

# DENNETT *TUDATOSSÁGÁNAK* *MAGYARÁZATA: A KRITIKUSAI ÉS* A VITA A A TUDAT "VALÓDI TERMÉSZETE"

BENJAMIN NEWMAN  
Swarthmore College

Lehetségesek a zombik? Nem csak lehetségesek, hanem valóságosak is. Mindannyian zombik vagyunk. (Dennett *Consciousness Explained*, 406)

Kétségbeesett intellektuális becstelenség lenne ezt az állítást a szövegkörnyezetéből kiragadva idézni! (lábjegyzet a fentiekhez).

A *MEGMAGYARÁZOTT TUDAT* című könyvében Daniel Dennett részletes vázlatot ad arról, hogy mi lehet, vagy mi nem lehet a tudatosság magyarázata. A vita arról, hogy Dennett pontosan mit *is* magyaráz meg, és mit hagy ki, ha egyáltalán kihagy, egyrészt abból fakad, hogy számos kortárs filozófus mélyen és őszintén nem ért egyet a nézeteivel, másrészt abból, hogy néhány kulcsfogalom (pl. "tudat" és "magyarázat") meghatározása nem egyértelmű. Az alapvetően ellentétes világnézetek között és a fogalmi kétértelműségek mögött vajon valódi filozófiai vitát találunk, vagy csak sok üres retorikát? Úgy vélem, hogy Dennett és kritikussai közötti interakcióban egy valódi és értelmes kérdés lappang, és ebben a tanulmányban az lesz a célom, hogy ezt a kérdést felszínre hozzam, és úgy fogalmazzam meg, hogy világos legyen, miről van szó.

Mi a tét? Dennett megközelítése a tudat problémájához határozottan reduktív és materialista. Úgy véli, hogy az agyra (és kisebb mértékben a környezetre, amelyben az agy található) vonatkozó fizikai tények elegendőek ahhoz, hogy megmagyarázzák az elmével kapcsolatos összes tényt, beleértve a tudatot is. Dennett néhány kritikusja ezt az utolsó mondatot a "tudatosság kizárva" szavakkal fejezné be, és valóban van igazuk. Dennett beszámolóját eliminatívnak tekinthetjük, a tudat *elmagyarázásának* egyik esetének. Ő maga kissé kétértelmű e tekintetben. Ha a "tudat" alatt valamilyen különleges, kimondhatatlan, mindent vagy semmit tulajdonságot értünk, akkor Dennett magyarázatában semmi *ilyesmit* nem találunk. Ha viszont a "tudatot" úgy értelmezzük, hogy az a körülöttünk lévő világra való kifinomult érzékenységet, az én-érzést és a gondolatmenet meglétét jelenti, akkor Dennett elmélete tartalmazza *ezeket*.

dolgokat, és meggyőző válaszokat ad arra a kérdésre, hogy ezek hogyan valósulhatnak meg az agyban.

Szóval, melyik legyen az? Dennett és kritikusai között alapvető nézeteltérés van abban, hogy mi az, amit meg kell magyarázni. Dennett ("Living on the Edge", 147) emlékeztet bennünket, hogy nem elég, ha közös intuíciónkra hivatkozunk, hogy mire gondolunk, amikor azt mondjuk, hogy egyik vagy másik entitás tudatos vagy nem tudatos. Lehet, hogy nem vitatható, ha egy ilyen attribúciót egyszerűen úgy értelmezünk, hogy a szóban forgó entitásnak az a tulajdonsága, ami nekünk is van, amikor tudatunk van, de ez nyilvánvalóan körkörös - Dennett (ibid.) amellett érvel, hogy annak a kérdésnek, hogy valami tudatos-e vagy sem, "nincs egyértelmű, elmélet előtti jelentése", és ezért *először* ki kellene dolgoznunk egy elméletet, és utána megnézni, hogy maga az elmélet világít-e rá az ilyen meghatározási kérdésekre vagy sem. Pontosan ezt tette Dennett is, aki az elméletéből vont le a következtetéseit arra vonatkozóan, hogy mi a tudat, ahelyett, hogy előzetesen bedugta volna őket. Így nem meglepő, hogy következtetései markánsan eltérnek a legtöbb filozófus által előzetesen kialakított elképzelésektől arról, hogy mi is kell, hogy legyen a tudat.

A Dennett és kritikusai közötti legfontosabb különbség, amennyire ez az előzetes elemzésből megállapítható, úgy tűnik, a definíciók kérdése. Nem arról van szó, hogy Dennett a tudat egy bizonyos definícióját vallja, míg kritikusai egy vagy több másikat. Inkább Dennett hajlandó elhalasztani annak kérdését, hogy mi a tudat, amíg a tudat működésének kérdése nem tisztázódik. Ez utóbbi kérdésre kínál valódi választ a *Consciousness Explained*, és erre a válaszra térünk ki most részletesebben.

### **Dennett magyarázata a tudatosságról: Diszkurzus a módszerekről**

Mivel az ebben a tanulmányban kiemelt legfontosabb nézeteltérés módszertani jellegű, helyénvaló, hogy Dennett a magyarázatát a módszertani döntéseiről szóló egyértelmű nyilatkozatával kezdi. Az első ilyen választása a dualizmus abszolút elutasítása. Hogyan lehet ez a döntés módszertani kérdés? Dennett megjegyzi, hogy a legtöbb ünnevelt dualista meg sem próbálja megmagyarázni, hogyan kellene működnie a nem fizikai elmének (vagy az agy nem fizikai mentális tulajdonságainak) - úgy tűnik, nem kényszerülnek arra, hogy az elme lényegi természetéről olyan módon adjanak számot, amely fényt derít az elme működésére. Bár a dualizmus nyitva hagyja az elme részletes empirikus tanulmányozásának lehetőségét, kevés bevallott dualista foglalkozik ilyen kutatással, és egyetlen dualistának sincs kutatási programja arra a kérdésre, hogy az elme hogyan viszonyul a testhez - a legtöbben nyíltan kijelentik, hogy szerintük a kérdésre nem lehet választ adni. Bár kiderülhet, hogy a dualistáknak ebben a kérdésben igazuk van, Dennettnek tökéletesen indokolt, hogy elfogadja azt a munkafeltevést, hogy a projektje

lehetséges. Ami őt illeti, "a dualizmus elfogadása feladás" (*Consciousness Explained* 37). Így, amikor a dualizmus elutasítása mellett dönt, Dennett azt a módszertani döntést hozza, hogy nem adja fel egy könnyű (de üres) válasszal, valahányszor nehézségekbe ütközik.

Lényegében ugyanez az oka annak, hogy Dennett *el*utasítja az "altatás színlelésének" behaviorista taktikáját - tagadva, hogy léteznek tudatos tudatállapotok. Azt mondani, hogy a tudat nem létezik, nem több magyarázat, mint azt állítani, hogy a tudat valódi természete introspektíve nyilvánvaló, és a kérdést lezártnak tekinteni (ahogy a legtöbb dualista teszi). Dennett tehát elkötelezi magát amellett a feltevés mellett, hogy valóban léteznek tudati állapotok. "Ha tagadni akarom a tudatosság valamely vitatott tulajdonságának létezését" - írja - "akkor rám hárul annak bizonyítása, hogy az valamilyen módon illuzórikus" (40). Mivel Dennett alaposan le akarja rombolni a tudatról alkotott józan ész szerinti felfogást, hogy helyet csináljon a sajátjának, elég gyakran fogja vállalni ezt a bizonyos bizonyítási terhet. Fenomenológiai világunk bármely nyilvánvaló vonásának kihagyása olyan lépés, amely igazolásra szorul. Másrészt Dennett határozottan emlékeztet minket arra, hogy néhány ilyen jellemzőt ki kell hagyni egy végső magyarázathoz - a fájdalom olyan magyarázata (például), amely végső soron a fájdalom "eredendő szörnyűségére" hivatkozik, körkörös és kérdéses.

Hogyan bízunk az önvizsgálat adataiban, miközben elutasítjuk a különleges tulajdonságokra való ilyen hivatkozásokat? Dennett válasza erre a kérdésre módszertani álláspontjának lényege. Nem lehet egyszerűen

nem tekinthetjük mérvadónak az alanyok önvizsgálati beszámolóit, és nem tehetjükY

a "többes szám első személyű vélelmet", és feltételezzük, hogy mindenki egyetért az ilyen jelentések jelentésében. Az introspekción, állítja Dennett, sohasem a tiszta, elmélettől mentes megfigyelésről szól. Így ahhoz, hogy az introspektív adatokat a lehető legobjektívebben közelíthessük meg, bizonyos mértékű értelmezésre (vagy de-értelmezésre) lesz szükség. Az értelmezés Dennett által javasolt formája a *heterofenomenológia*. E módszer alkalmazása során az alany introspektív jelentését szöveggként kezeljük, és kritikátlanul elfogadjuk, hogy amit az alany mond, az az, amit az alany valóban mondani akart - anélkül, hogy feltételeznénk, hogy minden, amit az alany mond, igaz. Valójában az alany teljes felhatalmazást kap egy fiktív világ létrehozására, amelyet a szöveg határoz meg. Ahogyan az is objektív tény lehet, hogy Sherlock Holmes a Baker St. 221B-ben lakott, annak ellenére, hogy a nagy detektív fiktív volt, úgy kezelhetjük az alany önvizsgálati jelentését, mint egy fiktív világra vonatkozó objektív tényeket. E világ tartalma alkotja az alany heterofenomenológiáját, és a legapróbb részletekig objektív.

Itt van tehát egy módszertani válasz arra a kérdésre, hogy mi az,

amit meg kell magyarázni. A heterofenomenológiai adatok azok, amik

meg kell magyarázni, és ezek magyarázata az elme "vázlatos" elméletének formáját ölténé. Állítása szerint az egységes tudatfolyam érzetünk attól az időskálától függ, amelyben önmagunkat szemléljük. Egy magasabb nagyítási szinten ez az áramlat sok párhuzamos, tartalmi rögzítésszerű eseményfolyamra bomlik szét, amelyeket az agyat alkotó számos specializált alrendszer valamelyike hajt végre. Ezek a tartalomrögzítések együttesen egyfajta narratívát alkotnak, azzal a különbséggel, hogy ennek a narratívának nincs kanonikus formája - mindig többféle munkavázlat létezik, amelyek mindegyike folyamatos szerkesztői átdolgozásnak van alávetve, és a meg nem írt jövőbe nyúlik. Az egyetlen módja annak, hogy egy tartalom egy "kanonikus" narratíva részévé váljon, az, ha egy *szondára* - egy közvetlen kérdésre, például a "Mire gondolsz most?" - válaszolunk, és az ilyen szondára válaszul megjelenő tartalom rendkívül érzékeny a szonda jellegére és időzítésére. Az, hogy mit fogunk ki a "tudatfolyamból", nagyban függ attól, hogy mikor megyünk horgászni, és hogyan csalizunk a horogra.

Mivel a tudat tartalma az imént megfogalmazott értelemben meghatározhatatlan, ha soha nem vesszük a fáradságot, hogy ellenőrizzük, az egységes tudatfolyamhoz hasonló dolgok létezése az állandó ellenőrzés vagy önbizonyítás szokásából nő ki - valójában több szokásból, beleértve, de nem kizárólagosan, az olyan ismert trükköket, mint az önmagunkkal való beszélgetés és annak néma leszármazottja, a privát deliberatív gondolkodás. Hogyan sajátítjuk el ezeket a szokásokat? Némelyikük beépített, mint például a csillapíthatatlan kíváncsiságunk, amely a legtöbb főemlős veleszületett adottsága, míg mások tanultak, de a legnagyobb hozzájárulást, beleértve az önbeszédet is, a nyelv és a kultúra adja. Nemcsak maga a nyelv az egyik legfontosabb szokásunk, hanem a nyelv teszi lehetővé az ilyen szokások elsajátítását is, szándékos gyakorlás és kulturális átadás révén. Az ókori görögöknek elképesztő emlékezőtehetségük volt - mesemondóiknak olyan eposzokat kellett emlékezetből előadniuk, mint az Odüsszeia, amit mi nagyon nehezen tudnánk. Az ő agyuk nem különbözött a miénktől, akkor miért különbözik ennyire az emlékezetünk? A megtanult mnemotechnika fejlett rendszerét használták, amit mi már nem tanítunk egymásnak, mert már nincs rá szükségünk - ha valamire emlékezni akarunk, ami olyan hosszú, mint egy könyv, egyszerűen csak leírjuk. A kultúra változásai tehát a mentális tájkép változásához vezethetnek, azáltal, hogy befolyásolják a mentális szokások szükségességét és elérhetőségét egyaránt.

E magasabb szintű, önellenőrző tevékenységek szintjén végül is létezik egy, kanonikus belső narratíva, bár ez időnként (amikor elfelejtjük ellenőrizni) elmarad, és értelmetlen lenne finom időbeli megkülönböztetéseket kérni tőle, amelyek rövidebbek, mint az őt létrehozó tevékenységek időskálája. Így egy tömegesen párhuzamos információs folyamat tevékenységeiből egy soros, egyszerre csak egy dologra kiterjedő információs folyamat jön létre,

gyakorlatilag egy "virtuális gép" létrehozása - a számítástechnikai szakzsargonban ez egy olyan számítógépet jelent, amelyet egy másik típusú számítógépen programként hajtanak végre. Ez a virtuális gép nagymértékben kiterjeszti az agy képességeit, mivel biztosítja a hosszú távú projektek (például filozófiai szakdolgozatok) elkezdéséhez szükséges szervezetséget, de nem túl hatékony - ezért olyan lassú a szándékos gondolkodás, az a fajta gondolkodás, amely hasonlít az önmagunkkal való beszélgetéshez, az olyan keményebben bedrótozott mentális tevékenységekhez képest, mint például a szűnyog lecsapása. Amellett, hogy nem hatékony, a "Joyce-i gépezet" megbízhatatlan is - könnyen elterelődik és letér a helyes útról, gyakran esik zökkenőbe vagy ismétli önmagát, sőt, ismétli önmagát. Ha ezeket a buktatókat el akarjuk kerülni, a gépnek valamilyen módon szabályoznia kell saját viselkedését. Az egyik megoldás a már meglévő érzékelőberendezések felhasználása, amelyek igen hatékonyan bizonyultak a környezet összetett, kiszámíthatatlan jellemzőinek szemmel tartásában. A rendszernek csak eggyel több összetett, kiszámíthatatlan, életbevágóan fontos dologgal kell majd saját magát szemmel tartania.

Egy ilyen rendszerben a legtöbb mechanizmus *ad hoc* zsűri-berendezés lesz, de a legtöbb (beleértve sok zsűri-berendezést is) csodálatosan kifinomult lesz. Idővel a szabályozó rendszerek fokozatosan alkalmazkodni fognak az általuk szabályozott rendszerek sajátosságaihoz, és végül a virtuális gép egésze könnyedén megismeri saját sajátosságait. Ez a virtuális gép, amely expliciten nyomon követi saját belső állapotait és a körülötte lévő világot, szervezi és irányítja a több párhuzamos folyamat tevékenységét, amelyekből létrejön, és olyan magánjelleggel ismeri saját tevékenységeinek furcsaságait, ahogyan egyetlen hasonló, de különálló rendszer sem lehet, képes lenne az emberi elme minden tevékenységét megismerni. "De vajon tudatos-e?" - kérdezzük. "IGEN" - jelenti ki Dennett bátran -, "az én elméletem a tudatosság elmélete. Bárki vagy bármi, amelynek egy ilyen virtuális gép a vezérlőrendszere, a legteljesebb értelemben tudatos ... mert ilyen virtuális géppel rendelkezik" (281). Ez Dennett magyarázatának lényege, amelyet továbbfejleszt, és arra használja fel, hogy feldolgozzon néhány, a tudatot körülvevő kényesebb filozófiai problémát. Bár Dennettnek éppen az ezekre a problémákra adott válaszai miatt oly sokszor a kritikusai szemére vetették, én nem maguk a problémák köré szervezem az ellenzékkel kapcsolatos kritikámat, amelyek közé tartoznak a zombik, a qualia és az én. Ennek az az oka, hogy a problémák átfedik egymást, és sok azonos (vagy mélyen hasonló) érvet használnak mindegyikben. Ezért a következő fejezet a Dennett kritikusai által használt érvelési formák köré szerveződik, és ebben a kontextusban fogunk találkozni magukkal a problémákkal.

**A másik oldal: Mi a baj a dualizmussal?**

John Foster ("Dennett's Rejection of Dualism") egyenesen támadja Dennett elismerten *a priori* elutasítását a dualizmusnak. Jellemzése Dennett saját okairól meglehetősen találó - egyrészt a dualizmus elfogadása a magyarázatok keresésének feladását jelenti, míg másrészt minden értelmes dualista nézet sérti a fizikai világegyetem zárttságát. Foster válasza az első pontra az, hogy még a fizikai törvényekkel kapcsolatos magyarázatoknak is "alul kell maradniuk", és meg kell maradniuk néhány olyan entitás tekintetében, amelyeket alapvetőnek tekintünk. Miért kellene többet kérnünk a dualistáktól? Dennett azt válaszolja ("Living on the Edge" 141), hogy ő nem kér többet - csupán egy mentális jelenség dualista fogalmakkal való *bármilyen* magyarázatát kéri, amely nem üres vagy kérdéseket felvető. Egy dolog azt mondani, hogy az alapvető fizikai részecskék egyszerűen olyanok, amilyenek, hiszen "olyanok, amilyenek", az egyszerű és könnyen leírható. Egészen más dolog azt mondani egy olyan összetett és finoman változó entitásról, mint az emberi elme, hogy "csak úgy van", különösen, hogy a dualista ritkán tesz nagy erőfeszítéseket ennek a "módnak" a leírására.

Foster válasza a második pontra sokkal aggasztóbb - egyszerűen elutasítja, hogy a fizikai világegyetem zárt rendszer lenne (25). Miért kellene - kérdezi - egyszerűen feltételeznünk, hogy az agyak ugyanazoknak a fizikai törvényeknek engedelmeskednek, mint az univerzum többi része, amíg nem erősítettük meg újra ezeket a törvényeket élő agyszövetekkel? Egyetért azzal, hogy *ceteris paribus* ez egy jó munkafeltevés, de tagadja a c. p. feltételt - más dolgok *nem* egyenlőek, mert a hagyományos antimaterialista érvek mindegyike a dualizmus mellett szól. Amíg ezeket az érveket nem cáfolják meg kerek-perec, addig jó okunk van arra, hogy elutasítsuk azt a hipotézist, hogy az agyak engedelmeskednek a fizika törvényeinek. Az én válaszom Foster érvére itt, amit Dennett csak burkoltan jegyez meg (*Consciousness Explained*, 7. fejezet), az, hogy egy ilyen különbség fizikai különbség lenne - az emberi agyak kivételével minden a világegyetemben a fizika törvényeinek engedelmeskedik. Ha ez így van, akkor a helyesen megfogalmazott fizikai törvényeknek ezt figyelembe kell venniük, inkább az " $E = mc^2$ ", kivéve az élő emberi agyak belsejét" módjára. Ekkor jogos a kérdés, hogy honnan *tudná* egy fizikai részecske, hogy egy emberi agyban van, és ezért meg kellene szegnie a fizika törvényeit, tekintettel arra, hogy az agy egy bonyolult szerkezet, amelynek felismerése nem triviális. Világos, hogy Foster nem szándékozik választ adni erre a kérdésre, sőt még csak nem is akarja megvizsgálni azt, megerősítve ezzel Dennett első ellenvetését a dualizmussal szemben - a kutatási program hiányát.

### **Ez egyszerűen nyilvánvaló!**

Míg Searle tagadja az antimaterializmussal való rokonságot, Sprigge ("A Disillusioned Zimbo") pedig arra utal, hogy az idealizmust

részesíti előnyben a dualizmussal szemben, a



dualista zászlót Dennett sok más kritikusa is lengetni fogja, ha nem is veszi fel egyenesen. Legtöbbjük azonban nem célozza közvetlenül Dennett dualizmusának elutasítását, mint Foster. Miközben azt állítják, hogy az elme természetében van valami, ami sérthetlenné teszi azt a Dennettéhez hasonló materialista magyarázatokkal szemben, úgy tűnik, úgy érzik, hogy nem szükséges érvekhez folyamodni az elme természetére vonatkozó ezen állítások alátámasztásához. Inkább megelégszenek azzal, hogy az elmével kapcsolatos közös intuíciónkra hagyják a dolgot, mintha minden fontos tény egyszerűen nyilvánvaló lenne, és mindenki egyszerűen egyetértene velük. Dennett erre hívja fel az olvasó figyelmét a módszertanról szóló fejezetének bevezetőjében (67), "többes szám első személyű feltételezésnek" nevezve ezt, és azzal az indokkal tiltakozik ellene, hogy még mindig sok vita van, még azok között is, akik azt állítják, hogy mindannyian egyetérthetünk a tudat természetével kapcsolatban, ahogyan azt az introspekció révén megközelítjük. Erre adott válasza, mint már említettük, annak tagadása, hogy az introspekció valaha is egyszerűen csak a belenézés és annak meglátása lenne, ami ott van - minden introspekció konstrukció, amely az elmélettől függ, és kategorikus, valamint partikuláris hibáknak is ki van téve. Kritikusai persze ebből semmit sem fognak elfogadni.

John Searle azt mondja, hogy nem nehéz a tudatosságnak egy józan ésszel használható definíciót adni, olyat, amely nem analitikus, és nem "a jelenség mögöttes lényegének elemzésére törekszik", és becsületére legyen mondva, ad is egy ilyen definíciót: egyszerűen fogalmazva, a tudatosság az, amivel akkor rendelkezünk, amikor - köznapi értelemben - nem vagyunk öntudatlanok. Mint Searle elismeri, ez a definíció nem vág a dolog lényegébe, de ebből a következő következtetéseket vonja le, teljesen érvek nélkül: a tudatosságnak vagy minden vagy semmi, jelen vagy nem jelenlévőnek kell lennie. Ez egy "belső, első személyű, minőségi jelenség". Searle sem bizonyítékot, sem érvet nem szolgáltat a tudat ezen állítólagos jellemzőinek egyikére sem, és nem is definiálja gondosan a jelzők egyikét sem. Amint azt ő maga is szabadon elismeri, nem adott olyan definíciót, amely a szóban forgó jelenség elemzéséhez releváns lenne, ennek ellenére mégis folytatja az elemzés nagy részét ebből kiindulva. Mire alapozza elemzését? Milyen alapon állítja, hogy tudja, miről beszél, amikor a "tudat" szót használja? Úgy tűnik, Searle azt állítja, hogy kritikátlanul, közvetlenül a közös, józan intuíciónkból meríthetünk fontos filozófiai eredményeket - más szóval, hogy a tudat minden fontos jellemzője egyszerűen nyilvánvaló, mindenki számára.

Searle meggyőződése, hogy az elme természete nyilvánvaló, csodálatosan megmutatkozik leghíresebb gondolatkísérletében, a Kínai szobában ("Elmék, agyak és programok"). Searle megadja ellenfeleinek, hogy van egy számítógépes programunk, amely átmegy a Turing-teszten (kínaiul). Ezután ezt a programot egy sor keresőprogramként valósítja meg.

táblázatok és főkönyvek, amelyek műveleteit kézzel végzi. Magára a programra úgy hivatkozik, mint egy "utasításkészletre", amelyet "kap", ami talán egy viszonylag szerény papírcsomagra utal, amely könnyen követhető, lépésről-lépésre történő, angol nyelvű utasításokat tartalmaz.

Mint állítja, a program ugyanolyan biztosan átmegy a Turing-teszten

Searle implementációja alatt is, mint korábban, így kísértésbe eshetnénk, hogy azt mondjuk, hogy érti a kínai nyelvet; de Searle

tiltakozik. Ő maga nem ért kínaiul - soha nem is

értett -, és az általa leírt rendszerben *nyilvánvalóan* nincs máshol a

megértés helye. Nehogy azt állítsuk, hogy valahogyan az egész

rendszer az, ami megérti, Searle azt javasolja, hogy nyeljük le a

rendszert. Egyszerűen fogalmazva, memorizálni fogja a programot. Ez

az a pont, ahol Dennett és én is felkiáltunk. Searle átverte az

olvasóit, hiszen a program, amelyet arra kérte őket, hogy képzeljék el,

minden valószínűség szerint meglehetősen bonyolult, és most arra kéri

őket, hogy képzeljék el, hogy ő már mindent megjegyzett. Ha tényleg

annyi fáradságot veszünk, hogy elképzeljünk egy olyan programot,

amely átmegy a Turing-teszten, papíron megvalósítva, akkor

valószínűleg egy nagy könyvtárat kellene elképzelnünk, nem pedig

néhány kis papírcsomagot. Mégis a képzelet eme csalása

ellenére Searle azt állítja, hogy még mindig *nyilvánvaló*, hogy a

rendszer nem érti, és ezzel számos ellentmondásos premisszát hagy

érveknélkül elsiklani. Báróvatosan

elkülöníti az olyan fogalmakat, mint a megértés, a szándékosság és a

tudatosság, ez a minta tipikusan ilyen

érvelésének mindhárom témában.

Egy másik fontos filozófus, aki a legfontosabb premisszáinak

"nyilvánvalóságára" támaszkodik, Thomas Nagel. A "Milyen érzés

denevérenek lenni?" című írásában azt állítja, hogy ahhoz, hogy egy

organizmus tudatos legyen, "kell lennie valaminek, ami olyan, mintha

az az organizmus *lenne* - valami olyannak, ami az organizmus számára

olyan". Nem hoz fel érveket erre az álláspontjára, és nem is veszi a

fáradságot, hogy részletesen kifejtse, mit ért ez alatt. Viszont hosszasan

bizonygatja, hogy egyetlen materialista elmélet sem képes befogadni a

tudatnak ezt a tulajdonságát - ezzel kerek pereg megcáfolva minden

(vagy minden jelenlegi) materialista elméletet, ha mindannyian érvek

és bizonyítékok nélkül egyetértünk azzal, amit ő állít, feltehetően azon

az alapon, hogy ez egyszerűen nyilvánvaló. Végül is Nagel nagyon is

hajlandó "feltételezni, hogy mindannyian hisszük, hogy a denevéreknek

van tapasztalata" (393), mielőtt esettanulmányként használná őket.

Nyilvánvalóan azt is feltételezi, hogy mindannyian egyetértünk az ő

definíciójával arról, hogy mit jelent tudatosságnak lenni. John Foster,

akivel már találkoztunk a dualizmus védelmében, hasonló állítást tesz

Dennett azon követelésére adott válaszában, hogy egy dualista adjon

valamilyen magyarázatot a nem fizikai elme természetéről. "Magam

sem látom - írja -, hogy a dualistának miért kell elismernie, hogy az

elme természetét illetően van valami több, mint amit az introspekció képes feltárni" (1993, 29. o.). Dennett válaszát már megjegyeztük - hogy egyszerűen egy olyan elme tételezése, amely minden

az introspekció által "feltárt" tulajdonságok valójában semmit sem magyaráznak meg, és nem hagyja, hogy bármilyen jobban megalapozott elmélethez folyamodjunk, amikor az intuíciók ütköznek. Végül is ez az alapvető probléma azzal, ha bármit is "nyilvánvalónak" fogadunk el az elmével kapcsolatban - ha nem értünk egyet abban, hogy mi a nyilvánvaló az elmével kapcsolatban, akkor nem várhatjuk el, hogy a "nyilvánvaló" intuíciókon alapuló következtetésekben is egyetértsünk.

### A visszafejlődés kínjai

Egy másik, Dennett kritikusai által gyakran használt érv az, hogy a mentális tulajdonságoknak valamilyen forrásból kell eredniük, amely azokat eredendően megtestesíti - ha egy rendszer által létrehozott elbeszélésnek "valódi jelentése" van, akkor valahol a rendszerben kell lennie egy eredendő jelentés forrásának, mert nyilvánvalóan semmilyen egyszerű szöveg, egyszerű szintaktikai művelet vagy egyszerű fizikai interakció nem adhat értelmet (vagy megértést, vagy tudatot). Ennek a nézetnek egy változatát Searle fejti ki, amint azt fentebb részleteztük, valamint Fellows & O'Hear ("Consciousness Avoided"). Azt írják (Searle védelmében), hogy "a számítógépes programoknak elme kell ahhoz, hogy olvassák őket, ezért az elme nem lehet pusztán számítógépes program" (82). Később, Dennettnek az énről mint a narratív súlypont feltételezett központjáról szóló beszámolójára reagálva azt kérdezik, hogy "ki az, aki jelentést tulajdonít a szavaknak és a szövegeknek", ragaszkodva ahhoz (és nem meglepően), hogy "a szavaknak és a szövegeknek nincs jelentésük önmagukban". De aztán azt követelik, hogy "a körkörösség veszélye miatt Dennett nem mondhatja, hogy csak egy másik szöveg ... az, ami a szövegeknek értelmet ad. Legalábbis néhány szövegnek [rendelkeznie kell] az én és az ágensek jelentést kölcsönző tulajdonságaival" (87). Más szóval, valamilyen entitásnak kell lennie az elsőnek, amely jelentéssel rendelkezik, és így a jelentéssel önmagában kell rendelkeznie.

Dennett szerint (és szerintem is) ez az egyik legnépszerűbb kategóriátévesztés a filozófia területén. Igaz, a hétköznapi angolban nincs értelme úgy beszélni egy entitásról, hogy "majdnem" vagy "egy kicsit" tudatos, vagy hogy "majdnem" valódi szemantikai tartalommal rendelkezik, de ez nem biztos útmutatás a szóban forgó jelenségek természetéről - csupán aköznyelvünk korlátairól. A végtelen regresszió vagy valamilyen belső forrásban való megalapozottság dilemmájából való menekülésként Dennett ("Living on the Edge" 144) egy véges, homályos regressziót javasol, a következő példával: Az emlősök története mégis véges. A látszólagos paradoxon megoldása annak felismerése, hogy a nem emlősökkel kezdődő és az emlősökkel végződő genealógia "csak önkényesen tagolható". Cseréljük ki az érvelésben az "emlős" kifejezést "valódi jelentést hordozó tartalmat hordozó járműre", és Dennett nézete arról, hogyan keletkezhet jelentés

és tudat egy értelmetlen és tudattalan szubsztrátból, világossá válik. I

Fellows & O'Hearnak válaszolva azt a további észrevételt tenném, hogy a körköröség remélhető, ha az ént saját narratívájának értelmezőjeként azonosítjuk, mert ha az én narratíva, akkor - a nem-létezés veszélye esetén - körkörösön önértelmezőnek kell lennie.

### Nem lehet megmondani, ha ránézel

Ebben a szakaszban a kvaliák és a zombik kapcsolódó problémáival foglalkozom. Gyakran nehéz megmondani, hogy egy érv mely problémákkal foglalkozik, és ez nem meglepő, hiszen a lényegük ugyanaz - a hagyományos filozófus azt javasolja, hogy a tudatosságunk vannak olyan jellemzői, amelyek a mentális tartalmak funkcionális kölcsönhatásán túlmutatnak; így ezek a jellemzők megváltoztathatók vagy teljesen eltávolíthatók anélkül, hogy ez a funkcionális kölcsönhatás sérülne, és így az ágens külső viselkedését *semmilyen módon* nem befolyásolnák. Dennett és én úgy véljük, hogy ez a felfogás inkoherens, de sok kritikusa még mindig ragaszkodik hozzá a legnépszerűbb formáinak egyik vagy mindkét formájában - a kvaliák és a zombik -.

A tudat azon jellemzői között, amelyek kihagyásával a materialistákat leggyakrabban vádolják, a tudatállapotaink nyers, minőségi aspektusai, vagyis a *kvalitások* foglalnak el előkelő helyet. Dennett szinte minden kritikusa azzal vádolja őt, hogy nem tesz igazságot a kvalitásoknak, és hogy hitelt adjon ezeknek a kritikusoknak, Dennett egyenesen azt állítja, hogy ilyen kvalitások nem léteznek (*Consciousness Explained* 372). Ez alatt azt érti, hogy a mentális állapotoknak nincsenek belső, kimondhatatlan tulajdonságaik, amelyek különböznek az általuk betöltött funkcionális szerepektől, ehelyett azt javasolja, hogy ezek a funkcionális diszpozíciók elegendőek ahhoz, hogy a kvaliákkal kapcsolatos összes állított tény, *beleértve* a feltételezett kimondhatatlanságukat is, megmagyarázzák. Sok filozófus azt is állítja, hogy el tud képzelni "zombikat", olyan lényeket, amelyek rendelkeznek a Dennett által elképzelt összes funkcionális tulajdonsággal - olyannyira, hogy fizikailag vagy legalábbis külsőleg megkülönböztethetetlenek a tudatos emberi lényektől -, de valódi tudatosság nélkül. Azt, hogy a két állítás szorosan összefügg, legjobban az jelzi, hogy a qualia vezeti a listát arról, hogy a zombiknak mi az, amivel nem rendelkeznek. Dennett felvetésére, *miszerint* zombik vagyunk, és a tudatunk, amely különbözik a funkcionális alapjainktól, egy illúzió, azt kérdezik, hogy ha nem egy tudatos én, akkor kit téveszt meg ez az illúzió, visszhangozva a regresszióval kapcsolatban már említett érveket. A kvaliák és a zombik kérdésére adott válaszom azonban más lesz. A bizonyíték fogalma körül forog.

Mit jelent az, hogy bizonyítékunk van valamire? Feltételezhetően a "P-re vonatkozó bizonyíték" alatt olyan információt értünk, amely *ceteris paribus* valószínűbbé teszi P igazságát - tehát olyan információt,

amelynek a számunkra való hozzáférhetősége oksági kapcsolatban áll P igazságával. Amikor azt mondjuk, hogy valamilyen kijelentést teszünk P-re vonatkozó bizonyíték alapján, akkor azt értjük, hogy a kijelentésünket ezen információ rendelkezésre állása okozza. A mi definíciónk a

a bizonyítékok a következő kontrafaktumot támasztják alá: *ceteris paribus*, a P eltávolítása legalább bizonyos körülmények között a P-re vonatkozó bizonyítékok eltűnését okozza. Ellenkező esetben a P-re vonatkozó bizonyítékok és a P valószínűsége közötti kapcsolat nem áll fenn. Ha tehát valamilyen kijelentést P-re vonatkozó evidencia alapján tettünk, nem menekülhetünk meg a "ha nem lett volna P, nem tettem volna meg ezt a kijelentést" ellenténytől. Az ellentényállás tagadása azt jelenti, hogy tagadjuk, hogy a kijelentés bármelyik oka valóban bizonyítéknak nevezhető P-re. Most pedig legyen a P tétel: "Valóban tudatosak vagyunk, és nem csak zombik, akik *minden* tekintetben pontosan úgy viselkednek, mint a tudatos emberek". Az ellentételezés ebben az esetben egyértelműen hamis, mert a *hipotézis* szerint egy zombi-filozófus pontosan úgy viselkedik, mint tudatos társa, egészen a tudatosságra vonatkozó állításokig, és nem csak zombiként. Tehát nem áll fenn az a helyzet, hogy egy filozófus zombival való helyettesítése valaha is azt eredményezné, hogy más véleményt fogalmazzon meg a saját tudatosságáról. Bármi legyen is az oka annak, hogy a filozófusok a tudat létezésével kapcsolatos nézeteiket bármilyen értelemben zombiként koherens módon lehetségesnek tartják, azt *nem lehet* értelmesen evidenciának nevezni. Ugyanez az érv alkalmazható a qualia-ról szóló kijelentésekre is.

A fenti érvelés kapcsolódik Richard Rorty "Holism, Intrinsicity, and the Ambition of Transcendence" című írásához, amely még erősebb érvet hoz fel az intrinzikális tulajdonságok hívei ellen, és azt kérdezi, hogy az intrinzikális, nem-relációs tulajdonságok fogalma egyáltalán koherens-e? Azt sugallja, hogy nem csak az ének, hanem minden tárgy, amennyiben diskurzusunkban szerepelnek, nem jelentenek semmit a narratív (vagy legalábbis leíró) gravitáció központjain túl. Más szavakkal, egy X tárgynak nincs semmi lényeges azon túl, hogy az, amire a legtöbb X-mondat igaz. Hogyan állíthatnánk mást? Ha ezt tennénk, akkor egyszerűen csak egy újabb X-mondatot adnánk hozzá a listához. Ha Dennett ezt a (Wittgensteinnek tulajdonított) metafizikai tézist olyan messzire vinné, mint Rorty, akkor nem is kellene heterofenomenológiai entitásait "puszta fikcióknak" neveznie, hiszen ha *minden* csak egy leíró súlypont, akkor az egész tény-fikció, valós-intentív megkülönböztetés teljesen feloldódik. Sajnos (Rorty szerint) Dennett nem megy ilyen messzire - ő még mindig ragaszkodik az ontológia lehetőségéhez, mint az episztemológiától megkülönböztetett lehetőséghez. Dennett azonban elég messzire megy ahhoz, hogy azt állíthassa, hogy a nem-relációs tulajdonságoknak nincs értelme. Kritikusai, mondja Rorty, ezen a mély, metafilozófiai szinten nem értenek vele egyet - ők (Rorty kiemeli Searle-t és Nagelt) mind az intrinzikális tulajdonságok hívei. Bizonyára Searle és Nagel maguk is ezt mondják. Rorty azonban Dennettet is szidja, amiért a heterofenomenológiát semleges talajnak állítja be ebben a vitában, mert nincs semleges talaj - ha azt állítjuk, hogy egy jelenségnek azok az



aspektusai, amelyekről beszélni lehet, azok a fontos aspektusok, az a kérdés felvetése.

a nyelv hatókörén kívül eső belső tulajdonságok lehetőségével szemben. A szakadék, amely Dennettet elválasztja a Searle-hez és Nagelhez hasonlóktól, olyan szakadék, amelyet csak hitbeli ugrással lehet átlépni, nem pedig érvekkel. Rorty mégis biztosított egy fontos pontot Dennett számára - az esszencialisták az érvek erejével sem tudnak érvényesülni, és így, bár lehetnek okok arra, hogy higgyünk (vagy ne higgyünk) a kvalitásokban, nem lehet *bizonyíték* egyik irányba sem.

Mindez hangsúlyozottan nem azt jelenti, hogy tagadjuk, hogy léteznek tudatos állapotok, vagy hogy vannak olyan tulajdonságaik, amelyeket tulajdonságoknak nevezhetnénk - csupán azt tagadjuk, hogy ezek a tulajdonságok elkülönülnek attól a szereptől, amelyet ezek az állapotok a szervezet működésében játszanak. Dennett nézetének egyik zavara úgy tűnik, hogy különös erővel vonja magára a kvalitás-erősítőket. Dennett azt állítja, hogy nem létezik valódi látszat, eltekintve attól, *hogy megítéljük*, hogy valami tényleg így van, és hogy minden mentális tartalom *kijelentéses* tartalom. Kritikusai közül sokan félreértik e kifejezések használatát. Sprigge például úgy képzei, hogy Dennett fenomenális világának nem érzéki minőségekből, hanem "briliáns verbális dialektikából" (49) kell állnia, Lockwood pedig azt kifogásolja, hogy a látás nem lehet olyan, mint a hit, mert "az ítéletek ... túl vérszegények, túl magas szintű, túl *intellektuális* feladat az érzékelés és az észlelés szubsztanciájához" ("Dennett's Mind" 69). Dennett gyorsan reagál, és egy újabb kategóriahiba miatt ítéli el ezeket a kritikusokat. Emlékezteti őket, hogy *tételekről* beszélünk. Nincs korlátja annak, hogy egy tétel milyen tartalmú lehet - például az, hogy a szemem előtt lévő látvány pontosan úgy néz ki, mint {egy vizuális jelenet ábrázolása}, tökéletesen jó tétel. Nem *mondatokról* beszélünk, *így* nem lesz elég az az ellenvetés, hogy "a látás olyan, mint a képek, a hit pedig olyan, mint a mondatok, tehát mivel egy kép felér ezer szóval, a látás nem lehet hit" (49-150). Talán van elvi okunk kételkedni abban, hogy az észlelés több, mint különböző tartalom-áramlatokban hozott, tudatot megelőző ítéletek, de az, hogy az ítéletek nem elég *gazdagok*, vagy nem tudnak *elég* tartalmat hordozni, nem ilyen ok.

### **Mi történik valójában**

A bevezetőmben azt sugalltam, hogy a Dennett és kritikusai közötti valódi különbség a definíciók között van - Dennett megelégszik azzal, hogy a tudatosságot utólag találja meg elméletében, míg kritikusai a tudatosság előre megfontolt definícióiból indulnak ki. Elsőre úgy tűnhet, hogy ez ésszerű módja annak, hogy valamit megmagyarázzunk. Elvégre hogyan magyarázhatnál meg valamit, ha még azt sem tudod megmondani, hogy mi az? Az előző részekben felajánlottam néhány homályos utalást arra, hogy a tudat ilyen megközelítésének problémája valójában a nyelvvel kapcsolatos probléma. Most ezt a felfogást fogom tisztázni. Van értelme, hogy

a nem mentális jelenségek leírása, mielőtt megkísérelnénk megmagyarázni őket, mert ugyanabból a harmadik személyű nézőpontból írjuk le és magyarázzuk meg őket, és általában nagyjából ugyanazokkal a kifejezésekkel. A mentális jelenségek esetében ez nem így van, mivel a nyelvünk által a tudat leírására használt kifejezések számos előfeltevést hordoznak magukban, amelyek mind a definícióikban, mind a használatukban megtestesülnek. Az, hogy *nyelvileg* kínosnak tűnik egy entitásról úgy beszélni, mint "majdnem tudatos", csak egy példa erre. Ha a tudatosság vizsgálatát azzal kezdjük, hogy a megmagyarázandó jelenséget ilyen terhelt kifejezésekkel írjuk le, azzal a legfontosabb kérdések közül sokat felvetünk, talán anélkül, hogy ennek tudatában lennénk. Például azzal, hogy Searle elismeri a tudat "józan ész szerinti" definícióját, felveti azt a kérdést, hogy a tudat egy mindent vagy semmit jelenség-e vagy sem. Ezt úgy teszi, hogy olyan nyelvet fogad el, amelyben az "Ez az entitás egy kicsit tudatos" mondat rosszul hangzik, mert *nem grammatikus*. Természetesen a kérdésfeltevés még akkor is lehetséges, ha az ilyen nyelvi csapdákat elkerüljük - végül is Kennett elismeri, hogy a dualizmus ellen felveti a kérdést -, de ha valaki úgy kezdi a vizsgálatot, hogy a tudatot bármilyen módon, különösen a köznyelvben definiálja, az valószínűleg legalábbis véletlen kérdésfeltevést eredményez. A legtöbb filozófus egyetért abban, hogy ez a legjobb esetben is hanyagság, a legrosszabb esetben pedig számos lehetőséget kizár, és megalapozatlan következtetésekhez vezet. Ha már fontos kérdéseket teszünk fel, legalább szándékosan tegyük.

Persze Dennett nem tehet róla, hogy néhány metafizikai feltevésből indul ki, és így néhány fontos kérdést felvet (ahogy Rorty megjegyzi), de mindent megtesz, hogy ezt elkerülje azzal, hogy nem indul ki a tudat semmilyen definíciójából. Ehelyett azokra az interszjektíven ellenőrizhető jelenségekre összpontosít, amelyeket általában a tudattal társítunk, kidolgoz egy koherens elméletet, amely képes megmagyarázni ezeket a jelenségeket, és csak amikor a modell elkészült, nézi meg, hogy tartalmaz-e olyan horgokat, amelyekre felakaszthatjuk a tudattal kapcsolatos korábbi elképzeléseinket. Íme a horgok, amelyeket kiválaszt: egy mentális tartalom akkor és csak akkor *tudatos*, ha rendelkezésre áll, hogy a későbbi feldolgozás során felhasználható legyen, ami a nyílt (különösen verbális) viselkedésben való kifejeződéséhez vezet. Mivel ez a kritérium gyakran nem jól meghatározott abban az időpontban, amikor egy adott tartalom rögzül, bármely tartalom tudatos vagy nem tudatos státuszának ténye valóban a jövőbeli megismeréstől függ - nincs tényszerűség arról, hogy mely mentális tartalmak tudatosak bármely adott pillanatban. Egy mentális tartalom tulajdonságai, vagyis a kvalitásai pontosan az összes reaktív diszpozíció összessége, amely az adott tartalom jelenlétének eredményeként a szervezetben rejlik, és e tulajdonságok kellemes vagy visszataszító jellege annak köszönhető, hogy e reaktív diszpozíciók

között nagyon ősi, ősi törekvés jelen van.

és elkerülő reakciók. Az *én* egy magyarázó pozíció, a narratív súlypont, amelyet egy szervezet (ebben az esetben az emberi) hoz létre a narratívák folyamatos fonásának folyamatában, amely a magasabb szintű megismerés. Az *én* azonban nem illúzió, mert vannak olyan igaz tények, amelyek a legkönnyebben az *én* fogalmaiban állapíthatók meg – például maga az a tény, hogy az elbeszélések egy bizonyos halmaza, amellett, hogy mindannyian beszédhangokként ugyanazon emberi test ugyanazon nyílásából származnak, logikailag is egyetlen entitás önéletrajzoként áll össze.

Ahogy Dennett a *Consciousness Explained (A tudat megmagyarázva)* című könyvben végig érvel, ezek az azonosítások nem állnak ellentétben a tudattal kapcsolatos olyan tényekkel, amelyeket az interszjektíven nem vitatott premisszákból kiinduló érvekkel alá lehet támasztani. Ebben a tanulmányban megpróbáltam azonosítani és megkérdőjelezni néhány olyan előfeltevést, amelyek a *Consciousness Explained-re* a filozófiai közösségen belül érkező túlnyomóan negatív kritikai reakciók alapjául szolgálnak. Ezek a premisszák a dualizmus életképességétől a belső tulajdonságok létezéséig terjednek, de messze a leggyakoribb és legelterjedtebb az az egyszerű állítás, hogy *már tudjuk*, mi a tudat, hogy már rendelkezünk egy leírással róla, és hogy a magyarázat feladata az, hogy az empirikus tudomány adatait felhasználva igazolja *ezt* a már meglévő leírást. Ezért tanulságos, hogy Dennett kritikusi közül oly sokan vidáman kétségbeesnek azon, hogy valaha is találnak ilyen magyarázatot. Dennett és az *én* véleményem szerint az, hogy egy jelenség egy bizonyos leírása garantálja, hogy soha nem lesz rá magyarázat, *kizárja* ezt a leírást, hacsak nem lehet meggyőzően bizonyítani, hogy minden leírás ezt garantálja. Mivel ilyen érv nincs (hiszen Dennett kritikusi közül senki sem veszi figyelembe a tudat alternatív *leírásait*), Dennett joggal javasolja, hogy a magyarázathoz egy teljesen új leírásra van szükség, de egy ilyen új leírás elfogadása azt jelenti, hogy el kell engednünk néhány legkedvesebb intuíciónkat. Ezt a lépést Dennett hajlandó megtenni, dacolva kritikusával. Egyes kritikusok azt állítják, hogy nemesebb, hősiesebb és az emberi természethez igazabb, ha kitartunk valamilyen különleges lényeg mellett, amely tudatossá és emberré tesz bennünket. Dennett azt kérdezi, hogy nem hősiesebb-e elengedni az előítéleteket, és vállalni a tudat természetének őszinte vizsgálatából fakadó következményeket? És *én* azt válaszolom: igen, az.

#### HIVATKOZÁSOK

- "Szimpózium: Dennett: A tudatosság magyarázata." *Inquiry* 36 (1993): 3-160.  
 Dahlbom, B. Ed. *Dennett és kritikusi*. Cambridge: Blackwell Publishers, 1993.  
 Dennett, D. *A tudatosság magyarázata*. Boston: Brown & Co., 1991.  
 ----- "Living on the Edge." *Inquiry* 36 (1993).

- Fellows, R. és A. O'Hear. "A tudatosság elkerülése." *Inquiry* 36 (1993). Foster, J. "Dennett elutasítása a dualizmusról". *Inquiry* 36 (1993).
- Hofstadter, D. és D. Dennett. *Az elme énje: Fantasies and Reflections on Self and Soul*. New York: Bantam Books, 1981.
- Lockwood, M. "Dennett elméje". *Inquiry* 36 (1993).
- Nagel, T. *Más elmék: Nagel 1969-1994* :Kritikai esszék  
University Press, 1995. :Oxford
- "Milyen érzés denevérenek lenni?" in Hofstadter, D. és D. Dennett. *Az elme énje* (1974).
- Rorty, R. "Holizmus, intrinzikalitás és a transzcendencia ambíciója" in Dahlbom, B. Ed. *Dennett és kritikusai* (1993).
- Searle, J. "Minds, Brains, and Programs." in Hofstadter, D. and D. Dennett. *Az elme énje* (1980).
- Sprigge, T. "Egy kiábrándult zimbó?" *Inquiry* 36 (1993).

