

# Reprezentationalismul și scenariile de tipul pământului inversat

Prep.drd. Bogdan Rusu

## Abstract

*The explanation of the phenomenal consciousness depends a great deal on the explanation of qualia, the phenomenal features of our experiences. One good established theory about qualia is the representationism, strongly supported by philosophers like Fred Dretske or Michael Tye. Among the most powerful critiques of the representationism we find Ned Block. His counterarguments come mainly from the direction of putative counterexamples, best known as "the inverted spectrum" and "the swampman". The present paper proposes a representationist theory of qualia and tries to defend it mainly against the inverted spectrum scenario, the most dangerous way of attacking representationism about qualia. The paper provides an analysis of other philosophers strategies of defence, all based on the attempt of proving that the hypothesis of the inverted spectrum is incoherent. We try to provide an argument against the logical impossibility of a spectral inversion, based on the physical relations between the different length-waves of the visible electro-magnetic radiation. The argument shows that a spectral inversion is not possible because of the asymmetry of the visible spectrum of the light.*

## 1. O teorie funcționalist-reprezentationalistă

"Qualia"<sup>1</sup> este un nume care desemnează ceea ce Locke numea "calități secundare" ale obiectelor. De fiecare dată când iau contact senzorial cu obiectele, am experiențe legate de culoarea acestora, de mirosul sau gustul lor etc. De fiecare dată eu sunt subiectul unei stări mentale cu un puternic caracter subiectiv. Așa cum spune Nagel<sup>2</sup>, există ceva precum felul cum este pentru mine să am astfel de experiențe. Caractere fenomenale ale experiențelor, caractere subiective, calități fenomenale, calități secundare – acestea sunt numele sub care au apărut în istoria filosofiei aceste entități denumite generic acum, de către majoritatea filosofilor, qualia (sg. quale).

După o sinteză a lui Dennett<sup>3</sup>, qualia sunt proprietăți ale stărilor mentale ale unui subiect, caracterizate în mod tradițional de următoarele atribute:

- (1) sunt inefabile;
- (2) sunt intrinseci;
- (3) sunt private;
- (4) sunt direct sau nemijlocit aprehensibile în conștiință.

Așa cum se prezintă lucrurile, este greu de negat că există qualia, tocmai datorită faptului că sunt direct accesibile introspectiv. Din punct de vedere filosofic însă se pune problema criteriilor de identitate

ale acestor entități. Cu alte cuvinte, ce poate să însemne că există qualia?

Întrebarea aceasta este fundamentală, întrucât explicația conștiinței fenomenale<sup>4</sup> este în mare măsură explicația acestor quale. Din această cauză toate teoriile metafizice ale minții au trebuit să se confrunte cu problema de a explica qualia, în afara cazului când au negat că există qualia, precum Dennett<sup>5</sup>. De fiecare dată qualia s-au dovedit a fi niște hapuri greu de înghițit. Searle<sup>6</sup> evidențiază acest lucru atunci când trece în revistă materialismul recent. Din tabelul său vedem că tocmai cele mai puternică teorie este cea care are și cele mai mari probleme cu qualia: funcționalismul și mai ales funcționalismul reprezentationalist (Turing machine functionalism).

Ceea ce vreau să susțin în această lucrare este o teorie a calităților fenomenale care comportă următoarele trei componente:

- (1) teoria reprezentatională a minții (computationalism);
- (2) psihofuncționalism nonreductiv;
- (3) reprezentationalism moderat despre qualia.

RTM (Representational Theory of Mind) constă din următoarele teze<sup>7</sup>:

- (A) Teoria reprezentatională a gândului: Pentru fiecare atitudine propozițională A, există o relație psihologică unică și distinctă R, și pentru orice propoziție P și

<sup>1</sup> Părți din acest articol au fost publicate sub titlul „Qualia și scenariile inversiunii spectrale” în revista de filosofie *Eidos*, nr. 3/2003, pp. 44-58.

<sup>2</sup> NAGEL [1974]

<sup>3</sup> DENNETT [1990]

<sup>4</sup> vezi diferența dintre conștiință fenomenală și conștiință psihologică în CHALMERS [1996], cap. 1.

<sup>5</sup> *ibidem*

<sup>6</sup> SEARLE [1992], p.53

<sup>7</sup> conform FODOR [1987]

subiect S, S are atitudinea A față de P dacă există o reprezentare mentală #P# astfel încât

- (i) S poartă R la #p# și
- (ii) #P# înseamnă că P.

(B) Teoria reprezentatională a gândirii: Procesele mentale constau din secvențe cauzale de simboluri ale reprezentărilor mentale.

De notat că Fodor consideră relația R ca fiind de natură funcțională. Însă funcționalismul lui Fodor este unul materialist. Cum eu nu doresc să susțin teoria limbajului gândirii, mă pot dispensa de ea pentru a adopta ipoteza metafizică a psihofuncționalismului.

**Psihofuncționalismul** este o teorie cu privire la individuarea stărilor mentale, expusă cel mai clar într-o lucrare a lui Ned Block<sup>8</sup>. Conform acestei teorii stările mentale sunt identice cu stările psihofuncționale ale creierului. Stările mentale sunt explicabile psihofuncționalist în termeni de input-uri (stimuli periferici), output-uri (comportamente) și cauzalitate între stări mentale definite abstract. Psihofuncționalismul nonreductiv pe care îl susțin eu este o variantă a teoriei lui Chalmers. Astfel, consider că stările mentale sunt determinate de organizarea funcțională, dar ele nu pot fi reduse la funcțional. Stările mentale, după această teorie, supervin pe funcțional. Voi arăta mai încolo că această teorie este necesară pentru a păstra coerența unor experimente mentale foarte relevante pentru ceea ce pot spune despre conceptele noastre.

**Reprezentationalismul** este o teză despre caracterul fenomenal al experienței, susținută cel mai puternic de Michael Tye. Teza are două versiuni: una tare și una slabă. După reprezentationalismul tare, qualia sunt identice cu conținuturile reprezentationale non-conceptuale. În versiunea moderată, reprezentationalismul susține superveniența calităților fenomenale ale stărilor mentale pe conținutul lor reprezentational non-conceptual. Reprezentationalismul este de asemenea și externalist cu privire la conținutul reprezentational. El susține că ceea ce reprezintă o experiență este metafizic determinată, cel puțin parțial, de factori din mediul exterior. Adică, se susține de obicei, dacă am avea doi duplicați microfizici și aceștia ar avea conținuturi reprezentationale diferite, atunci, dacă diferențele ar fi de tipul care trebuie, cei doi duplicați vor diferi necesarmente și în ce privește caracterele fenomenale ale experiențelor lor. Diferențele între conținuturi reprezentationale se datorează diferențelor externe, conexiunilor diferite ale subiecților cu diferiți itemi din mediul înconjurător. Această conexiune generică este denumită "urmărire" (tracking), dar nu există o înțelegere precisă a ce înseamnă asta. Datorită diferențelor de mediu, cei doi duplicați microfizici diferă în ce privește conținuturile

reprezentationale *extinse* ale stărilor lor mentale. Ei au în schimb aceleași conținuturi reprezentationale *restrânse*, determinate de rolul funcțional al acelor stări mentale în sistemele de reprezentări ale subiecților.

Reprezentationalismul tare nu este compatibil cu funcționalismul, deoarece el este o teorie despre natura calităților fenomenale (pe care le reduce la conținuturi reprezentationale), la fel cum este și funcționalismul; pe când reprezentationalismul moderat lasă loc pentru o teorie despre natura calităților fenomenale ale experiențelor<sup>9</sup>.

Teoria pe care o susțin afirmă, așadar, că cogniția este reprezentatională, că stările mentale sunt determinate de organizarea funcțională a creierului și că proprietățile fenomenale ale stărilor mentale supervin pe conținuturile reprezentationale non-conceptuale. Asta înseamnă că de fiecare dată când am o experiență, în cutia reprezentatională este înscris un simbol cu un anumit rol funcțional, simbol care determină caracterul fenomenal al experienței respective. Acest simbol constituie output-ul modulului perceptiv, care va servi ca input pentru modulul cogniției superioare<sup>10</sup>.

Teoria pe care tocmai am schițat-o trebuie să facă față următoarelor obiecții:

- (i) externalismul despre conținuturile mentale duce la probleme pentru superveniență, și astfel este eliminată orice posibilitate de a explica cauzalitatea mentală;
- (ii) reprezentationalismul despre qualia întâmpină probleme deoarece nu poate menține superveniența;
- (iii) psihofuncționalismul are probleme cu qualia, încălcând principiul indiscernabilității identicilor.

Aceste obiecții apar de obicei sub forma contraexemplilor furnizate prin experimente mentale. Astfel, obiecția (i) apare din celebrul Pământ Geamă al lui Putnam, obiecția (ii) se naște din Pământul Inversat al lui Block, iar (iii) din Spectrul Inversat al lui Shoemaker (+ Pământul Inversat).

În continuare voi încerca să văd cum poate teoria pe care o susțin să depășească aceste obiecții.

## 2. PĂMÂNTUL GEAMĂ ȘI PROBLEMELE PENTRU CONȚINUTURILE MENTALE

### 2.1 Pământul geamă

Așa cum se știe, Pământul Geamă este o planetă la fel ca și Pământul, cu deosebirea că în loc de apă (H<sub>2</sub>O), pe Pământul Geamă există apă<sub>2</sub>, compusă din substanțele chimice XYZ. Apa<sub>2</sub> are exact aceleași proprietăți macrofizice ca și apa: este un lichid transparent, insipid, incolor, inodor, care dizolvă

<sup>8</sup> BLOCK [1980]

<sup>9</sup> conform TYE [1997]

<sup>10</sup> Se subînțelege că accept o teorie modulară a minții.

zahărul, fierbe la 100 grade Celsius și îngheață la 0 grade. Fiecare locuitor al Pământului are un doppelganger pe Pământul Geamăn, doppelganger care este un duplicat moleculă cu moleculă al omologului său terestru. Locuitorii țării omoloage României de pe Pământul Geamăn vorbesc un dialect denumit română2. Atunci când Oscar – un locuitor al Pământului – crede că apa este udă, el se referă la substanța H<sub>2</sub>O. Doppelganger-ul său, crezând aceeași propoziție, crede în realitate ceva despre apa2 (desemnată prin cuvântul „apa2”, cuvânt ce nu diferă la pronunție de „apă”) – adică despre substanța XYZ.

Intuiția pe care acest experiment mental ne invită să o împărtășim este aceea că Oscar2 nu poate avea nici un fel de opinii despre apă.

O variantă a „Pământului Geamăn” care extinde morala și asupra termenilor care nu stau pentru genuri naturale este cea a lui Burge. Adaptată la limba română, această variantă apare în felul următor:

Cuvântul „potroace” înseamnă în românește măruntaie de pasăre (în special măruntaiele acestor păsări folosite ca hrană). În română2, „potroace2” înseamnă măruntaie de găină. Oscar crede că „potroace” înseamnă doar măruntaie de găină. După Oscar, gășca, de exemplu, este lipsită de potroace. Prima intuiție pe care ne-o trezește acest scenariu este că Oscar poate avea opinii despre potroace. Chiar și opinia că gășca nu are potroace este o opinie despre potroace. A doua intuiție care ne este indusă este că Oscar2 nu poate avea atitudini propoziționale despre potroace, întrucât lui îi lipsește conceptul „potroace”. El are în schimb conceptul „potroace2” care, explicat în românește, înseamnă „măruntaie de găină”.

## 2.2 Problema

Aceste două scenarii, așa cum sunt ele interpretate de către Fodor, pun în joc ceva în legătură cu felul cum sunt individuate atitudinile propoziționale. Asume, intuițiile induse de acestea sugerează că individuirea atitudinilor propoziționale se face cu privire la proprietățile lor relaționale. (Avem de-a face cu două tipuri de criterii de individuație: relaționale și individualiste. Criteriile relaționale sunt cele după care opiniile lui Oscar diferă în conținut de cele ale lui Oscar2. După standardele individualiste, Oscar și Oscar2 au aceeași stare mentală).

Dat fiind că intuițiile noastre ne conduc să acceptăm că standardele de individuire ale atitudinilor propoziționale sunt relaționale, Fodor consideră că scenariul Pământului Geamăn spune următorul lucru despre atitudinile propoziționale: că, așa cum sunt înțelese în mod obișnuit de către simțul comun, ele nu supervin stărilor creierului.

Conceptul de supervinență se definește de obicei astfel: Fie două tipuri de proprietăți: A-proprietăți și B-proprietăți, dintre care primele sunt

proprietăți de nivel jos, iar ultimile sunt proprietăți de un nivel mai înalt. Se spune că B-proprietățile supervin peste A-proprietăți dacă nu există două situații care să fie identice cu privire la A-proprietățile lor și să difere în același timp cu privire la B-proprietățile lor (cf. Chalmers). În cazul nostru A-proprietățile sunt proprietățile fiziologice ale creierului, iar B-proprietățile sunt stările mentale (atitudinile propoziționale). Întrucât Oscar și Oscar2 sunt duplicați moleculari, ei au aceleași A-proprietăți. Însă, așa cum arată experimentul, având atitudini propoziționale diferite, B-proprietățile lor diferă.

Prin urmare, scenariul Pământului Geamăn implică faptul că individuirea atitudinilor propoziționale se face relațional. În același timp presupune că stările creierului sunt individuate individualist, i.e. că, potrivit simțului comun, Oscar și Oscar2 au aceleași stări ale creierului. Dacă admitem această presuposiție, atunci „Pământul Geamăn” va ridica cu adevărat o problemă pentru cauzalitatea mentală, deoarece furnizează un exemplu care sparge supervinența stărilor mentale pe stările creierului.

Așa cum arată Fodor, însă, nu avem nici un motiv să respingem această presuposiție; mai mult, Fodor va argumenta că în știință în general individuirea entităților din ontologia științei trebuie să se facă individualist. Odată stabilit acest lucru este ușor de văzut că nu avem nici un temei de a individua și stările mentale altfel decât individualist.

Trebuie făcută aici o precizare terminologică: „individualist” nu înseamnă la Fodor „non-relațional”. O entitate este individuată individualist atunci când se ține cont doar de puterile ei cauzale. Aici se poate foarte ușor strecura confuzia între „individualismul metodologic” și „solipsismul metodologic”. Această confuzie are drept rezultat pretenția ca entitățile (în cazul de față: atitudinile propoziționale) să fie individuate non-relațional. Aceasta vrea să spună că aparatul taxonomic al științei nu trebuie să includă nici un termen care desemnează o proprietate relațională.

Acest lucru, însă, nu este de susținut. De exemplu, taxonomia astronomică conține termenul „a fi planetă”, care desemnează o proprietate relațională. Ce se întâmplă de fapt?

Există două tipuri de proprietăți relaționale: unele influențează puterile cauzale ale entităților, altele nu le influențează. Să considerăm în această privință exemplul lui Fodor. Am o monedă dintre acelea care nu cad niciodată pe cant. Oricum aș arunca-o, moneda nu ateriza fie pe partea cu „marca” (stema), fie pe partea cu „banul”. Definesc acum două predicate: „a fi B-particulă” și „a fi M-particulă”. Un atom, să spunem, este o B-particulă dacă la momentul T moneda mea este cu „banul” în sus; dacă moneda este cu „marca” în sus, atunci atomul satisface predicată „a fi o M-particulă”. Astfel, printr-o simplă aruncare a monedei, eu pot

schimba starea tuturor particulelor din univers. Însă aceste proprietăți relaționale ale particulelor nu afectează în nici un fel puterile lor cauzale. Ar fi, deci, nebunește ca în aparatul taxonomic al fizicii să includem aceste predicate, deoarece fizica se ocupă cu explicația cauzală.

Principiul individualismului metodologic nu împiedică însă includerea în taxonomii a proprietăților relaționale. Prevede însă că taxonomiile științifice pot include numai acele proprietăți relaționale care determină puterile cauzale ale entităților în cauză. Astfel, proprietate a unui corp de a fi planetă determină ce traiectorie în spațiu va avea corpul, ce alte corpuri poate să lovească, cu ce forță etc.

„Solipsismul metodologic” este o teorie empirică a minții care susține că procesele mentale sunt computaționale, deci sintactice și că, prin urmare, stările psihologice sunt individuate fără legătură cu evaluarea lor semantică. Fodor este un externalist; așa că evaluarea semantică a atitudinilor propoziționale se face pentru el în legătură cu ceva extra-psihic. Confuzia solipsismului metodologic cu individualismul metodologic poate, astfel, implica afirmația că stările psihologice trebuie individuate non-relațional. Pe baza acestei confuzii se poate argumenta contra individualismului arătând că taxonomia psihologică este deseori relațională, adică non-solipsistă (datorită abundenței explicațiilor teleologice din psihologie).

Dacă individualismul metodologic este, însă, înțeles corect, atunci este limpede că taxonomia psihologică nu trebuie să cuprindă termeni ce desemnează proprietăți relaționale ale entităților din ontologia acestei științe.

Conform simțului comun se pot afirma următoarele:

- (1) stările creierului sunt individuate individualist (presupoziție a „Pământului Geamăn”);
- (2) atitudinile propoziționale sunt individuate relațional (implicație a „Pământului Geamăn”).

Aceste două teze implică spargerea supervenienței stărilor mentale pe stările creierului.

Observațiile lui Fodor, însă, arată că nu avem nici un temei pentru a individua stările mentale altfel decât individualist. Cerința menținerii supervenienței necesită astfel violarea taxonomiei psihologiei populare.

Dacă acceptăm acum că stările mentale supervin stărilor creierului, cădem în paradoxul lui Putnam: trebuie să acceptăm că gândurile nu au condiții de adevăr *esențialmente*. Două instanțe ale *aceluiași* gând pot avea condiții de adevăr *diferite*, deci valori de adevăr diferite. Ceea ce înseamnă că extensiunea atitudinilor propoziționale nu este determinată de conținutul lor.

Ceea ce lipsește acestui comentariu asupra Pământului Geamăn pentru a releva adevărata problemă

pe care o ridică pentru conținuturile mentale este următoarea diagnoză:

- (a) Înainte de „Pământul Geamăn” aveam o teorie a individuării conținuturilor: Identitatea conținuturilor depindea de identitatea extensiunilor. Adică puteam spune că extensiunea supervine pe conținut.
  - (b) Acum vedem că este fals că diferența extensiunilor implică diferența conținuturilor, deoarece două instanțe ale aceluiași gând („Apa este udă” la Oscar, „Apa2 este udă” la Oscar2) pot avea *condiții de adevăr* diferite, deci *valori de adevăr* diferite.
  - (c) Psihologia nu poate renunța la superveniență.
- (a), (b) și (c) arată că, dacă vrem să păstrăm superveniența, „Pământul Geamăn” pune probleme deosebite conținuturilor mentale, deoarece sparge conexiune dintre acestea și extensiuni, adică ne lasă fără criterii de identitate pentru conținuturile mentale.

Jerry Fodor consideră că diagnoza de mai sus se bazează pe o eroare: „Pământul Geamăn” nu sparge conexiunea dintre conținuturi și extensiuni, ci doar o relativizează la context. Tratamentul pe care-l aplică Fodor conținuturilor mentale este similar tratamentului kaplanian al indexicalilor: conținuturile mentale sunt, astfel, funcții de la gânduri și contexte la condiții de adevăr, care condiții sunt funcții de la lumi posibile la valori de adevăr. Prin urmare „Pământul Geamăn” nu pune probleme conținuturilor mentale. În schimb îl ajută pe filosoful american să ofere o definiție a noțiunii de „conținut restrâns” cu care trebuie să opereze psihologia pentru a păstra atât superveniența, cât și explicațiile belief/desire.

În opinia mea, însă, Fodor nu are dreptate. Analiza pe care o întreprinde el este viciată de faptul că „Pământul Geamăn” suferă el însuși de o eroare de construcție, pe care o voi discuta în continuare.

### 2.3 Eroarea

Scenariul Pământului Geamăn presupune că Oscar și Oscar2 sunt duplicații moleculari. Ei trăiesc pe planete care diferă între ele doar prin faptul că apa ( $H_2O$ ) de pe Pământ este înlocuită pe Pământul Geamăn cu apă2 ( $XYZ$ ). Între Oscar și Pământ există, însă, o relație care ține de ecologia acestei planete: anume, corpul lui Oscar este compus, să zicem, din 80% apă, adică  $H_2O$ . Această proprietate relațională este una care afectează în mod cert puterile cauzale ale corpului lui Oscar. Prin urmare, Oscar este individuat și de proprietatea de a fi alcătuit în proporție de 80% din apă. El are această proprietate *esențialmente*; este metafizic imposibil ca el să nu fie alcătuit în acest procent din apă. Nu este metafizic imposibil ca Dumnezeu să fi creat o lume în care oamenii să nu aibă în compoziție apă; pe această bază și „Pământul Geamăn”. Însă este

imposibil ca, în lumea actuală, Oscar să nu fi conținut 80% apă.

Aceeași relație care există între Oscar și Pământ există, *mutatis mutandis*, și între Oscar2 și Pământul Geamăn: Oscar2, prin urmare, conține 80% apă2, adică XYZ, *esențialmente*.

Aceasta înseamnă că Oscar și Oscar2 nu pot fi duplicați moleculari. Posibilitatea „Pământului Geamăn” este pur metafizică. Imaginația lui Putnam s-a întins aici mai mult decât poate acoperi intelectul. Odată acesta pus la lucru, reiese clar că „Pământul Geamăn” este empiric imposibil. Este practic imposibil să existe o lume în care relația dintre un X și ecologia planetei să nu țină. Adică este imposibil ca, pe o planetă în care există XYZ, Oscar2 să fie compus 80% din H<sub>2</sub>O, substanță care nu există în altă parte pe planetă.

Cu siguranță că ne putem imagina o lume pe care Dumnezeu a creat-o în fiecare detaliu. Într-o astfel de lume în care există XYZ, Dumnezeu l-ar fi putut crea pe oscar2 astfel încât numai el să încalce relația cu ecologia planetei. Dacă am vedea lucrurile așa și în cazul nostru, ar însemna să renunțăm la știință în general și la biologia evoluționistă în special. Nu știu dacă Putnam este dispus la așa ceva; în ce mă privește, nu pot accepta aceste pierderi.

Dacă observația pe care am făcut-o este justă atunci, indiferent cum am individua stările creierului și stările mentale, „Pământul Geamăn” nu implică nici o problemă pentru superveniență și, deci, nu impune violarea taxonomiei psihologiei populare.

#### 2.4 Reconstrucția

Cele spuse mai sus implică încă o observație: „Pământul Geamăn” este în continuare coerent dacă susținem că Oscar și Oscar2 nu sunt duplicați moleculari, ci duplicați *funcționali*. În mod firesc suntem conduși să acceptăm că „Pământul Geamăn”, amendat în felul de mai sus, este posibil empiric numai dacă presupunem că o moleculă compusă într-o anumită proporție din XYZ poate avea în organismul lui Oscar2 același rol funcțional cu molecula omoloagă bazată pe apă în organismul lui Oscar.

Ca să păstrăm superveniența trebuie, deci, să acceptăm că stările mentale supervin peste stările funcționale ale creierului, la limită *sunt* stări funcționale ale creierului. „Pământul Geamăn” presupune, după părerea mea, *psihofuncționalismul*, dar pe cel *nonreductiv* susținut de Chalmers.

Conform psihofuncționalismului, stările mentale sunt individuate de totalitatea relațiilor cauzale și inferențiale, actuale și potențiale, în care pot intra în cadrul unui sistem de reprezentări<sup>11</sup>. Individuarea atitudinilor propoziționale este, deci, *strict relațională*, fără a viola prin aceasta individualismul metodologic,

deoarece proprietățile inferențiale și ale atitudinilor determină puterile cauzale ale lor.

Dacă reconstruim scenariul Pământului Geamăn în acord cu observațiile de mai sus, remarcăm imediat că:

(1\*) Stările creierului sunt individuate individualist.

(2\*) Stările mentale sunt individuate relațional, dar individualist.

Or, cele două afirmații nu ridică nici o problemă pentru superveniență.

Experimentul Pământului Geamăn arată că, atunci când gândim conceptul de conținut mental, acceptăm implicit teza psihofuncționalistă. Acceptând psihofuncționalismul, Pământul Geamăn nu mai pune nici o problemă pentru noțiunea de conținut mental, deoarece nu mai violează superveniența naturală a mentalului pe funcțional.

### 3. PĂMÂNTUL INVERSAT ȘI PROBLEMELE PENTRU CARACTERELE FENOMENALE

#### 3.1 Experimentele și problema

Scenariul spectrului inversat<sup>12</sup> prezintă aspectul unei inversiuni intersubiective a qualia. Problema a fost sesizată, se pare, prima dată de Locke, dar formularea contemporană cea mai des citată este cea a lui Shoemaker. Acest experiment mental se bazează pe intuiția că este foarte plauzibil că lucrurile ar putea să arate pentru cineva exact invers decât arată pentru mine. De pildă, el ar putea să vadă în locul fiecărei culori a spectrului exact culoarea complementară: acolo unde eu văd roșu, el ar putea să vadă verde și invers; acolo unde eu văd albastru, el ar putea să vadă galben și viceversa etc. Datorită caracterului privat al experiențelor noastre coloristice, datorită deci conectării lor esențiale cu un punct de vedere de persoana I, eu n-aș putea niciodată să îmi dau seama că celălalt percepe culorile exact invers decât mine. Persoana în cauză ar putea să opereze aceleași discriminări pe care le operează și eu și, datorită caracterului public al limbajului, el va folosi cuvintele care se referă la culori exact în același fel în care le folosesc și eu.

Problema metafizică pe care o ridică acest scenariu constă în încălcarea legii lui Leibniz care se obține dacă interpretăm qualia funcționalist. În cadrul acestui experiment este pus în evidență faptul că două stări mentale diferite (percepția roșului și percepția verdelui) pot avea același rol funcțional:

„Pare că posibilitatea inversiunii intersubiective a spectrului este de asemenea incompatibilă cu ceea ce mulți consideră a fi cel mai respectabil descendent al behaviorismului, anume funcționalismul, unde acesta este înțeles ca viziunea după care stările

<sup>11</sup> BLOCK [1980]

<sup>12</sup> SHOEMAKER [1982]

mentale sunt definibile în termenii relațiilor lor cauzale cu input-urile senzoriale, output-urile comportamentale și alte stări mentale. Dacă spectrul inversat este o posibilitate logică, atunci acea quale implicată curent în percepția mea a lucrurilor albastre nu poate fi definită prin rolul funcțional pe care îl joacă în cei asemenea mie, deoarece în cineva al cărui spectru a fost inversat față de al meu o quale diferită (poate acea quale implicată în percepția mea a lucrurilor galbene) ar juca acel rol funcțional și această quale ar juca un rol diferit. Aceasta este o versiune a ceea ce a fost numit "obiecția qualei inversate la funcționalism – inversiunea spectrului fiind un caz special de "inversiune a qualia"<sup>13</sup>.

Complicarea acestui experiment cu inversiunea intrasubiectivă a spectrului a rezultat, drept cel mai cunoscut caz, în celebrul scenariu al Pământului Inversat al lui Block<sup>14</sup>:

„Pământul Inversat diferă de Pământ în două privințe. În primul rând, orice lucru are culoarea complementară culorii de pe Pământ. Cerul este galben, iarba este roșie, hidranturile sunt verzi și așa mai departe. [...] În al doilea rând, vocabularul locuitorilor de pe Pământul Inversat este de asemenea inversat: dacă întrebi ce culoare are cerul (galben), ei îți vor spune (cinstit) "albastru!". Dacă îi întrebi ce culoare are iarba (roșie), ei spun "verde". [...] O echipă de savanți nebuni te adorm. Cât timp ești inconștient, îți inserează lentile care inversează culorile în ochi și îți schimbă pigmenții ca să nu ai un șoc neplăcut când te trezești și te uiți la picioarele tale. Te transportă pe Pământul Inversat, unde înlocuiești un reprezentant care a ocupat o nișă pe Pământul Inversat care corespunde exact (cu excepția culorii lucrurilor) nișei tale de acasă. Te trezești și, deoarece lentilele inversează anulează culorile inversate, nu observi nici o diferență. "Felul cum este" pentru tine să interacționezi cu lumea și cu alți oameni nu se schimbă deloc. [...] În măsura în care este vizat aspectul calitativ al vieții tale mentale, nimic nu este diferit de felul în care ar fi fost dacă ai fi rămas acasă"<sup>15</sup>.

Acest scenariu nu numai că pune mai acut probleme funcționalismului (un exemplu similar este prezentat în BLOCK [1982] în scopul expres de a respinge funcționalismul) dar pune probleme deosebite și reprezentationalismului. Așa cum arată același Ned

Block<sup>16</sup>, trăsăturile cheie ale acestui scenariu sunt următoarele:

- (1) Caracterul fenomenal al experiențelor tale rămâne același (pentru că nu percepi nici o diferență atunci când te trezești).
- (2) Conținutul reprezentational, fiind externalist, se modifică odată cu condițiile de mediu și cu comunitatea lingvistică.

Din cele de mai sus rezultă că există un gol (gap) între conținutul reprezentational și caracterul fenomenal al experiențelor. Iar acest gol arată faptul că trăsăturile fenomenale nu supervin peste conținuturile reprezentational. Prin urmare reprezentationalismul este fals.

### 3.2 Soluția lui Michael Tye

Apărarea tezei reprezentationaliste nu este un lucru ușor, deoarece intuiția noastră că trăsăturile fenomenale ale experienței nu variază odată cu emigrarea pe Pământul Inversat este foarte puternică. Michael Tye<sup>17</sup> propune însă un mod de a scoate reprezentationalismul la liman.

Filosoful britanic propune o interpretare a covarianței cauzale a experiențelor perceptuale de bază. El susține că pentru orice stare senzorială S a unei creaturi c, din setul relevant de stări senzoriale alternative ale lui c, putem defini ce reprezintă S după cum urmează:

S reprezintă că P =df dacă condițiile optime ar fi întrunite, S ar fi simbolizat în c ddacă ar fi cazul că P; mai mult, în aceste împrejurări, S ar fi simbolizat în c pentru că este cazul că P.

Condițiile optime care apar în definiția de mai sus sunt la Tye condiții de bună funcționare a unui organism. Ele sunt întrunite atunci când există o potrivire între comportament și stările urmărite (tracked) în mediul înconjurător. Dacă o creatură își poate satisface necesitățile și se poate dezvolta înfloritor într-un mediu, atunci pentru acea creatură sunt întrunite condițiile optime.

Aceste condiții optime (sau normale) pot varia în funcție de creatură cu care ne confruntăm. Această interpretare este una contrafactuală. Ceea ce contează este urmărirea care s-ar obține sub anumite condiții, nu cea care se obține actual. Întrebarea critică ce apare din această perspectivă este dacă sau nu, odată cu călătorul, călătoresc și contrafactualii relevanți pe Pământul Inversat. Tye susține plauzibilitatea unui răspuns afirmativ.

<sup>13</sup> *ibidem*

<sup>14</sup> BLOCK [1990]

<sup>15</sup> *ibidem*

<sup>16</sup> BLOCK [forthcoming]

<sup>17</sup> TYE [forthcoming]

Implantarea lentilelor convertoare alterează operațiunea transductorilor senzoriali. Deci procesul de transducție nu este în sine normal sau optim.

Intuitiv este deci adevărat despre starea senzorială a călătorului că, dacă n-ar fi avut loc nici o interferență, starea lui senzorială când se uită la cer pe Pământul Inversat ar fi fost corelată cauzal (în el) cu lucrurile albastre. În conformitate, prin covariația cauzală propusă, starea senzorială a călătorului continuă să reprezinte cerul ca albastru.

De fapt propunerea lui Tye înseamnă că calificarea "condiții optime" în acest context este o clauză *ceteris paribus*. Or, odată cu emigrarea pe Pământul Inversat *ceteribus* nu mai sunt *paribus*.

La același rezultat ajunge Tye în privința conținuturilor reprezentationale non-conceptuale adoptând teoria lui Fodor a dependenței nomice asimetrice<sup>18</sup>. După acest autor, o reprezentare mentală R reprezintă o propoziție P dacă

- (a) există o lege *ceteris paribus* că instanțe ale lui P (P-uri) cauzează instanțe ale lui R (R-uri), și
- (b) generalizarea nomică că P-urile cauzează R-uri nu depinde asimetric de nici o generalizare de tipul că P\*-urile cauzează R-uri (unde P\* nu este identic cu P).

Ideea aceasta, aplicată pentru conținuturi non-conceptuale pe Pământul Inversat, dă rezultatul că emigrantul continuă să-și reprezinte cerul ca albastru. Iată, însă, și cum se aplică această idee.

Pe Pământul Inversat experiențele mele cauzate de galben vor fi încă experiențe care, *ceteris paribus*, sunt cauzate de albastru pentru că, așa cum s-a observat mai sus, *ceteris* nu mai sunt *paribus*. Să numim aceste experiențe "experiențe ALBASTRU". Să considerăm relația nomică albastru – ALBASTRU și relația galben – ALBASTRU care se obține pe Pământul Inversat. Este plauzibil că ultima relație depinde asimetric de prima. Pentru a observa acest fapt, considerăm doi contrafactuali care trebuie să fie ambii adevărați pentru a se obține dependența asimetrică:

- (1) Dacă s-ar obține relația albastru – ALBASTRU, s-ar sparge relația galben – ALBASTRU.
- (2) Dacă s-ar obține relația galben – ALBASTRU, nu s-ar sparge relația albastru – ALBASTRU.

Intuitiv, după cum remarcă Tye, cea mai apropiată lume posibilă în care antecedentul lui (1) este adevărat este o lume în care nu este o lege că albastru cauzează ALBASTRU. În acea lume, dacă mă duc pe Pământul Inversat cu lentile convertoare, nu va mai fi adevărat că galbenul cauzează ALBASTRU. Deci (1) este adevărată.

Cea mai apropiată lume posibilă în care antecedentul lui (2) este adevărat este ce în care îmi îndepărtez lentilele convertoare. Astfel, prin

îndepărtarea lentilelor, relația albastru – ALBASTRU nu este afectată. Prin urmare (2) este adevărată.

Odată stabilită această dependență asimetrică, rezultă că conținuturile reprezentationale relevante rămân aceleași și după călătoria pe Pământul Inversat. Această soluție a lui Tye este un mod alternativ de a susține, în definitiv, că qualia sunt bazate în istoria evolutivă a organismelor și, deci, nu pot fi afectate de nici o călătorie. Dacă conținuturile reprezentationale rămân aceleași și la fel se întâmplă și cu qualia, atunci Pământul Inversat nu mai pune probleme pentru superveniență și, deci, pentru reprezentationalismul moderat pe care îl susțin.

### 3.3 Replica lui Block. Smârcomul

Block contestă valabilitatea soluției reprezentationaliste apelând la un alt experiment cauzator de probleme, anume la **scenariul Smârcomului** (Swampman). Această creatură apare ca urmare a unui trăsnet într-o mlaștină, care are drept consecință organizarea moleculelor de smârc într-un duplicat molecular al unui om obișnuit. Se poate discuta dacă Smârcomul are sau nu conținuturi reprezentationale. Este însă rezonabil să considerăm că dacă Smârcomul ar avea un copil, ar exista ceva care este felul cum este pentru acest copil când are experiențe. Trebuie însă să negăm aceasta, deoarece, conform reprezentationaliștilor, caracterele fenomenale decurg din evoluție. Cum Smârcomul nu are istorie evolutivă (el nu face parte din nici o specie), rezultă că Smârcomul nu are caractere fenomenale și că zombietatea este ereditară.

Isolat, acest experiment mental nu pune mari probleme reprezentationalismului. Răspunsul lui Tye este următorul: Smârcomul poate să aibă sau poate să n-aibă caractere fenomenale. Faptul că un Smârcom are sau n-are caractere fenomenale depinde de condițiile optime. În momentul când un mediu întrunește aceste condiții, atunci acel mediu înconjurător poate furniza Smârcomului conținuturi reprezentationale. Rezultă că Smârcomul poate avea caractere fenomenale.

Problema apare atunci când Block evidențiază faptul că strategiile lui Tye pentru cazurile Smârcomului și al Pământului Inversat intră în conflict. Pentru a releva acest conflict, Block sintetizează cele două experimente, propunând **scenariul nepotului de Smârcom care se duce pe Pământul Inversat**.

Ei bine, ce caractere fenomenale putem să-i atribuim acestei creaturi atunci când contemplă cerul (galben)? Să presupunem că atât mediul de pe Pământ cât și cel de pe Pământul Inversat întrunesc condițiile optime. Atunci ambele furnizează conținuturi reprezentationale.

„Deci pe ce bază ar putea Tye să atribuie nepotului Smârcomului caracterul fenomenal

<sup>18</sup> în FODOR [1990]

care merge cu reprezentarea cerului Pământului Inversat ca albastru (precum unui emigrant normal, după Tye) mai degrabă decât caracterul fenomenal care merge cu reprezentarea cerului ca galben (precum unui locuitor normal al Pământului Inversat)? O alegere aici ar fi arbitrară. Să presupunem că Tye alege caracterul fenomenal pământean. Dar ce face acest caracter privilegiat pentru nepotul Smârcomului? Faptul că bunicii lui s-au materializat pe Pământ și nu pe Pământul Inversat? Dar acesta e un temei slab. Să presupunem că nepotul Smârcomului este născut pe Pământul Inversat în timpul unei vizite a părinților săi și rămâne acolo. Sunt caracterele lui fenomenale determinate de locul lui de naștere sau de locul de naștere al bunicilor săi? Nu există bune temeiuri pentru nici una dintre aceste alegeri...<sup>19</sup>

Tye, spune Block, nu are altă alternativă decât să devină "ambiental" și să accepte că conținuturile reprezentationale se schimbă odată cu călătoria și, deci, se schimbă și caracterele fenomenale, ceea ce înseamnă că ele nu sunt bazate în istoria evolutivă a călătorului și, prin urmare, reprezentationalismul se dovedește din nou fals.

Reprezentationalismul poate trata separat atât Smârcomul, cât și Pământul Inversat. Problema apare când cele două scenarii se contopesc. Rezultă de aici că, pentru un reprezentationalist, cea mai la îndemână strategie de contraatac este să arate că acest scenariu mixt este imposibil. Cum acest scenariu poate fi socotit fie o variantă a Smârcomului, fie o variantă a Pământului Inversat, problema este rezolvată dacă se arată că unul dintre aceste scenarii este imposibil. Scenariul cel mai contestat dintre cele două este, bineînțeles, cel al Pământului Inversat. Criticile care i se aduc de obicei vizează imposibilitatea inversiunii intrapersonale (și chiar interpersonală) a spectrului. Probabil că cele mai interesante argumente de acest tip sunt cele ale lui Dennett și Chalmers, pe care le voi trata în continuare. De remarcat că genul acestor de contraargumente (care atacă Spectrul Inversat și nu Smârcomul) sunt, de altfel, și singurele care îmi folosesc, dat fiind că teoria pe care o susțin comportă psihofuncționalismul, iar Smârcomul nu pune decât probleme relativ minore funcționalismului. În schimb, un argument pentru imposibilitatea inversiunii spectrului exonerează și funcționalismul și reprezentationalismul.

#### 4. POSIBILITATEA SPECTRULUI INVERSAT

##### 4.1 Dennett: modelul multiple drafts și economia cognitivă

La fel cum Ryle a discreditat definitiv dualismul cartezian, denunțând mitul "stafiei din mașinărie", Dennett ridiculizează o altă idee carteziană, depistând și el un astfel de mit filosofic – mitul Teatrului Cartezian. Descartes susținea că sediul conștiinței (al sufletului) este glanda pineală. Așadar susținea că există un fel de centru de comandă al creierului, un centru funcțional, un Birou Oval sau un cartier general – oricum ar fi să-l numim – unde are loc tot ceea ce este actual în conștiință.

„Să numim ideea unui astfel de loc central în creier *materialism cartezian*, deoarece este concepția la care ajungi când te distanțezi de dualismul lui Descartes dar nu reușești să te dispensezi de imageria unui tetru central (dar material) unde totul "se înjgheabă" [...] Materialismul cartezian este concepția că există o linie de sosire crucială sau o frontieră undeva în creier, care demarchează un loc în care ordinea sosirilor este ordinea "prezentării" în experiență deoarece ești conștient de *ceea ce se întâmplă acolo*."<sup>20</sup>

Dennett susține așadar falsitatea unui model al conștiinței care presupune un loc de prezentare, loc în care ordinea prezentării este de tip FIFO. Modelul alternativ pe care îl propune acest filosof este cel numit "Multiple Drafts".

„După modelul Multiple Drafts, toate varietățile de percepție – într-adevăr, toate varietățile de gândire sau activitate mentală – sunt realizate de creier prin procese paralele, multidirecționale de interpretare și elaborare a impulsurilor senzoriale. Informația care intră în sistemul nervos este într-o continuă "revizuire editoriaală".<sup>21</sup>

Există experimente care confirmă această ipoteză, experimente prezentate de Dennett. În special este vorba despre experimente în legătură cu experiențele vizuale – de exemplu, interpretarea adâncimii în cazul stereogramelor din puncte aleatorii. Ceea ce demonstrează aceste experimente este că procesele editoriale durează câteva lungi fracțiuni de secundă, după care survin modificări, amendări, suprascrieri ale conținuturilor. Nu experiențiem direct ceea ce se întâmplă pe pielea sau retina noastră.

„Ceea ce experiențiem este de fapt un produs al multor procese de interpretare – procese editoriale, de fapt. Ele preiau reprezentări relativ brute și unilaterale și produc

<sup>19</sup> BLOCK [forthcoming]

<sup>20</sup> DENNETT [1990], p. 107

<sup>21</sup> *ibidem*, p. 111

reprezentări revizuite și mărite, și au loc în fluxuri de activitate care apar în diferite părți ale creierului [...] Noua trăsătură a modelului Multiple Draft: detectările de trăsături sau discriminările *trebuie făcute doar o dată*. Adică, odată ce "observația" particulară a anumitor trăsături a fost făcută de către o porțiune specializată, localizată a creierului, conținutul informațional astfel fixat trebuie să fie trimis altundeva pentru a fi rediscriminat de un anumit discriminator "șef". Cu alte cuvinte, discriminarea nu conduce la o reprezentare a trăsăturii deja discriminate pentru beneficiul unei audiențe în Teatrul Cartezian – pentru că nu există nici un Teatru Cartezian.<sup>22</sup>

Modelul Multiple Drafts postulează că în creier se petrece ceva de genul unei secvențe sau flux narativ, care este subiectul editărilor continue prin multe procese distribuite în diverse părți ale creierului. Acest flux este ca o narațiune din pricina multiplicității sale. În orice moment există drafturi multiple de fragmente narative, aflate în diferite etape ale editării în diferite locuri din creier. Modelul Multiple Drafts nu consideră că există ceva de genul unei narațiuni canonice singulare (draftul publicat, cum ar veni) – narațiune care să fie *actualul* flux al conștiinței unui subiect, fie că această are sau nu are acces la el.

În lumina acestui model, Dennett va considera că qualia inversate nu sunt posibile. El va aborda problema spectrului inversat în aspectul său epistemologic: cum am putea fi conștienți de o astfel de inversiune a spectrului, fie ea inter- sau intrasubiectivă? Strategia lui este ryleiană. Dennett va arăta că ideea spectrului inversat, în lipsa unui Teatru Cartezian, este auto-contradictorie. Cum un asemenea Teatru nu există, rezultă că "qualia inversate" este un non-sens.

Așadar, Dennett ia mai întâi în considerare un caz de inversie interpersonală a qualia vizuale.

Să presupunem că lucrurile roșii ne apar la fel mie și ție. Este această ipoteză irefutabilă și inconfirmabilă? Unii au zis că da. Alții s-au întrebat dacă nu cumva tehnologia ar putea confirma sau infirma inversia interpersonală a spectrului. Un astfel de exemplu citat de Dennett aparține filmului SF *Brainstorm*: Un aparat neuroștiințific poate fi în filmul respectiv conectat la capul tău și poate transmite toate experiențele tale vizuale în capul meu printr-un cablu. Cu ochii închiși eu raportez tot ce vezi tu, doar că mă minunez cum vezi tu cerul ca fiind galben, iarba verde etc. Ar putea confirma o astfel de mașină ipoteza spectrului inversat? Este evident că, așa cum argumentează Wittgenstein, citat de Dennett, experiențele vizuale care îmi vin prin cablu nu sunt experiențele *tale* vizuale, ci tot ale mele. Eu știu acum,

vorba lui Nagel, cum este pentru mine să fiu în pielea ta, dar nicidecum cum este pentru tine să fii subiectul unor experiențe "colorate".

Morala generală a acestui experiment mental este că nici cu cea mai performantă tehnologie n-ar fi posibilă o comparare intersubiectivă a qualia. Acest lucru sprijină ideea că qualia inversate intersubiectiv sunt un nonsens.

Dar ce înseamnă aceasta? se întreabă Dennett. Că qualia sunt rele dar inefficiente? Sau că nu există deloc qualia? Majoritatea filosofilor au afirmat realitatea caracterelor fenomenale ale experienței, chiar dacă nu s-ar putea detecta niciodată și în nici un fel diferența dintre qualia. Așa au stat lucrurile până ce a apărut ideea inversiunii intrapersonale a spectrului (ex. BLOCK [1990]). În acest caz experiențele de comparat sunt în același cap, așa că nu avem nevoie de inefficienta mașină Brainstorm.

Inversiunea trăsăturilor fenomenale ale experiențelor vizuale are loc prin interschimbarea firelor (neuronilor) care duc de la celulele conice fotosensibile la retină. Ai putea în cazul acesta să-ți dai seama că lucrurile s-au scimbat și arată altfel și ai dispune și de o explicație științifică: fasciculele de neuroni din cortexul vizual care se ocupă cu culoarea ar primi stimuli de la un set de receptori retinali sistematic interscimbați. Deci o diferență în qualia ar fi detectabilă dacă scimbarea ar fi intrapersonală și cu asta bătălia ar fi pe jumătate câștigată.

Dar ar fi câștigată doar pe jumătate, deoarece farsorul care și-a schimbat firele și-a inversat și dispozițiile comportamentale, așa încât comportamentul tău legat de culoare s-a inversat.

Pentru a câștiga bătălia, "qualofili" (sau "qualia freaks") au nevoie de un experiment mental care să demonstreze că felul în care arată lucrurile poate fi independent de orice dispoziție comportamentală. Deci vom complica scenariul introducând o modificare care restaurează dispozițiile de reacție, lăsând însă neatinsse qualia inversate. Cum se poate acest lucru? Să presupunem că în aceeași noapte în care echipa A de savanți nebuni îți inversează firele, o a doua echipă (B) îți reinversează firele la loc. Dar ceva tot este în neregulă:

„Asta restaurează vechile dispoziții de reacție (putem presupune), dar, din păcate, restaurează și vechile qualia. Celulele din cortex care "au grijă" de culoare, de exemplu, vor primi din nou semnalele originale, mulțumită reparării rapide de către echipa B a pagubelor. A doua schimbare s-a întâmplat *pe drumul spre* experiența conștientă. Deci va trebui să spunem povestea altfel, cu a doua schimbare petrecându-se mai târziu, *după* ce ce qualia inversate și-au luat locul în conștiință, dar *înainte* să se poată instala vreo reacție

<sup>22</sup> *ibidem*, pp. 112-113

inversată la ele. Dar este posibil acest lucru?  
Nu și dacă argumentele pentru Multiple Drafts  
sunt corecte.<sup>23</sup>

Lucrul nu este posibil deoarece modelul lui Dennett acceptă existența unei rețele cauzale și nu a unui lanț cauzal unic. Povestea qualofililor ar face sens dacă ar exista un Teatru Cartezian, "un loc special în creier unde are loc experiența conștientă"<sup>24</sup>. Dacă ar exista un astfel de loc, l-am putea izola între două comutatoare, păstrând qualia inversate în Teatrul Cartezian și menținând în același timp dispozițiile reactive normalizate.

„Deoarece nu există un astfel de Teatru Cartezian, oricum, experimentul gândit nu face sens. Nu există nici un fel coerent de a spune povestea necesară. Nu există nici o modalitate de a izola proprietățile prezentate în conștiință de reacțiile multiple ale creierului față de discriminările lui, deoarece nu există nici un astfel de proces adițional de prezentare.”<sup>25</sup>

Economia cognitivă nu permite creierului să facă de două ori aceeași operație: mai întâi discriminarea unei trăsături, apoi reprezentarea acesteia în Teatrul Cartezian de unde să fie rediscriminată de alte părți ale creierului care participă la cogniția superioară. Operația de discriminare este unică și este accesată de alte părți ale creierului simultan în locul unde se produce, nu într-un Teatru Cartezian.

Dennett susține că nu există nimic în plus față de dispozițiile de a reacționa, astfel încât dacă aceste dispoziții ar fi readuse la normal, qualia să poată rămâne inversate. Din această cauză, ideea de a tine qualia constante în timp ce modifici dispozițiile comportamentate este auto-contradictorie.

Motivul pentru care argumentarea lui Dennett nu mi se pare convingătoare este faptul că acest filosof nu arată imposibilitatea *in principium* a inversiunii calităților fenomenale, ci imposibilitatea *de facto*. Nu este metafizic relevant faptul că în realitate nu există qualia inversate, ci faptul că în principiu acest fenomen se poate produce. Deocamdată îi putem replica lui Dennett astfel: Eu îmi pot imagina o lume posibilă în care ține materialismul cartezian, în care există Teatrul Cartezian și în care, deci, se poate concepe o inversiune a qualia. Așadar, soluția lui Dennett nu rezolvă problema metafizică pe care o ridică spectrul inversat; de aceea, teoria funcționalist-reprezentationalistă prezentată în această lucrare nu se poate apăra de argumentele de tip Inverted Earth urmând strategia lui Dennett.

<sup>23</sup> *ibidem*, p. 392

<sup>24</sup> *ibidem*, p. 393

<sup>25</sup> *ibidem*

#### 4.2 Chalmers: imposibilitatea empirică a spectrului inversat

Aspectul posibilității spectrului inversat este luat în discuție și de către Chalmers. Filosoful menționat intenționează să arate imposibilitatea empirică sau naturală a spectrului inversat, printr-un argument de tip *reductio ad absurdum*. Astfel, Chalmers va arăta că dacă inversiunea intrasubiectivă a spectrului este posibilă, atunci se va petrece fenomenul "dancing qualia", care fenomen este extrem de neplauzibil să fie empiric posibil. Va urma că inversiunea spectrului nu este posibilă empiric.

Să presupunem, urmându-l pe Chalmers<sup>26</sup>, că qualia inversate sunt posibile. Conform principiului invarianței organizaționale<sup>27</sup> susținut de funcționalismul non-reductiv rezultă că putem avea doi izomorfi funcționali, aflați în aceeași stare funcțională dar cu qualia diferite. Să presupunem că acești izomorfi sunt eu (care am qualia roșu) și duplicatul meu de silicon, Robotul (care are qualia inversată, verde).

Chalmers urmează strategia construcției de cazuri intermediare între mine și Robot. Esențialul este că între cele două extreme se vor găsi două cazuri care să difere foarte puțin în alcătuirea internă, dar la care să se manifeste deosebiri semnificative în ce privește experiențele lor conștiente. Fie aceste două cazuri eu și Bill. Bill este un izomorf funcțional al meu care are un circuit de silicon acolo unde eu am unul neuronal. Din punct de vedere al experiențelor colorate, ipoteza prevede că acolo unde eu văd roșu, el vede albastru.

Chalmers face apoi trecerea de la intersubiectiv la intrasubiectiv, astfel: Să presupunem că mi se instalează în creier un circuit de silicon identic cu cel prin care Bill se deosebește fizic de mine. Acest circuit va fi funcțional izomorf cu unul neuronal deja prezent în creierul meu și va fi un circuit de rezervă. El nu va fi conectat direct la creierul meu, ci la două comutatoare prin care se poate face trecerea de la circuitul de silicon la cel neuronal. Când comutatoarele conectează circuitul de silicon la creier, circuitul neuronal izomorf devine irelevant. Când se inversează comutatoarele, circuitul de silicon devine irelevant, în timp ce se activează cel neuronal.

Problema care se pune este ce se întâmplă cu experiențele mele atunci când acționez comutatoarele. Înainte vedeam roșu. Și după instalarea circuitului de silicon voi vedea tot roșu, atât timp cât acesta nu este activ, deoarece, fiind irelevant cauzal, el nu afectează

<sup>26</sup> CHALMERS [1996]

<sup>27</sup> Acest principiu afirmă că două sisteme conștiente izomorfe din punct de vedere al organizării funcționale vor avea necesarmente aceleași experiențe conștiente. Experiențele respective sunt determinate de organizarea funcțională, dar nu sunt stări funcționale. (CHALMERS [1996], pp. 247-249)

organizarea mea funcțională. După activarea circuitului de rezervă, va trebui să am aceeași experiență ca Bill, care prin ipoteză vedea verde. Deci și eu voi vedea verde. Se va petrece astfel o schimbare a experienței în fața ochilor mei. Mărul roșu de pe masă va fi perceput brusc drept verde. Dacă acționăm comutatorul de mai multe ori la rând, ne putem imagina cum experiențele verzi și roșii dansează în fața ochilor mei.

Dar se petrece un lucru mai ciudat. Deoarece organizarea mea funcțională nu se modifică prin acționarea comutatoarelor, eu nu îmi voi da seama de acest "dans" al calităților fenomenale. Nu voi observa nimic neobișnuit, deoarece orice observație de acest fel ar implica o schimbare în organizarea mea funcțională – care, prin ipoteză, rămâne aceeași. Este evident că prin acțiunea comutatoarelor nu se pot dobândi noi opinii de genul "qualia mele tocmai s-au schimbat". Pur și simplu nu este loc pentru ele în organizarea mea cognitivă (decât dacă apelez la o minte carteziană, separată de corp).

„Pare complet neplauzibil să presupunem că experiențele mele se pot schimba într-un fel atât de semnificativ, eu fiind cât se poate de atent la ele, fără să fiu în stare să observ nici o schimbare. Ar sugera din nou o disociere radicală între conștiință și cogniție.”<sup>28</sup>

Așadar, "dancing qualia" sunt empiric implauzibile, chiar dacă ar fi din punct de vedere logic posibile. Dacă le-am presupune posibile natural, ar trebuie să facem față îngrijorării că așa ceva ni se poate întâmpla tot timpul, ceea ce ar însemna să acceptăm posibilitatea permanenței acestei ciudate rupturi între conștiință și cogniție. Or, aceste probleme nu provin decât din asumția susținută de intuiție că un duplicat funcțional ar putea vedea roșu acolo unde eu văd verde. Prin urmare trebuie să renunțăm la această asumție, pe plauzibilitatea căreia se bazează toate argumentele de tip qualia inversate.

Intuiția că o schimbare drastică a qualia nu este posibilă fără să o observ dacă mi se întâmplă este fructificată și de către Ned Block<sup>29</sup>. Acesta susține că este o trăsătură *necesară* a calităților fenomenale faptul că dacă o schimbare a lor este suficient de mare și se petrece suficient de repede eu o pot observa. Or, în cazul călătoriei pe Pământul Inversat, schimbarea este de o jumătate de disc al culorilor și se întâmplă quasi-instantaneu.

Cea mai importantă obiecție pe care o anticipează Chalmers ia în calcul posibilitatea dublei inversiuni a firelor, discutată de Dennett, unde Teatrul Cartezian este compus din zonele centrale ale creierului:

„Ne putem imagina un experiment înrudit în care reinversăm firele de la input-urile roșii și albastre la zonele centrale ale creierului astfel încât input-urile albastre să joace rolul pe care l-au jucat odată input-urile roșii și în care schimbăm de asemenea sistematic conexiunile *descendente* dinspre zonele centrale în compensație. Când un input albastru face zona centrală să intre într-o stare asociată anterior cu roșul, conexiunile de la zona centrală la restul creierului sunt reschimbate astfel încât restul creierului funcționează la fel cum ar fi funcționat dacă n-ar fi existat nici o reinversare. În felul acesta, experiența mea se va schimba cu siguranță de la roșu la albastru, dar dispozițiile mele comportamentale vor rămâne constante. În acest caz, o comutare repetată ar conduce la dancing qualia. Deci nu sunt până la urmă dancing qualia rezonabile?”<sup>30</sup>

Răspunsul lui Chalmers este următorul: Mai întâi, nu putem schimba decât input-urile între ele. Output-urile pot fi atât de variate, încât nu există nici un mod de simplă schimbare a firelor care să îndeplinească sarcina. Interschimbarea output-urilor este atât de complicată, încât este irealizabilă. Input-ul va afecta creierul într-o mulțime de locuri diferite, nu există o singură secvență cauzală de la input la output.

În al doilea rând, chiar dacă s-ar putea încrucișa output-urile, ar rămâne nemodificate doar dispozițiile de comportament, nu și organizarea funcțională. În acest caz *observa* schimbarea și *aș încerca să mă comport corespunzător*, dar n-aș reuși<sup>31</sup>. Chalmers consideră că feedback-ul acțiunilor mele ar zgudui într-un asemenea hal zona centrală a creierului, încât probabil că ar conduce la colaps, dat fiind faptul că nu există nici o posibilitate de a compensa sistematic aceste feedback-uri. În orice caz, faptul că ar exista o diferență semnificativă în organizarea funcțională arată că cele două cazuri nu sunt analoge.

Nemulțumitoare la argumentația lui Chalmers sunt următoarele două lucruri: (1) este prinsă în Teatrul Cartezian; (2) fenomenul dansului calităților fenomenale nu poate fi folosit la demonstrarea

<sup>30</sup> CHALMERS [1996], p. 273

<sup>31</sup> Un asemenea caz este imaginat de către Searle. Implantarea progresivă a unor chip-uri de silicon în creier produce în scenariul lui Searle incapacitatea de a te comporta în conformitate cu input-urile primite, fără a afecta în nici un fel experiențele tale. Astfel se poate ajunge la paralizia ta completă, în timp ce viața ta conștientă nu se modifică deloc. Eventual, în scenariul sealeian, medicii te vor considera în moarte cerebrală. Problema cu experimentul lui Searle este, însă, că încalcă principiul invarianței organizaționale.

<sup>28</sup> CHALMERS [1996], p. 269

<sup>29</sup> BLOCK [forthcoming]

imposibilității conceptuale a inversiunii spectrului, așa cum recunoaște însuși Chalmers.

Teoria propusă în capitolul 1 are nevoie pentru a menține reprezentationalismul de un argument pentru imposibilitatea conceptuală a scenariilor de tip Pământul Inversat. Voi încerca să aduc un astfel de argument în secțiunea următoare.

#### 4.3 Un argument pentru imposibilitatea conceptuală a spectrului inversat intrasubiectiv

Argumentul pe care îl voi oferi în continuare este tot de forma *reductio*. Voi arăta că, dacă acceptăm scenariul Pământului Inversat, va trebui să acceptăm, într-un fel sau altul, concluzii care contrazic ipoteza. Așadar Pământul Inversat este fie contradictoriu, fie incoerent.

Să presupunem că Pământul Inversat este posibil. Atunci rezultă că trebuie să acceptăm ca ipoteză următoarele afirmații:

- (1) Pe Pământ călătorul percepe culorile normal.
- (2) Pe Pământul Inversat, datorită lentilelor care anulează inversiunea culorilor, el nu va percepe nici o diferență în experiențele lui colorate.

Acum să presupunem că emigrantul era pe Pământ un pictor. El lucra în acuarelă și își pregătea singur culorile, amestecând cele șase culori principale ale spectrului, pe care și le procura su forma unor pastile de culoare dintr-un magazin în sistem auto-servire. În seara de dinainte de a fi răpit, pictorul nostru trebuia să umple o suprafață a unui tablou cu violet, culoare pe care o prepara amestecând în proporții egale roșu cu albastru. Să presupunem în continuare că, din cauza oboselii, pictorul se duce la culcare după ce a pictat numai o jumătate din suprafață cu violet.

Noaptea, o echipă de savanți nebuni îi implantează în ochi lentile convertoare, execută o copie a tabloului său în culori complementare și îl transportă pe Pământul Inversat unde îl substituie *doppelgänger*-ului său.

A doua zi dimineată, pictorul nostru se va reîntoarce la tablou. Își va prepara ca de obicei violetul, amestecând ceea ce el percepe ca roșu cu ceea ce el percepe ca albastru. Datorită faptului că se află pe Pământul Inversat, însă, ceea ce va amesteca el va fi, în realitate, verde cu galben. Violetul de pe Pământ fusese replicat pe tabloul - copie cu culoarea sa complementară, portocaliul. Ceea ce obține pictorul și întinde cu pensula în prelungirea portocaliului de pe pânză este, însă, galben-verzui.

În acest punct al experimentului nostru se pot întâmpla, la prima vedere, două lucruri: sau pictorul poate suferi un șoc văzând cum combinația de roși și albastru dă ceva diferit de violet, sau pictorul nu va observa absolut nici o diferență între cele două culori de pe pânză. Însă conform propoziției (2) din ipoteză, numai ultima situație este posibilă. În mod similar,

amestecând roșu cu galben pentru a obține portocaliu, pictorul va amesteca de fapt verde cu albastru și va obține albastru-verzui, nu violet (complementara portocaliului). Dar de asemenea, prin ipoteză, el nu va remarca nici o deosebire între violet și albastru-verzui.

Acesta înseamnă că din mulțimea experiențelor colorate au dispărut o serie de experiențe, fără ca pictorul să poată observa. Ceea ce înseamnă că simpla călătorie pe Pământul Inversat l-a transformat într-un zombie parțial, întrucât această îngustare a câmpului lui de experiențe colorate este semnificativă. Mergând mai departe cu imaginația putem întrevedea cum, datorită faptului că pictorul este brusc incapabil să distingă galbenul de galben-verzui el poate combina galbenul verzui cu albastrul, crezând că amestecă violet cu roșu pentru a obține indigo. În realitate va obține un fel de verde, pe care nu-l va distinge de portocaliu-roșcat, complementara indigoului. Încet-încet, dacă operăm toate combinațiile de culori posibile, așa cum le percepe pictorul, ajungem la următoarea concluzie:

- (i) pictorul percepe prin ipoteză toate culorile spectrului perceptibile de un ochi normal distinct, chiar dacă inversat;
- (ii) ipoteza implică faptul că pictorul nu este în stare să perceapă decât culorile principale ale spectrului;
- (iii) mai mult, este plauzibil că pictorul nu poate percepe nici măcar toate culorile principale ale spectrului.

După cum se observă, (ii) și (iii) contrazic (i). Ipoteza Spectrului Inversat intrasubiectiv implică afirmații care o contrazic, ceea ce arată că Pământul Inversat este auto-contradictoriu. Or, a fi conceptibil (sau logic posibil) înseamnă cel puțin a nu fi auto-contradictoriu.

Rezultă din considerațiile de mai sus *quod erat demonstrandum*, anume că Pământul Inversat este o imposibilitate logică, un non-sens. Prin urmare el nu poate pune probleme nici reprezentationalismului, nici funcționalismului.

## BIBLIOGRAFIE:

1. BLOCK [1980] – Block, N., "Troubles with Functionalism", in *Readings in the Philosophy of Psychology*, Cambridge, Mass.: Harvard University Press
2. BLOCK [1990] – Block, N., "Inverted Earth", *Philosophical Perspectives*, Northridge: Northridge Publishing Co.
3. BLOCK [forthcoming] – Block, N., "Is Experiencing Just Representing?", <http://www.nyu/gsas/dept/philo/faculty/block>
4. CHALMERS [1996] – Chalmers, D., *The Conscious Mind*, Oxford: Oxford University Press
5. DENNETT [1990] – Dennett, D., "Quining Qualia", in *Mind and Cognition*, Oxford: Blackwells
6. DENNETT [1991] – Dennett, D., *Consciousness Explained*, Boston: Little, Brown and Co.
7. FODOR [1987] – Fodor, J., *Psychosemantics: The Problem of Meaning in the Philosophy of Mind*, Cambridge, Mass.: MIT Press
8. FODOR [1990] – Fodor, J., *A Theory of Content, and Other Essays*, Cambridge, Mass.: MIT Press
9. NAGEL [1974] – Nagel, T., "What is it like to be a Bat?", *Philosophical Review*, 83
10. SEARLE [1990] – Searle, J., *The Rediscovery of the Mind*, Cambridge, Mass.: MIT Press
11. SHOEMAKER [1982] – Shoemaker, S., "The Inverted Spectrum", *Journal of Philosophy*, 79
12. TYE [1997] – Tye, M., "Qualia", *Stanford Encyclopedia of Philosophy*, <http://www.plato.stanford.edu>
13. TYE [1998] – Tye, M., "Inverted Earth, Swampman and Representationism", final version on <http://www.nyu/gsas/dept/philo/faculty/block>