

Un argument probabilistic pentru teza identității minte-creier

István Aranyosi
Bilkent University, Ankara

In this paper I offer a new, probabilistic argument for the mind-brain identity thesis, put forward by U.T. Place, H. Feigl, and J.J.C. Smart in the 1950s. After considering the epistemic, or conceivability based arguments against physicalism, I build an argument to the effect that naturalistic dualism - the view that phenomenal properties do not metaphysically supervene on physical properties, but they are nomically connected - is probabilistically incoherent. The conclusion will be that physicalism, in the form of the identity thesis, is almost surely true.

Keywords: physicalism, mind-brain identity thesis, conceivability arguments, zombies, qualia inversion, Herbert Feigl, indifference principles, probability.

Teza identității minte-creier își începe cariera - lăsând deoparte aserțiunile anterioare, frugale argumentativ și conceptual, ale unor filosofi și oameni de știință - în a doua jumătate a anilor 1950, cu operele lui Ullin Place (1956), Herbert Feigl (1958) și Jack Smart (1959). Din perspectiva prezentului, studiul lui Feigl, "The 'Mental and the Physical'", este cel mai remarcabil dintre cele trei *loci classici* mai sus amintite, prin faptul că este de largă respirație, atât teoretic, cât și istoric, și de pionierat în mai multe privințe; putem identifica în acest studiu un număr mare de subiecte care sunt amplu dezbătute în zilele noastre în filosofia minții, cum ar fi: intenționalitatea, qualiile, corelatele neurale ale conștiinței, realizabilitatea multiplă, cauzarea mentală, reducția slabă și tare etc.

În acest studiu îmi asum sarcina de a reconsidera noțiunea de "ciucuri nomologici" (*nomological danglers*), introdusă de către Feigl, în lumina discuțiilor recente despre viabilitatea acomodării proprietăților fenomenale, a qualiilor, într-o imagine fizicistă asupra realității. Voi construi un argument având drept concluzie că ciucurii nomologici, inclusiv modul în care qualiile sunt înțelese ca intrând în relație cu proprietățile fizice de către dualiștii contemporani, sunt extrem de improbabili. Ultimul pas al argumentului, cel către plauzibilitatea postulării identității minte-creier, va coincide cu rațiunea principală oferită de către Feigl, anume parcimonia ontologică sau briciul lui Ockham. Însă, spre deosebire de Feigl și alți teoreticieni care au discutat noțiunea de ciucuri nomologici, subliniez că problema dualismului naturalist (adică nomologic) ține de partea "nomologică", nu de cea "atârnată". Voi începe cu o foarte scurtă trecere în

revistă a noțiunii lui Feigl și a rolului pe care-l joacă în argumentul său pentru teza identității minte-creier, după care trec în revistă evoluția dialecticii post-Feigl. Apoi ofer un argument menit să demonstreze că ciucurii nomologici sunt extrem de improbabili, singurele opțiuni coerente probabilistic fiind “ciucurii anomici” (nici măcar nomic corelați) și “ciucurii necesari” (corelați mai mult decât nomic, adică logic). După ce arăt, pe baza unui raționament probabilistic similar, că primul termen al disjuncției (ciucurii anomici) este extrem de improbabil, conchid, pe baza mai sus menționatului principiu al parcimoniei, ca teza identității este singurul candidat extrem de probabil pentru relația mental-fizic. Noutatea argumentului constă în faptul că aduce considerații probabilistice în favoarea fizicalismului, o strategie neglijată în literatura contemporană tot mai voluminoasă asupra acestui subiect.

1. Noțiunea de ciucur nomologic

Student al lui Moritz Schlick, fondatorul Cercului de la Viena, Feigl a fost primul membru al cercului care avea să emigreze în SUA, în 1930, devenind lector (1931), și apoi profesor (1933) la University of Iowa. A fost, de asemenea, singurul membru al cercului simpatizant de la bun început al realismului științific și al metafizicii sau ontologiei. Studiul “The ‘Mental and the ‘Physical’”, publicat în 1958, reflectă multe dintre angajamentele sale doctrinare. Însuși subiectul – problema minte-corp – nu obișnuia să figureze în scrierile colegilor de cerc mai mult decât ca un exemplu de pseudo-problemă, un candidat pentru eliminare. Obiectivul principal al lui Feigl este de a argumenta pentru coerența și plauzibilitatea tezei identificării empirice a minții cu creierul, pusă pe agendă ceva mai devreme de către Place.¹ Feigl începe prin a argumenta că paralelismul mental-fizic, adică existența unor legi de corelație între evenimente sau proprietăți mentale și fizice, este superior dualismului interacționist. Paralelismul este pur și simplu teza că există conexiuni nomice între cele două domenii, mental și fizic, astfel că este o noțiune presupusă a fi neutră între o interpretare dualistă, caz în care este echivalentă cu epifenomenalismul, și una fizicistă. Acceptarea paralelismului ca plauzibil este un prim pas către teza identității. Următorul pas este de a observa că aceste legi, dacă sunt ireductibile, sunt foarte diferite față de alte generalizări nomice prezente în științe. Astfel, Feigl scrie în capitolul 5, secțiunea B:

“Aceste legi de corelație sunt complet diferite de orice alte legi ale științei (de tip fizic), prin faptul că, în primul rând, ele sunt “ciucuri” nomologice, i.e. relații ce

¹ După cum nota David Armstrong (comunicare orală), “Place a fost cel de la care a pornit totul, dar, din păcate, și-a publicat studiul într-un loc nepotrivit.” Teza identității minte-corp își începe cariera în mod vizibil odată cu articolul lui J.J.C. Smart, publicat în 1959, și culminează cu scrierile lui Armstrong [1968] și David Lewis [1966, 1970, 1972], care oferă un argument funcționalist, bazat pe considerații semantice.

conectează evenimente intersubiectiv confirmabile cu evenimente ce *ex hypothesi* nu sunt, în principiu, intersubiectiv și independent confirmabile. Astfel, prezența sau absența datelor fenomenale nu reprezintă o diferență ce ar putea fi concepută a crea o diferență în datele confirmatorii fizici-observaționale, i.e. în comportamentul public observabil, sau, dacă vrem, în procesele neurale observate sau inferate de către neurofiziologi. În al doilea rând, aceste legi de corelație, spre deosebire de alte legi ale științelor naturale, ar fi (din nou *ex hypothesi*) absolut nederivabile din chiar premisele postulatelor celor mai cuprinzătoare și mai îmbogățite ale oricărei viitoare fizici sau biologiei teoretice.” (*traducere proprie*)

Feigl definește două noțiuni ale fizicului în capitolul 5, secțiunea A. “Fizic₂” se referă la tipul de concepte și enunțuri teoretice care sunt suficiente pentru explicarea enunțurilor observaționale privind domeniul anorganic (fără viață) al naturii, în timp ce “fizic₁” se referă la conceptele și enunțurile utilizate de către toate științele, care implică conexiuni logice și probabilistice cu limbajul observațional intersubiectiv. De fapt, potrivit lui Feigl, cele două domenii ale fizicului coincid în lumea actuală, ceea ce înseamnă că nu există proprietăți cu adevărat emergente.

Îndreptându-ne atenția acum asupra noțiunii de ciucuri nomologici, observăm că pentru Feigl ea se aplică la conexiuni nomice, și nu la termenii acestor conexiuni.² Deci, partea “nomologică” a sintagmei se referă la faptul că aceste corelații psiho-fizice reprezintă, într-adevăr, o formă de conexiune nomică. Pe de altă parte, partea “atârnată” se referă la faptul că, spre deosebire de alte legi științifice, legile psiho-fizice sunt stranii prin faptul că ele relaționează fenomenele standard, public observabile și confirmabile (stările cerebrale), cu fenomene ce sunt *ex hypothesi* exclusiv subiective (qualiile), astfel că ele nu fac (și nu pot face) nicio diferență explicativă în procesul de confirmare a unei ipoteze despre o potențială conexiune nomică.³

Noțiunea de atârnată nomologic nu este una incoerentă logic. Singura problemă cu ea, potrivit lui Feigl și Smart, este cu partea “atârnată”, însă aceea nu este o problemă de coerență logică; este mai curând o problemă de neadecvare cu *Zeitgeist*-ul naturalist și reduționist. Acestea fiind spuse, Feigl își completează argumentul pentru teza identității susținând că, odată ce teza paralelistă sau corelaționistă este acceptată ca plauzibilă, considerații de parcimonie ontologică și constrângerea metodologică de a repudia ciucurii nomologice ca redundanți fac din identificarea empirică (i.e. non-analitică) dintre tipurile de stări mentale și cele neurofiziologice singurul candidat plauzibil pentru relația

² Spre deosebire de Smart [1958, p. 142], care adoptă sintagma de la Feigl, dar îi schimbă înțelesul, astfel că ea se aplică qualiilor, presupuse a atârna nomologic, adică a fi ontologic distincte, dar nomic conectate cu ceea ce ar cuprinde o imagine științifică asupra lumii.

³ A se compara cu David Chalmers [1996], unde autorul folosește expresia “irelevanța explicativă a experienței” pentru a exprima aceeași idee.

mente-corp. Odată efectuată identificarea, modul în care mentalul este relaționat cu fizicul devine cu nimic diferit de modul în care anumite aspecte ale fizicului sunt relaționate cu fizicul, iar aceste conexiuni nomice au loc între termeni intersubiectiv disponibili și confirmabili. Totuși, faptul că ciucurii dispar în teoria finală nu înseamnă că putem să-i evităm în cursul construirii argumentului în favoarea tezei empirice de identificare a mentalului cu fizicul. Astfel, în postfața la ediția din 1967 a lucrării pe care o discutăm, Feigl scrie:

“Chestiunea crucială este acum: implică oare forma ipotezei identității oferite de mine asumarea “ciucurilor nomologice”? Cu ocazia multor discuții, bunul meu prieten, profesorul J.J.C. Smart, a încercat să mă convingă de faptul că mă pot lipsi de “ciucuri”. Însă am fost îndărătnic, deoarece nu vedeam cum ar putea cineva menține caracterul empiric (sintetic) al identificării, pe care Smart însuși, în felul său, îl accentuează, și care trebuie într-un fel sau altul să reflecte corelațiile și izomorfismele ce sunt gradual descoperite de către psiho-(neuro)-fiziologie. Totuși, de la bun început, în reflecțiile mele asupra dilemei tradiționale, am fost convins că “ciucurii” sunt metafizic inofensivi. Smart are dreptate, în esență, cred eu, când spune că aceștia nu vor apărea și nu pot să apară în concepția științifică “finală” despre lume.” (Feigl [1967], p. 139)

Înainte de a-mi expune argumentul pentru fizicalism, bazat pe ideea că ciucurii nomologici sunt, chiar dacă nu incoerenți din punct de vedere logic, incoerenți din punct de vedere probabilistic, merită să luăm în considerare modul în care acesta se încadrează și se implică dialectic în dezbaterile asupra fizicalismului de după Feigl, Smart și Place.

2. Dialectica post-Feigl

După marea breșă reprezentată de scrierile celor trei gânditori fizicaliști mai sus amintiți, teza empirică a identității minte-corp a fost transformată, în a doua jumătate a anilor 1960, prin scrierile lui David Lewis [1966] și David Armstrong [1968], într-o teză cu o componentă analitică esențială. Motivul pentru care Feigl fusese descurajat de ideea de analiticitate în legătură cu relația minte-corp a fost faptul că nu dădea – pe bună dreptate – nicio șansă de izbândă unei posibile sinonimii între vocabularul fenomenal și cel neurofiziologic. Cu toate acestea, Smart a fost de fapt primul⁴ care a intuit faptul că, în pofida lipsei unei sinonimii directe între cele două vocabulare, s-ar putea formula o

⁴ Pentru a fi corecți vizavi de Feigl, să menționăm că și el ajunge foarte aproape de teza identității, mijlocită de analiza neutră, dat fiind că de fapt utilizează un proces de identificare în doi pași, unul între qualii și denotațiile conceptelor corelate inferențial cu conceptele logic-comportamentale, iar al doilea între acestea din urmă și denotațiile conceptelor neurofiziologice. Vezi Feigl [1958/1967], cap. 5, sect. A.

analiză, sau cel puțin o glosă asupra conceptelor fenomenale apelând la un așa-nunit vocabular *neutru* (*topic-neutral*):⁵

“Sugestia mea este după cum urmează. Când o persoană spune “Văd o aparență iluzorie gălbui-portocalie”, spune ceva de tipul: “*Se întâmplă ceva ce este același lucru cu ce se întâmplă atunci când am ochii deschiși, sunt treaz, și există o portocală bine iluminată în fața ochilor mei, adică atunci când realmente văd o portocală*”. (...) Observați că termenii scriși înclinat, și anume “*Se întâmplă ceva ce este același lucru cu ce se întâmplă atunci când*” sunt toți quasi-logici sau neutri. Asta explică de ce descrierile verbale ale țăranilor din Grecia antică despre propriile senzații sunt neutre între o metafizică dualistă și metafizica materialistă pe care o propun eu.” (Smart [1959], pp. 149-150)

Analiza neutră avea să devină mai târziu, în doctrina funcționalismului analitic, pasul intermediar standard spre identificarea proprietăților mentale cu cele neurofiziologice. Primul pas îl reprezintă o analiză a termenilor mentali în termeni funcționali, ai rolului cauzal. Apoi realizatorul rolului cauzal în lumea actuală este identificat drept o proprietate neurofiziologică. Astfel, teza identității minte-creier este argumentată fără a se face apel la briciul lui Ockham (cf. David Lewis [1966]).

Deceniile '70 și '80 sunt apoi martore ale apariției unor importante argumente antifizicaliste: argumentul lui Saul Kripke împotriva tezei timpurii, empirice a identității, bazată pe noțiunea de identitate contingentă (1972), argumentul bazat pe conceptibilitatea lumilor *zombie* (Robert Kirk [1974a], [1974b], resuscitat și dezvoltat mai târziu de către David Chalmers [1996]), argumentul bazat pe subiectivitate (Thomas Nagel [1974]),⁶ argumentul bazat pe existența unei prăpastii explicative (Joseph Levine [1983]),⁷ argumentul cunoașterii (Frank Jackson [1986]). Nu voi intra în detaliile acestor argumente, ci le voi reuni, urmându-l pe Chalmers, sub denumirea de “argumente epistemice” împotriva fizicalismului. Potrivit lui Chalmers [2004, p. 108]:

“Forma generală a unui argument epistemic împotriva materialismului este următoarea:

Există o prăpastie epistemică între adevărurile fizice și cele fenomenale.

Dacă există o prăpastie epistemică între adevărurile fizice și cele fenomenale, atunci există o prăpastie ontologică între ele, iar materialismul este fals.

⁵ La drept vorbind, așa cum sublinia și Colin McGinn [2001, p. 286], găsim două teorii esențial diferite în unul și același articol al lui Smart, sub denumirea “teoria identității minte-corp”: fizicalismul a posteriori și fizicalismul a priori, ca să folosim o terminologie recentă (cf. Daniel Stoljar [2001]).

⁶ Deși ar trebui să adăugăm că Nagel însuși nu consideră că argumentul său este unul care demonstrează falsitatea fizicalismului.

⁷ Să adăugăm, din nou, că Levine nu consideră argumentul ca demonstrând falsitatea fizicalismului.

Materialismul este fals.

Bineînțeles, acest mod de a privi lucrurile este prea simplificator, neglijând diferențele dintre argumente (...). Totuși, această analiză oferă o lentilă utilă prin care să vedem ce au în comun aceste argumente și prin care să analizăm diverse răspunsuri la aceste argumente.”

Argumentele epistemice se aplică, desigur, nu doar proprietăților fizice, ci și celor neutre caracteristice versiunilor analitice ale tezei identității. Se argumentează, în prima premisă, că instanțierea, atât a proprietăților fizice, cât și a celor funcționale, este compatibilă epistemic cu lipsa instanțierii oricăror proprietăți fenomenale, apoi se inferează că, odată ce posibilele rațiuni de a ne îndoi de faptul că această compatibilitate epistemică, în cazul nostru particular, este un bun ghid către compatibilitatea metafizică, au fost eliminate, respectiva compatibilitate metafizică este justificată. Deci, fizicalismul – teza că totalitatea proprietăților fizice și funcționale instanțiate în mod actual implică în mod necesar instanțierea qualiilor (a proprietăților fenomenale) – trebuie să fie fals.

Dacă aceste argumente epistemice sunt acceptate drept corecte, atunci, în combinație cu teza lui Feigl potrivit căreia ciucurii nomologici, deși metodologic stranii, sunt coerente din punct de vedere logic, obținem teza dualismului naturalist ca o doctrină perfect coerentă și plauzibilă. Potrivit dualismului naturalist, atât proprietățile mentale, cât și cele fizice sunt fundamentale în lumea actuală, adică niciunul din aceste tipuri de proprietăți nu survine metafizic pe celălalt.⁸ Cu alte cuvinte, aceste tipuri de proprietăți sunt distincte din punct de vedere ontologic. Pe de altă latură, partea de „naturalism” a dualismului naturalist afirmă că aceste proprietăți sunt puse în relație în lumea actuală de către legi ale naturii a căror formă o putem deduce pornind de la corelațiile mental-fizic pe care le observăm în propriul caz (adică în cazul subiectiv, la persoana întâi). În fine, aceste conexiuni nomice sunt contingente, la fel ca alte legi ale naturii, potrivit acestei doctrine. Printre suporterii doctrinei se numără, printre alții, Chalmers [1995, 1996], Tim Crane [2001], precum și eu însumi (Aranyosi [2008]), rațiunea fiind, de fapt, observația că, în ciuda credinței din ultimii 50 de ani, potrivit căreia singurul mod în care qualiile pot fi luate în serios din punct de vedere științific este acela de a le reduce, într-un fel sau altul, la proprietăți neurofiziologice, care pot figura fără probleme în formulări ale unor legi ale naturii, nu e nimic incoerent din punct de vedere logic în a asuma existența unor legi naturale psiho-fizice ireductibile, *sui generis* și fundamentale. După cum am

⁸ Versiunea de superveniență pe care o adopt aici este cea globală: Proprietățile actuale de tip **A** survin metafizic-global pe proprietățile de tip **B** dacă, și numai dacă, orice lume posibilă, *W*, care este un duplicat-A al lumii actuale, @, este și un duplicat-B al lui @.

văzut, Feigl însuși, deși incomodat de “estetica” acestor legi, și sceptic în privința *indispensabilității* lor, nu considera că ele sunt incoerente logic.

Deși am fost eu însumi în trecut un suporter al ideii coerenței dualismului naturalist, acum cred că a fost o greșală și că, de fapt, cei care în ultimii 50 de ani au considerat că dualismul naturalist nu este o opțiune plauzibilă în soluționarea problemei minte-corp au avut dreptate, deși dintr-un alt motiv decât cel presupus de aceștia. Aceasta mă aduce, în sfârșit, la argumentul pe care vreau să-l supun atenției, având drept concluzii că dualismul naturalist este incoerent din punct de vedere probabilistic și că fizicalismul, în forma tezei identității minte-creier, este cel mai probabil candidat pentru natura relației mental-fizic.

3. Incoerența probabilistică a dualismului naturalist

Ideea ciucurilor nomologice, înțeleasă după modelul lui Feigl, implică ideea afirmării existenței unei conexiuni nomice între fenomene intersubiectiv confirmabile și fenomene ce nici măcar în principiu nu sunt disponibile intersubiectiv. Cum ajungem să cunoaștem că un anumit model de corelație nomică între acestea este instanțiat în lumea actuală, dacă este într-adevăr instanțiat? Un lucru asupra căruia atât Feigl, cât și dualiștii naturaliști cad de acord este că nu putem fi siguri de această instanțiere, însă tot ei spun că putem infera probabilistic existența termenilor fenomenali ai relației, pornind de la date comportamentale și alte date disponibile obiectiv, astfel că în final inferăm existența modelului nomic. Dar acest raționament este prea abrupt, sau chiar repezit. Chiar dacă putem deduce existența qualiilor în lumea actuală din datele intersubiectiv disponibile plus datele subiective din perspectiva persoanei întâi, acest lucru încă nu e suficient pentru a deduce că modelele de corelație ce caracterizează cazul propriu – de exemplu, de forma: “de câte ori am o durere intensă, urlu” etc. – sunt generalizabile la cazul celorlalți. Totuși, să presupunem pentru moment că avem o justificare probabilistică pentru a infera ipoteza existenței unui profil nomic corespunzând modului în care credem că acel profil se prezintă.

Să numim profilul nomic psiho-fizic actual obținut prin generalizare universală de la cazul propriu “profilul nomic normal” (PNN). Dacă lumea actuală este în PNN, atunci, întotdeauna când un subiect vede un obiect roșu, el este dispus să exclame “acest obiect este roșu”, întotdeauna când un subiect este într-o stare de durere intensă este dispus să o manifeste comportamental prin a exclama “auuu!” ș.a.m.d. Dacă acestea sunt legile de corelație, atunci PNN este modul în care dualistul naturalist concepe lumea actuală. Nu există subiecți inverși în privința spectrului cromatic fenomenal (de pildă, subiecți cărora vederea unui obiect roșu le cauzează aparența subiectivă de verde și invers.), nu există subiecți *zombie* (fizic identici cu mine, dar fără vreo stare subiectivă, fenomenală) etc. După cum am subliniat, nu putem fi siguri, *ex hypothesi*, niciodată că ceilalți sunt în

aceleași stări fenomenale ca și noi, atunci când sunt în aceleași stări comportamentale, neurale, fizice ca și noi. Am presupus, cu toate acestea, că avem o justificare probabilistică pentru a infera că lumea actuală este în PNN. Problema cu dualismul naturalist este că însăși premisa care asigură coerența logică a dualismului, anume prima premisă a argumentului epistemic canonic, este în același timp cea care slăbește până la nulitate coerența probabilistică a ipotezei că lumea actuală este în PNN. Deci, dacă acea premisă este adevărată, atunci inferența noastră potrivit căreia lumea actuală este în PNN este extrem de implauzibilă.

Să explic. Prima premisă a argumentului epistemic afirmă că există o prăpastie epistemică între faptele fenomenale și cele fizice. Tradusă în termeni de conceptibilitate, premisa afirmă că este conceptibil că există o lume fizic identică cu lumea actuală, dar fenomenal diferită de aceasta. Formula “fenomenal diferită” este foarte generală, așa că numărul de scenarii conceptibile de combinații ale proprietăților fenomenale este infinit numărabil, asumând că proprietățile fenomenale sunt discrete și că există proprietăți fenomenale străine.⁹ Mai mult, numărul scenariilor anomice va fi mult mai mare decât numărul profilelor nomice, dat fiind că distribuțiile aleatorii vor fi funcții de combinații chiar la nivelul unui singur particular sau al unui singur punct temporal. Bineînțeles, literatura actuală asupra argumentelor bazate pe conceptibilitate ne-a condiționat în sensul concentrării atenției asupra unor cazuri retoric frapante, adevărate puncte Schelling ale spațiului logic, și anume scenariul inversiunii de spectru sau scenariul *zombie*. Când discutăm argumentele bazate pe conceptibilitate, ne concentrăm în mod tipic atenția asupra acestor cazuri. Dar nu ar trebui să o facem. Acestea sunt dispozitive pur retorice menite a înfățișa argumentul antifizicalist într-o lumină intuitivă. Dacă inversiunea qualiilor și *zombies* sunt conceptibile, atunci nu există nicio rațiune de principiu împotriva conceptibilității tuturor celorlalte scenarii, infinite la număr. Aceasta ne aduce la principiul pe care-l voi numi “principiul exploziei”:¹⁰

(EXPLOZIA) Dacă un scenariu *S* este conceptibil, atunci toate scenariile similare în mod relevant sunt conceptibile.

⁹ Proprietățile străine sunt definite ca acele proprietăți care sunt instanțiate în alte lumi posibile decât lumea actuală.

¹⁰ Principiul are ca sursă de inspirație literatura asupra lumilor imposibile. Un exemplu de explozie este *ex falsum sequitur quodlibet* în logica standard, potrivit căruia putem deriva orice propoziție din orice contradicție. Principiul implică faptul că, într-o lume în care o contradicție este adevărată, toate propozițiile sunt adevărate. Am împrumutat termenul “explozie” de la Daniel Nolan [1997], care numește o lume în care toate propozițiile sunt adevărate o “lume de tip explozie”. Nolan argumentează contra principiului exploziei; Lewis [1988] oferă un argument în favoarea acestuia.

Bineînțeles, nu vom fi întotdeauna capabili să decidem dacă o pereche de scenarii sunt similare în mod relevant, însă în cazul de față – cazul psiho-fizic – similaritatea se bazează pe (a) duplicare fizică și (b) diferență fenomenală. Datorită faptelor (a) și (b), putem acum să aplicăm principiul exploziei la cazul psiho-fizic:

(PSIHO-EXPLOZIA) Dacă un duplicat fizic al lumii actuale, diferit fenomenal de aceasta în privința R , este conceptibil, atunci orice duplicat fizic al lumii actuale, diferit fenomenal de aceasta în orice altă privință R^* , este conceptibil.

Însă dacă există infinit de multe scenarii non-PNN, cele mai multe dintre ele anomice, adică distribuții aleatorii de proprietăți fenomenale în mulțimea duplicatelor fizice ale lumii actuale, cum putem ajunge să deducem, probabilistic, că lumea actuală este în PNN? Știm, în urma discuției asupra lui Feigl, că proprietățile fenomenale nu sunt accesibile intersubiectiv. În mod similar, Chalmers argumentează îndelung pentru teza irelevanței explicative a proprietăților fenomenale, ceea ce înseamnă că ele nu determină nicio diferență la nivelul observației obiective. Dar aceasta înseamnă că, luând în considerare tot ceea ce știu în mod intersubiectiv, lumea actuală ar putea fi în oricare dintre profilurile nomice sau anomice conceptibile. Principiul indiferenței, care este o regulă de distribuire a probabilităților în cazuri de ignoranță, ne-ar spune că lumea actuală aflată în starea PNN este un rezultat echiprobabil cu oricare dintre scenariile conceptibile, infinite la număr. Asta înseamnă că probabilitatea ca lumea actuală să fie în PNN este de $1/\infty$, adică aproximativ zero.

În cazul în care condițiile de aplicabilitate ale principiului indiferenței sunt satisfăcute, probabilitatea ca lumea actuală să fie în PNN ($p(@_{pnn})$) este dată de proporția (f_{pnn}) a lumilor PNN în totalul lumilor duplicate fizice ale lumii actuale ($N_{pnn} + N_{non-pnn}$):

$$p(@_{pnn}) = f_{pnn} = N_{pnn} / (N_{pnn} + N_{non-pnn})$$

Observăm că $p(@_{pnn})$ se apropie de unitate (adică gradul nostru de credință că lumea este în PNN este aproape 1) atunci, și numai atunci, când diferența dintre ($N_{pnn} + N_{non-pnn}$) și $N_{non-pnn}$ se apropie de zero, adică atunci când $N_{non-pnn}$ este aproximativ zero.¹¹

¹¹ Obținem rezultate similare și într-un cadru Bayesian. Să denotăm prin H ipoteza că există un număr infinit de lumi conceptibile duplicate fizice ale lui $@$, dar cu distribuții fenomenale diferite față de cele actuale. Denotăm prin E propoziția că $@$ este în PNN și prin $\emptyset H$ ipoteza că aproape toate duplicatele fizice ale lui $@$ sunt în PNN. Regula lui Bayes ne spune: $p(H|E) = p(H)p(E|H) / [p(H)p(E|H) + p(\emptyset H)p(E|\emptyset H)]$. Să atribuim o probabilitate foarte ridicată lui H , să zicem 0,99999. Prin intermediul principiului indiferenței, obținem o probabilitate foarte scăzută pentru $p(E|H)$ și una foarte ridicată, și anume $1 - p(E|H)$, pentru $p(E|\emptyset H)$. Valorile numerice pentru parametrii noștri sunt după cum urmează:

$$p(H) = 0,99999$$

Dat fiind că numărul de scenarii cu distribuții aleatorii de proprietăți fenomenale este cu mult mai mare decât al celor nomice, obținem un rezultat și mai tulburător, anume că este aproape sigur că trăim într-o lume cu o distribuție aleatorie de proprietăți fenomenale.

Cu toate acestea, principiul indiferenței ne ajută încă o dată. În forma sa cea mai generală, principiul ne spune că, atunci când nu avem nicio rațiune a priori de a atribui o probabilitate mai mare vreunui eveniment în dauna altuia, se cere să atribuim probabilități (sau grade de credință) egale tuturor evenimentelor posibile. O versiune a acestui principiu este formulabilă făcând apel la noțiunile de observator și clasă de referință. O astfel de versiune, propusă de Nick Bostrom (2002), este “asumpția auto-eșantionării”:

(AAE) Un observator trebuie să raționeze ca și cum ar fi un eșantion selectat aleatoriu din mulțimea tuturor observatorilor din clasa sa de referință.

După cum am mai notat, Feigl susținea că putem deduce probabilistic, pornind de la cazul propriu, faptul că lumea actuală se află în PNN. Dualiștii sunt de acord, deoarece alternativa în viziunea lor ar fi un scepticism extrem în privința existenței altor minți (adică solipsismul), precum și concluzia inacceptabilă că, până și acceptând existența altor minți, acestea ar fi relaționate în mod aleatoriu cu stările fizice. Noi, pe de altă parte, putem argumenta pentru o astfel de inferență de la cazul propriu la ipoteza că lumea se află în PNN, apelând la AAE. Ceea ce observ este că eu însumi mă conformez cu ceea ce PNN ar prescrie unui observator să observe. Observ că stările mele cerebrale și comportamentale întotdeauna se potrivesc cu stările mele fenomenale. Dat fiind acest lucru, care este probabilitatea ca toți ceilalți observatori actuali să se conformeze acelorași corelații psiho-fizice? Dat fiind AAE, ar trebui să mă consider un eșantion aleatoriu din mulțimea tuturor observatorilor actuali. Dat fiind ceea ce observ în propriul meu caz, mulțimea observatorilor actuali trebuie să conțină un număr mult mai mare de observatori PNN-conformi decât numărul observatorilor cu distribuții fenomenale

$$p(\emptyset|H) = 0,00001$$

$$p(E|\frac{1}{2}H) = 0,0000001$$

$$p(E|\frac{1}{2}\emptyset H) = 0,9999999$$

Substituind în formula lui Bayes, obținem:

$$p(H|\frac{1}{2}E) = 0,99999 \cdot 0,0000001 / [0,99999 \cdot 0,0000001 + 0,00001 \cdot 0,9999999]$$

$$= 9,9999 \cdot 10^{-8} / (9,9999 \cdot 10^{-8} + 9,999999 \cdot 10^{-6}) = 0,009$$

Aceasta înseamnă că probabilitatea lui H trebuie actualizată de la 0,99999 la 0,009. Dacă numărul duplicatelor fizice ale lui @ care sunt non-PNN se apropie de infinit, atunci probabilitatea posterioară a lui H devine aproximativ zero.

aleatorii. Aceasta înseamnă că ar trebui să considerăm ca aproape sigur că lumea actuală este în PNN.

Am putea fi întrebați în acest punct: ce ar fi fost dacă am fi considerat drept clasă de referință mulțimea tuturor observatorilor conceptibili? Nu am fi obținut oare același rezultat ca și atunci când am considerat lumi întregi drept membrii clasei de referință, adică rezultatul că lumea actuală este aproape sigur într-o stare non-PNN?

În primul rând, să notăm că răspunsul este nu: dacă am fi considerat mulțimea tuturor observatorilor conceptibili, faptul că eu însumi mă conformez PNN ar fi fost o coincidență enormă dacă nu am fi considerat că majoritatea observatorilor conceptibili se conformează PNN. Pentru ca experiențele mele să fie cum sunt, frecvența lor în mulțimea tuturor observatorilor trebuie să fie extrem de mare, dat fiind că eu sunt doar unul dintr-un număr infinit de observatori, selectat aleatoriu. Asta înseamnă, în lumina principiului indiferenței, că aproape toți observatorii conceptibili trebuie să aibă experiențe PNN-conforme.

În al doilea rând, rațiunea de a alege drept clasă de referință mulțimea tuturor observatorilor actuali este simplă: ne interesa dacă ceilalți oameni *în lumea actuală* au aceleași stări fenomenale corelate cu aceleași stări fizice pe care le am eu. Clasa de referință fusese deja selectată prin faptul că, în propriul meu caz, posed dovada observațională directă în privința corelațiilor. Când, pe de altă parte, ne-am întrebat dacă lumea actuală ca atare se conformează PNN, am procedat pornind de la premisa ignoranței în privința a ce probabilități să atribuim fiecărei lumi conceptibile; astfel, în chip natural, a trebuit ca lumi întregi să figureze în clasa de referință.

Cu ce ne alegem din toată această discuție? Aparent, am obținut rezultate aflate în conflict. Date fiind PSIHO-EXPLOZIA și prima premisă a argumentului bazat pe conceptibilitate, avem temeuri pentru a crede în existența unui număr infinit de lumi conceptibile non-PNN și, deci, pentru a crede într-un grad foarte ridicat (de aproximativ 1) că lumea noastră este anomică în privința relațiilor psiho-fizice. Pe de altă parte, dat fiind AAE, ar trebui să credem într-un grad la fel de ridicat că lumea noastră se conformează PNN. Dar dacă este aproape sigur că lumea actuală este în PNN, atunci, printr-o nouă aplicare a primului principiu al indiferenței, ar trebui să deducem că este aproape sigur că lumi duplicate fizic ale lumii actuale având distribuții fenomenale diferite de distribuția actuală sunt inconceptibile, ceea ce este în conflict cu prima premisă a argumentului bazat pe conceptibilitate.

În orice caz, chiar și atât este suficient pentru a demonta ideea că dualismul naturalist este coerent din punct de vedere probabilistic. Asumând că principiul PSIHO-EXPLOZIEI este într-adevăr plauzibil, ajungem la o concluzie disjunctivă ce exclude, probabilistic, teza că ar exista legi psiho-fizice care sunt ciucuri pur nomologici:

(C) Ori este aproape sigur că lumea actuală este anomică din punct de vedere psiho-fizic, ori este aproape sigur că lumi duplicate fizic ale lumii actuale având distribuții fenomenale diferite de distribuția actuală sunt inconceptibile.

A spune că aceste scenarii non-PNN sunt inconceptibile este echivalent cu a spune că PNN este epistemic necesar. Ca atare, opțiunile ce rămân pe masă sunt: ciucuri anomici și ciucuri logic necesari, dar nu și ciucuri pur nomologici, contrar doctrinei dualismului naturalist.

4. Inferența către teza identității minte-creier

Aș vrea să merg mai departe și să argumentez pentru diverse moduri în care, din acest punct, putem deduce teza identității minte-creier. Există două traiectorii inferențiale disponibile, ambele părându-mi plauzibile.

Prima este de a asuma naturalismul sub forma tezei că lumea actuală se conformează PNN și, în consecință, de a nega termenul prim al disjuncției C, care afirmă că este aproape sigur că lumea actuală se află într-o stare psiho-fizică anomică. Aceasta ar implica inconceptibilitatea inversiunii de spectru, a *zombie*-ilor sau a oricăror distribuții anormale de proprietăți fenomenale și, în consecință, necesitatea profilului nomic actual. Opțiunile în privința relației mental-fizic cu care rămânem sunt, deci: (i) distincție ontologică, dar relație nomică necesară între proprietăți fizice și fenomene, și (ii) identitate.

Există trei argumente contra opțiunii (i) și în favoarea opțiunii (ii). Primul este bazat pe dictonul lui Hume cum că nu există conexiuni necesare între existențe distincte. Prin contrapozitie, din moment ce avem rațiuni să credem că relația necesară este prezentă, avem rațiuni de a identifica fenomenalul cu fizicul. Un alt argument este bazat pe briciul lui Ockham, care spune că nu este permis să multiplicăm entitățile dincolo de ceea ce este necesar. Cu alte cuvinte, dacă putem explica aceleași fenomene în mai multe chipuri, trebuie să optăm pentru modul cel mai parcimonios din punct de vedere ontologic de explicare a acestor fenomene. Aplicat cazului nostru, briciul infățișează teza identității drept cea mai plauzibilă, dat fiind că ea are același statut explicativ și generează exact aceleași predicții ca și alternativa bazată pe conexiuni necesare între tipuri ontologic distincte de proprietăți. În fine, al treilea argument se bazează pe principiul repudierii necesităților brute, inexplicabile, care afirmă că orice necesitate trebuie să se bazeze explicativ pe necesitatea logică. Profilul nomic necesar postulat este în mod clar un caz de

necesitate brută, însă dacă teza identității este adoptată, atunci necesitatea respectivă va avea un suport în proprietatea logică a necesității identității.¹²

A doua traiectorie inferențială către teza identității se bazează nu pe asumarea naturalismului, adică a ipotezei că lumea actuală este în PNN, ci pe punerea în balanță a celor două rațiuni aflate în conflict, rațiunea în favoarea naturalismului (și împotriva premisei conceptibilității) și rațiunea în favoarea conceptibilității unui număr infinit de duplicate fizice ale lumii actuale având diverse combinații de distribuție a proprietăților fenomenale. Dacă rațiunea în favoarea naturalismului este mai puternică, atunci putem, apelând la cele trei argumente explicate mai sus, să derivăm din nou concluzia că teza identității este preferabilă. Întrebarea este, deci: care dintre rațiuni este mai puternică?

Argumentul meu în favoarea necesității PNN și împotriva conceptibilității *zombie*-ilor, a inversiunii de spectru sau a altor scenarii non-PNN, se bazează pe faptul că raționamentul întemeiat pe principiul indiferenței poate fi considerat fără nicio problemă drept o extensie a unui raționament a priori, propriu evaluării serioase a ipotezei conceptibilității *zombie*-ilor, a inversiunii etc., astfel că din moment ce rezultatul final al acestui raționament a priori este în conflict cu intuiția inițială în privința conceptibilității acestor scenarii, intuiția inițială trebuie respinsă. Dacă argumentul meu este corect, atunci este doar *prima facie* conceptibil să existe lumi *zombie*, inversate sau altele neconforme cu PNN. Începem prin a aplica principiul indiferenței la propriul caz (pentru care avem dovezi directe), având drept clasă de referință mulțimea observatorilor actuali, și deducem că lumea actuală trebuie, aproape sigur, să se conformeze PNN. Apoi aplicăm din nou principiul indiferenței, considerând lumea actuală drept un eșantion aleatoriu din mulțimea tuturor lumilor inițial conceptibile, duplicate fizic ale lumii actuale, dar diferite fenomenal de aceasta. În fine, conchidem că, dat fiind că altfel ipoteza că lumea noastră se conformează PNN ar fi o coincidență enormă, trebuie să fie aproape sigur că toate duplicatele fizice ale lumii actuale sunt și duplicate mentale, ceea ce este echivalent cu a conchide că intuiția noastră inițială în privința conceptibilității a fost greșită și că intuiția trebuie revizuită în lumina raționamentului a priori ulterior.

5. Obiecții

¹² Am putea fi întrebați în acest punct: care este explicația identității însăși? David Papineau [2002, p. 114] argumentează că identitatea, în general, nu are nevoie de explicație; nu are niciun sens să întrebăm, odată ce știm că o identitate este adevărată, de ce este adevărată. În contextul de față, totuși, nu este nevoie să ne sprijinim pe o astfel de teză. Ceea ce explică identitatea minte-creier este faptul că acea identitate explică ea însăși teza noastră asupra corelației necesare între proprietățile fizice și cele fenomenale. Se întâmplă frecvent în științe ca angajamentul nostru ontologic în privința existenței unei entități, x , să fie explicat de faptul că x însuși explică, în cel mai bun mod, un fenomen y . De exemplu, angajamentul ontologic în privința existenței genei se explică prin faptul că existența genei înseși explică o serie de fenomene observabile legate de ereditate.

Înainte de a concluziona, aş vrea să iau în considerare câteva potențiale obiecții la argumentul meu.

Obiecția 1: *Este știut faptul că principiul indiferenței conduce la inconsistențe logice.*

Într-adevăr, există o dezbatere aprinsă privitoare la cum să formulăm principiul de indiferență fără a genera distribuții inconsistente de probabilități. Putem genera astfel de distribuții inconsistente de probabilități prin aspirarea spațiului logic al rezultatelor posibile. Pentru a exemplifica, să considerăm că tot ceea ce știm este că există trei găleți duplicate și că una dintre ele conține o bilă, dar nu putem vedea conținutul găleților. Ni se cere să oferim o distribuție de probabilități asupra celor trei găleți pentru evenimentul "bila este prezentă în găleată". Principiul indiferenței ne spune că, deoarece nu există niciun motiv pentru care să favorizăm o găleată în defavoarea oricărei alte găleți atunci când considerăm evenimentul, trebuie să atribuim probabilități egale tuturor găleților în privința ipotezei că ar conține bila, și anume $1/3$. Deci, de exemplu, presupunând că am numit gălețile G_1 , G_2 , și G_3 , probabilitatea ca bila să fie în G_1 este de $1/3$. Putem acum aspri spațiul logic al rezultatelor posibile, prin redescrierea unora dintre rezultate, ca de exemplu:

Rezultatul 1: Bila este în G_1 .

Rezultatul 2: Bila este în G_2 -sau- G_3 .

Dat fiind că numărul de rezultate posibile este acum de 2, principiul indiferenței va prescrie o probabilitate de $1/2$ pentru evenimentul "bila este în G_1 ", ceea ce este inconsistent cu distribuția probabilistică precedentă. Asta în condițiile în care am apelat la același principiu al indiferenței.

Aplicat cazului nostru, putem aspri spațiul lumilor posibile duplicate fizic ale lumii actuale, dar cu distribuții fenomenale diferite de aceasta, prin redescrierea unora dintre ele, după cum urmează:

Rezultatul 1: Lumea actuală este în PNN.

Rezultatul 2: Lumea actuală este în non-PNN.

Din nou, deoarece avem două rezultate posibile, principiul indiferenței ne dictează să atribuim o probabilitate de $1/2$ fiecărui rezultat, ceea ce înseamnă că, în pofida a ceea ce am argumentat în secțiunile precedente, ar trebui să ne suspendăm judecata, adică să fim agnostici în privința celor două ipoteze: că dualismul naturalist este adevărat sau că disjuncția dintre ciucurii anomici și cei necesari este adevărată.

În replică, am putea argumenta că noțiunea de lume posibilă ca mulțime maximală consistentă de propoziții este destul de clară încât să excludă posibilitatea aspririi spațiului logic al rezultatelor prin redescoperire disjunctivă. O expresie de tipul "lumea posibilă W -sau- W^* " nu se referă deloc la o lume posibilă, dacă W și W^* sunt ele însele mulțimi maximale consistente de propoziții, iar o disjuncție de forma "lumea posibilă W sau lumea posibilă W^* " se va referi întotdeauna la una dintre cele două lumi, niciodată la ambele sau la o fuziune a lor. În același timp, putem de asemenea apela la o restrângere consistentă a principiului indiferenței. Paul Castell (1998) oferă o astfel de restrângere, pe care o numește "principiul irelevanței". În loc să considerăm numărul total de rezultate din spațiul logic al rezultatelor posibile și să le atribuim probabilități egale, vom lua în considerare un sistem fizic, P , și un rezultat sau stare particulară, O , în care P are posibilitatea de a fi, după care afirmăm că probabilitatea ca P să fie în O este aceeași la orice timp t , sau că probabilitatea oricărui duplicat al lui P de a fi în O este constantă, dat fiind un punct temporal. Apoi repetăm același raționament pentru toate celelalte rezultate posibile. Probabilitatea unui rezultat în particular va fi dată de frecvența adevărului propoziției care afirmă rezultatul în mulțimea tuturor propozițiilor care descriu toate duplicatele sistemului, sau sistemul însuși la timpi diferiți.

Pentru a exemplifica, să luăm în considerare exemplul nostru cu gălețile și cu bila. Sistemul nostru este găleata și bila, iar gălețile G_1 , G_2 și G_3 sunt asumate a fi duplicate. Rezultatul sau starea relevantă este bila ca prezentă în găleată, stare ce o denotăm cu expresia "1" (starea în care bila este absentă o denotăm cu "0"). Acum putem reprezenta problema în felul următor:

A_1 : sistemul 1 (i.e. G_1) este în starea 1.

A_2 : sistemul 2 (i.e. G_2) este în starea 1.

A_3 : sistemul 3 (i.e. G_3) este în starea 1.

Ceea ce ne spune principiul irelevanței este că propozițiile A_1 - A_3 sunt echiprobabile. Valoarea numerică în particular ne este dată prin atribuirea valorii ADEVĂRAT unei propoziții de forma "sistemul x , pentru un anume x , este în starea 1", și a observării frecvenței acestui adevăr în mulțimea propozițiilor A_1 - A_3 :

ADEVĂRAT: sistemul 1 (i.e. G_1) este în starea 1.

FALS: sistemul 2 (i.e. G_2) este în starea 1.

FALS: sistemul 3 (i.e. G_3) este în starea 1.

Cu alte cuvinte, potrivit descrierii problemei noastre, ori de câte ori unul dintre sistemele noastre este în starea 1, celelalte sisteme trebuie să fie în starea 0. Astfel, obținem probabilitățile de $1/3$ pentru ca sistemul să fie în starea 1, și $2/3$ pentru ca

sistemul să fie în starea 0, pentru un anumit sistem x . Problema atribuirii de distribuții inconsistente de probabilități prin aspirarea disjunctivă a spațiului logic al rezultatelor posibile dispare, deoarece mai sus menționatul Rezultat 2 (adică "bila este în găleata G_1 -sau- G_2 ") nu este el însuși un duplicat al sistemului nostru.

Aplicând acest tip de raționament problemei noastre din acest articol, sistemul fizic pe care-l considerăm relevant este sistemul format din totalitatea faptelor fizice din lumea actuală, să-l numim \odot , iar starea relevantă supusă atenției este PNN. \odot va avea o infinitate de duplicate, dacă prima premisă a argumentului bazat pe conceptibilitate și PSIHO-EXPLOZIA sunt adevărate, fiecare duplicat corespunzând unei combinații de proprietăți fenomenale. Atunci putem descrie problema noastră în felul următor:

A_1 : sistemul \odot_1 este în starea PNN.

A_2 : sistemul \odot_2 este în starea PNN.

...

A_n : sistemul \odot_n este în starea PNN.

Propozițiile A_1 - A_n vor fi echiprobabile. Mai departe, probabilitatea ca sistemul să fie în starea PNN va fi întotdeauna $1/n$, în timp ce probabilitatea ca sistemul să fie în starea non-PNN va fi întotdeauna $(n-1)/n$, deoarece atribuind ADEVĂRAT propoziției A_1 va genera falsitatea tuturor propozițiilor A_2 - A_n . O redescoperire de forma "sistemul \odot_1 or \odot_2 " nu va fi acceptabilă deoarece nu se referă la un duplicat al sistemului \odot , adică nu este un duplicat fizic al lumii actuale. Nu e nimic surprinzător în acest rezultat, dat fiind că A_1 - A_n este de asemenea o împărțire a spațiului duplicatelor fizice ale lumii actuale, astfel că asemenea disjuncții nu vor corespunde unor sisteme acceptate în contextul definirii lumilor posibile drept mulțimi maximale consistente de propoziții.

În fine, dat fiind că n este un număr foarte mare (adică infinit), probabilitatea ca lumea actuală să fie în PNN este practic nulă. De aici concluzia: ori premisa conceptibilității este falsă, ori este aproape sigur că nu trăim într-o lume psiho-fizic nomologică.

Obiecția 2: Argumentul pare a se generaliza asupra oricărei forme de aparentă cunoaștere inductivă sau abductivă.

Este universal acceptat faptul că este conceptibil ca lumea să aibă regularități diferite în viitor față de cele din trecut și, bineînțeles, există un număr infinit de asemenea posibile regularități viitoare. Dacă argumentul meu este acceptat, continuă obiecția, atunci nu putem cunoaște faptul că regularitățile din viitor sunt de fapt aceleași cu cele din trecut. Însă majoritatea teoreticienilor cred că putem cunoaște asemenea fapte. Deci, argumentul meu ar implica scepticism inductiv, ceea ce îl face implauzibil. În mod

similar, raționamentul abductiv ar apărea potrivit obiecției drept nejustificat dacă argumentul de față ar fi acceptat.

În replică, aș sugera în primul rând că în privința inducției există o lipsă de analogie între cazul generalizărilor de tip trecut-viitor și cel al generalizării de la cazul propriu la existența unor legi ale naturii aplicabile la nivelul tuturor subiecților relevanți. Inducția, în general, implică derivarea unei propoziții de tip “dacă p , atunci probabil q ” din existența unui număr mare de observații de tip “ p și q ”. Însă, în cazul proprietăților fenomenale, singurul caz în care astfel de conjuncții se observă în număr mare este cel la persoana întâi și, într-adevăr, am asumat existența acestor legi la persoana întâi, pe baza observării unui număr mare de corelații p - q , unde p reprezintă proprietăți neurofiziologice și funcționale, iar q reprezintă proprietăți fenomenale. Parte a explicației rezonabilității acestei asumptii ar putea desigur fi faptul că în trecut s-a observat un număr mare de astfel de corelații. Dar în cazul încercării derivării inductive a legilor care să caracterizeze toți subiecții din lumea actuală ne lipsește condiția existenței unui număr mare de observații privind instanțierea proprietăților fenomenale specifice într-un număr mare de subiecți; dată fiind teza irelevanței explicative a experienței (Chalmers), sau teza că legile psiho-fizice sunt ciucuri nomologici (Feigl), nu putem avea vreodată baza evidențială acceptabilă pentru un potențial raționament inductiv prin care să stabilim existența acestor legi. Totuși, putem deriva aceste legi, așa cum am făcut-o, indirect, prin aplicarea raționamentului bazat pe principiul indiferenței în varianta asumptiei auto-eșantionării. Deci, nu există nicio rațiune pentru a crede că argumentul meu ar implica scepticismul în privința inducției în general. Tot ce arăt este că, în cazul psiho-fizic, spre deosebire de cazul fizic-fizic, de exemplu, raționamentul inductiv nu este fezabil. Aceasta nu ar trebui să ne surprindă, deoarece este o implicație logică a proprietăților pe care Chalmers sau Feigl le atribuie experienței fenomenale în contextul teoriei explicației.

În ceea ce privește abducția, numită și “inferență spre cea mai bună explicație”, din nou, nu văd vreo rațiune pentru care să credem că argumentul este angajat în direcția scepticismului, nemaivorbind de faptul că de fapt unele părți ale argumentului se bazează pe un astfel de raționament. Raționamentul abductiv se referă la derivarea unei propoziții de tip “ p explică q ” din observația că p reprezintă cea mai bună explicație – din punctul de vedere al probabilității inițiale, al simplității și al puterii explicative – din mulțimea de explicații posibile ale q . În cazul nostru, p reprezintă totalitatea legilor care constituie PNN, în timp ce q reprezintă corelațiile observate în cazul propriu. Potrivit părții secunde a argumentului, când am aplicat asumptia auto-eșantionării, cea mai bună explicație a lui q este într-adevăr p , deoarece altfel existența lui q ar fi o coincidență enormă. În acest fel, am eliminat termenul “ciucuri anomici” din disjuncția (C) ca fiind a priori improbabil în condițiile asumptiei auto-eșantionării. În fine, raționamentul de tip abductiv apare încă o dată, și anume atunci când inferăm teza identității minte-creier, dat

fiind că este mai economică din punct de vedere ontologic, pentru a explica corelațiile necesare rezultate în urma argumentului principal.

Deci, luând în considerare toate aceste date, nu avem temeiuri să credem că argumentul meu ar implica scepticismul în privința inducției sau a abducției.

Obiecția 3: Principiul indiferenței se aplică atunci când sistemele fizice necesare pentru a afirma propozițiile spațiului logic al rezultatelor posibile există; lumile posibile, însă, nu există. Deci, principiul indiferenței nu este aplicabil.

Bineînțeles, de multe ori aceste sisteme există actual. De exemplu, în modelul nostru din problema cu gălețile și cu bila, gălețile există. În mod similar, să luăm în considerare problema atribuirii unei probabilități pentru ca o anumită carte de joc de poker, selectată aleatoriu dintr-un pachet de cărți de poker, să fie as de treflă, dată fiind ignoranța în privința oricărui alt fapt relevant în ce privește pachetul de cărți. Atribuim o probabilitate de 1/52 (e vorba de jocul de poker englezesc standard), iar celelalte 51 de cărți există. Dar important nu este ca sistemele fizice care susțin rezultatele neactualizate să existe, ci ca sistemele să fie reprezentate, adică să existe o reprezentare abstractă a tuturor acestor sisteme. În exemplul problemei legate de asul de treflă am obține exact același rezultat chiar dacă toate cărțile de joc cu excepția celei pe care o avem în față ar fi distruse. Ne putem chiar imagina Dumnezeu creând lumea astfel încât să conțină o singură carte de joc de poker, cu aceleași rezultate ale rezolvării problemei ca mai sus. Singurul lucru de care avem nevoie este o reprezentare abstractă a jocului de poker, cu 52 de cărți, asul de treflă fiind una dintre acestea.¹³

¹³ Uneori principiul de indiferență este utilizat tocmai pentru a deduce existența sistemelor fizice care susțin posibilitățile neactualizate. Argumentul pentru existența Multiversului este un astfel de exemplu. Aici variabila este faptul, sau nu, că un univers conține viață și observatori conștienți. Teoreticianul multiversului argumentează în felul următor. Date fiind (a) reglajul fin al universului nostru (adică sensibilitatea extremă a variabilei "viață" în raport cu valorile magnitudinilor și constantelor fizice ale stării inițiale a Universului, de exemplu cele din timpul Big Bang-ului) și (b) faptul că trăim într-un astfel de univers, am fi cu totul nesurprinși de faptul că trăim într-un astfel de univers dacă existența universului nostru ar fi un fapt brut, și extrem de surprinși de acest lucru dacă magnitudinile și constantele ar fi selectate aleatoriu din mulțimea extrem de numeroasă de valori posibile ale acestora. Dar ambele atitudini sunt extreme: nu suntem nici deloc surprinși, nici exagerat de surprinși că trăim într-un univers care conține viață. Singurul mod de a ne adecva la un nivel moderat de surprindere este să asumăm existența unui număr mare de universuri paralele, corespunzând diverselor valori posibile ale variabilelor fizice din starea inițială, unul dintre aceste universuri fiind cel în care trăim. Date fiind toate aceste universuri, nu este surprinzător că unul dintre ele conține viață, însă, pe de altă parte, este totuși surprinzător că trăim exact în universul care era compatibil cu evoluția vieții, dat fiind că majoritatea universurilor nu sunt capabile să genereze și să susțină viața. Pentru mai multe detalii asupra tezei multiversului, a se consulta John Leslie 1989.

Oricum, cazul care motivează teza multiversului este foarte diferit de cel care motivează și sprijină argumentul meu. În contextul primei aplicări a principiului indiferenței, i.e. când am considerat lumea actuală drept un eșantion aleatoriu din mulțimea duplicatelor fizice ale sale, condiția (b) nu este satisfăcută, deoarece,

Obiecția 4: *Ipoteza că lumea actuală ar fi în PNN este demonstrabilă prin experimente mentale de tipul “qualii pălînde” sau “qualii dansânde”.*

Chalmers (1996, cap. 5) oferă argumente, bazate pe experimentele mentale mai sus amintite, în favoarea a ceea ce el numește “principiul invarianței organizaționale”, care ar fi o lege a naturii ce afirmă că sistemele cognitive identice din punctul de vedere al organizării funcționale (stimul-centru-comportament) instanțiază aceleași proprietăți fenomenale, indiferent de ce diferențe există între sisteme la nivel neurofiziologic (sau fizic). În cazul qualiilor pălînde ne imaginăm că experiența noastră cromatică, inițial foarte vie, pălește gradual, ca urmare a înlocuirii treptate a neuronilor cu circuite de silicon, astfel că la sfârșitul procesului orice urmă de culoare dispare din câmpul fenomenal. Întrebarea este dacă organizarea funcțională se va menține constantă. Chalmers argumentează că ea trebuie să se schimbe drept consecință a faptului că, dacă qualiile ar păli, am observa acest lucru și l-am raporta verbal imediat. A presupune contrariul ar însemna să deconectăm complet experiența fenomenală de cogniție – un scenariu foarte ne-naturalist. Experimentul mental al qualiilor dansânde implică un dispozitiv implantat în creier cu rolul de comutator între baza neurofiziologică naturală a experienței fenomenale și o bază alternativă artificială, astfel încât prin comutarea în direcția activizării bazei artificiale are loc o inversiune de spectru la nivelul fenomenal. Dacă experiența se schimbă “în fața ochilor mei”, ca să folosim o exprimare a lui Chalmers, atunci aceasta va avea un efect asupra componentelor funcțional-definite ale sistemului meu cognitiv – voi recunoaște și voi raporta schimbarea. Deci, nu există schimbare la nivel fenomenal fără o schimbare la nivel funcțional.

În replică, aș vrea să scot în evidență două lucruri. Primul este că Chalmers însuși nu consideră aceste experimente mentale drept *demonstrații* ale faptului că lumea actuală ar fi în PNN. El afirmă explicit că:

“Aceste argumente bazate pe experimente mentale sunt doar argumente de plauzibilitate, ca de obicei, însă cred că au o forță considerabilă. A susține posibilitatea naturală a qualiilor absente și inversate în pofida acestor experimente mentale implică acceptarea unor teze implauzibile cu privire la conștiința fenomenală și în particular în ce privește relația dintre conștiință și cogniție. Date fiind unele asumptii naturale în ceea ce privește această relație, principiul invarianței este stabilit drept cea mai plauzibilă ipoteză.” [1996, pp. 250-251]

ex hypothesi, nu avem cum să observăm instanțierile (exceptând cazul propriu) de proprietăți fenomenale în lumea actuală. În contextul aplicării secunde a principiului indiferenței, adică atunci când am considerat cazul propriu drept un eșantion aleatoriu din mulțimea tuturor observatorilor actuali, în timp ce condiția (b) este satisfăcută, condiția (a) nu este, deoarece nu avem niciun motiv să credem că instanțierile de proprietăți fenomenale ar fi hipersensibile la vreo particularitate fizică a cazului propriu.

Cu alte cuvinte, ceea ce se poate stabili prin astfel de experimente mentale nu este că astfel de scenarii sunt imposibile în lumea actuală, ci că acestea nu sunt natural posibile, *presupunând că știm care sunt legile naturii în lumea actuală*, sau cel mult că aceste scenarii sunt implauzibile, *presupunând că știm ce legi ale naturii sunt de fapt plauzibile*. Astfel, asemenea considerații nu au niciun efect asupra argumentului pe care l-am propus.

În al doilea rând, și mai important, după cum observa recent Michael Pelczar (2008), experimentele mentale ale lui Chalmers nu dovedesc nimic mai mult decât că în cadrul *unei singure conștiințe*, de-a lungul diverselor stadii temporale ale sale, nu poate exista schimbare fenomenală fără schimbare funcțională. Această teză nu are niciun efect asupra cazului *interpersonal*, când avem de a face cu două sisteme identice funcțional, dar inversate spectral, sau unul dintre ele având qualii pălinde. Tot ce știm din cazul *intrapersonal* este compatibil cu ipoteza teoretică a existenței unei persoane identice funcțional cu mine, care are din naștere qualii fade în comparație cu mine. Persoana respectivă ar recunoaște orice schimbare în câmpul ei fenomenal, la fel ca și mine, însă aceasta nu schimbă faptul că are qualii fade în comparație cu cele pe care le instanțiez eu. Chiar propriile mele experiențe s-ar putea să fie fade, chiar când sunt cromatic intense, în comparație cu o persoană care se întâmplă să fie identică funcțional cu mine, dar care are experiențe cromatice hiper-vii din naștere.

Obiecția 5: *Argumentul pare să aibă o forță intuitivă în raport cu unele dintre argumentele epistemice, și anume argumentul conceptibilității lumilor zombie și argumentul conceptibilității lumilor inversate spectral, însă când vine vorba de un alt argument important, anume argumentul cunoașterii, își pierde puterea de atracție, deoarece implică afirmația că personajul Mary – super-cercetătoarea care cunoaște tot ceea ce este de cunoscut sub aspect fizic despre lume, dar și-a petrecut toată viața într-un mediu alb-negru – nu ajunge să cunoască nimic nou atunci când percepe un trandafir roșu pentru prima oară în viața ei.*

Verdictul intuitiv în cazul lui Mary este că ea ajunge, într-adevăr, să cunoască ceva nou atunci când vede un trandafir roșu pentru prima oară în viața ei. Dacă argumentul meu este corect, atunci proprietatea fenomenală instanțiată de către creierul lui Mary atunci când vede trandafirul este identică cu proprietatea neurofiziologică pe care creierul ei o instanțiază în acel moment, astfel că ea nu ar trebui să ajungă să cunoască nimic nou, dat fiind că ea cunoștea, înainte de a vedea trandafirul, totalitatea faptelor neurofiziologice. În ciuda aparențelor, însă, argumentul meu și raționamentul, în linii generale, pe care l-am propus nu sunt de fapt în conflict cu intuiția *prima facie* potrivit căreia Mary ajunge să cunoască ceva nou! La fel, în cazul lumilor *zombie*, raționamentul a pornit de la supoziția conceptibilității lor și a altor scenarii non-PNN. Ceea ce arată argumentul meu este că intuiția este în *ultimă instanță* eronată, și nu că ea este *prima facie*

eronată. Bineînțeles, toate acestea sunt consistente cu puterea de atracție inițială a intuiției.

Să vedem acum mai exact cum se aplică raționamentul meu argumentului cunoașterii. Vă amintiți că am subliniat anterior că atunci când discutăm argumentele bazate pe conceptibilitate suntem condiționați să ne concentrăm asupra unor scenarii retoric frapante, precum lumea *zombie* sau lumea inversată spectral. În cazul Mary, ne-am obișnuit să ne concentrăm asupra ipotezei că aceasta ajunge să cunoască cum e să vezi roșu atunci când vede un trandafir roșu pentru prima oară. Dar de ce tocmai *roșu*? De ce, cu alte cuvinte, credem că Mary ajunge să cunoască cum este să vezi roșu, și nu, să zicem, cum este să vezi verde, atunci când vede trandafirul roșu pentru prima oară? David Lewis a formulat, corect în opinia mea, argumentul cunoașterii apelând la ceea ce el numea "informație fenomenală", anume informație conținând posibilitățile ce sunt lăsate deschise de cunoașterea fizică completă posedată de Mary. Potrivit acestei formulări, înainte de a vedea trandafirul roșu, cunoașterea fizică completă aflată în posesia lui Mary lăsa deschise o infinitate de posibilități fenomenale despre lumea actuală. Vederea roșului pentru prima oară este echivalentă cu eliminarea tuturor acestor posibilități, cu excepția uneia singure. Însă, dat fiind tot ce știm până acum din datele problemei, nu suntem justificați în a crede că posibilitatea actualizată în câmpul fenomenal al lui Mary este roșul. Dimpotrivă, dat fiind că actualizarea roșului este doar una dintre infinitele posibilități fenomenale ce sunt lăsate deschise de către cunoașterea fizică completă posedată de Mary, ar trebui să atribuim o probabilitate extrem de mică (aproximativ zero) acestei propoziții. Cu toate acestea, pornind de la propriul caz, adică de la faptul că eu însumi am experiența de roșu la vederea trandafirului roșu, împreună cu principiul indiferenței, pot deduce, probabilistic, că este aproape sigur că lumea este în PNN, astfel că Mary are experiența de roșu atunci când vede un trandafir roșu, și nu experiența de verde sau orice altă culoare. Dar dacă așa stau lucrurile, atunci aproape sigur, în toate lumile care sunt duplicate fizice ale lumii actuale, Mary ajunge să cunoască cum este să vezi roșu, și nu orice altă culoare, atunci când vede trandafirul; altfel, faptul că lumea actuală este în PNN ar fi o coincidență enormă. Această ultimă afirmație este echivalentă cu afirmația că relația dintre starea cerebrală a lui Mary, atunci când vede trandafirul, și roșul fenomenal este necesară din punct de vedere epistemic, dat fiind că am ajuns la această concluzie printr-un raționament probabilistic a priori. Din moment ce necesitatea brută este de repudiat, putem acum postula relația de identitate între tipul de stare cerebrală a lui Mary și roșul fenomenal, deoarece identitatea ar explica de ce corelația există în mod necesar. Aceasta înseamnă că intuiția că Mary ajunge să cunoască ceva nou atunci când vede roșu pentru prima dată este, la finalul analizei, eronată, dar, deoarece am purces de-a lungul raționamentului nostru probabilistic asumând – de dragul unui *reductio* probabilistic – existența informației fenomenale, merită să elimine posibilități lăsate deschise de cunoașterea fizică, raționamentul nostru este perfect compatibil cu

existența intuiției că Mary ajunge să cunoască ceva nou, deși este, desigur, în ultimă instanță incompatibil cu *adevărul* acesteia.

Concluzia este la fel de radicală precum cea a lui Lewis, i.e. că nu există informație fenomenală, sau precum răspunsul lui Daniel Dennett (1991) la argumentul cunoașterii, potrivit căruia Mary nu ajunge să cunoască nimic nou. Diferența este că, în opinia mea, nici Lewis, nici Dennett nu oferă un argument destul de bun pentru această teză, dincolo de a o afirma pur și simplu. Argumentul oferit aici pare a fi cel puțin un bun candidat pentru *derivarea*, prin raționament probabilistic a priori, a acestei concluzii radicale.

Trebuie să admit, totuși, că argumentul cunoașterii este deconcertant tocmai din cauză că avem intuiția *prima facie* că Mary ajunge să dobândească o informație nouă. Trebuie doar să ne gândim la faptul că, probabil, Mary este foarte surprinsă atunci când vede lumea în culori pentru prima dată și se gândește cât de multe lucruri există în lume despre care ea nu știa. Cu toate acestea, dacă Mary ar parcurge raționamentul oferit în acest articol, și-ar revizui această opinie, potrivit căreia ea ar cunoaște informație nouă, chiar dacă ar continua să se minuneze de cât de colorată e lumea. Motivul pentru care găsim greu de acceptat răspunsul fizicist la argument este această îndărătnicie a sentimentului de surpriză chiar după ce revizuim intuiția inițială. Este greu de găsit o situație analoagă; cu toate acestea, voi oferi două parabole menite a exemplifica posibilitatea surprizei cauzate de informație, coexistând cu lipsa oricărei informații *noi*.

Parabola 1: Simourgh și cele treizeci de păsări. În fabula sa “Conferința păsărilor”, poetul persan Farid ud-Din Attar, reprezentant al tradiției Sufi în cultura islamică, înfățișează pelerinajul unor păsări în căutarea regelui lor (adică a lui Dumnezeu). La finalul călătoriei, doar treizeci de păsări ajung în fața porții lui Simourgh, regele pe care-l căutau. O “voce neauzită” le spune că ele însele, cele treizeci de păsări, sunt de fapt Simourgh. “Simourgh” înseamnă literal, în persană, “treizeci de păsări”, din “si” (treizeci) și “mourgh” (păsări). Vocea divină le spune:

“Though you traversed the Valleys’ depths and fought
With all the dangers that the journey brought
The journey was in Me, the deeds were Mine –
You slept secure in Being’s inmost shrine (...)”
(în traducerea lui Afkham Darbandi și Dick Davis 1984, p. 220)

Adoptând o interpretare panteistă sau panenteistă a povestirii,¹⁴ putem spune că, aparent, păsările află o informație nouă de la Dumnezeu, dar, dat fiind conținutul informației (anume, “voi sunteți Eu”, “faptele voastre au fost faptele Mele” etc.), este plauzibil să ne întoarcem la începutul povestirii și să revizuim credința cum că păsările sunt distincte de Dumnezeu, astfel încât chiar ideea de la început, că păsările pornesc să-l caute pe Simourgh, ca și cum Simourgh nu ar fi deja o cunoscută a povestirii în acel moment, ajunge să fie revizuită. Surprinderea păsărilor cauzată de informația al cărei autor este Simourgh coexistă cu faptul că, în termenii conținutului său, informația trebuie să existe deja într-o anumită formă, dat fiind că păsările sunt identice cu Simourgh (sau parte proprie a sa, dacă interpretăm povestea în mod panenteist).

Parabola 2: Oracolul fatal. Vrei să afli timpul, locul și circumstanțele exacte ale morții tale. Ți se spune că există un oracol care știe totul despre viitor. Te decizi, deci, să vizitezi oracolul, să-l plătești și să afli informația dorită. Întâlnești oracolul, îl plătești, iar în momentul următor oracolul te înjunghie în inimă. Mori în doar câteva secunde. Chiar înainte să mori, erai cât se poate de surprins și, aparent, primeai informația dorită, în modul cel mai realist cu putință. Totuși, nu ai aflat nicio informație nouă privind *moartea* ta, deoarece până în momentul în care ai murit informația încă nu era completă (încă nu muriseși, puteai fi salvat etc.), iar după ce ai murit nu mai ești prezent prin preajmă ca să fii în posesia informației respective, completă de data aceasta, din păcate. A existat o informație relevantă, a existat o stare de surprindere cauzată de transferul acesteia și, totuși, nu a existat nicio informație *nouă* pe care să o fi aflat în privința morții tale.

Aceste parabole nu sunt, desigur, analoage cu ce se întâmplă în cazul lui Mary, dar sunt, totuși, menite a sprijini ideea că intuițiile *prima facie* despre existența informațiilor noi, precum și starea de surprindere aferentă, nu sunt echivalente cu o demonstrație *bona fide* a adevărului dualismului.

6. Concluzie

Am pornit la drum prin analiza tezei timpurii a identității minte-creier, iar după un detur prin dialectica de după Feigl privind problema minte-corp, și luând în considerare cele mai puternice argumente împotriva tezei, am ajuns la aceeași concluzie ca și Feigl, Smart și Place, doar că într-un mod mai ocolit. Noțiunea lui Feigl de ciucuri nomologice ne-a ajutat să construim argumentul probabilistic împotriva coerenței dualismului

¹⁴ Ezgi Ulusoy Aranyosi (2009) – căreia îi datorez exemplul – oferă o astfel de analiză panteistă a textului lui Attar, punându-l în opoziție cu abordările ortodoxe asupra metafizicii sufiste, bazate pe textele metafizicienilor islamici clasici, cum ar fi Ibn al-'Arabi.

naturalist, ceea ce ne-a netezit calea spre postularea relației de identitate drept singurul candidat extrem de probabil pentru relația mental-fizic.

Bibliografie

- ARANYOSI, István – *Excluding Exclusion: the Natural(istic) Dualist Approach*; în *Philosophical Explorations*, 11/1, 2008, pp. 67-78.
- ARMSTRONG, D.M. – *A Materialist Theory of the Mind*; Londra, Routledge, 1968.
- ATTAR, Farid ud-Din – *The Conference of the Birds* [1177]; trad. Afkham Darbandi și Dick Davis, Londra, Penguin Books, 1984.
- BOSTROM, N. – *Anthropic Bias: Observation Selection Effects in Science and Philosophy*; New York, Routledge, 2002.
- CASTELL, P. – *A Consistent Restriction of the Principle of Indifference*; în *British Journal for the Philosophy of Science*, 49/3, 1998, pp. 387-396.
- CHALMERS, David J. – *Facing Up to the Problem of Consciousness*; în *Journal of Consciousness Studies*, 2(3), 1995, pp. 200-219; retipărit în Hameroff, S., Kaszniak, A. & A. Scott (ed.) – *Toward a Science of Consciousness*; MIT Press, 1996.
- Idem – *The Conscious Mind: In Search of a Fundamental Theory*; Oxford University Press, 1996.
- Idem – *Consciousness and Its Place in Nature*; în Stich, Stephen P. & Ted A. Warfield (ed.) – *The Blackwell Guide to Philosophy of Mind*; Blackwell Publishing, 2003.
- CRANE, T. – *The Significance of Emergence*; în Gillett, Carl & Barry Loewer (ed.) – *Physicalism and its Discontents*; Cambridge, Cambridge University Press, 2001, pp. 207-224.
- DENNETT, Daniel – *Consciousness Explained*, Boston, Little Brown and Company, 1991.
- FEIGL, H. – *The "Mental" and the "Physical"*; în Feigl, H., Scriven, M. & G. Maxwell (ed.) – *Concepts, Theories and the Mind-Body Problem*; Minneapolis, Minnesota Studies in the Philosophy of Science, Vol. 2, 1958; retipărit cu un *Postscriptum* în H. Feigl – *The 'Mental' and the 'Physical', The Essay and a Postscript*; Minneapolis, University of Minnesota Press, 1967.
- KIRK, R. – *Sentience and Behaviour*; în *Mind*, 83/329, 1974a, pp. 43-60.
- Idem – *Zombies vs. Materialists*; în *Proceedings of the Aristotelian Society*, 48, 1974b, pp. 135-152.
- KRIPKE, Saul A. – *Naming and Necessity*; Harvard University Press, 1999 [1972].
- LESLIE, J. – *Universes*; New York, Routledge, 1989.
- LEVINE, J. – *Materialism and Qualia: The Explanatory Gap*; în *Pacific Philosophical Quarterly*, 64/3, 1983, pp. 354-361.
- LEWIS, David K. – *An Argument for the Identity Theory*; în *Journal of Philosophy*, 63/1, 1966, pp. 17-25.
- Idem – *How to Define Theoretical Terms*; în *Journal of Philosophy* 67/13, 1970, pp. 427-446.
- Idem – *Psychophysical and Theoretical Identifications*; în *Australasian Journal of Philosophy*, 50/3, 1972, pp. 249-258.
- Idem – *Relevant Implication*; în *Theoria* 54/3, 1988, pp. 161-174; retipărit în D.K. Lewis – *Papers in Philosophical Logic*; Cambridge Studies in Philosophy, Cambridge University Press, 1998, pp. 111-124.
- Idem – *What Experience Teaches*; în William G. Lycan (ed.) – *Mind and Cognition: A Reader*; Blackwell, Oxford, 1990, pp. 499-519; retipărit în D.K. Lewis – *Papers in Metaphysics and Epistemology*, Vol. 2; Cambridge, Cambridge University Press, 1999, pp. 262-291.

- Idem – *How Not to Solve the Mind-Body Problem*; în Gillett, Carl & Barry Loewer (ed.) – *Physicalism and its Discontents*; Cambridge University Press, 2001, pp. 284-306.
- NAGEL, Thomas – *What is it Like to Be a Bat?*; în *Philosophical Review*, 83/4, 1974, pp. 435-450; retipărit în Th. Nagel – *Mortal Questions*; Cambridge, Cambridge University Press, 1979.
- NOLAN, D. – *Impossible Worlds: A Modest Approach*; în *Notre Dame Journal for Formal Logic*, 38/4, 1997, pp. 535-572.
- PAPINEAU, David – *Thinking about Consciousness*; Oxford University Press, 2002.
- PELCZAR, M. – *On an Argument for Functional Invariance*; în *Minds and Machines*, 18/3, 2008, pp. 373-377.
- PLACE, U.T. – *Is Consciousness a Brain Process?*; în *British Journal of Psychology*, 47/1, 1956, pp. 44-50.
- SMART, J.J.C. – *Sensations and Brain Processes*; în *Philosophical Review*, 68/2, 1959, pp. 141-156.
- STOLJAR, D. – *Physicalism*; în Edward N. Zalta (ed.) – *The Stanford Encyclopedia of Philosophy*, URL=<http://plato.stanford.edu/entries/physicalism/>, 2001.
- ULUSOY ARANYOSI, Ezgi – *An Enquiry into Sufi Metaphysics*, Senior Thesis, Bilkent University; în curs de apariție în *British Journal for the History of Philosophy*, 2009.