

## Interpretacionizmus, mentális okozás, fizikalizmus és Davidson szintézis-kísérlete

Az analitikus elmefilozófiában a logikai behaviorizmus viszaszorulása után az elme oksági-reprezentációs elméletei, a materialista típusazonosság-elmélet és a funkcionalizmus mellett megjelent egy olyan, Quine, Davidson és Dennett nevével fémjelezhető értelmezési irány, amit összefoglalóan „interpretacionizmusnak” nevezhetünk. Ezek az elképzelések számos részletben különböztek egymástól, mégis egy csoportba sorolhatók az intencionális terminusokkal kapcsolatos felfogásuk alapján. Az interpretacionisták egyfelől elfogadták Chisholm tézisét, miszerint az intencionális idióma, így a „valamit jelenteni”, illetve a hit, vágy és más propozicionális attitűd terminusokat tartalmazó mondatok redukálhatatlanok a fizikai idiómára,<sup>1</sup> továbbá közös kiindulópontjuk, hogy az intencionális idióma *lényegileg másfajta*, mint a fizikai idióma, amely tényeket ír le, illetve magyaráz. Ahogy például Sellars fogalmaz:

A „'...' azt jelenti, hogy ...” a magja egy olyan egyedülálló beszédmódnak, amely épp annyira különbözik az empirikus tények leírásától és magyarázatától, mint az előírás és az igazolás nyelve (Chisholm–Sellars 1957. 527).

Az interpretacionisták tehát nemcsak azt állították, hogy az intencionális idióma redukálhatatlan a fizikaira, hanem azt is, hogy az intencionális idiómának vannak olyan általános vonásai, amelyekben radikálisan különbözik a fizikai idiómától (éppen annyira, ahogy a leírás és a magyarázat „nyelve” vagy „logikai tere” különbözik az előírás és az igazolás „nyelvétől” vagy „logikai terétől”),

<sup>1</sup> Az intencionális idióma redukálhatatlanságának tézisét általánosan megfogalmazhatjuk úgy, mint ami elveti a logikai behavioristák javaslatát, amely az intencionális mondatok jelentését olyan ún. redukciós mondatokra kívánta visszavezetni, amelyekben kizárólag fizikai terminusok, tehát ingerekre és viselkedésre (beleértve a nyelvi viselkedést) vonatkozó terminusok szerepelnek. Más megfogalmazásban a redukálhatatlansági tézis azt tagadja, hogy inger–viselkedési válasz kondicionálisokból következtethetünk intencionális mondatokra (szemben például Carnap 1936, 1938, 1947/1956, 1955, Morris 1946, Ayer 1947 javaslataival). A redukálhatatlansági tézis mellett Chisholm tanulmányok sorában érvelt, lásd pl. Chisholm 1952, Chisholm 1955–56, Chisholm 1957. Chisholm–Sellars 1957.

és az intencionális idióma redukálhatatlansága a fizikai idiómára ezen a lényegi különbségen alapul.

Az interpretacionista elméletek közös feltevése, hogy az intencionális predikátumok tulajdonítása *normatív elvek* alkalmazását feltételezi; ha egy személyt intencionális terminusokkal írunk le, hiteket és vágyakat tulajdonítunk neki, akkor az illetőnek tulajdonított propozicionális attitűdöknek meg kell felelniük például „a jóhiszeműség elvének” (charity-principle)<sup>2</sup> vagy bizonyos racionalitási követelményeknek.<sup>3</sup> Ez azt jelenti: egy személy intencionális állapotait nem lehetséges ugyanolyan módon azonosítani, mint a fizikai állapotait. Míg egy személy fizikai állapotai elvben azonosíthatók pusztán a személy viselkedése és a rá ható ingerek alapján, intencionális állapotai, az, hogy mit hisz és mire vágyik, nem. Az intencionális állapotok azonosításához interpretációra van szükség, amely alkalmazza a normatív elveket (a racionalitási, illetve a jóhiszeműségi elveket) a személy viselkedésének értelmezésében. Következésképpen az intencionális állapotok, szemben a fizikai állapotokkal, nem interpretáció-függetlenek, abban az értelemben, hogy a meghatározott tartalmú hittel rendelkező szükséges feltétele az interpretációs elvek alkalmazása.

Ez nem jelenti azt, hogy szubjektív vagy értelmező-relatív kérdés lenne, hogy egy személynek milyen hitek tulajdoníthatók. A jóhiszeműségi elv, illetve a racionalitási elvek alkalmazása objektíve rögzíti egy személy hiteit, a személynek tulajdonított hitek ugyanazok lesznek, akárki is alkalmazza ezeket az interpretációs elveket. Azonban az, hogy egy személy milyen hitekkel és vágyakkal rendelkezik, nem olyan interpretáció-független tulajdonsága, mint az, hogy mennyi a testsúlya, vagy hogy milyen lelkiállapotban van, ezeknek a fizikai tulajdonságoknak a meghatározottsága ugyanis nem függ interpretációs elvek alkalmazásától.

Az intencionális idióma normativitását minden interpretacionista elfogadja. Ugyanakkor különböző állásponton vannak az intencionális állapotok ontológiájával kapcsolatban, különbözőképpen vélekednek arról, hogy az intencionális és a fizikai idiómák különbözősége mit implikál az intencionális és a fizikai jelenségek viszonyával kapcsolatban. Az alábbiakban Quine, Dennett és Davidson reprezentatív nézeteit emelem ki.

Quine (1960) nézete az intencionális állapotok létezéséről kettős. Egyfelől azt állítja, szigorú értelemben nem léteznek hitek, vágyak és más propozicionális attitűdök. Emellett alapvetően a létezéssel kapcsolatos nézeteire támaszkod-

<sup>2</sup> A jóhiszeműség elve röviden azt állítja, hogy amikor egy személy nyelvi megnyilvánulásait interpretáljuk, fel kell tételeznünk, hogy az illető hiteinek a többsége igaz. Vö. Davidson 1973.

<sup>3</sup> Mint például Dennett 1981/1996. A racionalitási elvek között szerepelnek bizonyos elemi logikai következtetések, például a modus ponens, illetve a cselekvés racionalitását jellemző olyasféle egyszerű elvek, hogy ha valaki  $p$ -t szándékozik elérni, és azt hiszi, hogy ehhez  $q$ -t kell tennie, akkor *ceteris paribus*  $q$ -t fogja tenni.

va érvel. Általánosan szólva, Quine szerint csak azok az entitások valóságosak, amelyeket a világ legjobb tudományos elmélete posztulál. A világ legjobb tudományos elmélete Quine fizikalizmusa szerint a (legjobb) mikrofizikai elmélet, tehát a mikrofizikai entitások léteznek, valamint mindaz, amit meghatároznak. A mikrofizikai tények meghatározzák a makrofizikai tényeket, beleértve a viselkedésekkel kapcsolatos tényeket is, ezek tehát léteznek. A viselkedési tények azonban aluldeterminálják a beszédmegnyilvánulások jelentését, így jelentés mint entitás nem létezik. A jelentés aluldetermináltsága ugyanakkor öröklődik a propozicionális attitűdökre is: a hitek propozicionális tartalmát nem határozzák meg a fizikai tények, azaz a nyelvi viselkedés és a stimulusok (szemben bizonyos logikai behaviorista javaslatokkal, így például Carnap, Ayer vagy Morris bizonyos elképzeléseivel).<sup>4</sup> Mivel a fizikai tények nem határozzák meg egyértelműen a propozicionális attitűdök tartalmát, így hitek, vágyak és más propozicionális attitűdök szigorú értelemben nem léteznek.

Másfelől viszont Quine szerint az intencionális idióma mégiscsak nélkülözhetetlen az emberi viselkedés leírásában. Ezért kettős mérce alkalmazását javasolja: az intencionális idióma, fogalomrendszer tudományos elméletként, amely a világ végső létezőit ragadja meg, nem jó. Másfelől viszont nem is kell tudományos elméletként értelmezni, az intencionális idióma másfajta célokat szolgál.<sup>5</sup>

Nem mondanék le az intencionális idióma hétköznapi használatáról, és nem állítanám, hogy az intencionális idióma gyakorlati szempontból nélkülözhető volna. Azonban a kanonikus jelölés megkettőzése szükséges. (...) Amennyiben a valóság végső és igaz szerkezetét akarjuk leképezni, akkor a kanonikus jelölés az a szigorú, aszkétikus séma, amely nem ismer (...) propozicionális attitűdöket, kizárólag az organizmusok fizikai alkotórészeit és viselkedését. (...) De ha a kanonikus jelölést úgy értjük, mint ami csak arra szolgál, hogy felszámoljuk vele a verbális zűrzavarokat (perplexities) és logikai dedukciókat eszközöljünk velük, akkor gyakran tanácsos a propozicionális attitűdök idiómáját tolerálni (Quine 1960. 221).

Dennett másképp értelmezi az intencionális állapotok létezését; felfogása Ryle-éhoz áll közel. Mint Haugeland írja Dennetről:

<sup>4</sup> Vö. az 1. lábjegyzettel.

<sup>5</sup> Vö. Quine 1960. 216–221.

»Dennett valószínűleg egyetértene Ryle-lal: Teljesen szabályszerű és helyénvaló dolog azt mondani – egy bizonyos logikai hanghordozásban –, hogy létezik szellem, és azt mondani – másfajta logikai hanghordozásban –, hogy létezik test. De ezek a kifejezések nem a létezés két különböző fajtát jelzik, mert a 'létezés' nem olyan generikus fogalom, mint a 'színes' vagy az 'ivaros'«” (Ryle 1949/1974. 29).<sup>6</sup>

„Haugelandnak igaza van – írja Dennett –, Dennettnek egyet kell értenie, és egyet is értett Ryle-lal, éppen ebben a kérdésben.”<sup>7</sup>

Az intencionális idióma által megragadott jelenségek igenis valóságosak; igaz, a hitek és vágyak azonosítása feltételezi az interpretációs elvek alkalmazását, amit a fizikai entitásoké nem, de ettől még létezik a hitek és vágyak világa, csak éppen másféle értelemben, mint a fizikai entitásoké. Tévedés azt gondolni, hogy a létezés generikus fogalom, kategóriahiba, hogy a test és az elme, a fizikai és az intencionális jelenségek ugyanabban az értelemben valóságosak.<sup>8</sup>

Davidson nézete az intencionális állapotok természetéről az interpretacionizmust egy erőteljesebb realizmussal igyekszik ötvözni. Első megközelítésben úgy tűnik, mintha Dennett nézetéhez állna közel, azaz azt állítaná, az intencionális állapotok valóságosak, csak éppen másféle értelemben, mint a fizikai állapotok. Davidson ugyanis hangsúlyozza a mentális jelenségek valóságosságát és önállóságukat, tehát hogy az intencionális jelenségek normativitásuk és holisztikus jellegük miatt alapvetően különböznek a fizikai jelenségektől. Másfelől azonban felfogása különbözik Dennettétől abban, hogy sokkal inkább névértéken veszi a fizikalista ontológiát. Noha hangsúlyozza a mentális és a fizikai jelenségek alapvető különbözőségét, emellett a mentális és a fizikai *események* (példány)azonosságát is állítja.

A mentális jelenségek interpretacionista megközelítésének legfőbb képviselői az analitikus filozófiában, Quine, Dennett és Davidson mindannyian megpróbálták ötvözni az interpretacionizmust a fizikalizmus valamilyen, egymástól is különbözően értelmezett formájával. A fizikalizmus Quine esetében azt jelenti: azok az entitások léteznek (szigorú értelemben), amelyeket a legjobb tudományos elmélet posztulál, illetve amelyek a legjobb tudományos elméletre redukálhatók, és mivel a világ legjobb tudományos elmélete a mikrofizikai elmélet, ezért a mikrofizikai entitások léteznek, illetve ami ezekre redukálható. Dennett fizikalizmusa abban a tézisében ragadható meg, hogy minden valóságos „viselkedő rendszer” leírható fizikai hozzáállással is (a tervezéses és az intencionális

<sup>6</sup> Ryle így folytatja: „Az említett kifejezések a »létezik« két különböző jelentését mutatják, némiképp úgy, ahogy az »emelkedik«-nek különböző jelentése van a »vízszint emelkedik«-ben, a »hangulat emelkedik«-ben és az »átlagos életkor emelkedik«-ben (...). Épp ilyen jó vagy éppen rossz vice lenne azt mondani, hogy léteznek prímszámok és szerdák, vélemények és hajóhadak; vagy hogy létezik emberi test és szellem.” Ryle 1949/1947. 29.

<sup>7</sup> Vö. Dennett 1987/1989. 348–349.

<sup>8</sup> Dennett 1987/1989. 348–349.

hozzállás mellett), Davidsoné pedig abban, hogy szerinte minden entitás (így a mentális entitások is) azonosíthatók fizikai lefrással is. Az analitikus interpretacionisták e tekintetben tehát gyökeresen szemben állnak azzal a Dilthey-től kezdődő hermeneutikai hagyománnyal, amely (az analitikus interpretacionistákhoz hasonlóan) az interpretációt a szellemi jelenségek meghatározásának, avagy megértésének konstitutív elemeként értelmezi, erre alapozva azonban a mentális és a fizikai jelenségek természetének alapvető különbségét állítja.<sup>9</sup>

\*

Az alábbiakban arra a kérdésre fogok koncentrálni, hogy az analitikus interpretacionizmus hogyan értelmezi a mentális okozás jelenségét. Ez a kérdés önmagában is érdekes, de általánosabb jelentősége is van, tudnillik áttételesen arra a kérdésre is kihat, hogy az interpretacionizmus és a fizikalizmus együtt tartható-e. A fizikalizmus egyik neuralgikus pontja ugyanis a mentális jelenségek oksági hatásának összegegyeztetése a fizikalizmussal. Ilyenformán, ha sikerül alátámasztanom, hogy ennek a problémának a kezelésére az interpretacionistáknak sincs elfogadható javaslatuk, azzal egyben aláásom a mentális jelenségek analitikus interpretacionista értelmezését is.

A problémát, amit a mentális okozás felvet a fizikalisták számára, bevett szokás a túldetermináció vs. mentális jelenségek epifenomenalitásának dilemmájaként megfogalmazni. Ha a mentális jelenségek valóban okságilag hatékonyak fizikai jelenségek előidézésében, ugyanakkor a fizikalizmus alapfeltevése szerint minden fizikai eseménynek van elégséges fizikai oka (a fizikai világ oksági zártságának tézise), akkor úgy tűnik, a mentális jelenségek által okozott fizikai jelenségek túldetermináltak. Az alternatíva az, hogy a mentális jelenségek okságilag hatékonysága illúzió, a mentális jelenségek epifenomének.

Az antinaturalista hermeneutikai és fenomenológiai hagyományban a 19. század vége óta bevett stratégia ezt a problémát arra a tézisre hivatkozva kikerülni, hogy a szellemi jelenségek magyarázatának logikája, ideértve a tudatos emberi cselekvéseket és a kulturális jelenségeket (tehát az objektív vagy szupraindividuális szellem manifesztációit) is, más típusú, mint a természeti jelenségeké. Ennek az általános episztemológiai-tudományelméleti tézisnek egy változata az a nézet, amely szerint a mentális jelenségeknek nem lehet ugyanabban az értelemben oksági képességeket tulajdonítani, mint a fizikaiaknak. E nézet egy jellegzetes megfogalmazása Husserltől:

<sup>9</sup> A „szellemi” és a „mentális” (avagy „pszichológiai”) jelentése ebben a hagyományban nem esik egybe: a szellemi jelenségekhez tartoznak az egyéni psziché jelenségei mellett az „objektív szellem” vagy kevésbé hegelizálón a „szupraindividuális szellem” (a társadalmi intézmények, műalkotás etc.) jelenségei is.

(Pszichikai) fenoméneknek természetet tulajdonítani, reális alkotórészeiket, kauzális összefüggéseiket kutatni – éppoly tiszta képtelenség, mintha a számok kauzális tulajdonságai, összefüggései iránt érdeklődnénk. Képtelenség, mert olyasvalamit naturalizálnánk, aminek lényege az, hogy léte nem lehet természeti jellegű (Husserl 1911/1972. 147).

Illetve:

Ami az oksági viszonyt illeti, azt kell mondanunk, amennyiben *okságon* azt a függvényjellegű vagy törvényjellegű függést értjük, ami a Természet típusába tartozó perzisztens reális dolgok perzisztens tulajdonságai konstitúciójának korrelátuma, akkor *a lélek vonatkozásában egyáltalán nem beszélhetünk okságról...* A (lélek) olyan létező, amely feltételesen kapcsolódik az élő test bizonyos körülményeihez, a fizikai természet körülményeihez. És, hasonlóképpen, a lélek jellemezhető azon ténnyel, hogy bizonyos lelki események, rendszeres módon, következményekkel járnak a fizikai természetben. De sem az élő test, sem a lélek nem tesz szert ezáltal »természeti tulajdonságokra«, a logiko-matematikai természet értelmében (Husserl 1989. 140).

Úgy is meg lehet ezt fogalmazni, hogy a mentális jelenségek oksági hatását a fizikai világra nem értelmes a természeti okság kategóriáiban leírni, ilyenformán a mentális oksági fizikalista problémája fenti formájában értelmetlen, tudniillik a mentális jelenségekre alkalmazni a fizikai jelenségekre érvényes okságfogalmat értelmetlenség.<sup>10</sup>

Kétséges azonban, hogy a fizikalista analitikus interpretacionisták számára is ilyen könnyen elintézhető-e ez a probléma. Ha ugyanis az interpretacionizmust az intencionális állapotok metafizikai realista felfogásának tekintjük (ha mégoly „puha” realistának is), akkor mondanunk kell valamit a mentális okozásról, tehát arról, hogy az intencionális állapotok és események hogyan járulnak hozzá a cselekvés okozásához. Ez csak akkor kerülhető el, ha az interpretacionizmust semmilyen értelemben nem tekintjük metafizikai realista felfogásnak, ahogy leginkább Quine értelmezhető. Ha azonban úgy tartjuk, az intencionális állapotok valóságosak, noha más értelemben, mint a fizikaiak, ahogy Dennett és Davidson is gondolja,<sup>11</sup> akkor felmerülhet, hogy oksági hatásokról számot kel-

<sup>10</sup> Általános vonásait tekintve hasonló nézeteket képviselt már korábban Dilthey 1894; vagy egy másik hagyományból Ryle 1949/1974 is.

<sup>11</sup> Davidson realista a mentális eseményekkel és állapotokkal kapcsolatban, e tekintetben egyértelműen elhatárolja magát Quine álláspontjától. Vö. pl.: „Az anomáliás monizmus nem azt állítja, hogy amikor egy ágensnek mentális eseményeket és állapotokat tulajdonítunk, akkor ezeket pusztán projektáljuk az ágensre; éppen ellenkezőleg, az anomáliás monizmus szerint a mentális események éppen olyan valóságosak, mint a fizikai események, lévén hogy azonosak velük, és a mentális állapotok tulajdonítása éppen ilyen objektív. Quine ábrázolása, miszerint az attitűdtulajdonítások drámai ábrázolások, nem implikálja, hogy ne lenne valami, amit ábrázolnak.” Davidson 1997/2001. 72.

lene adni valahogy. Dennett semmilyen elmélettel nem szolgál a mentális okozásról. A dennetti felfogás távolabbi előzményei, mint például Ryle (1949/1974) nézetei, persze értelmetlennek tekintették az oksági fogalmak alkalmazását a mentális jelenségekre; a naturalista Dennett azonban vélhetően nem osztja ezt az elképzelést.

Quine és Dennett tehát egyáltalán nem foglalkoztak a mentális okozás problémájával, Davidson azonban kidolgozott egy sajátos elméletet. Az alábbiakban ezt az elképzelést fogom részletesen vizsgálni. Tézisem az, hogy Davidson javaslata a mentális okozás értelmezésére nem kielégítő: alapvetően azért, mert a mentális jelenségek természetének interpretacionista jellemzése nem illeszkedik azzal a képpel, ahogy a mentális jelenségeknek a fizikai világra gyakorolt oksági befolyását értelmezi.

Davidson mentális okozás elméletének a fő elemei a következők. (1) Az oksági kapcsolat események közötti redukálhatatlan metafizikai viszony, (2) amelynek fennállása olyan metafizikai tény, amely független attól, hogyan írjuk le az ok- és az okozat-eseményt. (3) A mentális okozás valóságos: bizonyos mentális események okoznak fizikai (vagy más mentális) eseményeket, és (2)-nek megfelelően az ilyen relációk fennállása vagy nem fennállása független attól, hogyan írjuk le az eseményeket. Továbbá (4) minden mentális esemény példányazonos valamilyen fizikai eseménnyel. Ez azt jelenti, minden mentális leírással azonosítható esemény azonosítható fizikai leírással is. (5) A mentális és a fizikai események közötti oksági kapcsolatok, amennyiben a mentális eseményeket fizikai leírással határozzuk meg, fizikai törvények alá vonhatók.<sup>12</sup>

Ez a koncepció a következőképpen kezelné a mentális okozás fizikalista problémáját. (5) A mentális és a fizikai események közötti oksági kapcsolatok, amennyiben a mentális eseményeket fizikai leírással határozzuk meg, fizikai törvények alá vonhatók, így a mentális és fizikai események közötti oksági viszonyok beilleszthetők a fizikai folyamatok oksági-magyarázati rendjébe, ilyenformán a fizikai jelenségek oksági zártsága nem sérül. Másfelől a túldetermináció avagy mentális epifenomenalizmus dilemmát Davidson azzal szándékozik kikerülni, hogy azt állítja, a mentális események nem – mentális vagy fizikai – *tulajdonságaik* révén fejtik ki oksági hatásukat, hanem a mentális események metafizikailag egyszerű *esemény-partikuláréként* okoznak más *esemény-partikulárékat*. (Ha tulajdonságaik révén fejtenék ki oksági hatásukat, akkor nem lehetne kikerülni a túldetermináltság vagy epifenomenalizmus problémát, tudniillik azt a kérdést, hogy a mentális eseményeknek a mentális vagy a fizikai tulajdonságai hatékonyak okságilag.)<sup>13</sup>

<sup>12</sup> Vö. Davidson 1970a/1980.

<sup>13</sup> Lásd erről Kim és Davidson vitáját. Vö. Davidson 1963/1980, 1970a/1980, 1973a/1980, 1993 és Kim 1990, 1992.

Nos, ez a javaslat véleményem szerint nem tartható. Alapvetően azért, mert a mentális okozás ilyen, a mentális és fizikai események példányazonosságára támaszkodó értelmezése, valamint Davidson interpretacionista felfogása a mentális attitűdökről nem illeszkedik jól egymáshoz: a példányazonossági tézis és „az indokok okok” tézis, amely szerint a cselekvés hitekre és vágyakra hivatkozó magyarázata oksági magyarázat, nem alkotnak koherens koncepciót.<sup>14</sup> Tömören a következők miatt. Davidson szerint (1) az indokok hit-vágy komplexumok, (2) a hitek és vágyak nem események, (3) az okok események. Tehát: az indokok nem okok. Részletesebben, a hitek és vágyak, amelyek a cselekvés indokai, attitűdök. Davidson szerint azonban az attitűdök nem entitások.<sup>15</sup> Az attitűdök a *személy* viselkedésének leírásmódjai, azt jellemzik, ahogy egy személy (mint egység) viselkedését interpretáljuk, illetve ahogy ő maga leírja a viselkedését.

Azonban az, hogy egy személynek milyen hitei és vágyai vannak, nem olyan interpretáció-független tulajdonsága, mint az, hogy milyen testmozgásokat végez, vagy hogy milyen agyállapotban van. Más szóval az, hogy egy személy milyen hitekkel és vágyakkal rendelkezik (pontosabban fogalmazva, hogy milyen hitek és vágyak tulajdoníthatók neki), nem *empirikus tény* a személy állapotával kapcsolatban, Davidson felfogása szerint. Miért?

Egyfelől pusztán a közvetlenül megfigyelhető empirikus adatok, a nyelvi viselkedés és a stimulusok nem elégségek ahhoz, hogy meghatározott tartalmú attitűdöket tulajdoníthassunk egy személynek, ehhez szükség van az interpretációs elvek alkalmazására. Másfelől maguk az interpretációs elvek nem empirikus tényeket fejeznek ki. Nem arról van szó, hogy *az interpretációs elvek tényeket fejeznének ki* az intencionális állapotokkal rendelkező lények *általános természetével* kapcsolatban. Azt, hogy az emberek hiteinek többsége igaz, illetve hogy követnek bizonyos elemi racionalitási elveket, nem empirikus módszerekkel, megfigyelések, induktív következtetések alapján fogadjuk el Davidson szerint. Nem az emberek (nyelvi és nem-nyelvi) viselkedésének és a rájuk ható ingereknek a megfigyelésére alapozva fogadjuk el ezeket az általános kijelentéseket a hitekről és vágyokról, hanem a viselkedés értelmezésében, a viselkedést magyarázó hitek és vágyak meghatározásához *előfeltételezzük* őket. A hitekkel és más attitűdökkel rendelkezés tehát nem olyan tulajdonsága a személyeknek, mint a fizikai tulajdonságaik, például az agyi-idegrendszeri tulajdonságaik.<sup>16</sup> Ezért Da-

<sup>14</sup> Vö. Davidson 1963/1980.

<sup>15</sup> Vö. pl. Davidson 1999. 654.

<sup>16</sup> Vö. „Nincsenek szigorú pszichofizikai törvények, a mentális és a fizikai séma lényegileg különböző (disparate) elköteleződése miatt. A fizikai valóság egyik jellegzetessége, hogy a fizikai változást meg lehet magyarázni olyan törvényekre hivatkozva, amelyek azt más fizikailag leírt változásokkal és feltételekkel kapcsolják össze. A mentális szféra egyik jellegzetessége, hogy a mentális jelenségek tulajdonítása meg kell hogy feleljen az individuum indokai, hitei és intenciói hátterének. Nem lehetséges szoros kapcsolat a szférák között, ha mindkét szféra esetében fenn akarjuk tartani evidenciáik megfelelő (egymástól különböző) forrását” (Davidson 1970a/1980. 222). Részletesebben Davidson 1970a/1980. 219–223.

vidson felfogása alapján *nem is értelmes* az attitűdök és az agyállapotok *korrelációjáról* beszélni, ez, ha fogalmazhatunk így, kategóriahiba.

De hogyan értjük akkor a mentális és a fizikai okok példányazonosságának a tézisé, miszerint vannak mentális ok-események, amelyek azonosak fizikai ok-eseményekkel? Milyen eseményekre vonatkozhat Davidson tézise?

Felmerülhet, hogy noha a hitek és vágyak nem események, a hitek és vágyak *manifesztálódása* lehet esemény. Gondolhatjuk úgy, hogy ezt az elképzelést ki lehet bontani az állandó (standing) és a manifesztálódó (occurrent) hitek és vágyak megkülönböztetésére támaszkodva. Eszerint az állandó hitek és vágyak diszpozíciók, amelyek bizonyos körülmények között manifesztálódnak, például valamilyen észlelési esemény hatására, vagy amikor hirtelen eszünkbe jut valami. De közelebbről hogyan értelmezzük a hiteket és vágyakat mint diszpozicionális állapotokat? A materialista típusazonosság-elmélet a hiteket mint diszpozíciókat úgy értelmezi, ahogy a diszpozicionális fizikai tulajdonságokat, tudniillik valamilyen mögöttes kategoriális tulajdonsággal azonosítja őket. A törekenység diszpozicionális tulajdonságot például valamilyen anyagszerkezeti tulajdonsággal azonosítja (egy tárgy összetörését, azaz törekenységének manifesztációját a tárgyra gyakorolt valamilyen erőhatás, például egy ütés és az anyagszerkezeti tulajdonság együtt okozza). Ez az elképzelés azonban nem egyeztethető össze Davidson nézetével, hiszen Davidson szerint nem értelmes a(z állandó) hiteket és vágyakat agyállapottal azonosítani. Mint írja: „A hitek és vágyak (véleményem szerint) nem kicsi entitások az agyban” (Davidson 1970a/1980. 222). Ha pedig az állandó hitek nem azonosíthatók meghatározott agyállapottokként, akkor a hitek oksági szerepe a cselekvések okozásában sem értelmezhető a fenti analógiával, a hitekkel és vágyakkal azonos agyállapotok oksági képességeinek manifesztálódásaként.

Egy másik elképzelés szerint, ahogy a mentális eseményeket cselekvések okaiként értelmezhetnénk, hit- és vágy-attitűdjeink maguk nem események, *megváltozásuk* azonban esemény. Tegyük fel például, hogy az a szándékom, elutazom valahova a 6 órakor induló vonattal. Továbbá azt hiszem, ehhez hamarosan el kell indulnom a pályaudvarra, és ezért szándékozom is hamarosan elindulni a pályaudvarra. Kisvártatva azonban észlelem, hogy már túl késő van, vagy eszembe jut, hogy valamit még el kell intéznem, mielőtt elutaznék, és ezen mentális események hatására rájövök, hogy már nem érem el a vonatot. Emiatt megváltozik az a szándék-attitűdöm, hogy hamarosan a pályaudvarra menjek. A probléma az, mint Davidson is rámutat, hogy ezt az attitűdváltozást csak metaforikusan érthetjük eseményként. Tudniillik ha maguk az attitűdök nem entitások, akkor az attitűdök megváltozása sem lehet entitás. Mint írja: „mivel a hitek és vágyak nem entitások, csak metaforikus értelemben beszélhetünk a megváltozásukról, és ennek a metaforának a kiterjesztése, mondhatni, okként és okozatként beszélni róluk” (Davidson 1970a/1980. 222). De ha így van, akkor az attitűdváltozás nem egy szó szerinti értelemben vett mentális esemény, amire vonatkozhatna a példányazonossági tézis.

Egy harmadik lehetőség lehetne esetleg, hogy bár az attitűdök nem események, de az, hogy a személy megfelelő attitűdökkel rendelkezzen (azaz olyan állapotban legyen, aminek az alapján az interpretációs elveket alkalmazva megfelelő – tudniillik a cselekvést magyarázó – hitek és vágyak tulajdoníthatók neki), szükséges feltétele annak, hogy mentális események (tudniillik észlelések vagy gondolati aktusok) fizikai eseményeket, cselekvéseket okozhassanak.

Ezen elképzelés mellett szól, hogy pusztán a mentális eseményekre hivatkozni nem elégséges a pszichofizikai oksági kapcsolatok leírásában, a megfelelő hit-, vágy-attitűdökre is szükség van. Vegyünk az előzőhöz hasonló példát. Tegyük fel, az a szándékom, hogy elutazom a 6 órás vonattal, és azt hiszem, ennek kivitelezéséhez  $\frac{1}{2}$  6-kor el kell indulnom a pályaudvarra.  $\frac{1}{2}$  6-kor észlelem, hogy  $\frac{1}{2}$  6 van (ránézek az óramra), és elindulok a pályaudvarra. Nyilvánvaló, hogy pusztán az az esemény, hogy észlelem, mennyi idő van, nem elégséges annak az oksági magyarázatához, hogy elindultam a pályaudvarra; a magyarázatban hivatkozni kell arra a szándékomra is, hogy a 6 órás vonattal elutazzak, illetve arra a hitemre is, hogy ehhez  $\frac{1}{2}$  6-kor el kell indulnom a pályaudvarra.

Azonban ez a javaslat is hasonló okokból problematikus, mint az előzők. Nevezetesen: a cselekvés pszichofizikai oksági magyarázatában szereplő „mentális ok” fogalma, ti. a hit–vágy–észlelés komplexum olyan „kores fogalom”, amelynek az elemei nem ugyanabba a kategóriába tartoznak; így homályos, valójában miféle dologról is van szó.

Felvethető esetleg, hogy bár a hit- és vágy-attitűdök nem azonosíthatók (mert nem korrelálhatók értelmesen) agyállapotokkal, ettől még fennáll, hogy a partikuláris pszichofizikai okozási folyamatokban a személy azon állapota „mögött” vagy „mellett”, amelynek az alapján a meghatározott hitek és vágyak tulajdoníthatók neki, valamilyen agyállapotban is van, és ez az agyállapot okságilag közreműködik a cselekvés okozásában.

Nos, egy ilyen minimális fizikalizmust mindenképpen el kell fogadni, ha ki akarjuk zárni, hogy a pszichofizikai oksági kapcsolatokat tisztán mentális okok és fizikai okozatok közötti közvetlen kapcsolatokként értelmezzük. De ha csak ennyi Davidson fizikalizmusának a tartalma, akkor nem világos, mennyiben nyújtja a pszichofizikai okozás fizikalista értelmezését.

Ugyanis: a cselekvés pszichofizikai oksági magyarázata és a hit–vágy–észlelés komplexum „mögötti” agyi állapotokra hivatkozó fizikai oksági magyarázata között a mentális ok és a fizikai ok közötti korreláció értelmetlensége miatt nincs kapcsolat. Ez azt jelenti, a minimális fizikalizmus egyáltalán nem járul hozzá a pszichofizikai oksági kapcsolat magyarázatához. A minimális fizikalizmus olyan természetlen doktrína, amely nem alapoz meg többet, mint ami explicite benne van a meghatározásában, tudniillik hogy a pszichofizikai oksági kapcsolatok kivitelezésében valamilyen agyállapotok valahogyan közreműködnek.

De Davidson felfogása, amely szerint az indokok okok, valamint a mentális események példányazonosak fizikai eseményekkel, ennél tartalmasabb értel-

mezést ígért a pszichofizikai oksági kapcsolat fizikalista értelmezésére. Hogy világosabb legyen a minimális fizikalizmus terméketlensége, vessük össze más javaslatokkal a pszichofizikai okozás fizikalista értelmezésére. Egyfelől a redukcionista típusazonosság-elmélet szerint a mentális állapot típusok azonosak agyi állapot típusokkal, következésképp a pszichofizikai oksági folyamatok típusazonosak bizonyos agyi folyamatokkal. Másfelől az olyan nem-redukcionista fizikalista elméletek szerint, amelyek csak a pszichofizikai példányazonosságot tartják fenn, a típusazonosságot nem, a pszichofizikai oksági folyamatok példányazonosak bizonyos agyi oksági folyamatokkal. Davidson minimális fizikalizmusa szerint azonban sem típus-, sem példányazonosság nem áll fenn a pszichofizikai oksági kapcsolatok és az agyi oksági folyamatok között. A cselekvés pszichofizikai és fizikai oksági magyarázata között így nincs semmilyen kapcsolat.

\*

Foglaljuk össze az eddigieket, illetve idézzük fel újra kissé a kontextust is. Davidson az „Actions, Reasons and Causes” (1963) című tanulmányában amellelt foglalt állást, a hermeneutikai eredetű és a késői Wittgensteinhez is kapcsolódó hagyománnyal szemben, hogy „az indokok okok”, a hitekre és vágyakra hivatkozó magyarázatok lehetnek oksági magyarázatok. Ennek a davidsoni javaslatnak a következményei szerteágazóak, az egyik mindenestre az volt, hogy megnyitotta az utat a hitek és vágyak redukcionista materialista értelmezése számára,<sup>17</sup> amennyiben hatásosan érvelt amellelt, hogy a racionalizáló magyarázatok lehetnek oksági magyarázatok, és így nem eleve értelmetlen (fogalmi ellentmondás) a hitek és vágyak és a fizikai állapotok kapcsolatára vonatkozó kérdés. Davidson azonban nem volt redukcionista: az anomáliás monizmus tanával (1970a/1980) olyan nézetet akart kifejtteni, amelyik nem redukcionista a mentális jelenségekkel kapcsolatban, ugyanakkor összeegyeztethető a fizikalizmusmal; a mentális-fizikai típusazonosság elvetése egyfelől (a mentális és a fizikai predikátumok azonosítási kritériumainak „lényegi különbözősége” miatt),<sup>18</sup> a mentális és fizikai eseménypéldányok azonosságának a fenntartása másfelől, ezt a célt próbálta beteljesíteni. A fenti vizsgálat, azt hiszem, megmutatta, hogy ezt a célkitűzést nem sikerült elérni. Mégpedig azért nem, mert Davidson más nézetei, jelesül a mentális attitűdök interpretacionista értelmezése, nem teszi lehetővé a hitek és vágyak példányazonosítását sem agyi állapotokkal. Emiatt viszont homályos, mi is lenne a pszichofizikai oksági kapcsolatok fizikalista értelmezésének a tartalma. A minimális fizikalizmus, ami csak annyit állít, hogy

<sup>17</sup> A redukcionista materializmus legrészletesebben kidolgozott korabeli elmélete Armstrong 1968 volt.

<sup>18</sup> Vö. a 15. lábjegyzettel.

a pszichofizikai oksági kapcsolatoknak van valamilyen fizikai-agyi bázisuk, összefér Davidson interpretacionizmusával, azonban semmilyen magyarázóerővel nem bír a pszichofizikai oksági kapcsolatok és az ezek bázisát képező fizikai-agyi folyamatok viszonyáról.

## IRODALOM

- Ayer, Alfred Jules 1947. *Thinking and Meaning*. London, H. K. Lewis.
- Armstrong, David 1968. *A Materialist Theory of Mind*. London, Routledge.
- Carnap, Rudolf 1936/1972. Jelentés és ellenőrizhetőség. Ford. Altrichter Ferenc. In Altrichter Ferenc (szerk.) *A Bécsi Kör filozófiája*. Budapest, Gondolat.
- Carnap, Rudolf 1938. The Logical Foundations of the Unity of Science. *International Encyclopedia of Unified Science* Vol. I. Chicago, The University of Chicago Press. 42–62.
- Carnap, Rudolf 1947/1956. *Meaning and Necessity*. Chicago, The University of Chicago Press.
- Carnap, Rudolf 1955a. Meaning and Synonymy in Natural Languages. *Philosophical Studies* 6/3. 33–47.
- Chisholm, Roderick M. 1952. Intentionality and the Theory of Signs. *Philosophical Studies* 3/4. 56–63.
- Chisholm, Roderick M. 1955–1956. Sentences About Believing. *Proceedings of the Aristotelian Society* (New Series) 56. 125–148. Újranyomva in *Minnesota Studies in the Philosophy of Science* 2. 125–148.
- Chisholm, Roderick M. 1957. *Perceiving. A Philosophical Study*. Ithaca, Cornell University Press.
- Chisholm, Roderick M. – Sellars, Wilfrid S. 1957. Intentionality and the Mental: A Correspondence. *Minnesota Studies in the Philosophy of Science* 2. 507–539.
- Davidson, Donald 1963/1980. Actions, Reasons, and Causes. In *Essays on Action and Events*. Oxford, Clarendon Press.
- Davidson, Donald 1970a/1980. Mental Events. In *Essays on Action and Events*. Oxford, Clarendon Press.
- Davidson, Donald 1973a/1980. The Material Mind. In *Essays on Action and Events*. Oxford, Clarendon Press.
- Davidson 1973b/1980. Radical Interpretation. In *Essays on Action and Events*. Oxford, Clarendon Press.
- Davidson, Donald 1993. Thinking Causes. In John Heil – Alfred Mele (szerk.) *Mental Causation*. Oxford, Clarendon Press.
- Davidson, Donald 1997/2001. Indeterminism and Antirealism. In *Subjective, Intersubjective, Objective*. Oxford, Clarendon Press.
- Davidson, Donald 1999. Reply to Bruce Vermazen. In Lewis E. Hahn (szerk.) *The Philosophy of Donald Davidson*. Chicago–La Salle, Open Court.
- Dennett 1981/1996. Az igazhívők és az intencionális stratégia sikerének forrásai. In Pléh Csaba (szerk.) *Kognitív tudomány*. Ford. Thuma Orsolya. Budapest, Osiris.
- Dennett, Daniel C. 1987/1989. Mid-Term Examination: Compare and Contrast. In *The Intentional Stance*. Cambridge MA, MIT Press.
- Husserl, Edmund 1911/1972. *A filozófia mint szigorú tudomány* (részletek). Ford. Baránszky Jób László. In *Válogatott tanulmányok*. Budapest, Gondolat.
- Husserl, Edmund 1913. *Ideen zu einer reinen Phänomenologie und phänomenologischen Philosophie*, Erstes Buch. Halle, Max Niemeyer.

- Husserl Edmund 1989. *Ideas Pertaining to a Pure Phenomenology and to a Phenomenological Philosophy, Second Book. Collected Works*, Vol. 3. Ford. Richard Rojcewicz – André Schuwer. Dordrecht, Kluwer.
- Kim, Jaegwon 1990. Explanatory Exclusion and the Problem of Mental Causation. In Enrique Villanueva (szerk.) *Information, Semantics, and Epistemology*. Oxford, Blackwell.
- Kim, Jaegwon, 1992. Multiple Realization and the Metaphysics of Reduction. *Philosophy and Phenomenological Research* 52. 1–26.
- Morris, Charles W. 1946. *Signs, Language and Behavior*. New York, Prentice-Hall.
- Quine, Willard Van Orman 1948/2002. Arról, hogy mi van. Ford. Eszes Boldizsár. In *A tapasztalattól a tudományig*. Budapest, Osiris.
- Quine, Willard Van Orman 1960. *Word and Object*. Cambridge MA, MIT Press.
- Ryle, G. 1949/1974. *A szellem fogalma*. Ford. Altrichter Ferenc. Budapest, Gondolat.