamiféle alapintuíciója mellett szállnak síkra.

3. Pragmatista attitűd: a versengő nyelvjátékok és az intencionális alapállás sikeressége

A következő egyezés, mely Dennettet a Vizsgálódások Wittgensteinjével rokonítja, az instrumentalizmusának természetével kapcsolatos. Mint már többször utaltam rá, Dennett legfőbb érve az intencionális alapállás mellett pragmatista színezetű, annak sikerességén alapul. Így például az etológia területén a pusztá adatok gyűjtése a mentalista terminusok mellőzésével mondjuk majmok megfigyelése esetén egyszerűen nem alkalmas a kapott adatok rendszerezésére, a statisztikák nyers adathalmazként feldolgozatlanok maradnak. Az adaptacionizmus, mely a túlélési esélyek növelésének normatív elméletét adja és az intencionális rendszerek elméletének mentalista kifejezései azonban olyan stratégiákat nyújtanak, melyekkel hatékonyak lehetünk ezeken a területeken a viselkedések előrejelzésében és megértésében. *„Mindkét taktikával azonban az a helyzet, dióhéjban, hogy működik. Nem mindig, de kielégítően gyakran. Meglehetősen jók vagyunk a megfelelő vélekedés- és vágytulajdonításban”²⁵.*

<http://www.jstor.org/stable/1342976> .

²⁵ Dennett, Daniel: Intencionális rendszerek a kognitív etológiában:a „panglossi paradigma” védelmében, in: id. mű 141.p.

Wittgenstein nyelvjátékelméletéről nagyon sokan azt gondolják, hogy az egyfajta relativizmust támogat, amelynek alapján tetszőlegesen válogathatunk, semmiféle mércét nem adhatunk meg a felől, éppen melyiket használjuk. A következő idézet talán elég támpontot szolgáltat arra, hogy ezt ne így gondoljuk, és arra is, hogy belássuk, mennyire lényegi pontja a dennetti gondolkodásnak is ez a hozzáállás:

„569. A nyelv- szerszám. Fogalmi szerszámok. Az ember mármost olyasmit gondol, hogy nem jelenthet nagy különbséget, hogy milyen fogalmakat használunk. Mint ahogy végül is lábbal és hüvelykkel ugyanúgy lehet fizikát művelni, mint m-rel és centiméterre; csak kényelmi szempontból van közöttük különbség. De még ez sem igaz, ha például az egyik mértékrendszerben a számítások több időt és fáradságot igényelnek, mint amennyit rájuk tudunk fordítani.²⁶”

A nyelvjátékok ezen értelmezésével nem mindenki ért egyet, jómagam Hilary Putnamhez csatlakozom, aki *Was Wittgenstein a pragmatist?*² című írásában úgy véli, Rorty nagyon rosszul értelmezi Wittgensteint, amennyiben relativizmusát látja igazolhatónak vele. Úgy néz ki, Wittgenstein értelmezése döntő nem csak etikailag, de elmefilozófiaiag is, hisz ahhoz hasonló érveket, melyeket Putnam etikai megfontolásokból Rorty ellen szegezett, Dennett is hozott az intencionális alapállás kapcsán. Rorty úgy gondolta, hogy mivel versengő leírások lehetségesek a mentális életről, elvben még azt sem mondhatnánk ki, hogy például az asztrológia rosszabb stratégia, mint a mentalizmus. Dennett érve ez ellen a durva relativizmus ellen pragmatikus: nem egyenrangúak a leírásaink, hisz az asztrológia nem tesz ellenőrizhető előrejelzéseket, sikere csak véletlen, az intencionális alapállás viszont a sikerének okai felől is garántál, Rortynak *„el kell fogadnia, hogy a természetes ontológiai attitűdön belül időnként a sikert a megfeleléssel magyarázzuk²⁷.”* Olyan ez Dennett szerint mintha két térképet vetnénk össze: az a jobb amelyik jobban ábrázolja ugyanazt a területet.

²⁶ FV 569. 220. p.

²⁷ Dennett . id mű. 235. p.

Konklúzió

Dolgozatomban annak igyekeztem utánajárni, hogy milyen hatások következtében foglalhatta el Daniel Dennett azt a nagyon egyedi pozíciót az amerikai filozófiai életben, melyet mai napig betölt. Egyedülálló helyzete akkor válik leginkább nyilvánvalóvá, ha abból a szempontból tekintjük, hogy az elme természetének vizsgálata során milyen elmefilozófiai nézőpontot alakított ki. Szemben áll egyrészt ugyanis azokkal a „reakciós”gondolkodókkal, akik a mesterséges intelligencia kutatása és a neurofiziológia fejlődésének hatására az emberi elme egyfajta kitüntetettsége mellett kardoskodnak és az elme megismerhetetlenségének valamifajta misztériumát hangsúlyozzák, a tudomány eredményeire pedig a legkisebb mértékben sem tartanak számot. Ez a fajta gondolkodás sokféle formában nyilvánul meg a kortárs elmefilozófiai diskurzusban, a leggyakoribbak az emberi elme eredendő intencionalitásának hirdetői (pl. Searle), a megismerés naturalizálhatóságának lehetetlenségét vallók (Putnam), a különböző magyarázati szakadékok tételezői (Levine, Kripke), illetve a kválérealista vagy a tudat nehéz problémáját feltételezők(Chalmers). A másik oldalon pedig Dennett a kritikátlan redukció utópisztikus vízióitól is elhatárolja magát (pl. Churchland.). Helyzete azonban nem csak a filozófián belül érdekes, hanem a tudományos életben betöltött funkciója miatt is. Ahogy Richard Griffin és Simon Baron-Cohen cikkükben²⁸ megjegyzi, Dennett egyfajta „filozófiai rendőrség” szerepét játszotta a legtöbb tudományos kooperáció alkalmával, de közel sem abban az értelemben, hogy a tudósok figyelmét korlátaikra hívta volna fel, sokkal inkább a közös munka igényével közelített munkájukhoz, az együttműködés reményében. Tudományos ambíciói pedig nem is maradtak beteljesületlenek, a viselkedés-magyarázat terén úttörő jelentőségűnek számít az intencionális alapállás elméletének kidolgozása, a gyermekviselkedés előrejelzése, legfőképp az elméletória kialakulásának elméletei pedig nagyrészt abból indultak ki, amit Dennett létrehozott, azt módosították a kísérletek, eredmények fényében.

A történeti hatások vizsgálatánál arra jutottam, hogy Dennett gondolkodása irányát és alapvonalait tekintve a legjobban tanára, Quine filozófiájának az örököse. Ezt ő maga egy interjúban ekképp ismerte el: *„Annak ellenére, hogy soha nem foglalkoztam azokkal a*

²⁸ RICHARD GRIFFIN & SIMON BARON-COHEN: The Intentional Stance: Developmental and Neurocognitive Perspectives In (2002) A. Brook & D. Ross (Eds.), *Daniel Dennett: Contemporary Philosophy in: Focus*. Cambridge, UK: Cambridge University Press, pp. 83-116

*technikai-logikai témákkal, amelyeken véleménye szerint egész filozófiája nyugszik, perspektívájához való ragaszkodásom igen erős.*²⁹” Azért is lehet fontos ez a megjegyzése, mert amellet, hogy Dennett gondolkodásának quine-i pragmatista szellemiségére, tehát tudomány és filozófia egységének gondolatára ugyanúgy magyarázatot szolgáltat, mint filozófiájának néhány túlkapására. Azzal talán Dennett is egyetértene, hogy Quine gondolkodását követve a metafizikai állítások tudományos irrelevanciájának vagy értelmetlenségének doktrínájáról is lemondunk. Ebből adódhat aztán Dennett gondolkodásának az a furcsasága, mely gondolkodásának leginkább későbbi szakaszában vált nyilvánvalóvá: az evolúció neo-darwinista felfogása egyfajta metafizikai alapként is megjelenik, melytől akár az élet értelmére is választ remélhetünk. A másik ilyen, ehhez kapcsolódó irányultság a harcos ateizmus hirdetése, melyet Richard Dawkins társaként vállal fel. Ezen a profetikus attitűdön a leginkább akkor csodálkozhatunk, ha eszünkbe jut, a különböző reakciós elmefilozófiai tendenciákkal szemben mennyire demisztifikáló irányultságú kritikákat fejt ki. Ehhez képest harcos ateizmusa és evolúciós világképének totalista jellege egyfajta új vallás propagálójának benyomását kelti.

A kettőségre véleményem szerint akkor derülhet fény, ha Wittgenstein hatását is tekintetbe vesszük. Úgy gondolom, hogy ahogy az előző fejezetben kimutattam: a rivális filozófusokkal szemben Dennett alapvetően Wittgenstein szellemében szeretne (és általában tud is) válaszolni, innen annak demisztifikáló ereje. Amikor azonban saját tevékenységének világnézeti implikációról van szó, akkor inkább Quine, és az ő „neo-pragmatista” tendenciája érvényesül, melyet legjobban a már egyszer megidézett hajó-metaphora szemléltet, tudomány és filozófia kerek egész egységének víziójával. Dennett életműve ilyen kerek egész benyomását szeretné keltetni. Elképzelhetőnek tartom, hogy ha gondolkodásának világnézeti pereme tekintetében is Wittgenstein szellemében járna el, akkor elkerülné ezen túlkapásokat. Hisz Wittgenstein gondolkodásának alapját képezte mindkét főművében azok önkritikus irányultsága. A Tractatusban az értelem határainak szeretne belülről határt szabni, a Vizsgálódások pedig minden esetben csak az adott nyelvjátékot belülről meghatározó szabályok vizsgálatát végezte el. Ez lehet az oka annak is, hogy Wittgenstein nem igazán rokonszenvezett az őt követő pozitivistákkal, például Carnappal sem, hisz az a fajta gondolkodás mintegy metafizikaellenes metafizikába torkollik. Összefoglalva tehát úgy gondolom, hogy egyrészt Dennett rendkívül egyedi filozófiai helye nagyon sokat köszönhet

²⁹ Elmagyarázta-e Daniel Dennett a tudatot? Boros János beszélgetése Daniel Dennettel, in: Boros János: A megismerés talánya, Áron 2009, Budapest, 370. p.

Wittgenstein hatásának, ahogy ezt ő maga is elismeri sok helyütt, az általánosabb metafizikai vagy világnézeti tendenciái esetén viszont döntően Quine követő marad.

Irodalomjegyzék

Primér

Kristin Andrews :Our Understanding of Other Minds Theory of Mind and the Intentional Stance
Journal of Consciousness Studies, 7, No. 7, 2000, pp. 12-24

Davidson, Donald: What metaphors mean?
Critical Inquiry, Vol. 5, No. 1, Special Issue on Metaphor (Autumn, 1978), pp. 31-47
Published
<http://www.jstor.org/stable/1342976>

Dennett, D:Az igazhívők: az intencionális stratégia és sikerének forrásai, in: Az intencionalitás filozófiája

Dennett, Daniel: Consciousness explained, BackBay Books, 1991

Dennett, D: Darwin veszélyes ideája, Typotex, 1998.

Dennett, D.: Micsoda elmék, Kulturtrade 1996.

Dennett, D: Quinning Qualia, in: Consciousness in Contemporary Science, 1988, OUP

Dennett, Daniel: Sweet Dreams, MIT Press 2005.

Dennett, Daniel: Újfent az instrumentalizmusról In: Az intencionalitás filozófiája, Osiris, 1998. 82. p. Pléh Csaba ford.

Fodor, Jerry: „Fodor kalauza a mentális reprezentációhoz: az intelligens nagynéni segédlete”
in, Kognitív tudomány, Osiris 1996, Pléh Csaba szerk.

Quine, W. V.: Arról, hogy mi van, in: A tapasztalattól a tudományig, Osiris 2002, 115-137.

Quine, W. V.:Ontológiai relativitás, in: A tapasztalattól a tudományig, Osiris 2002,137-177.p.

Quine, W. V: Az empirizmus két dogmája. In: Forrai Gábor – Szegedi Péter (szerk.), Tudományfilozófia: Szöveggyűjtemény. Budapest: Áron Kiadó, 1999. 131-153. p

RICHARD GRIFFIN & SIMON BARON-COHEN: The Intentional Stance: Developmental and Neurocognitive Perspectives In (2002) A. Brook & D. Ross (Eds.), Daniel Dennett: Contemporary Philosophy in: Focus. Cambridge, UK: Cambridge University Press, pp. 83-116

Sellars, Wilfrid: Empiricism and the philosophy of mind, XII: Our Rylean Ancestors.
<http://www.ditext.com/sellars/epm.html>

Stephen Stich-Ian Ravenscroft: Mi a nép pszichológia? In: Elmefilozófia szgy., L'harmattan, 2008.

Ryle, Gilbert: A szellem fogalma, Osiris 1999

Wittgenstein, Ludwig: Észrevételek, Atlantisz, 1995

Wittgenstein, Ludwig: Filozófiai vizsgálódások, Atlantisz 1992

Szekunder

Boros János: Elmagyarázta-e Daniel Dennett a tudatot? Boros János beszélgetése Daniel Dennettel, in: Boros János: A megismerés talánya, Áron 2009, Budapest

Forrai Gábor: Az intencionalitás (Bevezető tanulmány) in: Elmefilozófia szgy., L'harmattan 2008

Goldman, Alvin I. : Theory of Mind Oxford Handbook of Philosophy and Cognitive Science (2012)
Edited by Eric Margolis, Richard Samuels, and Stephen Stich

Tózsér János: Általános bevezetés: a test-lélek probléma (Bevezető tanulmány)

Tartalomjegyzék

Bevezetés.....2.

Elődök, mesterek és módszerek

(Quine, Ryle, Sellars).....3.

Wittgenstein szelleme

1. Behaviorizmus-kritika és a természetes gondolkodás rehabilitálása.....10.

2. A metafora „atmoszférája”, mint filozófiai intuíció és az „eredeti intencionalitás” esete.....17.

3. Pragmatista attitűd: a versengő nyelvjátékok és az intencionális alapállás sikeressége.....20.

Konklúzió.....22.