

Połączone Biblioteki WFIS UW, IFiS PAN i PTF

U.84172



39084172000000

Walter Bremer

JAK TO JEST BYĆ ŚWIADOMYM



PRACE Z FILOZOFII I KOGNITYWISTYKI

JAK TO JEST BYĆ ŚWIADOMYM

ANALITYCZNE TEORIE UMYŚLU
A PROBLEM ŚWIADOMOŚCI

JÓZEF BREMER

JAK TO JEST BYĆ ŚWIADOMYM
ANALITYCZNE TEORIE UMYŚLU
A PROBLEM NEURONALNYCH PODSTAW
ŚWIADOMOŚCI

Wydawnictwo IFiS PAN
Warszawa 2005

Projekt okładki
Marta i Zdzisław Kwiatkowsy

Redaktor
Elżbieta Morawska

© 2005 by Wydawnictwo IFiS PAN & Józef Bremer

ISBN 83-7388-069-0

Wydawnictwo Instytutu Filozofii i Socjologii PAN
00-330 Warszawa, ul. Nowy Świat 72, tel. 65 72 897
Obj. 25 ark. wyd., 23,25 ark. druk.
Druk: Drukarnia Wydawnictwa WAM
ul. Kopernika 26, 31-501 Kraków

SPIS TREŚCI

WSTĘP	11
ROZDZIAŁ PIERWSZY	
Współczesna analityczna filozofia świadomości	19
1. Wprowadzenie: cel pracy i przyjęta metoda	19
1.1. „Trudny” problem świadomości i jego neurologiczne znaczenie	21
1.2. Świadomość zjawiskowa – przykłady występowania	23
1.3. Filozofia świadomości – podstawowa terminologia	27
1.3.1. Świadomość z psychologicznego punktu widzenia	30
1.3.2. Świadomość od strony metafizycznej	32
1.3.3. Holizm ontologiczny i metodologiczny	36
1.3.4. Holistyczne rozumienie świadomości	38
1.3.5. Debata: eksternalizm–internalizm	44
1.3.6. Eksperymenty myślowe	49
1.4. Subiektywność świadomości w historii filozofii	53
1.5. Podsumowanie	56
ROZDZIAŁ DRUGI	
Filozoficzno-psychologiczny opis subiektywności świadomości	59
2.1. Wprowadzenie: cechy świadomości	59
2.1.1. Subiektywność ujęta ontologicznie i epistemologicznie	60
2.1.2. Perspektywa pierwszo- i trzecioosobowa	63
2.1.3. Perspektywa czasowa i przestrzenna	67
2.1.4. Odczuwanie własnego ciała	68
2.2. Subiektywność wyrazem jedności świadomości	70

2.3.	Model świadomości zjawiskowej	73
2.4.	Inne przejawy świadomości	77
2.4.1.	Qualia	77
2.4.2.	Samoświadomość	81
2.5.	Podsumowanie	84

ROZDZIAŁ TRZECI

Miejsce świadomości w wybranych teoriach umysłu	87	
3.	Wprowadzenie: subiektywna świadomość a mózg	87
3.1.	Dualizm <i>versus</i> materialistyczny monizm	91
3.2.	Behawioryzm logiczny i nomologiczny	94
3.3.	Teoria identyczności typów	97
3.3.1.	Materializm ontologiczny według Paula Churchlanda	99
3.3.2.	Johna Smarta teza o ontologicznej neutralności	102
3.4.	Teoria identyczności zdarzeń	108
3.4.1.	Saula Kripkego krytyka teorii identyczności	111
3.4.2.	Teorie superweniencji	117
3.5.	Funkcjonalizm	119
3.5.1.	Funkcjonalistyczna analiza „trudnego” problemu	123
3.5.2.	Relacyjne wyjaśnianie stanów jakościowych	128
3.5.3.	Teoria zmysłu wewnętrznego	130
3.6.	Teorie emergencji	134
3.6.1.	Podstawowe rodzaje emergencji	138
3.6.2.	Wrażenia jako cechy emergentne	141
3.7.	Podsumowanie	146

ROZDZIAŁ CZWARTY

Daniel Dennett – świadomość aktywnością mózgu	151	
4.	Wprowadzenie: jaka teoria świadomości?	151
4.1.	Heterofenomenologia – perspektywa trzecioosobowa	153
4.1.1.	Postulowana jaźń	155
4.1.2.	Ocena heterofenomenologii	159
4.2.	Kartezjański Teatr	161
4.3.	Wiele Szkiców jako model subiektywnych przeżyć	164
4.3.1.	Orwellowski i stalinowski model świadomego umysłu	166
4.3.2.	Model Wielu Szkiców	168
4.4.	Jedność świadomości	170
4.4.1.	Jedność świadomości w ramach Funkcjonalnego Teatru	171

4.4.2.	Wielość szkiców a świadomość zjawiskowa	174
4.5.	Kulturowa ewolucja świadomości – darwinizm społeczny	177
4.6.	Komputerowo-funkcjonalistyczna struktura świadomego umysłu	179
4.7.	Stany zdawania sobie sprawy	183
4.8.	Neuronalne podłoże subiektywności	186
4.9.	Podsumowanie	189

ROZDZIAŁ PIĄTY

Thomas Nagel – subiektywne własności świadomości	193
--	-----

5.	Wprowadzenie: jakiej filozofii świadomości poszukujemy? ...	193
5.1.	Fizykalna koncepcja obiektywności	195
5.1.1.	Obiektywna fenomenologia	198
5.1.2.	Obiektywność zjawiska psychicznego	200
5.2.	Neuronalne podłoże jedności świadomości	203
5.3.	Argumenty przeciwko fizykalizmowi	208
5.3.1.	Nieredukowalność przeżyć	209
5.3.2.	Argument z poznania	210
5.3.3.	Fizykalistyczno-redukcyjnistyczne kontrargumenty	212
5.4.	Panpsychiczna teoria dwu aspektów	216
5.5.	Poszerzona teoria stanów mentalnych	222
5.6.	Podsumowanie	227

ROZDZIAŁ SZÓSTY

W poszukiwaniu nowej teorii świadomości	231
---	-----

6.	Wprowadzenie	231
6.1.	Neurokognitywistyka a problem świadomości	233
6.1.1.	Pytanie o kryterium świadomości	234
6.1.2.	Świadome przetwarzanie informacji	238
6.1.3.	Nowe techniki neuroobrazowania	241
6.1.4.	Koneksjonizm	245
6.1.5.	Neuroplastyczność	249
6.2.	Ujęcie mózgu w materialistycznych teoriach świadomości	253
6.3.	Wybrane neurokognitywne teorie świadomości	255
6.3.1.	Neuropsychologia	255
6.3.2.	Neurofenomenologia	259
6.3.3.	Neurofilozofia	263
6.4.	Podstawowe założenia poszukiwanej teorii świadomości	267
6.5.	Podsumowanie	271

ROZDZIAŁ SIÓDMY

Neuronalne korealty świadomości	275
7. Szkic programu badawczego	275
7.1. Definicje neuronalnych korelatów świadomości	276
7.1.1. Uszczegółowienie podanych definicji	279
7.1.2. Dwa rodzaje kwestii metodologicznych	283
7.2. Neuronalne korelaty marzeń sennych	289
7.3. Wpływ wad mózgu na stany świadomości	290
7.3.1. Przykłady niedorozwoju mózgu	294
7.3.2. Określanie korelatów świadomości w zdeformowanym mózgu	294
7.4. Stopień zależności między systemami neuronalnymi a świadomością	298
7.4.1. Warunek konieczny i wystarczający wystąpienia stanu świadomości	298
7.4.2. Minimalny system neuronalny	301
7.5. Przykłady rodzajów korelatów świadomości	306
7.5.1. Neuronalne korelaty świadomych przeżyć	306
7.5.2. Neuronalne korelaty jedności świadomości	309
7.6. Neurokognitywna prototeoria świadomości	312
7.6.1. Redukcja bez identyfikacji: teoria emergencji jako surogat teorii naukowej	316
7.6.2. Wyjaśniająca wartość proponowanej prototeorii	322
7.7. Podsumowanie	326
Zakończenie	329
Literatura	334
Indeks nazwisk	357
Indeks rzeczowy	362
Summary	368

Jak to jest być piskorzem?

Rozmowa pomiędzy chińskimi filozofami Czuang-tsy (taoista) i Huei-tsy (logiczny sceptyk), którzy spacerując po moście na rzece Hao obserwują ławice ryb.

Czuang-tsy powiedział: „Jak ochoczo te piskorze pływają [w wodzie]”. „Nie jesteś, panie, rybą; skąd [możesz wiedzieć na czym polega radość ryb]?” „Nie jesteś mną, więc skąd możesz wiedzieć, co ja wiem o radości ryb?” „,[Wprawdzie] nie jestem tobą i oczywiście nie wiem, co ty wiesz, ale i ty tak samo nie jesteś rybą, więc twoja niewiedza o radości ryb jest całkowita”. „Ale, jeśli łaska, wróćmy do początku” — rzekł Czuang-tsy — „Powiedziałeś mi: «Skąd wiesz, jaka jest radość ryb?» [To znaczy], że sam wiesz, że ja mogę coś o tym wiedzieć i dlategoś mnie spytał. A ja to wiem [dzięki podobieństwu naszej radości] podczas przechadzania się po moście” (Czuang-tsy 153: 191).

Jak to jest być nietoperzem?

„Jak to jest być nietoperzem?
Nie mów: «Nie wiem,» Ja chcę wiedzieć!
Jak to jest spać głową na dół?
Nietoperskich mieć sąsiadów?
Słyszeć domy, drzewa, skały?
Lub świat widzieć czarno-biały?
Powiedz: Jak to jest nie-wiedzieć,
jak to jest być nietoperzem?” (JKW)

WSTĘP

„[...] żaden problem filozoficzny nie może zostać rozwiązany, zanim nie zostaną rozwiązane wszystkie problemy filozoficzne; znaczy to, że do tego czasu każda nowa trudność będzie stawiała pod znakiem zapytania wszystkie wcześniejsze wyniki” (Wittgenstein 1984: 74).

Świadomość jest wprawdzie ściśle związana z naszym rozumieniem świata i samych siebie, jednak wszelkie próby jej naukowego wyjaśnienia wydają się być skazane na niepowodzenie. Świadomość wymyka się bowiem wysiłkom podejmowanym w celu jej zobiektywizowania, a tym samym jej teoretycznego ujęcia. Odpowiedzialna za ten stan rzeczy jest subiektywność stanów świadomości.

Subiektywność łączy się z naszym odbieraniem siebie i otoczenia poprzez pryzmat własnej osobowej indywidualności i przez pryzmat naszej przynależności do danego gatunku. To właśnie subiektywna natura świadomości powoduje, że jest ona zasadniczo odmienna od innych przedmiotów badań naukowych.

W teoretycznym dyskursie o subiektywności świadomości odróżniamy dwie kategorie obiektów, stanów i cech: mentalne i fizyczne. Przykładami stanów mentalnych są: akty spostrzegania, ból, pragnienia, myślenie, wrażenia. Przykładami stanów fizycznych są procesy badane przez fizykę, chemię czy neurofizjologię. Obiekty i stany fizyczne identyfikujemy przez odwołanie się do kryteriów czasoprzestrzennych. Trudniej podać obiektywne kryteria identyfikacji stanów mentalnych i ich własności.

W potocznym użyciu obydwie kategorie wydają się być jasno określone. Trudności pojawiają się wtedy, gdy zapytamy o relacje zachodzące pomiędzy tymi dwiema kategoriami stanów. Zagadnienie to konstytuuje jeden z aspektów problemu nazywanego „problemem umysł–ciało” (*mind-body problem*, das *Leib-Seele-Problem*) lub „problemem psychofizycznym”¹. Problem ten, który pojawił się w filozofii wraz z dualizmem Kartezjańskim, będziemy rozpatrywali pod kątem pytania o filozoficzno-empiryczne związki pomiędzy świadomym umysłem podmiotu a stanami neuronalnymi jego mózgu.

Mówiąc w tej pracy o filozofii mamy na myśli jej nurt określane mianem filozofii analitycznej. Z jednej strony świadomość stanowiła przedmiot badań wielu filozofów analitycznych (por. prace C.D. Broada, B. Russella, E. Moore’a, L. Wittgensteina), z drugiej strony część z nich – na przełomie XIX i XX w. – przestała zajmować się świadomością. Podobne dążenia pojawiły się w tym czasie w wielu gałęziach psychologii, inspirowanych analitycznymi metodami badań². Poza analitycznym nurtem badań dominował w tym czasie introspekcjonizm w psychologii, a w ramach filozofii powstawała fenomenologia.

Zagadnienia dotyczące świadomości są dzisiaj szeroko badane nie tylko w filozofii analitycznej, ale przede wszystkim w tak zwanej kognitywistyce (w sensie *cognitive science*)³. Ta ostatnia zajmuje się naukowym wyjaśnianiem procesów poznawczych, traktowanych jako procesy przetwarzania informacji. Jest ona nauką interdyscyplinarną, obejmującą niektóre dziedziny fizyki, biologii, psychologii, informatyki, logiki i filozofii⁴.

Z historii filozofii znane są dwa podstawowe typy rozwiązania problemu umysł–ciało: rozwiązania monistyczne (materialistyczne lub spiritualistycz-

¹ Por. Ajdukiewicz 1960: 318–320. W dalszych analizach, jako polskiego tłumaczenia angielskiego „mind-body” używamy terminu „umysł–ciało”, rezygnując ze zwrotu „problem psychofizyczny”.

² Według neurologa O. Sacksa, w nauce o człowieku dokonał się wówczas podział na „bezduszną neurologię i bezcielesną psychologię” (Sacks 1996: 123). Zdaniem Sacksa, Torurette, Freud i Babiński byli ostatnimi, którzy łączyli obrazy ciała i duszy, którzy łączyli „[...] «To» i «Ja», neurologię i psychiatrę” (s. 123). Od strony historycznej problem subiektywności omawia Hagenbüchle (por. 1998: 82–88). Wallace łączy „naukowo-materialistyczne” opisy subiektywności z opisami spotykanyymi w buddyzmie zen (por. 2000: 21, 26, 128).

³ W roku 1950 (szczytowy okres rozwoju behawioryzmu) terminu „świadomość” użyto w pięciu biomedycznych artykułach, zaś w roku 2000 w około 1400 (por. Baars 2003: 102).

⁴ Zob. odpowiedź na pytanie: czy nauka kognitywna jest paradygmatem dla badań procesów poznawczych? (Zegleń 2003: 117–132). Obszerne przedstawienie zagadnień kognitywistycznych zawiera *The MIT Encyclopedia of the Cognitive Sciences* (por. Wilson i Keil 2001: passim).

ne) i dualistyczne. W filozofii analitycznej dominowały i dominują rozwiązania monistyczno-materialistyczne. W ostatnich dziesięcioleciach pojawiła się w jej ramach grupa rozwiązań pośrednich. Przedstawiciele tej grupy uważają, że uwzględniające subiektywność świadomości rozwiązanie problemu umysł-ciało wymaga nowego teoretycznego podejścia, różniącego się od podejścia monistycznego i dualistycznego.

Należący do tej grupy D. Chalmers odróżnia tak zwane łatwe (*easy*) problemy świadomości i trudny (*hard*) problem świadomości⁵. Mówiąc ogólnie, problemy „łatwe” są lub mogą być rozwiązywane metodami kognitywistycznymi i dotyczą one stanów intencjonalnych, do których zaliczamy myśli oraz akty woli. Odnoszące się do tych stanów zdania mogą być uznane za prawdziwe lub fałszywe⁶. Stany intencjonalne posiadają tak zwaną treść semantyczną i są wyrażalne językowo (na przykład znak „444” oddziałuje na nas, ponieważ posiada znaczenie). Semantyczna treść znaku określa rodzaj jego oddziaływania na podmiot. „Syntaktyczne”, tzn. fizyko-chemiczne cechy znaku (ilość zużytej farby, rodzaj papieru, na którym jest wydrukowany) są wówczas najczęściej pomijane.

Problem „trudny” dotyczy świadomych stanów „przeżyciowych” lub „doświadczeń” (*experiences*)⁷. Należą do nich wrażenia (*sensations*), spostrzegania (*perceptions*), bóle (*pains*) i inne podobne. Stany przeżyciowe, zwane także zjawiskowymi, posiadają „treść jakościową” i są przez nas subiektywnie przeżywane. Kluczowe pytania dotyczące subiektywności świadomości brzmią: [A] Czy, a jeżeli tak, to w jaki sposób fizyczne procesy w mózgu mogą wywołać subiektywne stany świadomości? Czy tego rodzaju stany mogą się pojawić w jakimś innym, np. nieorganicznym, systemie fizycznym? Są to pytania o wyjaśnianie świadomości.

Nawet całkowita znajomość wszystkich fizycznych procesów zachodzących w mózgu świadomego podmiotu oraz znajomość fizjologicznej różnicy pomiędzy „świadomym” i „nieświadomym” (na przykład będącym pod

⁵ Por. Chalmers 1995: 1. Do „łatwych” problemów należą przykładowo: kwestie związane ze zdolnością do odróżniania przez podmiot bodźców, do ich kategoryzowania i reagowania na nie, ze zdolnością do zdawania sprawozdań o stanach wewnętrznych, do skupiania uwagi, do odróżniania pomiędzy jawą a snem.

⁶ „Thoughts, beliefs, intentions, and so forth are typically specified in terms of subordinate sentential clauses, that-clauses, which may be judged as true or false. Pains, feels, tickles, and so forth have no special semantical relation to sentences or to truth or falsity” (Burge 1979: 74).

⁷ Dwuznaczny angielski termin „experience” tłumaczymy jako „przeżycie”. Dziarnowska tłumaczy „experience” jako „doznanie”, termin „qualia” zaś – jako „jakości doznaniowe” (por. Dziarnowska 2001: 174).

wpływem narkozy) mózgiem, zdają się nie dostarczać wyjaśnienia tak rozumianej świadomości zjawiskowej. Krótka refleksja nad terminami „łatwy” i „trudny” uświadamia nam, że chodzi tutaj raczej o etykiety i że obydwie problemy należą do trudnych.

Przyjmuje się, że stany intencjonalne, ponieważ odnoszą się do dyspozycji poznawczych podmiotu, można łatwiej opisać i wyjaśnić w terminach nauk kognitywnych (stąd „łatwe” problemy) aniżeli stany zjawiskowe. Prowadzone w niniejszym opracowaniu analizy świadomości skupiają się na jej „trudnym” problemie, który sprowadzamy do pytania: W jaki sposób materialno-fizyczny mózg jest skorelowany ze świadomością zjawiskową?⁸ Analizując „trudny” problem bada się zazwyczaj:

(a) rozróżnienie na podejście pierwszo- i trzecioosobowe: subiektywna (pierwszoosobowa) perspektywa przeżywania świata jest niedostępna dla innych osób, zaś perspektywę trzecioosobową może przyjąć każda osoba;

(b) jakościowe elementy naszych przeżyć zwanych „qualiami” albo „jakościami doznaniowymi”: przeżycia te zdają się nie należeć do fizycznych struktur lub funkcji ciała.

Problem „trudny”, jak jest on opisywany w tej pracy, podpada wprawdzie pod aspekt (a) oraz (b), dotyczy jednak głównie ustalenia zależności pomiędzy subiektywnym przeżywaniem przez każdego z nas siebie samego wraz ze swoimi stanami wewnętrznymi, a obiektywnie ujmowalnymi procesami zachodzącymi w naszych mózgach. Takie ujęcie sprawy stawia mniejsze wymagania poszukiwanej teorii świadomości aniżeli wymagania określone w powyższych pytaniach [A].

Poszczególne teoretyczne ujęcia „łatwego” problemu umysł–ciało zostały przez nas omówione w książce *Problem umysł–ciało* (por. Bremer 2001: passim). W niniejszym opracowaniu, w którym nie zakłada się znajomości wspomnianej książki, przedstawiamy wyniki badań chociażby takich współczesnych filozofów analitycznych, jak: D. Chalmers, T. Nagel, J. Searle, C. McGinn. Prowadzone przez nich badania oraz programy tych badań cechują się wysokim standardem metodologicznym, do którego należy między innymi formalnologiczna analiza pojęć.

Umysł świadomego podmiotu cechuje się jednością świadomości. Jej stany intencjonalne i zjawiskowe są ze sobą ściśle związane. Pojęciowy podział na świadome stany intencjonalne i stany zjawiskowe wskazuje jedynie na aspekty świadomości i ułatwia jej teoretyczne ujęcie. Parafrazując słowa Wittgensteina, można powiedzieć: Gdy chodzi o rozwiązywanie filozoficz-

⁸ Zdaniem W. Ducha, „trudnego” problemu da się uniknąć, gdy mózg i umysł potraktujemy jako dwa sprzężone aspekty tej samej rzeczywistości (por. Duch 1996: 3–6).

nych problemów związanych ze świadomością, to nie rozwiążemy zadowalająco żadnego z nich, dopóki nie zostaną rozwiązane oba typy problemów („łatwe” i „trudny”).

Celem ogólnego przybliżenia treści książki, którą zaliczamy do prac z zakresu kognitywistyki i epistemologii, podajemy jej skondensowany plan.

W rozdziale pierwszym nakreśliłyśmy główne linie metodologicznego postępowania i historię pojawienia się współczesnego problemu subiektywności świadomości. Omawiamy w nim również ogólne filozoficzno-naukowe ramy, w których problem ten należy badać. Istotne elementy tychże ram tworzą filozoficzno-psychologiczne i neurologiczne spojrzenie na świadomy umysł.

W rozdziale drugim przedstawiamy filozoficzno-psychologiczny opis świadomości. Założono w nim pewien model świadomości zjawiskowej, mający pomóc w zbudowaniu nowej wielowymiarowej teorii świadomości. Model ten różni się od modeli, jakimi operujemy w teoriach z dziedziny fizyki czy chemii. Jest on jedynie modelem w sensie metaforycznym (intuicyjnym). Odwołujemy się do niego w pozostałych rozdziałach, a zwłaszcza w szóstym i siódmym, w których szkicujemy metodologiczne podstawy wspomnianej, nowej teorii świadomości. Chcemy tym samym pokazać, w jaki sposób założenia takiej teorii należy szukać w określeniu filozoficznych i neurologicznych warunków, których spełnienie jest wymagane dla mówienia o tak zwanych neuronalnych korelatach świadomości (*neural correlates of consciousness*, NCCs)⁹. Opisując korelaty, odwołujemy się do wyników badań neurologicznych i do filozoficznych analiz pojęciowych i definicyjnych.

W rozdziale trzecim zestawiono trylemat tez, odwzorowujący rodzaje zależności pomiędzy subiektywną świadomością podmiotu a jego mózgiem. W zależności od przyjęcia, bądź odrzucenia, którejs z tych tez, otrzymamy albo dualistyczne, albo monistyczne teorie świadomości. Powstała tą drogą systematyzacja teorii świadomości stanowi dodatkowe wprowadzenie do pozostałych rozdziałów pracy.

W tym samym rozdziale poddajemy krytycznej analizie klasyczne już dzisiaj, materialistyczne próby rozwiązania problemu świadomości (od logicznego behawioryzmu po funkcjonalizm). W ramach współczesnych rozwiązań materialistycznych odróżniamy dwie skrajne pozycje i omawiamy je szczegółowo w rozdziałach czwartym i piątym.

W rozdziałach szóstym i siódmym, jak już wspomniano, przedstawiono metodologiczne kwestie związane z poszukiwaniem neuronalnych korelatów

⁹ W tym sensie zgadzamy się z przekonaniem R. Piłata, że integralną część stanów i procesów składających się na bycie świadomym, stanowią struktury niemające nic wspólnego z potocznie rozumianą świadomością (por. Piłat 1993: 9).

świadomości. Czy i na ile znalezienie takich korelatów pozwoli na zbudowanie nowej filozoficzno-neurologicznej teorii świadomości? Dyskusja nad kwestiami ontologicznymi (dotyczącymi bytowego statusu cech i obiektów mentalnych) ma w filozofii świadomości długą historię, podobnie jak nad kwestiami epistemologicznymi (dotyczącymi chociażby naszej wiedzy o świadomości innych istot). Od tych dyskusji należy odróżnić zagadnienia metodologiczne dotyczące filozoficzno-naukowego sposobu badania zjawisk mentalnych.

Dodatkową strukturę omawianym tematom oraz całemu opracowaniu nadają wyprowadzone twierdzenia i przyjęte założenia (por. I.1, I.2, ... II.1, II.2, ...). Eksponują one reprezentowane przez nas stanowisko, pozwalając je odróżnić od innych stanowisk. Twierdzenia te nie mają charakteru dowodów.

Programy badawcze podobne do przedstawionych w tej książce są współcześnie realizowane w neuropsychologii, neurofilozofii i neurofenomenologii (por. Walsh 2000: 30–40, 143–160; por. Borkowska i Szepietowska 2000: *passim*). W naszym podejściu rozgraniczamy dociekania filozoficzne od opisów badań neurologicznych pokazując, że poszukiwanie neuronalnych korelatów świadomości nie odbywa się bezzałożeniowo. Zakłada ono pytania natury filozoficznej i filozoficzną interpretację wyników badań neurologicznych. Dlatego zamiast do „neurofilozofii” (gdzie „neuro” stoi na pierwszym miejscu), ograniczamy się: (a) do takiego filozoficznego postawienia problemu subiektywności świadomości, który jest aktualny także dla neurologii, oraz (b) do filozoficznych analiz przebiegu i wyników badań neurologicznych.

Do badań neurologicznych wykorzystuje się aparaturę rejestrującą zdarzenia czasoprzestrzenne. W tym sensie mówimy o fizycznym poziomie zjawisk neurofizjologicznych, nie wchodząc głębiej w samą fizykę (np. próbując wytłumaczyć świadomość podmiotu prawami kwantowymi, jak to czynią inni¹⁰).

Zgodnie z zakreślonym programem badań, powołujemy się głównie na literaturę filozoficzno-psychologiczną (powstałą w tradycji filozofii analitycznej) oraz na literaturę neurologiczną¹¹. Jedyne sporadycznie odwołujemy się do „tradycji” kontynentalnej czy amerykańskiej, traktujących o problemie subiektywności świadomości¹². Tradycja ta nie zamykała się na problemie

¹⁰ S. Hameroff i R. Penrose wyjaśniają świadomość odwołując się do praw fizyki kwantowej (por. Gabora 2001: 132–140).

¹¹ Omawianej przez nas tematyce jest poświęcone specjalne wydanie „Świata Nauki” (1/2003) zawierające chociażby artykuły: Damasio 2003: 4–9; Crick 2003: 10–17; Gazzaniga 2003: 26–31; Chalmers 2003: 102–112.

¹² Należy chociażby wspomnieć istotne dla badań nad świadomością prace E. Husserla, R. Ingardena, E. Stein, M. Merleau-Ponty’ego, W. Jamesa i H. Bergsona.

subiektywności świadomości, lecz starała się przewyciężyć podział subiektywne–obiektywne. Czynili to np. W. James, a potem C. H. Mead, a w Europie M. Merleau-Ponty. Referując prace współczesnych amerykańskich kognitywistów pamiętamy, chociaż ich bliżej nie omawiamy, o osiągnięciach fenomenologów czy pragmatystów.

Powstanie tego opracowania było możliwe dzięki różnorodnej pomocy, której doświadczyłem od bardzo wielu osób. Wszystkich nie da się wymienić. Niemniej jednak pragnę serdecznie podziękować panu prof. dr. hab. Jerzemu Perzanowskiemu i panu doc. dr. hab. Krzysztofowi Krzyżewskiemu (Uniwersytet Jagielloński) oraz panu dr. Robertowi Poczobutowi (Uniwersytet w Białymstoku) za otrzymane od nich cenne uwagi oraz wskazówki. Dziękuję panu prof. dr. Chrisowi Maloneyowi (University of Arizona, Tucson), za umożliwienie mi w roku 2001 semestralnego pobytu jako *visiting scholar* w Uniwersytecie Arizony. Dziękuję panu prof. dr. Davidowi Chalmersowi (wówczas profesorowi Uniwersytetu Arizony) za poświęcony mi czas na „filozoficzne dyskusje” na temat subiektywności świadomości. Panu doc. dr. hab. med. Dariuszowi Adamkowi (Zakład Neuropatologii, Collegium Medicum, Uniwersytetu Jagiellońskiego) dziękuję za liczne sugestie dotyczące omawianych przeze mnie zagadnień związanych z neurologią mózgu. Prof. dr. hab. Stanisławowi Ziemiańskiemu i dr. Robertowi Januszowi (Wyższa Szkoła Filozoficzno-Pedagogiczna Ignatianum w Krakowie) dziękuję za korekty jednego z „draftów” maszynopisu. Dziękuję wszystkim uczestnikom prowadzonych przeze mnie w Wyższej Szkole Filozoficzno-Pedagogicznej seminariów na temat filozofii umysłu.

Wydawcom ogólnopolskiej serii „UMYSŁ. Prace z filozofii i kognitywistyki” dziękuję za przyjęcie mojego opracowania do wydawanej przez nich serii. Paniom pracującym w Wydawnictwie IFiS PAN dziękuję za prowadzenie prac wydawniczych.

W lipcu 2003 roku zmarł nagle pan prof. dr. hab. med. Józef Kałuża (Zakład Neuropatologii, Collegium Medicum, Uniwersytetu Jagiellońskiego). Przez ostatnie dwa lata swego życia Pan Profesor z wielką cierpliwością wspierał mnie swoją fachową wiedzą, przez co na zawsze pozostanie w mojej pamięci.

ROZDZIAŁ PIERWSZY

WSPÓŁCZESNA ANALITYCZNA FILOZOFIA ŚWIADOMOŚCI

1. Wprowadzenie: cel pracy i przyjęta metoda

A. W niniejszym opracowaniu chcemy zrealizować dwa ogólne cele badawcze:

1. Opisać w sposób systematyczny aspekt zjawiskowy świadomości (rozdział: II), a następnie określić zasadnicze kroki metodologiczne, pozwalające na ujęcie tego aspektu od strony neurologicznych badań mózgu (rozdziały: VI i VII). Pozostałe rozdziały zawierają krytyczne przedstawienie – szeroko rozumianych – analitycznych teorii umysłu i świadomości.

2. Drugi cel wynika z pierwszego i jest z nim ściśle związany. Chodzi o ukazanie filozoficzno-metafizycznych i naukowych założeń analitycznych teorii świadomości. Teorie te odwołują się często do twierdzeń o konieczności łączenia płaszczyzny neurologicznej i filozoficznej. Adekwatne przedstawienie tego połączenia domaga się zbudowania wspomnianej już wielowymiarowej teorii świadomości, którą będziemy nazywali „nową teorią świadomości”. Wychodzimy przy tym z założenia, że:

I.1 Filozofowie i psycholodzy zajmujący się świadomością muszą uwzględnić neurologiczne badania struktur mózgu, a neurofizjologowie muszą dostrzegać analizy filozofów i psychologów, jeżeli dzięki tym analizom staje się zrozumiałe funkcjonowanie mózgu jako całości oraz poszczególnych jego części.

W związku z I.1 pojawiają się dwa rodzaje pytań:

a. Pytanie o teoretyczno-empiryczną podstawę poszukiwanej teorii świadomości: Na ile szczegółowe wyniki uzyskiwane w naukach o systemie nerwowym powinny być brane pod uwagę przy tworzeniu tej teorii? Teoria ta ma bowiem pozwolić na systematyzowanie i wyjaśnianie kolejnych faktów gromadzonych w neurologicznych i filozoficznych ujęciach świadomości.

b. Ogólne pytanie o typ relacji pomiędzy zdarzeniami w mózgu a zdarzeniami mentalnymi oraz naukowe pytanie szczegółowe o korelację pomiędzy świadomością przejawiającą cechy subiektywne a zdarzeniami w mózgu.

Znalezienie wspomnianych korelacji należy traktować jako pierwszy krok do zrozumienia, czym są stany świadomości oraz do zbudowania wspomnianej teorii świadomości. Przyjmujemy przy tym, że rodzaj tej korelacji nie może zostać ustalony *a priori*, jak to ma na przykład miejsce w niektórych analitycznych teoriach świadomości (por. rozdział III). Kolejne badania naukowe określają nie tylko ścisły związek, ale i rozdział pomiędzy świadomością i fizycznymi warunkami, w których jest ona realizowana.

Należy zauważyć, że samo znalezienie korelatów świadomości jeszcze jej nie wyjaśnia. Wciąż bowiem brak odpowiedzi na zasadnicze pytania typu: Czy i jak stany neuronalne wywołują świadomość? W jaki sposób stany neuronalne wywołują stany zjawiskowe? Nie potrafimy ani teoretycznie wyjaśnić, w jaki sposób subiektywnie rozumiana świadomość zależy od struktury mózgu, ani empirycznie „pokazać” jej w mózgu (choćby za pomocą odpowiedniej aparatury). Dla uzyskania tak rozumianego wyjaśnienia potrzebujemy dalszego doprecyzowania pojęć, kolejnych badań naukowych oraz zbudowania bardziej adekwatnej, naukowej teorii świadomości. Uwzględniając dzisiejszy stan badań neurologicznych i analiz filozoficznych, będziemy się starali pokazać, że poszukiwana teoria należy do grupy teorii emergencji (por. 3.6).

B. Stosowana w tej pracy metoda odpowiada metodom filozofii analitycznej. Badamy nie tylko umysłowo-cielesne aspekty rzeczywistości, lecz także język teorii, za pomocą których mówimy o umyśle i o ciele. Jakiś stan nie musi sam w sobie być fizyczny lub mentalny. Termin „stan mentalny” rozumiemy zasadniczo jako „stan opisany jako mentalny”. Od metod tradycyjnej filozofii analitycznej odbiegamy w tym, że poszukując teorii świadomości odwołujemy się zarówno do wyników uzyskanych w szczegółowych naukach o mózgu (por. rozdział VII), jak i do opisów świadomości opartych na jej filozoficzno-psychologicznych ujęciach (por. rozdział II).

Przyjmujemy: (a) że świadome stany zjawiskowe są przez podmiot przeżywane. W tym sensie, mówiąc ogólnie, subiektywność jest rozumiana jako różnica pomiędzy „świadomością w ogóle” a „moją świadomością” (jestem i przeżywam siebie jako kogoś odrębnego od innych podmiotów).

Uwzględniając założenie (a) przeprowadzamy krytyczną analizę dotychczasowych monistyczno-materialistycznych teorii świadomości: ich punktu wyjścia, przyjętej metody postępowania oraz uzyskanych wyników. Należy dodać, że krytyka tych teorii nie jest prowadzona, przykładowo, z pozycji dualistycznych. W naszych badaniach rozróżniamy dwa pytania:

1. Czy świadomość może być wyjaśniona przez teorie materialistyczne: redukcjonistyczne i niereducjonistyczne? Teorie materialistyczne stawiają sobie za cel zamknięcie luki wyjaśniającej, która pojawia się chociażby pomiędzy filozoficzną, a neurologiczną płaszczyzną opisu i wyjaśniania.

Oraz: 2. Czy świadomość jest czymś materialnym?

W niniejszej pracy będzie nas głównie interesowała odpowiedź na pytanie 1. Należy jednak dodać, że zarówno od strony aparatury pojęciowej, jak i od strony empiryczno-ontologicznej, odpowiedzi na te pytania są ze sobą powiązane.

1.1. „Trudny” problem świadomości i jego neurologiczne znaczenie

„[...] subiektywność jest dla materialistycznych teorii umysłu swoistym «Trójkątem Bermudzkiem»: materialistyczne argumenty za jakąś teorią po prostu znikają z ekranu radaru, gdy mniej lub bardziej poważnie zaczyna się traktować subiektywne intuicje i je przedstawiać. Subiektywność rozumiana w jej topornym sensie nie jest dostępna dla wyjaśnień materialistycznych” (Hershfield 1998: 236).

Główna część niniejszego opracowania nosi charakter metateoretyczny, w sensie badania teoretycznych ujęć podstawowej cechy świadomości zjawiskowej – jej subiektywności. Jeżeli J. Hershfield twierdzi (por. wyżej), że współczesne, materialistyczne teorie świadomości nie potrafią wyjaśnić subiektywności, to musimy w miarę szczegółowo pokazać, dlaczego tak jest.

Jedną z przyczyn zawodności rozwiązań materialistycznych zdaje się leżeć w tym, że rezygnują one z uwzględniania filozoficzno-psychologicznych opisów świadomości, rezygnując tym samym z pytania dotyczącego „trudnego” problemu świadomości. Drugą przyczyną jest to, że zarówno przedstawiciele nauk ścisłych, jak i filozofowie używają terminu „świadomość” w różnych, często odmiennych znaczeniach.

Przykład „bogactwa znaczeniowego” tego terminu znajdujemy chociażby w wydanej przez T. Metzinger książce o neuronalnych korelatach świadomości¹.

¹ Por. Metzinger 2000: passim. Książka zawiera 22 opracowania napisane przez filozofów, psychologów i neurologów. Zainteresowanie badaniami nad korelatami świadomości

(1) Przez jednych filozofów termin „świadomość” jest odnoszony do organizmów jako całości, przez innych do konkretnych stanów jakiegoś organizmu (mówi się wtedy, że pewien stan jest stanem świadomym, a inny nie jest)². Twierdzi się, że dany stan mentalny jest świadomy, gdy jest on stanem, którego ktoś jest świadomy, ale także że stan mentalny jest świadomy, gdy ktoś dzięki niemu jest czegoś świadomy (choć sam ten stan nie musi mu być uświadomiony).

(2) Dla neurologów „świadomość” oznacza „przytomność” pacjenta, czyli jego zdolność do komunikowania. Neurologowie nie posługują się zwrotami „trudny” problem świadomości czy zwrotem „teorie emergencji” i raczej nie pytają o neuronalne korelaty stanów świadomych. Tego rodzaju korelatów poszukują filozofowie, a wśród neurologów zajmują się nimi głównie neuroradiolodzy. Jednym z pytań jest, jak wspomnieliśmy: czy interesująca filozofów tematyka subiektywności świadomości może mieć swe uzupełnienie w neurologii?

(3) W naukach o sztucznej inteligencji pod terminem „świadomość” rozumie się najczęściej rozbudowaną zdolność systemu do wykonywania operacji obliczeniowych. System jest tutaj rozumiany jako układ fizyczny, który można opisać dynamicznie. Przyjmuje się, że gdy uda się zaprogramować funkcje świadomości (na przykład w procesie przetwarzania informacji), wtedy nie będzie potrzeby odwoływać się do świadomości, aby owe funkcje wyjaśnić. Procesy, które zachodzą w świadomym podmiocie, mogą zachodzić w systemach mechanicznych albo elektronicznych, bez udziału świadomości. Czy takie rozumienie świadomości wystarczy jednak do uchwycenia jej subiektywności?

Dostrzegając powyższe rozbieżności w rozumieniu „świadomości” można podać dwa sposoby przewyżczenia problemu jej filozoficznych i neurologicznych określeń. Jednym z nich jest próba tworzenia dokładniejszych definicji świadomości. Świadomość definiuje się wtedy albo w terminach psychologicznych (podając, rolę jaką stany świadome odgrywają przy podejmowaniu przez podmiot decyzji, przy przetwarzaniu informacji o jego

mości wywołały prace N. K. Logothetisa i J. D. Schalla, którzy omawiając badania przeprowadzone na małpach stwierdzają: „[...] the results of this study suggest the possibility of experimentally relating the activity of single neurons in the visual system to the internal perceptual state of the subjects” (Logothetis 1989: 763).

² Zdaniem Blocka: „The word «consciousness» connotes a number of different concepts and denotes a number of different phenomena. We reason about «consciousness» using some premises that apply to one of the phenomena that fall under «consciousness» other premises that apply to other «consciousness», and we end up with trouble” (Block 1995a: 227).

otoczeniu), albo w terminach neurologicznych (wskazując na odpowiednie zdarzenia w mózgu, skorelowane ze zdarzeniami świadomymi), albo w terminach chemicznych (mówiąc o wymianie substancji chemicznych w synapsach neuronów). Świadomość można wreszcie próbować definiować w terminach pochodzących z nauki o sztucznej inteligencji, odwołując się do komputerowego modelu jej materialnego podłoża.

Drugi sposób polega na podawaniu przykładów występowania świadomości oraz na systematyzowaniu rodzajów i funkcji świadomości, precyzując tym samym znaczenie terminu „świadomość”. W dalszych rozdziałach pracy przyjmujemy tę drugą metodę i stosujemy ją głównie przy analizach świadomości zjawiskowej.

1.2. Świadomość zjawiskowa – przykłady występowania

Brak jasnego pojęcia tego, czym jest świadomość stawia pod znakiem zapytania zarówno metody jej badania, jak i to, że uda nam się znaleźć jej neuronalne podłoże. Przy badaniach świadomości wyróżnia się zazwyczaj: (A) jej funkcje³, (B) jej rodzaje⁴.

(A) Do funkcji świadomości zalicza się: zapewnienie diachronicznej ciągłości podmiotu pomiędzy czasem t_1 a t_2 , monitorowanie stanów umysłu (na przykład „wiem, o czym myślę”), kontrolę i planowanie odpowiednich działań. Pytanie o funkcje świadomości nie może być oddzielone od pytania o przyczynowe oddziaływanie świadomości. Jeżeli bowiem świadomość podmiotu nie ma przyczynowego wpływu na aktywność neuronalną, to jaka jest jej rola w aktywności jego mózgu? Ukazanie wzajemnego powiązania tych pytań jest wymaganiem, przed którym stoi każda teoria świadomości.

(B) Mówiąc o podstawowych rodzajach świadomości mamy najczęściej na myśli zagadnienia związane z tożsamością podmiotowego „ja”: wiedza podmiotu o sobie samym, jego dostęp do własnych stanów umysłu, jego samoświadomość, stany subiektywne utożsamiane zazwyczaj z jakościami doznaniowymi (czyli „qualiami”).

³ Zdaniem Jamesa świadomość nie istnieje, istnieją jedynie funkcje poznawcze spełnianie przez myśli. „[...] I mean only to deny that the word [consciousness – J.B.] stands for an entity, but to insist most emphatically that it does stand for a function” (James 1904: 478).

⁴ Nienależący do filozofii analitycznej H. Bergson wyróżnia trzy cechy świadomości, która jest doświadczana jako coś charakteryzującego się jednością i jest rozumiana jako subiektywne przeżycie. Owe cechy to: pamięć, uwaga oraz antycypacja przyszłości. Por. Kulczycki 1996: 75–77, 158–160.

Utworzone w (A) oraz (B) listy nie są ani listami zupełnymi, ani zamkniętymi.

W niniejszym opracowaniu odwołujemy się do wprowadzonego przez N. Blocka rozróżnienia pomiędzy dwoma „pojęciami” (*concepts*) świadomości: świadomością zjawiskową, fenomenalną (*phenomenal consciousness*) i świadomością dostępu (*access consciousness*)⁵. Pierwsza z nich jest natury „przeżyciowej”, druga – „sprawozdawczej” (tzn. zawiera w sobie element językowy). Wspomniane dwa rodzaje świadomości Block odróżnia od samoświadomości i od świadomości kontrolującej zachowania. Stąd nasze założenie:

- I.2 Można wyróżnić co najmniej dwa aspekty świadomości: zjawiskowy i dostępu.

Świadomość zjawiskowa (przeżyciowa) i świadomość dostępu są ze sobą powiązane: tzn. normalnie występują one razem (gdy jestem przeżyciowo świadom czerwonej kartki, mam również do niej świadomy dostęp). Jednak pomimo naturalnego połączenia, możemy wprowadzić pojęciowe rozróżnienie pomiędzy świadomością dostępu, a świadomością zjawiskową (możemy sobie wyobrazić, że jakiś stan jest dostępne świadomy, bez odpowiadającego mu stanu przeżyciowego).

Block zaznacza, że kontrowersyjne w jego podziale na świadomość dostępu i zjawiskową, jest to, że przyjmuje on: „[...] cechy świadomości zjawiskowej jako różne od jakiejś kognitywnej, intencjonalnej czy funkcjonalnej cechy” (Block 1995a: 230).

Do wprowadzonego przez Blocka rozróżnienia odwołują się T. Bayne i D. Chalmers zakładając, że wspomniane dwa aspekty świadomości występują zawsze razem. Według nich, świadomość dostępu jest zdefiniowana w terminach przyczynowych ról, jakie dany stan odgrywa, zaś świadomość zjawiskowa – w terminach sposobu, w jaki dany stan się odczuwa (por. Bayne i Chalmers 2003: 5).

W dalszych częściach pracy przyjmujemy, że świadomość dostępu odgrywa aktywną rolę w racjonalnym działaniu podmiotu, w jego mówieniu i argumentowaniu. Natomiast świadomość zjawiskowa jest określona przez jakościowe cechy charakteryzujące przeżycia.

⁵ Por. Block 1995a: 230–245. Wprowadzone przez Blocka rozróżnienie opisuje Burge (1996: 584–592). Krytykę tego rozróżnienia przeprowadza Rosenthal (1998: *passim*).

I.3 Cechy zjawiskowe – to przeżywane cechy wrażeń, odczuć i spostrzeżeń typu: jak to jest być podmiotem? Stany intencjonalne mogą wprawdzie posiadać cechy zjawiskowe, lecz te nie służą do ich identyfikacji.

Cechy zjawiskowe mają przeżyciowy charakter, służący do ich identyfikacji. Według filozofa T. Burge'a, pojęcie świadomości zjawiskowej dotyczy centralnego aspektu świadomości: „Uważam, że istoty, które nie są świadome na sposób zjawiskowy (*phenomenal*), w ogóle nie są świadome” (Burge 1996: 585).

Zgadzając się z tezą Burge'a przyjmujemy, że warunkiem wystąpienia subiektywności jest istnienie świadomości zjawiskowej. Z podobnego założenia wychodzi neurolog A. Revonsuo: „Według mnie, świadomość zjawiskowa stanowi centrum naszego rozumienia świadomości” (Revonsuo 1995: 266.).

Filozoficzne dyskusje o świadomości zjawiskowej są często redukowane do dyskusji o „qualiach”⁶. Przyjmuje się, że qualia charakteryzują przeżycia podmiotu, ale prawdopodobnie nie odgrywają ani funkcjonalnej, ani przyczynowej roli w procesach poznania i w zachowaniach. W sensie rodzajowym, qualia są najczęściej rozumiane jako własności (cechy), do których tylko dany podmiot ma dostęp. Problematyka qualiów nie stanowi zasadniczej części niniejszego opracowania, chociaż pojawia się w nim (por. 2.4.1). Rozwiązanie „trudnego” problemu, rozumianego jako problem qualiów, jest – naszym zdaniem – mniej związane z samą świadomością, a bardziej z argumentacją odwołującą się do ich przedstawialności. Jako świadome podmioty czujemy się niejako „umiejscowieni” w świecie jakości zjawiskowych (qualiów), w którym całościowe przeżycie tworzy subiektywny świat, przedstawiając zewnętrzny świat i nasze ciało, jako składające się z już dalej niepodzielnych, ale określonych jakości. Zewnętrzny świat ukazuje nam się zwykle jako całość złożona z różnorodnych jakości, należących do przedmiotów będących częściami tego świata.

Uwzględniając rozróżnienie na świadomość dostępu i na świadomość zjawiskową możemy podać następujące przykłady występowania tej ostatniej. Powiedzmy, że patrzymy na otaczającą nas rzeczywistość, a następnie zamykamy oczy. Tym samym znika nasze świadome wzrokowe przeżywanie otoczenia. Po zatkaniu uszu znika świadome przeżywanie dźwięków docie-

⁶ W historii filozofii qualia rozumiano jako paradygmat cech mentalnych (niefizykalnych), jako materiał zmysłowy: odczucia (*sensations*, por. Locke 1955: passim), wrażenia (*impressions*, por. Hume 1963: passim), intuicje (*Anschauungen*, por. Kant 1957: passim). Współcześnie, terminem tym określa się również tak zwane surowe odczucia (*raw feels*, Feigl 1958: 416). Według Stubenberga, „To be conscious is to have qualia” (Stubenberg 1998: 12).

rających z otoczenia. Podobnie z innymi zmysłami. Nadal jednak jesteśmy świadomi, chociażby siebie samych, swojego oddechu, położenia swojego ciała. Świadomość tracimy, gdy zapadamy w głęboki sen (tzw. *slow wave sleep*, SWS), albo gdy jesteśmy pod wpływem całkowitej narkozy. Natomiast marzenia sennie podczas tak zwanej fazy snu REM (*rapid eye movement*) są świadomymi przeżyciami, chociaż często posiadają one nieco inną strukturę i są mniej spójne niż przeżycia na jawie⁷. Sny i halucynacje mogą być pod względem przeżyciowym bardzo intensywne. Osoba śniąca nie przyjmuje bodźców z zewnątrz (wykluczając przykładowo dzwonienie budzika), doświadcza jednak całej gamy subiektywnych przeżyć.

Zdolność do odpowiadania na bodźce zewnętrzne odróżnia stan snu od stanu czuwania. Stan czuwania cechuje się uważnym, skoordynowanym przeżywaniem naszej własnej obecności w świecie zewnętrznym.

Posiadamy jakieś poznanie tego, *jak to jest być świadomym* (gdy śnimy albo jesteśmy zbudzeni), jako przeciwieństwo do niebycia świadomym (głęboki sen, narkoza). Świadomość zjawiskowa jest obecna wtedy, gdy jest obecna przeżyciowa treść. Wiemy, *jak to jest widzieć* coś czerwonego, *jak to jest widzieć* coś zielonego. Czerwone ukazuje się jako subiektywnie odmienne od zielonego, a to jako subiektywnie odmienne od żółtego itd. Subiektywne ukazywanie się tych jakości pomaga nam uchwycić owo *jak to jest być* podmiotem z normalnym spostrzeganiem kolorów.

Nie wiemy jednak *jak to jest* dla nietoperza spostrzegać świat, ponieważ szeroko rozumiane „pojęcia”, o które tutaj chodzi, może mieć tylko ten, kto sam jest w stanie mieć takie spostrzeżenia (albo potrafi je sobie wyobrazić). Jakościowy charakter różnych wrażeń (czy też jakościowy charakter doznań cielesnych) jest w tym sensie *subiektywny*, gdyż pojęcie odnoszące się do takiego stanu, może posiadać tylko ten, kto posiada odpowiednią perspektywę przeżyciową.

Z problemem subiektywności spotykamy się głównie w obrębie świadomości zjawiskowej: Dlaczego stany świadome w ogóle występują i dlaczego są odczuwane w określony sposób? Jaka korzyść daje nam to, że coś przeżywamy? Możliwe wydaje się bowiem zbudowanie systemu (tzw. zombi), który będzie posiadał świadomość dostępu chociaż nie będzie nic przeżywał (będzie pozbawiony świadomości zjawiskowej).

Z kolei hipoteza o niemożliwości zbudowania takiego systemu znajduje poparcie w twierdzeniu, że świadomość (jako całość) posiada w sobie coś

⁷ Fenomenologiczne analizy „utruty przytomności” podczas snu (bez rozróżniania faz REM i SWS) oraz „jakiegoś doświadczenia o tym, że spaliśmy” podaje R. Ingarden (1948: 741).

niewypowiedzianego, a tym samym niemożliwego do jej obiektywnego, teoretycznego ujęcia.

Używając w tej pracy terminu „świadomość”, będziemy mieli na myśli głównie przeżyciowy stan *jak to jest być świadomym* oraz doznaniowo-jakościowe cechy stanów mentalnych. Termin „świadomość” może być zakresowo równy terminowi „subiektywność” wtedy, gdy wszystkie stany świadomości zjawiskowej są rozumiane jako stany przeżywane przez dany podmiot.

Świadomość zjawiskowa nie jest koniecznie świadomością jakiegoś przedmiotu materialnego, choć posiada ona jakąś intencjonalną treść. Gdy na przykład zamkniemy powieki i naciśniemy gałki oczu, „zobaczymy” różnego rodzaju kolorowe plamy, które, jak się zdaje, nie przedstawiają żadnego obiektu. Świadomość zjawiskowa nie wymaga posiadania języka i pojęciowego rozumienia przez podmiot siebie samego. Niemowlę przeżywa otaczający świat, chociaż nie posiada języka.

1.3. *Filozofia świadomości – podstawowa terminologia*

Filozofia świadomości (*philosophy of consciousness*) należy do filozofii umysłu (*philosophy of mind*). Ta ostatnia obejmuje zazwyczaj filozoficzną psychologię (analiza pojęć, których używamy mówiąc o zjawiskach mentalnych), filozofię psychologii (obejmującą tworzenie psychologicznych teorii i zagadnienia zwane kognitywnym zwrotem w psychologii) oraz tę część metafizyki, w której pytamy o ontologiczną naturę zjawisk mentalnych (por. Beckermann 1999: 1154–1156).

R. Kartezjusz jako pierwszy zbudował teorię umysłu, według której poprawne rozumienie stanów mentalnych zakłada posiadanie pojęć odnoszących się do świadomości. Podobnie twierdził J. Locke, przyjmując, że myślenie obejmuje bycie świadomym tego, że się myśli (por. Rosenthal 2002: 229–230).

W dalszej części pracy odróżniamy terminy „świadomość” (*consciousness*) i „umysł” (*mind*). Umysłem jest to, co potrafi myśleć, postrzegać, co J. Perzanowski definiuje jako: „[...] *jakiś zbiornik myśli, tj. coś, w czym lokuje się maksymalna liczba współlokowanych myśli*, albo też jako *minimalną kombinację myśli*”⁸.

⁸ Perzanowski 1995: 23. Perzanowski wypracował ontologię lokacyjną pozwalającą zarówno na opis lokacji czasoprzestrzennych, jak i lokacji mniej konkretnych, np. mentalnych (por. Perzanowski 1993: 43–45). Omówienie szeroko dzisiaj dyskutowanej komputacyjnej teorii umysłu można znaleźć w: Judycki 1997: 40–42. Dokładne przedstawienie problematyki umysłu w historii filozofii oraz w czasach współczesnych podaje U. Żegleń (2003: 27–98).

Myśli się „lokują” w zależności od tego, jak wiążą się z całością procesów i stanów poznawczych jakiegoś podmiotu oraz z jego działaniami. W tym sensie „umysł” możemy określić terminami: „umieścić” i „umyśleć” rozumianymi jako „generowanie” myśli oraz „przetwarzanie informacji”. Przetwarzanie informacji może zachodzić w podmiotach naturalnych lub sztucznych. Termin „informacja” traktujemy przy tym jako pierwotny.

Czy umysł jest materialny lub niematerialny? Według monistycznych materialistów, umysł z wszystkimi jego stanami mentalnymi i odczuwanymi jakościami jest czymś fizycznym. Dualiści przyjmują zaś, że umysł jest niematerialny i nieprzestrzenny. W powyższej definicji umysłu, jak Perzanowski sam zauważa, „lokacja” nie jest ograniczona do zwykłej lokalizacji czasoprzestrzennej. Jest ona rozumiana jako wystarczająco ogólna, aby można ją było wykorzystać „jako podstawę do badania przypadków lokacji ekstrasfizycznej, np. lokacji moich myśli w czasie pisania tego eseju” (Perzanowski 1995: 21).

Mówiąc w dalszej części pracy o „świadomości” mówimy o świadomym podmiocie albo, za Chalmersem, o „świadomym umyśle” (*conscious mind*). Zarazem negujemy ściśle behawiorystyczne rozumienie umysłu⁹. Według Chalmersa, świadome przeżycie jest pewnym podstawowym fenomenem umysłu, niesprowadzalnym do innych zjawisk, chociaż równocześnie jest ono traktowane jako zjawisko naturalne podpadające pod prawa natury (choć nie jest powiedziane, że są to te prawa natury, które dzisiaj znamy) [por. Chalmers 1996: xii–xiii].

Umysł nie jest tym samym co *mózg* – biofizyczny organ, w którym zachodzą procesy odpowiedzialne głównie za utrzymanie życia organizmu. Większość z tych procesów przebiega bez udziału świadomego umysłu. Z drugiej strony nie ulega wątpliwości, że istotną podstawę funkcjonowania umysłu znajdujemy w neurochemii mózgu. Omawiany w tym opracowaniu materialistyczny monizm przyjmuje założenie o prymacie struktury mózgu nad dynamiką jego powstawania i funkcjonowania.

Pojęcie „umysł” jest bardziej nieokreślone aniżeli pojęcie „mózg”. Rezygnując ze ściśle ontologicznej odpowiedzi na pytanie „czym jest umysł?”, można w dalszych analizach, oprócz jego lokacyjnej definicji, przyjąć jego „pragmatyczną” definicję: świadomy umysł niesie ze sobą wiele korzyści, służy nam bowiem do rozwiązywania problemów napotykanych w życiu, do zrozumienia innych podmiotów i otaczającego nas świata. Tak rozumiany umysł realizuje program skierowany na przetrwanie organizmu.

⁹ Zdaniem behawiorysty D. Hebba: „Broadly speaking, the mind is that which controls behavior” (Hebb 1980: 3).

Istnienie świadomego umysłu stanowi istotę problemu umysł–ciało, rozumianego jako pytanie o przyczynowe oddziaływanie pomiędzy umysłem a ciałem (pod „przyczynowe oddziaływanie” rozumiemy dynamiczną relację pomiędzy zdarzeniami w czasie). Według T. Nagła:

„Gdyby nie istniała świadomość, problem stosunku umysłu do ciała byłby mniej interesujący. Ze świadomością – wydaje się beznadziejny” (Nagel 1997: 204).

Celem wstępnego określenia stosunku świadomego umysłu do ciała niezbędne jest dokonanie dalszych podstawowych objaśnień i ustaleń terminologicznych.

W języku angielskim odróżniamy dwa terminy: *awareness*¹⁰ (przytomność, zdawanie sobie sprawy) i *consciousness* (świadomość), które czasami są traktowane jako synonimy. Zdawanie sobie sprawy – to posiadanie dostępu do informacji, którą można w jakiś sposób używać. Gdy ktoś jest świadom czegoś, zdaje sobie także z tego sprawę (w tym sensie zdawanie sobie sprawy jest składową świadomości). Zdawanie sobie sprawy zawiera w sobie elementy spostrzeżeniowe.

Jesteśmy świadomi (w sensie zjawiskowym), gdy coś jest w nas obecne na sposób doświadczenia tego „jak to jest być w danym stanie?”. Stany zjawiskowe nieświadome (*nonconscious*) posiadamy, ale bez owego momentu *jak to jest być* w tym stanie. Ktoś może posiadać nieświadome myśli o obiektach leżących w tle jego pola widzenia, które możemy traktować jako nieświadome zdawanie sobie z nich sprawy. W dalszym odwoływaniu się do zwrotu „jak to jest być?” rezygnujemy z pytajnika na końcu, a sam zwrot piszemy kursywą, co wskazuje, że jest on traktowany jako operator techniczny.

Angielski termin *self* (*selves*) odpowiada polskiemu „ja sam” i można go przetłumaczyć jako „jaźń”¹¹. Termin „jaźń” przypomina termin abstrakcyjny, utworzony analogicznie do terminu „osobosc” (od „osoba”, wtedy „ja-źń” odpowiadałaby terminowi „ja-ość”, gdzie „ość” jest abstraktorem dodanym do „ja”)¹². Termin „jaźń” odnosi się głównie do osób, a jego desygnat może być rozumiany jako „jądro” osoby.

Terminem obejmującym osoby, zwierzęta, sztuczne inteligencje – aktualne, a także potencjalne – jest „podmiot” (*subject*, *Subjekt*). „Podmiot przeżyć”

¹⁰ Chodzi o *awareness* rozumiane w sensie „to be aware of Z”. Por. Johnson i Reader 1997: 264–265.

¹¹ Czarnecki i Cieśliński tłumaczą „self” jako „jaźń” (por. Czarnecki 2002: 51; por. *Widok znikąd*, tłum. Cieśliński, *passim*).

¹² Takie rozumienie terminu „jaźń” zaproponował J. Perzanowski.

jest zakresowo szerszym aniżeli „osoba ludzka”. Wszystkie osoby są (przynajmniej potencjalnymi) podmiotami przeżyć, zaś nie wszystkie podmioty przeżyć są osobami. Niektóre wyższe gatunki zwierząt są podmiotami przeżyć, trudno zaś mówić, że są osobami. Osoby wchodzą w relacje międzyosobowe, potrafią posługiwać się językiem, poznać same siebie, swoją własną tożsamość, i to, że są świadomym bytem. Mogą one poznać własne myślenie, rozumowanie, odczuwanie. Tego rodzaju poznanie jest ściśle związane z posiadaniem pojęcia „pierwszej osoby” („ja”) i z odniesieniem tego pojęcia do podmiotu jako do siebie samego. Termin „subiektywność” („podmiotowość”, *subjectivity*) ma swoje etymologiczne korzenie w terminie „podmiot” (*subiekt*).

1.3.1. Świadomość z psychologicznego punktu widzenia

„We consider life without consciousness to be little or no better than death” (Siewert 1998: 329).

Termin „świadomość” został wyparty z akademickiej psychologii angloamerykańskiej w okresie zdominowanym przez teorie behawiorystyczne. Ponownie zaczął go używać T. Shallice, który w latach 70. minionego stulecia opisywał świadomość w terminach „przetwarzania informacji” oraz w terminach nawiązujących do superwizyjnego systemu, nakierowanego na określony cel (por. Shallice 1972: 385–392). Opisy przetwarzania informacji były następnie używane przez Shallice’a do wyjaśniania zależności pomiędzy przekonaniem, celami i zamiarami wspomnianego systemu. W psychologii – zarówno teoretycznej, jak i empirycznej – obserwujemy dzisiaj rosnące zainteresowanie badaniami nad świadomością. Stan ten znajduje swój wyraz chociażby w tworzeniu psychologicznych i neuropsychologicznych teorii świadomości, od których oczekuje się zarówno właściwego stopnia opisu, jak i wyjaśnienia powstawania świadomych myśli i celowych zachowań.

Psychologowie empiryczni skupiają się na obserwacjach świadomych podmiotów. Jeżeli pojęcie świadomości ma być użyteczne w teorii psychologicznej, to musi ono: (a) być w relacji do pojęć uzyskanych na drodze badań empirycznych, (b) umożliwiać prognozy mające behawioralne (czyli stwierdzalne eksperymentalnie) konsekwencje. Stąd wynika hipoteza: jeżeli świadomość ma odgrywać jakąś weryfikowalną rolę w teorii przetwarzania informacji, to musi mieć jakieś funkcjonalne oddziaływanie na odpowiednie procesy fizyczne. Aby można było empirycznie wykazać, że tego rodzaju funkcjonalne oddziaływanie zachodzi, trzeba umieć empirycznie chociażby wykazać, że procesy świadome różnią się od nieświadomych. Stąd w psycho-

logii zajmujemy się nie tylko budową teoretycznych modeli świadomości i nieświadomości, ale interesują nas także behawioralne wskaźniki używane do empirycznej oceny tych modeli. Oceny te mogą opierać się na danych otrzymanych z eksperymentów i być weryfikowalne za pomocą pomiarów pośrednich lub bezpośrednich.

Dwa przykłady wspomnianych obserwacji:

(a) Psycholodzy interesują się empirycznie stwierdzalnym wpływem efektów z przeszłości na subiektywne przeżycia w teraźniejszości. Stwierdzenie takiego wpływu jest możliwe dzięki odpowiednim eksperymentom. Na przykład, w eksperymencie przeprowadzonym przez L. L. Jacoby'ego, osoby słuchały znanych już zdań oraz nowych zdań wraz z tzw. białym szumem (*white noise*) w tle i oceniały poziom szumu (wszystkie częstotliwości tego szumu posiadają tę samą amplitudę). Szum tła był przez badane osoby subiektywnie oceniany jako cichszy w przypadku znanych zdań, a jako głośniejszy w przypadku nowych zdań (por. Jacoby 1988: 242–245). Osoby błędnie przypisywały swój odbiór znanych zdań do niższych poziomów szumu.

(b) Omawiając „punkt obserwacji”, który przypomina nagłowski *point of view*, psycholog U. Neisser opisuje eksperyment ukazujący powstawanie w dziecku pojęcia własnej jaźni¹³. Z eksperymentu wynika, że już trzyletnie dziecko potrafi zauważyć różnicę pomiędzy pojęciem jego jaźni, a pojęciem jego części ciała.

Dalszy rozwój eksperymentalnych technik w psychologii pozwoli nie tylko odkryć kolejne zależności umożliwiające lepsze zrozumienie roli świadomości w takich czy innych zachowaniach podmiotu, ale także empirycznie mierzalne wielkości w jego subiektywnym przeżywaniu bodźców wewnętrznych i zewnętrznych. Metodologiczną pomocą w ustalaniu naukowych programów takich badań może być precyzyjna odpowiedź na pytanie: Co jest

¹³ „In a game-like situation, young children may cover their own eyes with their hands and say «You can't see me!» This behavior is often interpreted in terms of Piaget's concept of egocentricity: when a child can't see, it just assumes that you can't see either. The results of an experimental study by Favell, [...] show, however, that this interpretation is wrong. When Favell et al. asked their eyes-covered subjects «Can I see you?», most two- or three-year-old children did indeed answer «No». But the very same subjects answered «Yes» to many other questions about what the experimenter could see. [...] «Can I see your leg?» Yes. «Your head?» Yes. «Can I see you?» No. These results suggest that «You can't see me» does not result from any childish egocentric misunderstanding; rather, it reflects the child's own conception of self. When my eyes are covered, you can't see me: what children call «me» is somewhere near what I have been calling the «point of observation»” (Neisser 1997: 22).

neuronalnym korelatem świadomości zjawiskowej, albo co jest takim korelatem dla poszczególnych stanów tej świadomości?

Należy zauważyć, że świadomość nie jest tylko czymś „wewnątrz” podmiotu. Jest tak chociażby dlatego, że ani mózg, ani umysł nie są systemami zamkniętymi. W przeciwnym razie musielibyśmy przyjąć, że obowiązuje solipsyzm („ja” w izolacji od otoczenia). Stany mentalne zależą od informacji pochodzących od i dochodzących do ciała. Zależą one nie tylko od ontologii i filogenezy ciała, od jego obecnych i przeszłych stanów fizycznych, lecz także od jego otoczenia społecznego. Trzeba więc zapytać, na ile możliwa jest indywidualizacja stanów świadomości bez uwzględnienia wspomnianych wpływów (por. 1.3.5).

1.3.2. Świadomość od strony metafizycznej

Rozważania metafizyczne we współczesnej filozofii świadomości należy odróżnić od rozważań ontologicznych, zajmujących się cechami bytów, stanów, zdarzeń oraz relacjami pomiędzy nimi. W uprawianiu ontologii chodzi nam o rodzaje obiektów, o których istnieniu aktualnie wiemy, lub też hipotetycznych obiektów, których istnienia jeszcze nie odkryliśmy. Ontologia stanów mentalnych zajmuje się ich podstawowymi kategoriami i ich relacjami do stanów fizycznych. W tym sensie problem umysł–ciało jest problemem ontologii obiektów mentalnych. Należy wspomnieć, że większość współczesnych filozofów analitycznych jest – szeroko rozumianymi – naturalistami. Świadomość jest dla nich cechą naturalnego świata (stąd jest ona często własnością „redukowalną”), a jej relacje do innych obiektów tego świata podlegają weryfikowalnym badaniom. Materializm jest najpierw tezą metafizyczną, następnie ontologiczną, a później metodologiczną. W dalszych badaniach przyjmujemy, że:

I.4 Metafizyka dotyczy fundamentalnej natury rzeczywistości świadomości, a jej metodologia jest (w zasadzie) *aprioryczna*.

Fizyka mówi nam *aposteriori*, jak własności fizyczne jakiegoś przedmiotu (takie jak chociażby zestawienie jego atomów), prowadzą do jego kształtu lub stałości. Metafizyka specyfikuje *aprioryczne* faktory, odpowiedzialne za posiadanie tychże własności.

Z jednej strony filozofia świadomości odwołuje się do metafizycznego statusu podmiotu oraz do jego miejsca wśród innych bytów. W tym sensie żadna z nauk szczegółowych nie może pretendować do miejsca zajmowanego przez metafizykę, ponieważ każda z nich zakłada ontologiczne ramy,

w których interpretuje wyniki swoich badań. Z drugiej strony, bez całościowego ujęcia rzeczywistości, do którego pretenduje metafizyka, nie potrafimy połączyć ze sobą interpretacji teorii naukowych z interpretacjami obserwacji przeprowadzanych w naukach szczegółowych.

Filozofowie nurtu empiryczno-pozytywistycznego kwestionowali zasadność stosowania metod *a priori*, twierdząc, że prawdziwe poznanie można osiągnąć tylko poprzez badania empiryczne. Jednak metafizyki nigdy nie udało się całkowicie wyeliminować z dociekań prowadzonych w ramach filozofii analitycznej (por. Wszolek 1997: 23–32, 98–158). Gdy przeciwnicy metafizyki twierdzą, że jest ona niemożliwa chociażby dlatego, że sama rzeczywistość jak i jej struktury są dla nas nieosiągalne, to formułują oni tym samym „metafizyczną” tezę. Rozumowanie metafizyczne nie jest ani niezawodne, ani łatwe do przeprowadzenia. Podobnie jak rozumowania empiryczne, nie dostarcza nam ono pewnych tez, gdyż takie tezy mogą być formułowane jedynie w matematyce¹⁴.

Zdaniem W. Sellarsa, nauka może pretendować do tego, że tylko ona określa, zgodnie z jej empirycznymi kryteriami, co w rzeczywistości istnieje¹⁵. Zapytajmy jednak, czy nauka także określa, co *może* albo *musi* istnieć? Tego typu pytania wychodzą poza empiryczną oczywistość, a odwołując się do tej oczywistości nauka potrafi jedynie potwierdzić to, co w świetle koherencyjnej koncepcji rzeczywistości faktycznie istnieje. Trudniej byłoby jej ustalić, co mogłoby, a co nie mogłoby istnieć. Empiryczna oczywistość dotyczy tych rzeczy, których istnienie jest *możliwe* w tym sensie, że one po prostu są.

Poszukiwanie najogólniejszej koncepcji rzeczywistości należy do zadań filozofii, której cel Sellars określa jako: „[...] rozumienie, jak rzeczy, w najbardziej szerokim sensie tego terminu są połączone ze sobą, w najbardziej szerokim sensie tego terminu” (Sellars 1963: 37).

Jego zdaniem osiągnięcie tego rodzaju zrozumienia rzeczy jest możliwe dzięki dostępnym nam dwóm ogólnym projekcjom świata i człowieka, które można nazwać projekcjami „metafizycznymi”. Są nimi manifestujący się obraz świata (*manifest image*) oraz metodologicznie od niego zależny nauko-

¹⁴ L. Wittgenstein zaleca ostrożność przy odwoływaniu się do tez metafizycznych w matematyce (por. Wittgenstein 1995a: 108).

¹⁵ Z jednej strony proponowana przez W. Sellarsa teza przypomina scjentyzm W.V.O. Quine’a. Sellars pisze: „[...] that in the dimension of describing and explaining the world, science is the measure of all things, of what is that it is, and of what is not that it is not.” (Sellars 1963b: 173). Z drugiej strony Quine i Sellars różnią się rozumieniem samych teorii oraz rozumieniem ontologii obiektów postulowanych w tychże teoriach (por. Bremer 1997: 158–159, 186).

wy obraz świata (*scientific image*). Pomijając mniej lub bardziej znaczące różnice, możemy powiedzieć, że manifestujący się obraz świata odpowiada naszemu potocznemu obrazowi świata (por. Bremer 1997: 9–43), albo obrazowi proponowanemu w tak zwanej psychologii potocznej (*folk psychology*). W tym obrazie osoby są „ucieleśnione” (*embodied*) oraz „posiadają duszę” (*ensouled*), i dlatego mogą myśleć i działać. Obecny w obrazie potocznym dualizm polega na tym, że świat osób jest światem intencjonalnych działań (działania te są wyrazem charakteru osób), zaś w świecie rzeczy obowiązuje wyjaśnianie przyczynowe. „Potoczny” dualizm nie implikuje, że tak rozumiane osoby są bytami złożonymi (por. Sellars 1963: 11). Innymi słowy, w obrazie potocznym nie musimy automatycznie postulować istnienia ducha i materialnego ciała¹⁶.

Filozoficzno-metafizyczne rozumowanie możemy traktować z jednej strony jako artykulację napięcia pomiędzy wspomnianymi obrazami świata, a z drugiej strony jako poszukiwanie empirycznych i pojęciowych kategorii przyszłego obrazu świata, który syntezywałby obydwa dotychczasowe obrazy. W syntezie tej nie chodzi o eliminację obrazu potocznego, ale o jego połączenie z obrazem naukowym, które może zajdzie w przyszłym „synoptycznym” (tzn. „zjednoczonym”) obrazie świata i człowieka.

Idąc po tej linii nie dążymy w niniejszej pracy do eliminacji pojęć potocznych lub pojęć *folk psychology*, lecz do doprecyzowania ich stosowania, tam gdzie jest ono niejasne (por. rozdział II). W tym sensie przedstawione tutaj rozumowania są prowadzone według myśli Wittgensteina: poprzez analizy językowe doprowadzić do lepszego zrozumienia nas samych jako myślących i działających podmiotów. Rozumowania te nie są natomiast prowadzone po linii Wittgensteina, gdy chodzi o radykalny rozdział filozofii od nauki. Wittgenstein uważa, że filozofia powinna się zajmować opisywaniem, wszelkie zaś wyjaśniania należy pozostawić naukom szczegółowym (por. Wittgenstein 2000: §§ 126–127).

W niniejszym opracowaniu zostają omówione filozoficzne próby rozwiązania problemu umysł–ciało, w szczególności propozycje rozwiązania problemu subiektywności świadomości. W rozdziale VII odwołujemy się do empirycznych wyników uzyskanych w ramach nauk dotyczących naszego

¹⁶ Filozoficzno-metodologicznego modelu pomagającego w zrozumieniu obydwu obrazów świata dostarcza teza Kanta o „regulatywnej zasadzie rozumu”, tzn. „[...] żeby wszystkie związki w świecie traktować tak, jak gdyby one wyływały z koniecznej przyczyny” (Kant 1957: A 619/B647). Sellars mówi o „regulatywnym ideale” uwzględniającym wyniki badań naukowych (Sellars 1963b: 185) i zaznacza, że zdeterminowany obraz świata jest „regulatywnym ideałem” gdy chodzi o badania prowadzone w poszczególnych naukach (Sellars 1962: 601).

systemu nerwowego. Uwzględniając wzajemne zależności pomiędzy argumentacją filozoficzną i naukową, należy zarazem wskazać na ich wyraźne rozgraniczenie. Dotychczasowe naukowe teorie świadomości nie potrafią bowiem zastąpić filozoficznych analiz dotyczących struktury i własności świadomości.

Należy przypomnieć, że metafizyczne analizy problemu świadomości przebiegają w ramach debaty realizm–antyrealizm. Według naukowego realizmu przedmioty poznania istnieją niezależnie od umysłu i postępowania naukowca, a teorie naukowe są prawdziwe w ramach obiektywnego opisu świata. Tak rozumiany realizm jest, z jednej strony, tezą metafizyczną, stwierdzającą niezależne od podmiotu istnienie określonych przedmiotów. Z drugiej strony jest on tezą epistemologiczną, mówiącą, że możemy wiedzieć, jakie przedmioty istnieją i możemy odkryć teorie i prawa rządzące tymi przedmiotami.

Reprezentowany we współczesnej analitycznej filozofii świadomości naukowy realizm jest konfrontowany z dwoma kierunkami: ze sceptycyzmem oraz z antyrealizmem. Przedstawicielem tego pierwszego jest C. McGinn, przedstawicielami drugiego są H. Putnam i B. van Fraassen (por. Tennant 1995: 14–18). Kontrowersje pomiędzy realizmem a antyrealizmem doprowadziły, w ramach tego ostatniego, do powstania dwóch kierunków: „wewnętrznego realizmu” (*internal realism*) Putnama oraz „konstruktywnego empiryzmu” (*constructive empiricism*) van Fraassena¹⁷.

Putnam od realizmu przeszedł do jego krytyki, zbliżając się zarazem do pozycji pragmatystycznych¹⁸. Odrzucając (jak sam go nazywa) „metafizyczny realizm” połączony z „God’s eye view”, Putnam proponuje ujęcie, w którym prawda jest relatywna do języka albo do schematu pojęciowego. Dopuszcza tym samym, że twierdzenia naukowe są prawdziwe jedynie w odpowiednim zakresie i odnoszą się do konkretnej dziedziny. Putnam neguje jednak, iż twierdzenia te mówią nam pełną prawdę o rzeczywistości, albo że pełna prawda jest do wyrażenia. W innych sprawozdaniach o świecie mogą istnieć inne prawdy, a każde z tych sprawozdań może być wiarygodne.

Stawiając tezę zwaną „konstruktywnym empiryzmem” van Fraassen stwierdza, że realista poszukuje prawdy jako celu i gdy akceptuje jakąś teorię, jest ona uznana jako prawdziwa. W przeciwieństwie do tego „konstruk-

¹⁷ Por. Putnam 1981: xi, 54; 1998: 431–433; Railton 1995: 436–438. Współczesny termin „antyrealizm” tylko w przybliżeniu odpowiada terminowi „idealizm” i ma sens zbliżony do agnostycyzmu i relatywizmu (por. Judycki 2001: 118–123).

¹⁸ W pragmatyzmie Putnam widzi formę realizmu, mogącą mediować w debacie: realizm–antyrealizm (por. Putnam 1999: 39–42; Żegleń 2001: 7–56).

tywny empirycyzm” przyjmuje empiryczną adekwatność (a nie prawdę) jako cel nauki i gdy akceptuje on jakąś teorię, to uznaje ją jako empirycznie adekwatną. Nie ma podstaw, aby wierzyć, że teorie naukowe są prawdziwe, wystarczy, że są empirycznie adekwatne (tzn. zgodne z danymi obserwacyjnymi) [por. v. Fraassen 1980: 18].

Antyrealistyczna tradycja filozoficzna dostrzega „trudny” problem świadomości nie w pytaniu Chalmersa: Jak świadomość powstaje z materii i w jakiej relacji stoi ona do materii? ale w pytaniu: Czy istnieje świat niezależny od świadomości? U podstaw tego pytania leży Kantowska teza, że nie posiadamy jakiegoś niezależnego od umysłu dojścia do tego rodzaju „przypuszczalnego” świata. Tym samym każda oczywistość, która mogłaby popierać tezę o istnieniu takiego świata, zakłada od samego początku obecność ludzkiego umysłu. W tym sensie istnienie ludzkiego umysłu wyprzedza istnienie świata zewnętrznego¹⁹. Kwestie związane z antyrealistycznym rozumieniem świadomości nie będą nas w tej pracy zajmowały. Należy im poświęcić osobne opracowanie.

Reasumując można powiedzieć: trudno sobie wyobrazić rozwój filozofii umysłu oraz psychologii empirycznej przy równoczesnym ignorowaniu metafizyki (rozumianej we wspomnianym słabym sensie). Metody i osiągnięcia współczesnych nauk szczegółowych są istotne w badaniach nad świadomością, nie są one jednak substytutem tez metafizyczno-filozoficznych.

1.3.3. Holizm ontologiczny i metodologiczny

Doświadczenie takiej czy innej struktury określanej mianem całości jest nam znane zarówno z codzienności, jak i z różnych dziedzin nauki (kosmos jest przez naukę traktowany jako całość, związki zachodzące w naturze traktujemy jako całościowe). Na podstawie własnych przeżyć odkrywamy jedność i całościowy charakter otaczającej nas rzeczywistości, oraz nasze miejsce w tej całości²⁰. Tak rozumiana całość rzeczywistości, ze względu na wspomnianą przeżyciową niepodzielność, wydaje się być czymś innym, ani-

¹⁹ Argumenty za teoriopoznaczym idealizmem znajdujemy u J.G. Fichtego: „Kannst du sagen: ich bin mir äußerer Gegenstände bewußt? [...] Keineswegs, wenn ich es genau nehme: denn das Sehen und Fühlen usw., womit ich die Dinge umfasse, ist [...] dasjenige, dessen ich mir am ersten und unmittelbarsten bewußt bin. Der Strenge nach könnte ich nur sagen: ich bin mir meines Sehens oder Fühlens der Dinge bewußt”. Cyt. wg: v. Kutschera 1982: 205.

²⁰ Pomijamy tutaj chorobowe przypadki rozszczepienia osobowości (*multiple personality disorder*, MPD), w których numeryczna identyczność zjawiskowego świata lub samego podmiotu zdaje się być zniesiona.

żeli tylko jednością w sensie numerycznej, podmiotowej jedności. Pozostaje pytanie, czy możemy owo holistyczne przeżywanie rzeczywistości przedstawić za pomocą metod eksperymentalnych?

Z jednej strony, nie potrafimy, na przykład wolitywnie, podzielić swojej całościowej przestrzeni zjawiskowej ani też przeżywania naszej tożsamości jako podmiotu przeżyć. Treści naszych przeżyć wskazują na dużą selektywność. Z drugiej strony, otaczający nas świat, który wydaje nam się być bezpośrednio danym, jest zawsze podzielony, „zatomizowany” na wielość przedmiotów. Jedne z nich są widoczne, inne zasłonięte, jedne są bliżej, drugie dalej. Kierując naszą uwagę na otoczenie, możemy te przedmioty odróżnić od tła, możemy je sobie wyobrażać w subiektywnym przeżyciu jako oddzielne przedmioty. Jest to możliwe, chociaż, jak wspomnieliśmy, nie potrafimy podzielić całości przeżywanej rzeczywistości aktami naszej woli. Rozumiany w ten sposób całościowy (holistyczny) charakter rzeczywistości, zdaje się być czymś mocniejszym aniżeli czysto numeryczna identyczność. Nasze świadome przeżywanie świata i nas samych cechuje się jednością, której nie potrafimy ująć w kategoriach potocznego obrazu świata. Często mówimy, że całość ta jest nam dana w jakimś „teraz” i „tutaj”. Holistyczna różnorodność przeżyć staje się cechującym się jednością światem ze względu na owo „teraz” i „tutaj”.

Przeżywanie świata jako całości wprowadza nas w kwestie związane z „holizmem”, pod którym najczęściej rozumie się tezę, że całość jest czymś więcej aniżeli sumą jej części. U podstaw holizmu leży idea, że elementy jakiegoś systemu uzyskują swoje znaczenie dzięki ich wzajemnym powiązaniom. W filozofii umysłu holizm jest dyskutowany w związku z: (i) treściami postaw propozycjonalnych oraz (ii) znaczeniem konstrukcji językowych. Holista znaczeniowy przyjmuje, że znaczenie danego zdania w języku zależy od jego połączeń z innymi zdaniami. Holista treściowy przyjmuje, że treść czyjegoś przekonania zależy od jego połączenia z innymi przekonaniem (por. Guttenplan 1995a: 347). Przeciwnieństwem holizmu jest „molekularyzm” (*molecularism*) i „atomizm” (*atomism*). Molekularyzm przyjmuje, że czyjeś przekonanie otrzymuje treść na podstawie jego relacji do jakiejś grupy innych przekonań²¹. Atomizm twierdzi, że zdania i przekonania mają znaczenie i treść niezależnie od ich relacji do innych zdań lub przekonań (a tym samym mają znaczenie niezależnie od teorii, w których występują).

Tezy holistyczne odwołują się do twierdzenia W.V.O. Quine’a, że zdania o świecie nie są potwierdzane jako jednostkowe, lecz tylko w powiązaniu

²¹ Jako przykład Block podaje: „[...] the meaning of «bachelor» might be said by a molecularist to be man who has never married” (Block 1998a: 2).

z teoriami, do których należą. Nie możemy zrozumieć potocznych lub naukowych twierdzeń o rzeczywistości bez zrozumienia „sieci”, w skład której owe twierdzenia wchodzi²². Takie rozumienie holizmu, gdy chodzi o świadomość, niesie ze sobą tę trudność, że uniemożliwia dokonywanie uogólnień w psychologii i w filozofii. Jeżeli bowiem treść jakiegoś stanu zależy od wszystkich innych stanów, to staje się bardzo mało prawdopodobne, że kiedykolwiek dwa podmioty będą doświadczać stanów o tych samych treściach. Ponadto, holizm trudno pogodzić z naszym potocznym rozumieniem uzasadniania, gdyż przyjęcie jakiegokolwiek zdania już wpłynie na zdanie, które z niego wyprowadzimy²³.

Tezy holistyczne dotyczące świadomości, mogą być natury ontologicznej i metodologicznej.

Według holizmu ontologicznego niektóre przedmioty lub systemy (np. świadomy mózg) nie są jedynie złożone z podstawowych części (np. z neuronów), lecz posiadają cechy, które nie są określone poprzez fizyczne cechy ich części (neuronów). Ponadto przedmioty te podlegają prawom, które nie są określone poprzez prawa determinujące strukturę i zachowania ich części podstawowych.

Zgodnie z tezą holizmu metodologicznego – zrozumienie i wyjaśnienie pewnego rodzaju złożonych systemów jest trudniejsze, gdy rozpoczniemy je od analizowania struktury i zachowań części tego systemu. Łatwiejsze wydaje się rozpoczęcie analiz od zasad określających zachowanie się całego systemu (w naszym przypadku oznacza to wyjście od holistycznie rozumianej świadomości). Takie postępowanie rzuci światło na naturę i funkcjonowanie badanego systemu, którym jest świadomy umysł.

W dalszej części opracowania będzie nas interesował holizm metodologiczny, który przyjmujemy w celu dokonania właściwego opisu i wyjaśnienia świadomości zjawiskowej.

1.3.4. Holistyczne rozumienie świadomości

Na dzień dzisiejszy brak teoretycznego ujęcia subiektywnej świadomości i jej miejsca w świecie rzeczy materialnych. Aby jakaś cecha mogła spełniać swoją rolę w teorii świadomości, musi zostać zidentyfikowana (tzn. musimy

²² „Die Gesamtheit unseres sogenannten Wissens oder Glaubens, angefangen bei den alltäglichen Fragen der Geographie oder der Geschichte bis hin zu den grundlegenden Gesetzen der Atomphysik oder sogar der reinen Mathematik und Logik, ist ein von Menschen geflochtenes Netz, das nur an seinen Rändern mit der Erfahrung in Berührung steht” (Quine 1979: 47).

ją albo opisać, albo wskazać). Stąd poszukiwanie opisów świadomości, a następnie nowej, wielowymiarowej teorii świadomości. W efekcie teoria ta będzie albo zakładać inne niż dotychczasowe kategorie ontologiczne, albo kategorie dzisiaj nam dostępne zostaną w odpowiedni sposób zmodyfikowane. W związku z tym należy zwrócić uwagę na dwie rzeczy:

A. Patrząc od strony metodologicznej dostrzegamy dwa podstawowe programy badawcze.

A.1. Możemy założyć, że świadomość jest z zasady odmienna od jakichkolwiek innych przyrodniczych „obiektów” badań, a tym samym wszelkie poszukiwania jej redukcyjnego wyjaśnienia, w terminach neurofizjologicznych lub w terminach fizykalnych (czyli w terminach nieodwołujących się do świadomości), są skazane na niepowodzenie. Wynikający stąd program badawczy traktuje świadomość jako ontologicznie podstawową cechę rzeczywistości.

A.2. Alternatywny, naturalistyczno-metafizyczny program wychodzi z założenia, mówiąc ogólnie, że świadome przeżycia, zachowania i działania w jakiś sposób odpowiadają (są skorelowane) zdarzeniom i stanom, które obserwujemy empirycznie. Świadomość próbujemy wtedy wyjaśnić odwołując się do tych samych albo podobnych praw, jakich używamy do wyjaśnienia innych zdarzeń²⁴. Wyjaśnienie świadomości prowadzi wówczas do nowego określenia dzisiejszego rozumienia świadomości i połączenia go z wyjaśnieniami świadomości proponowanymi przez nauki szczegółowe.

W rozdziale VII niniejszej pracy zostaje postawione pytanie o warunki, jakie powinna spełnić argumentacja za programem A.2. Za przyjęciem tej metodologii przemawia chociażby historyczno-filozoficzny rozwój spojrzenia na wyjaśnianie naukowe. Zgodnie z nim, wyjaśnianie tego rodzaju „obiektów” jak świadomość jest kompleksowe i może się zmieniać w zależności od różnych stopni analiz. Przyjęcie programu A.1. przemawia przeciwko wspomnianej metodologii i w zasadzie nie oferuje nic na jej miejsce.

B. Analizy świadomości są związane z badaniami świadomego podmiotu albo takie badania zakładają. Podmiotowi, jak jest on rozumiany w filozo-

²³ Funkcjonalistyczne teorie umysłu oraz zarzuty wobec holizmu przedstawia Block (1998a: 3–6).

²⁴ Z metafizycznego punktu widzenia nie jest konieczne, aby różne rodzaje stanów mentalnych podpadały pod tę samą teorię filozoficzną. Wprawdzie często wychodzi się z założenia, że musi istnieć jakaś jednolita teoria, dotycząca wszystkich typów stanów mentalnych, nikt jednak nie mówi, dlaczego tak ma być (np. dlaczego jakaś funkcjonalistyczna teoria nie mogłaby wyjaśniać stanów intencjonalnych, a teoria egzemplarycznej identyczności wyjaśniać stanów zjawiskowych).

fii analitycznej, należy poświęcić osobne opracowanie²⁵. Nie będziemy się zajmować relacjami pomiędzy podmiotem i jego stanami mentalnymi, traktując te ostatnie jako „wewnętrznie świadome” (poniżej wyjaśniamy, co rozumiemy pod tym pojęciem).

Przy analizach świadomości możemy wyjść z założenia, że podstawowym pojęciem w nauce o świadomości jest pojęcie prostego („elementarnego”) stanu zjawiskowego (*jak to jest być w stanie x*), a następnie próbować te stany „złożyć” w jedną całość. Możemy jednak, jak to proponują między innymi T. Bayne i D. Chalmers, wyjść od podstawowego, całościowego stanu świadomości (*basic total state of consciousness*), a następnie wyróżnić jego poszczególne (elementarne) stany. Całościowym stanem może być na przykład: *Jak to jest być świadomym podmiotem (powiedzmy w czasie t)*.

I.5 Stan *jak to jest być świadomym* odpowiada całościowemu (holistycznemu) stanowi świadomości zjawiskowej.

Stan ten obejmuje zazwyczaj przeżycia związane z postrzeganiem, ze stanami emocjonalnymi oraz poznawczymi. Przyjęcie twierdzenia I.5. tłumaczy zarazem, dlaczego tak istotna jest teza o jedności świadomości. Przeciwnym do holistycznego ujęcia świadomości jest ujęcie „atomistyczne”²⁶ (treści świadomości lub jakiegoś stanu mentalnego nie zależą od związku tego stanu z innymi stanami mentalnymi).

Proponowane przez nas holistyczne rozumienie świadomości należy odróżnić od ujęcia, które można nazwać „ściśle holistycznym”. W ujęciu „ściśle holistycznym” coś pojedynczego nie jest rozumiane jako część, która jest złożona razem z innymi częściami tworząc w ten sposób system, lecz jedynie jako człon jakiejś całości. Wspomniana całość pozwala się w pewien sposób rozłożyć, a tym samym określić²⁷. Jakiś całościowy kompleks nie jest

²⁵ Takie opracowanie jest w przygotowaniu. Charakterystyki podmiotu w „psychologii jako nauce o duszy, świadomości i zachowaniu” oraz w neobehawioryzmie można znaleźć w: Krzyżewski 1989: 39–78, 162–188. Opisy jaźni odwołujące się do *reflexive consciousness*, *interpersonal aspect of selfhood* oraz *executive function* podaje Baumeister (por. Baumeister 1998: 683–711).

²⁶ „Atomistyczne” rozumienie świadomości zakładają między innymi Bartels i Zeki: „świadomość składa się z wielu mikroświadomości”, a rolą mózgu nie jest generowanie cechującej się jednością świadomości, lecz łączenie poszczególnych świadomych przeżyć (Bartels 1998: 2327). Tezy o „disunity” (braku synchronicznej jedności świadomości) bronią O’Brien i Opie (por. O’Brien 1998: 386–387).

²⁷ Tezę taką Stephan przypisuje A. Meyer-Abichowi, którego zdaniem: „[...] das Einzelne nicht als ein Teil zu verstehen ist, der mit anderen Teilen zu einem «System» zusammengesetzt ist, sondern nur als Glied eines Ganzen, das sich in bestimmter Weise zer-

jednak nigdy czymś złożonym i dlatego nigdy nie posiada części lub elementów. Podporządkowane «momenty», które można odróżnić w strukturze lub funkcjonowaniu takiej całości, nazywa się jej członami i organami. W stosunku do członów i organów obowiązuje, że ich cechy jako członów lub organów pozwalają się wyprowadzić tylko z każdorazowej całości, której są częściami. Ścisłe holistycznemu ujęciu można przypisać tezę, którą często błędnie przypisuje się emergentystom: „całość jest większa od sumy swych części”²⁸.

Mówiąc o holistycznej świadomości mamy na myśli przede wszystkim jej – przyjętą ze względów metodologicznych – własność, gwarantującą korzyści wyjaśniające (eksplanacyjne). Przyjmujemy, że:

I.6 Niezależność holistycznie rozumianej świadomości jest zagwarantowana przez jej eksplanacyjną, a nie przez jej ontologiczną (przyczynową) nieredukowalność.

Tym samym, od strony metodologicznej, stany przeżyciowe podmiotu w danym czasie mogą być widziane jako części wynikające z całościowego stanu świadomości: *jak to jest być świadomym w tymże czasie*²⁹. Całościowy stan jest przez nas traktowany jako podstawowy, ale nie jako pozbawiony cech szczególnych. Tym samym ma on złożoną strukturę.

legen und dadurch bestimmen läßt: «Ein ganzheitlicher Komplex aber ist niemals ein zusammengesetztes Ding und hat infolgedessen niemals Teile oder Elemente» (Stephan 1999: 153).

²⁸ Analizując różnice pomiędzy emergentyzmem a holizmem Stephan zaznacza, że żaden ze znanych emergentystów nie odwołuje się do wspomnianej tezy (por. Stephan 1999: 148–155). Zrównanie holizmu z emergentyzmem spotykamy przykładowo u E. Nagla (1970: 319). Holizm przyjmuje tezy o istnieniu systemowych, nieredukowalnych, nieprzewidywalnych cech, które to tezy odpowiadają tezom emergentyzmu. Oprócz tych tez holizm przyjmuje: tezę o tzw. *downward causation*, tezę o jedności wszystkiego z wszystkim (aby poznać coś pojedynczego, należy stale mieć na uwadze całość), tezę o „odwrotnej redukcji” (np. prawa fizyki mogą być wyprowadzane z praw biologii). Żadna z tych tez nie jest dla emergentyzmu konstytutywna oraz nie wszystkie są typowe dla poszczególnych rodzajów holizmu (por. Stephan 1999: 151–152).

²⁹ „In nature, it may be that the most basic sort of conscious state is the total phenomenal state, or the phenomenal field, or even the phenomenal world. These total states are basic, but they are not featureless: they come with a complex structure from which one can differentiate many aspects. (As an analogy, one can think of a quantum wavefunction, which is a basic state in physics but which nevertheless has a complex structure)” (Bayne i Chalmers 2003: 27). Powołując się na intersensoryczną integrację, T. Bayne pisze: „What we see influences what we feel and vice versa” (Bayne 2000: 253). Swe-go rodzaju emergentystyczno-holistyczne ujęcie świadomości proponuje także J. Searle (2000: 574–575).

Można w tym miejscu wprowadzić dwie ogólne definicje:

(1) Podmiot posiada świadomy stan sensoryczny, gdy stan ten jest przedstawieniem czegoś w świecie lub w samym podmiocie, i jest to coś takiego jak: *jak to jest* mieć ten stan.

(2) Podmiot jest w stanie samoświadomości, gdy stan ten jest pojęciowym stanem o własnych mentalnych przedstawieniach podmiotu, i jest coś takiego, jak: *jak to jest* być podmiotem świadomych stanów sensorycznych.

Holistyczny, czyli uwzględniający zjawiskowy poziom organizacji model świadomości, jest w tym sensie nierelacyjny, że nie musimy go bezpośrednio odnosić do podmiotu przeżyć. W relacyjnym rozumieniu świadomości przyjmuje się, że to, iż podmiot jest świadomy obiektu *x* oznacza, iż stoi on w jakiejś relacji do tego obiektu (pytaniem jest, o jaki rodzaj relacji i o jaki rodzaj obiektu tutaj chodzi)³⁰. W holistycznym rozumieniu świadomości nie traktujemy przeżyć jako świadomych dzięki relacji do podmiotu, lecz jako wewnętrznie świadome (tzn. przeżycia świadome nie mogą istnieć nie będąc świadomymi). Świadome przeżycia posiadają cechy – chociażby subiektywność – które występują tylko u nich.

Holistyczne rozumienie świadomości pozwala, przynajmniej do czasu, zrezygnować z mówienia o metafizyce świadomego podmiotu oraz o relacji: podmiot – przeżycia jakościowe. Postępowanie takie nie oznacza jednak redukcji podmiotu do jakichś innych „obektów” (np. do wiązki wrażeń)³¹.

Holistyczne rozumienie świadomości odpowiada naszemu potocznemu pojęciu subiektywności świadomości: gdy konkretny stan jest zjawiskowo świadomy, bycie w tym stanie zawiera jakiś rodzaj subiektywnego przeżycia. Takie rozumienie świadomości wskazuje na jakościowe cechy świadomości, na jej subiektywność, oraz na jej jedność. Należy zauważyć że:

I.7 Przyjęcie założenia o holistycznej naturze świadomości nie jest jej wyjaśnieniem, lecz opisem pewnego stopnia organizacji natury, którego na razie nie potrafimy naukowo wyjaśnić.

Sformułowanie „trudnego” problemu świadomości otrzymuje teraz postać: w potocznym sensie świadomość jest holistyczna i od strony zjawisko-

³⁰ Fenomenologiczne opisy relacji „Ja” do „pola świadomości” (owego dziwnego milieu, „[...] w którym łączą się przedmioty fizyczne, uczucia, fantazje, w którym wszystkie one podlegają wpływowi czasu [...]” – Piłat 1993: 12) podaje H. Ey (por. Piłat 1993: 94–100).

³¹ Podobne ujęcie samodzielnego „podmiotu przeżyć świadomych” spotykamy u R. Ingardena (1948: 830–836).

wej odznacza się jednością. W przeciwieństwie do tego mózg zdaje się nie odznaczać tego rodzaju jednością. Dlatego pytamy o rodzaj związku (identyczność, przyczynowość, paralelność) pomiędzy mózgiem, a odznaczającym się jednością zjawiskiem typu *jak to jest* być świadomym.

Do tak rozumianego holistycznego modelu świadomości odwołujemy się w całej pracy. Kończąc na obecnym etapie omawianie tego modelu, należy wskazać na dwie sprawy:

(1) Przyjmując holistyczny model świadomości podajemy w rozdziale VII prototeorię świadomości, uwzględniającą pojęciowe określenia neuronalnych korelatów świadomości oraz emergencyjnej zależności pomiędzy filozoficzno-psychologicznie rozumianą świadomością, a jej neurologicznym podłożem. Inne modele świadomości, to na przykład D. Dennetta Model Wielu Szkieł albo krytykowany przez niego Kartezjański Teatr (modele te zostaną przedstawione w rozdziale IV).

(2) Ideę holistycznego rozumienia świadomości znajdujemy u I. Kanta w jego „transcendentalnej jedności apercpepcji”. Kant rozróżnia świadomość empiryczną (apercepcja empiryczna) i świadomość transcendentalną (apercepcja czysta). Wspólna im jest funkcja syntetyzująca i jednocząca (por. Kant 1957: A 107, A 117). Świadomość empiryczna jest określana jako wewnętrzne spostrzeganie, jako refleksyjne poznanie własnych przedstawień. Jednak w strumieniu zjawisk wewnętrznych nie tworzy ona stałego i trwałego *Ja*. Według Kanta, wszystkie przedstawienia mają konieczny związek z możliwą świadomością empiryczną. Świadomość empiryczna obejmuje zmienne przedstawienia. U jej podstaw leży, jako zasada jedności wyprzedzająca wszelkie doświadczenie, świadomość transcendentalna: czysta i niezmienna świadomość samego siebie, czy też czyste przedstawienie *Ja*. Ten stan rzeczy Kant wyraża następująco:

„Apercepcja czysta jest samowiedzą (*Selbstbewußtsein*), która wytwarzając przedstawienie «Myślę» – które musi móc towarzyszyć wszelkim innym przedstawieniom i we wszelkiej świadomości jest jedno i to samo – nie może już sama być wyprowadzona z żadnego innego przedstawienia”³².

Każda świadomość empiryczna ma konieczne odniesienie do świadomości transcendentalnej, która tym samym warunkuje świadomość empiryczną

³² Kant 1957: B 132. Kantowskie „«Myślę» – które musi móc towarzyszyć wszelkim innym przedstawieniom” – przypomina Wittgensteina: „Ja pojawia się w filozofii przez to, że «świat jest moim światem»” (Wittgenstein 1997: 5.641), czyli jeśli jakieś przedstawienie ma być moim, to musi być możliwe połączenie go z jaźnią; wszystko, z czego zdaje sobie sprawę w percepcji lub działaniu, musi spełnić ten warunek.

i staje się podstawą konstytuowania przedmiotów. Świadomość, jako transcendentna jedność apercpcji, jest dla Kanta podstawowym warunkiem możliwości przedmiotów doświadczenia, a tym samymi i rzeczywistości w ogóle. Rosenthal zaznacza, że sama świadomość naszych różnych stanów mentalnych nie wystarczy (według Kanta), aby wyjaśnić sens, w jaki rozumiemy siebie samych, jako cechujących się – w naszym rozumieniu holistyczną – jednością podmiotów³³.

1.3.5. Debata: eksternalizm–internalizm

Czy przedstawiając to, co ktoś przeżywa, musimy się istotnie odnosić do obiektów z otoczenia przeżywającego? A jeżeli tak, to w jakiej relacji obiekty te stoją do świadomych przeżyć podmiotu? Czy umysł kończy się tam, gdzie mózg, czyli wewnątrz czaszki? Czy świadomość można ograniczyć do objętości mózgu, czy też przy jej opisie musimy uwzględniać fizyczne i społeczne otoczenie podmiotu? Próby odpowiedzi na te pytania włączają nas w debatę: internalizm–eksternalizm.

Eksternalista (albo inaczej „antyindywidualista”) twierdzi, że: (i) stany mentalne nie superwenują tylko na wewnętrznych, fizycznych stanach podmiotu, (ii) przy opisie i wyjaśnianiu stanów mentalnych należy uwzględnić zdarzenia z jego otoczenia.

Internalista („indywidualista”), mówiąc ogólnie, nie neguje, że otoczenie może przyczynowo wpływać na podmiot i na posiadanie przez niego jakiejś cechy, ale sama obecność tej cechy jest wewnętrznym, fizycznym stanem podmiotu³⁴. Internalista-fizyklista mówi wtedy o wewnętrznych stanach mózgu podmiotu, internalista-dualista o wewnętrznych stanach umysłu podmiotu. Internalistyczne (najpierw dualistyczne a później materialistyczne) spojrzenie na umysł było powszechne od czasów Kartezjusza do połowy XX wieku. Wzrost zainteresowania eksternalizmem spowodowały prace H. Putnama i T. Burge’a.

Najbardziej znana forma eksternalizmu dotyczy tzw. postaw propozycjonalnych (*propositional attitudes*) podmiotu twierdząc, że nie mogą być one opisane bez odniesienia się do obiektów znajdujących się w jego otoczeniu (eksternalista przyjmuje, że treści postaw propozycjonalnych są „szerokie”). Internalista twierdzi, że treści postaw propozycjonalnych są „wąskie”. Treść

³³ „Our sense that all our mental states occur in a single subject derives rather, from our «adding one representation to the other and being conscious of their synthesis» (B 133)” (Rosenthal 2002: 248).

³⁴ Na przykład promieniowanie radioaktywne może wywołać raka, ale bycie chorym na niego jest wewnętrznym stanem organizmu.

T jest „wąska” wtedy, gdy wszystkie możliwe mikrofizykalne duplikaty danego podmiotu, z jego postawą (przekonaniami, pragnieniami, ...) wobec treści T, posiadają także jako takie tę samą postawę (przekonania, pragnienia, ...). Wąska treść nie musi być w sposób konieczny identyfikowana z „treścią przeżyciową”.

Internalista uważa, że postawy propozycyjalne (a szczególnie „przekonania”), mogą być opisywane bez odniesienia do obiektów w otoczeniu podmiotu, gdyż treści własnych umysłów są podmiotom znane bezpośrednio³⁵. Gdy chodzi o treści przeżyć, internalista może się odwołać do pierwszoosobowego autorytetu. Jakiś intencjonalny stan mentalny jest stanem psychologicznego rodzaju, posiadającym określoną treść. Na przykład bycie przekonany, że pada deszcz, i nadzieja, że spadnie deszcz – są różnymi intencjonalnymi stanami z tą samą treścią. Być przekonany, że pada deszcz, i być przekonany, że świeci słońce, są przekonaniami, choć o różnej treści (nadal jednak mówimy tutaj o treściach postaw propozycyjalnych).

Mówiąc w niniejszej pracy o treściach przeżyciowych, wychodzimy z internalistycznego założenia³⁶:

I.8 Eksternalizm, dotyczący postaw propozycyjalnych, jest do pogodzenia z internalizmem, dotyczącym treści przeżyć jakościowych.

Twierdzenie I.8 ma uzasadnienie epistemologiczne: często możemy być w błędzie, gdy chodzi o nasze myśli przekonania (postawy propozycyjalne), rzadziej jesteśmy w błędzie względem naszych stanów jakościowych. Omyłność co do postaw propozycyjalnych jest do wytłumaczenia tym, że zawierają one odniesienie do obiektów zewnętrznych. Nasze zaufanie, dotyczące stanów jakościowych, jest związane z ich przeżywaniem (przeżywanie służy do ich identyfikacji).

³⁵ Mogę być w błędzie co do istnienia, powiedzmy ptaka, ale jak mogę być w błędzie co do mojego przekonania, że ten ptak istnieje (por. Guttenplan 1995: 289–290). Guttenplan omawia zaproponowany przez Putnama eksperyment myślowy z Twin Earth, opowiadający się za tak zwanym semantycznym eksternalizmem (argument ten omawia także Lau [por. 2002: 2–4]). Eksternalistyczną teorię świadomości podaje Tye (por. jego PANIC-theory, w: Tye 1996: 137–150), mówiąc, że przeżycia, podobnie do innych stanów intencjonalnych, posiadają intencjonalną treść. Eksternalistyczne i internalistyczne ujęcie psychologii potocznej przedstawia Ravenscroft (1997: 3–5). Według neurofenomenologa Thompsona „[...] individual human mind is not confined within the head, but extends throughout the living body and includes the world beyond the biological membrane of the organism, especially the interpersonal social world” (Thompson 2001: 2; por. s. 3–4).

³⁶ Przeciwniej tezy broni Tye, „phenomenology ain’t in the head” (1996: 151).

Klasyycznym już dzisiaj argumentem za tak zwanym semantycznym eksternalizmem jest zaproponowany przez Putnama eksperyment myślowy zwany „Twin Earth”³⁷. Wyobraźmy sobie, że w 1750 r. (czyli w czasach, w których nie znano jeszcze chemicznego składu wody) istnieje „Bliźniacza Ziemia” (B-Ziemia), która jest dokładnie podobna do naszej, z wyjątkiem, że zamiast wody (H_2O) posiada ona substancję o innym składzie chemicznym (XYZ). Makrocechy XYZ (smak, wygląd, ...) są takie same jak H_2O . Nikt na Ziemi ani na B-Ziemi nie potrafi odróżnić pomiędzy wodą a XYZ. Zamieszkujący Ziemię Piotr używając słowa „woda” będzie odnosił się do H_2O , a nie do XYZ. Oczywiście nie wie on, że woda to H_2O (zdaniem eksternalisty nie powinno go to powstrzymywać od odnoszenia się do H_2O , gdy używa terminu „woda”). Gdy natomiast wskazałby on na XYZ mówiąc „woda”, to wypowiedziałby zdanie fałszywe. Podobnie Piotr-bliźniak używając terminu „woda” będzie odnosił się do XYZ, a nie do H_2O .

Argument Putnama może być rozszerzony na treści stanów mentalnych. Piotr mówiąc: „woda gasi pragnienie” wyraża swoje przekonanie, że woda gasi pragnienie, które jest prawdziwe, tylko wtedy, gdy H_2O gasi pragnienie. Eksternalista proponuje nam następnie przenieść się na B-Ziemię, na której Piotr-bliźniak spotkał tylko bliźniaczą-wodę, a nigdy nie napotykał ziemskiej-wody ani o niej nie słyszał. Nasza intuicja powie nam wtedy, że Piotr-bliźniak nie wierzy, że woda gasi pragnienie. Gdy powie „woda gasi pragnienie”, to wypowiedział przekonanie, że bliźniacza-woda gasi pragnienie. Widzimy tutaj identyczność stanów psychologicznych bez identyczności stwierdzeń, stąd Putnam wyciąga wniosek, że znaczenie nie znajduje się „w głowie”. Tym samym okazuje się, że niektóre przekonania nie są superweniencyjne wobec wewnętrznych stanów podmiotu, a więc eksternalizm jest prawdziwy.

Argument Putnama pokazuje, że niektóre przekonania zawierające naturalny rodzaj pojęć („woda”) zależą od tożsamości pewnych chemicznych substancji w naszym otoczeniu. Rozumowanie Putnama opiera się na koncepcji rodzajów naturalnych i twardego desygnowania. Odwołując się do podobnych argumentów, T. Burge rozwinął wersję eksternalizmu zwaną

³⁷ Por. Putnam 1975a: 223–227. Putnam odrzuca założenie Fregego, że sens (*sense*) wyrażenia (który zawiera tak zwany modus prezentacji) determinuje znaczenie (*reference*) tego wyrażenia (czyli to, do jakiego obiektu się to wyrażenie odnosi). Putnam twierdzi, że „meanings» just ain't in the head!”, lecz są one zdeterminowane przez fizykalne otoczenie podmiotu (Putnam 1975a: 227). Przeciwniej tezy broni Searle: „Meines Erachtens sind Bedeutungen [*meanings* – J.B.] – jedenfalls in dem Sinn, der hier von Belang ist – genau im Kopf (sie könnten nirgendwo anders sein), und Putnams Argumente zeigen nichts Gegenteiliges” (Searle 1991: 250; por. s. 247–287).

„społecznym eksternalizmem”, w którym ważną rolę w determinowaniu treści niektórych przeżyć i myśli odgrywają instytucje społeczne³⁸. Krytycy eksternalizmu powątpiewają w to, że eksperymenty myślowe uzasadniają jego tezy i wskazują na konieczność badań empirycznych w tej sprawie. Jedni krytycy eksternalizmu społecznego nie zgadzają się z intuicją, że identyczni fizycznie osobnicy w odmiennych warunkach będą mieć różne przekonania. Inni twierdzą, że być może XYZ jest rodzajem wody, w zależności od tego, jak wyglądają szczegóły eksperymentu proponowanego przez Putnama. Klasyyczny epistemologiczny argument przeciwko eksternalizmowi wychodzi z założenia, że posiadamy specjalny dostęp do naszych postaw propozycyjalnych (nasze sądy o nich nie mogą być „błędne”).

Różniącą się od proponowanej przez Putnama wersji eksternalizmu, rozwijają A. Clark i D. Chalmers, nazywając go „aktywnym eksternalizmem” (por. Clark i Chalmers 1998: 9–18; por. Lau 2002: 9). W Putnama wersji eksternalizmu Piotr jest przekonany, że woda gasi pragnienie i jego bliźniak jest również przekonany, że bliźniacza woda gasi pragnienie. Zewnętrzne własności „wody”, odpowiedzialne za tę różnicę w przekonaniach, wynikają z pochodzenia Ziemi i B-Ziemi i leżą jakby na drugim końcu przyczynowego łańcucha. Natomiast własności dostrzegane tu-i-teraz na Ziemi nie odgrywają istotnego znaczenia w proponowanym przez Putnama eksperymencie. Gdyby Piotr nagle został przeniesiony na B-Ziemię, jego przekonania (ze względu na jego historię i biografię) dotyczyłyby standardowej wody. W takim przypadku zewnętrzne własności są pasywne, nie odgrywają one żadnej roli w kształtowaniu się procesu poznawczego, zachodzącego tu-i-teraz.

Clark i Chalmers proponują rozważyć eksperyment myślowy, w którym znaczące, zewnętrzne własności są aktywne i odgrywają istotną rolę poznaniu zachodzącym tu-i-teraz. „Ponieważ są one powiązane z ludzkim organizmem, mają one bezpośredni wpływ na organizm i na jego zachowania”³⁹.

Wyobraźmy sobie osobnika Otto, który z powodu choroby Alzheimera ma bardzo słabą pamięć, dlatego nosi z sobą notebook i zapisuje w nim wszystko, czego się dowiaduje. Gdy potrzebuje jakiejś informacji, zagląda po prostu do notebooka. Wiele przekonań Otta jest „ucieleśnionych” w jego notebooku. Tak jak każdy z nas, gdy chce zwiedzić muzeum przy 53. ulicy, naj-

³⁸ Por. Burge 1979: 80–83. Jakieś indywiduum, którego neurologiczne wyposażenie pozostaje wciąż to samo, może mieć różne myśli, jeżeli jest umieszczone w odmiennych społecznych otoczeniach. W zależności od aktualnego otoczenia jego myśli będą miały inne treści.

³⁹ Clark i Chalmers 1998: 8. Rozbudowane wersje eksternalizmu spotykamy u niektórych przedstawicieli neurofenomenologii (por. rozdział VI).

pierw „konsultuje” się ze swoją pamięcią, gdzie to muzeum jest, tak Otto „zagląda” na ekran swojego notebooka. Notebook odgrywa dla Otta podobną rolę, jaką pamięć odgrywa dla nas. Otto stale używa swojego notebooka, jest on istotny dla jego funkcjonowania w różnych kontekstach, tak jak dla nas istotna jest pamięć.

Według Clarka i Chalmersa, takie faktory, jak łatwość przenoszenia (*portability*), dostępność (*accessibility*) i zaufanie (*trust*), determinują to, czy jakiś zewnętrzny przedmiot może być uznany za część naszego umysłu. Rola, jaką dana informacja odgrywa w naszym życiu, powoduje, że staje się ona przekonaniem i nie ma żadnej racji aby twierdzić, dlaczego rola ta może być odgrywana tylko od wewnątrz. Podobnie jak w eksperymencie z B-Ziemią przypadek Otta i człowieka posługującego się pamięcią dotyczą nie tylko różnicy w odnoszeniu się do warunków prawdy, ale także różnicy w samej dynamice poznania⁴⁰.

Na zakończenie tego punktu należy zauważyć, co następuje:

1. We wszystkich wersjach eksternalizmu przyjmuje się, że treści stanów mentalnych podmiotu zależą od cech otoczenia, które jest zewnętrzne w odniesieniu do procesów poznawczych podmiotu. W eksternalizmie Putnama nie tworzą one jakiejś części mentalnego procesu, zachodzącego w podmiocie. „Aktywny eksternalizm” stwierdza, że otoczenie może odgrywać aktywną i konstytutywną rolę w procesie poznania. Kwestią dyskusyjną jest, na ile „proteza”, jaką jest notebook, może być traktowana jako część systemu składającego się z przekonań podmiotu. Argument Clarka i Chalmersa pokazuje raczej, że można poszerzyć fizyczne granice podmiotu poprzez wprowadzenie dodatkowego *hardware*, albo inaczej, że procesy zewnętrzne można czasem traktować jako procesy wewnętrzne (gdy Otto zapisuje coś w swoim notebooku). Wspomniani autorzy muszą jednak wziąć pod uwagę, że (przynajmniej w jakimś sensie) procesy zewnętrzne zależą od wewnętrznych, świadomych procesów.

2. Kognitywiści mają tendencję do traktowania zjawisk mentalnych (takich jak emocje, akty woli), w terminach restryktywnych modeli, ograniczonych do „głowy podmiotu”. Idea pierwszoosobowego, jakościowego przeżycia nadal stanowi wyzwanie dla kognitywistyki. Stąd próby zbudowania argumentacji o wewnętrznym przeżyciowym umyśle i skierowania jej prze-

⁴⁰ „[...] the twins in the Putnam and Burge cases differ only in their relational content, but Otto and his twin can be seen to differ in their notional content, which is the sort of content that governs cognition. Notional content is generally internal to a cognitive system, but in this case the cognitive system is itself effectively extended to include the notebook” (Clark i Chalmers 1998: 7).

ciwko tendencjom redukcjonistycznym i eliminującym, obecnym w naukach o świadomości i w neurologii. Chcąc dokonać swoistej deinterioryzacji poznawczych cech umysłu (poprzez wprowadzenie notebooka), Clark i Chalmers reinterioryzują świadomość zjawiskową. Zabieg ten rodzi jednak pytania związane z I.8: Czy pierwszoosobowe cechy typu *jak to jest być* są wykluczone z ich projektu „extended mind”? Czy nasze pojęcie umysłu jest jakąś nadrzędną kategorią, jakimś polem cechującym się istotnie swoją jednością? Czy też istnieje zasadniczy podział pomiędzy tymi cechami umysłu, które nadają się do – i które przeciwstawiają się – deinterioryzacji? Czy w całościowym przedstawieniu umysłu spotkamy – w jego ucieleśnieniu i w jego interakcji z ciałem – jakieś nieeliminowalne, wewnętrzne, zjawiskowe jądro? Clark i Chalmers nie udzielają na tego typu pytania odpowiedzi. Ich „extended mind” nie zajmuje się kwestią świadomości zjawiskowej.

3. Propozycję Clarka i Chalmersa trudno jednoznacznie zakwalifikować do eksternalizmu lub internalizmu (ich wypowiedzi wspierają zarazem intuicje dotyczące internalizmu, jak i eksternalizmu, a tym samym popierają I.8). Interesujące nas zagadnienie dotyczące tu-i-teraz procesu poznawczego, które wspomniani autorzy rozwiązują na drodze „aktywnego eksternalizmu”, można przeciwstawić eksternalizmowi Putnama, uznającemu, że mentalne treści nie są jedynie w głowie. Aktywny eksternalizm zakłada, że stany świadome posiadają jakąś wewnętrzną treść (można by ją nazwać „przeżyciową intencjonalnością”). W tym sensie Clark i Chalmers byliby internalistami, co odpowiadałoby zarazem naszemu holistycznemu rozumieniu świadomości.

1.3.6. Eksperymenty myślowe

Przedstawiony w poprzednim punkcie rodzaj argumentacji wprowadza nas zarazem w sposób „rozumowania”, często obecny we współczesnych teoriach świadomości. Chodzi o tak zwane eksperymenty myślowe (*thought experiments, Gedankenexperimente*)⁴¹, które należy traktować podobnie, jak traktujemy opowiadanie o jaskini Platona, opowiadanie o Kartezjańskim demonie czy też Hobbesa spojrzenie na umowę społeczną. Eksperymenty myślowe wprowadzają nowe perspektywy do spojrzenia na świadomość oraz na teorię świadomości.

⁴¹ Według T. Kuhna: „Eksperyment myślowy jest jednym z podstawowych narzędzi intelektualnych, z których korzysta się podczas kryzysu i które pomocne są w przeprowadzeniu reformy pojęciowej” (Kuhn 1985: 367; por. s. 336–369). Analizy eksperymentów myślowych podaje T. Burge (1979: 77–87), a omówienie ich roli w rozwoju nauki można znaleźć w: Rescher 1991: 32–41.

Zbudowane przy użyciu eksperymentów myślowych rozumowania pokazują jedynie logiczne (w sensie „niesprzeczne”) lub metafizyczne (w sensie „wyobrażalne”) „możliwości” (tzn. „możliwości” w bardzo słabym sensie). Rozwijane w tych rozumowaniach kroki są „wyobrażalne” i za pomocą odpowiednich założeń (odnośnie do tego, co jest możliwe) można z owej wyobraźności uzyskać argument za logiczną lub metafizyczną możliwością. Oczywiście w naukach szczegółowych bardziej jesteśmy zainteresowani argumentami empirycznymi albo nomologicznymi (w sensie „zgodne z tym, co mówią znane prawa natury”) niż argumentami ukazującymi logiczną lub metafizyczną możliwość. Sama wyobraźność nie potrafi sfalsyfikować argumentów empirycznych. Poniżej podajemy kilka przykładów eksperymentów myślowych, które pojawiają się w dalszych rozdziałach niniejszej pracy.

W krytyce funkcjonalizmu oraz tezy o mocnej sztucznej inteligencji⁴² J. Searle odwołuje się do argumentu Chińskiego pokoju, konkludując, że ponieważ program komputerowy składa się jedynie z reguł operowania symbolami, dlatego nie może on w ten sam sposób rozumieć znaczenia symboli, w jaki czynią to świadome umysły.

W eksperymencie „odwróconego spektrum” (*inverted spectrum thought experiment*) mówi się o podmiocie, który nagle widzi swoje otoczenie w kolorach innych niż normalnie. Rzeczy, które obydwaj (ja i ów podmiot) nazywamy „zielonymi”, podmiot ten spostrzega teraz w taki sposób, w jaki ja spostrzegam rzeczy, które obydwaj nazywamy „czerwonymi”⁴³.

Z eksperymentem odwróconego spektrum związana jest argumentacja, odwołująca się do tak zwanych nieobecnych qualiów. Może istnieć system, zombi, którego intencjonalne stany mentalne są typowe dla podmiotu ludzkiego, a który jest pozbawiony świadomości zjawiskowej⁴⁴. Zombi może być

⁴² Argument Chińskiego pokoju omawia J. Kloch (1996: passim).

⁴³ Por. Dennett 1991: 389–391. Przykład „odwróconego spektrum” można znaleźć w: Manzotti 1998: 1.

⁴⁴ Przypadek tego rodzaju zombi omawia Chalmers: „I am concerned with phenomenal zombies, which are physically and functionally identical, but which lack experience” (Chalmers 1996: 95). Można odróżnić dziewięć przypadków zombie, wynikłych z kombinacji ze sobą z jednej strony behawioralnej, funkcjonalnej i fizycznej identyczności, a z drugiej strony logicznej, metafizycznej i naturalnej możliwości. Najsilniejszym przypadkiem zombi byłby przypadek naturalnej możliwości fizycznie identycznych zombie. Najsłabszą formą byłaby logiczna możliwość behawioralnie identycznych zombie. Problem rozumieniem funkcjonalnego zombi będzie leżał w rodzaju przyczynowej struktury, jakiej wymaga się dla tego duplikatu. Możemy przyjąć dwie ekstremalne alternatywy: (i) że musi on podwajać moją przyczynową strukturę, aż do poziomu atomów, (ii) że przyczynowa struktura dotyczy jakichś wyższych poziomów organizacji.

moim fizycznym duplikatem (posiada tę samą fizykochemiczną strukturę), funkcjonalnym duplikatem (wykonuje te same funkcje, chociaż ma inną strukturę fizyczną), albo intencjonalnym duplikatem (posiada te same co ja stany intencjonalne – czyli stany posiadające znaczenie). Opowiadając się za lub przeciw możliwości istnienia zombi filozofowie nie argumentują, czy jest on możliwy w naszym świecie („nomologiczna możliwość”), lecz czy jest on możliwy w jakimkolwiek świecie (chodzi o tzw. metafizyczną lub logiczną możliwość: zombi jest możliwy, jeżeli istnieje chociażby jeden możliwy świat, w którym byłby możliwy chociaż jeden zombi)⁴⁵.

Można przyjąć jedno z dwóch założeń: internalistyczne założenie, że odpowiedni duplikat funkcjonalny będzie także koniecznym duplikatem intencjonalnym, albo eksternalistyczne założenie: duplikat funkcjonalny nie jest koniecznym duplikatem intencjonalnym (ponieważ stany intencjonalne osoby nie są jedynie determinowane przez jej stany funkcjonalne, lecz także przez wpływy otoczenia). Mówimy wtedy, że stany intencjonalne nie są determinowane lokalnie, lecz holistycznie. W dalszych rozdziałach niniejszej pracy przyjmuje się internalistyczne założenie, że funkcjonalny duplikat jest także koniecznym intencjonalnym duplikatem.

Z eksperymentem mówiącym o zombi jest związany eksperyment myślowy z panią neurofizjolog Mary, która od urodzenia przebywa w czarno-białym pomieszczeniu i wie wszystko o fizyce, chemii i neurofizjologii ludzkiego systemu nerwowego oraz jego reakcjach na widzialne kolory. Jednak dopiero po opuszczeniu swojego pomieszczenia Mary może przeżywać widzenie kolorów. Pytaniem jest, czy dowie się ona wówczas czegoś nowego o kolorach (por. 5.3.2).

Dennetta eksperyment myślowy z Modelem Wielu Szkiców pomaga, jego zdaniem, wyjaśnić w jaki sposób kompleksowa aktywność mózgu tworzy różnorodność myśli lub innych aktywności mentalnych. Wewnętrzne przeżycia są wynikiem różnych „procesów edytoryjnych” zachodzących w mózgu. Model Wielu Szkiców przedstawiono dokładnie w 4.3.

Prawa nauk empirycznych (może za wyjątkiem fizyki kwantowej) nie odgrywają w eksperymentach myślowych zasadniczej roli, dlatego ekspery-

⁴⁵ Hipoteza o zombi nie mówi, że jego istnienie jest faktycznie możliwe, mówi ona jedynie, że taka replika jest wyobrażalna (*imaginable*), a tym samym jest metafizycznie lub logicznie możliwa (por. Tye 1996: 23; Chalmers 1996: 94–95). Pisząc o zombim Chalmers i inni nie podają dowodu jego niesprzeczności (można sobie chociażby pomyśleć, że człowiek jest niesprzeczny, zombi zaś – sprzeczny). Logiczna spójność takiego tworu jak zombi wymaga dowodu. W tym samym sensie wszystkie eksperymenty myślowe wymagają dowodu niesprzeczności.

menty te znajdują ograniczone zastosowanie w naukach szczegółowych. Eksperymenty myślowe przedstawiają sytuacje, które mogą zaistnieć, jeżeli zostanie zrealizowany odpowiedni tok wydarzeń. Eksplikują tym samym problem i ułatwiają – w sensie wittgensteinowskim – bardziej precyzyjne stawianie pytań. To z kolei umożliwia jaśniejsze zrozumienie problemu, chociaż nie dostarcza ani jego empirycznego, ani czysto teoretycznego rozwiązania.

Odwoływanie się do eksperymentów myślowych spotykamy także w innych naukach. Znanymi eksperymentami myślowymi w fizyce są: (a) tak zwany kot Schrödingera, wskazujący na niektóre kontraintuicyjne tezy mechaniki kwantowej, (b) Einsteina eksperyment myślowy z pędzącym „pociągami” i stałą prędkością światła.

Reasumując należy podkreślić trzy rzeczy.

1. Nieformalne dowody często pojawiają się w formie eksperymentów myślowych. Odwołują się one do wyobraźności tego, że gdy zajdzie jedna sytuacja, to musi się także pojawić druga jako konsekwencja pierwszej. Dowody formalne zakładają werbalno-symboliczne definicje i skupiają się jedynie na tych wynikach, które mogą być logicznie wydedukowane. Również w eksperymentach myślowych muszą być uwzględnione kwestie niesprzeczności konstruktów myślowych.

2. Tworząc teorie świadomości, możemy się odwołać zarówno do metod nauk empirycznych oraz uzyskanych przez nie wyników, jak i do eksperymentów myślowych. Gdy z tych czy innych względów nie potrafimy przeprowadzić jakichś eksperymentów lub obserwacji, pomagamy sobie wynajdywaniem eksperymentów myślowych. „Eksperymenty” te powinny gwarantować, że przyjmowane w nich zdarzenia są (albo mogą być) zgodne z dostępnymi obserwacjami z danej dziedziny nauk. W eksperymentach myślowych z odwróconym spektrum, z wielością szkiców oraz z neurofizjolog *Mary* chodzi o zgodność z wynikami badań neurofizjologicznych (współczesnych lub przyszłych). Dobrze zbudowane eksperymenty myślowe ukazują niejasne miejsca w danej teorii świadomości, zakładają one bowiem przedteoretyczne idee i eksplikują je. Zdaniem Dennetta, aby wyeliminować „wierę” w używanie przedteoretycznych czy intuicyjnych pojęć, należy się odwoływać właśnie do eksperymentów myślowych (tzw. dźwigni intuicji), a nie do formalnych dowodów (por. Dennett 1988: 2). Dzięki eksperymentom myślowym można także wskazać na silne strony danej teorii świadomości.

3. Eksperymenty myślowe ułatwiają opisy i wyjaśnianie świadomości zjawiskowej. Ktoś może zaakceptować istnienie świadomości zjawiskowej bez akceptacji argumentów „odwróconego spektrum” i „nieobecnychQUALIÓW”, ponieważ potoczne podejście do świadomości zjawiskowej opiera się

na jej pierwszoosobowej znajomości. Akceptacja i skupienie się na świadomości zjawiskowej nie wyklucza z badań innych rodzajów świadomości. Można na przykład mówić o informacji jako o świadomej w takim sensie, że można jej używać jako założeń we wnioskowaniu, w formułowaniu planów, że może ona być dostępna dla sprawozdań. Informacja wyparta ze świadomości nie będzie używana we wnioskowaniach. Można także mówić o jakimś stanie jako świadomym, któremu towarzyszą myśli o treści mówiącej o tym, że ktoś posiada ten stan. Można również mówić o jakimś stanie jako o świadomym w sensie, że ktoś potrafi go monitorować (kontrolować).

1.4. Subiektywność świadomości w historii filozofii

Dyskutowany współcześnie, w ramach filozofii analitycznej, problem umysł–ciało i związany z nim problem świadomości pojawia się w historii filozofii wraz z R. Kartezjuszem i G. Leibnizem. Do czasów Kartezjusza nie kwestionowano istnienia prywatnego świata myśli i odczuć, odrębnego od świata zewnętrznego. Cechą kartezjańskiego sformułowania problemu umysł–ciało jest jego epistemologiczny punkt wyjścia: możemy wątpić w istnienie czegoś, co należy do świata zewnętrznego, zaś nie możemy wątpić w to, co należy do subiektywnego i prywatnego świata umysłu. Nie potrafię wątpić w moje własne istnienie, potrafię natomiast wątpić w istnienie mojego mózgu.

W znanym „eksperymentcie myślowym” o młynie Leibniz stwierdza, że sama obserwacja wewnętrznych fizycznych procesów w jakimś systemie (martwym czy ożywionym) nie dostarczy nam informacji o tym, czy system ten jest świadomy czy też nie (por. Leibniz 1995: § 17). Dusze działają według praw przyczyn celowych, a ciała działają według praw przyczyn sprawczych (czyli praw ruchu). Przyczyny sprawcze i przyczyny celowe pozostają ze sobą w systemie „wprzód ustanowionej harmonii”. System ten sprawia, zdaniem Leibniza, że ciała działają jak gdyby (założywszy tę niemożliwość) nie było dusz, a dusze działają jak gdyby nie było ciał, zaś jedno i drugie działają jak gdyby „wywierały nawzajem wpływ na siebie”⁴⁶. Pomiedzy mózgiem a umysłem brak jednak jakiegokolwiek interakcji, zjawiska mentalne i fizyczne odpowiadają sobie podobnie, jak odpowiadają sobie dwa zegary, wskazujące dokładnie tę samą godzinę. Z tezą o harmonii przedustawionej trudno jednak pogodzić to, że serie zdarzeń mentalnych, w odróżnieniu od zdarzeń neuronalnych, zdają się tworzyć ciągle, samozachowujące się se-

⁴⁶ Leibniz 1995: § 81. Bóg tak stworzył duszę i ciało, że harmonijnie współdziałały ze sobą.

kwencje. Eksperyment myślowy z młynem uzmysławia trudność w zrozumieniu tego, w jaki sposób system fizyczny, jakim jest mózg, może być równocześnie świadomym podmiotem. Istnieje „luka w wyjaśnieniu”, której nie pokonamy spacerując po mózgu powiększonym do rozmiarów młyna.

Leibniz odnosi się do mechanistycznej ontologii Kartezjusza, z jego interpretacją materialnej natury i duchowego umysłu. Zachodzące w umyśle zdarzenia mentalne – głównie zdarzenia intencjonalne – są czynnościami albo stanami niefizycznej (nierozciągłej) substancji. Podstawowe twierdzenie epistemologiczne kartezjanizmu brzmi: zjawiska umysłowe są dlatego bezpośrednio i nieomylnie prezentowane, bo zachodzą w naszej przestrzeni mentalnej i są świadome. Stąd w pokartezjańskiej filozofii umysłu „subiektywność świadomości” nie była traktowana jako „trudny” problem, lecz uchodziła za (kartezjańską) „jasną ideę”.

Stanowisko przeciwne do kartezjańskiego dualizmu prezentuje materialista T. Hobbes. Według niego, stany mentalne są wywoływane przez ruchy, a ruchy są materialno-cielesne i mogą przyczynować jedynie inne ruchy. Ponieważ zdarzenia mentalne są fizycznymi stanami substancji cielesnych, to kwestie związane ze świadomością nie przedstawiają dla Hobbesa jakiegoś „trudnego” problemu (por. Hobbes 1984: 11–12). Materialistyczna ontologia służy mu do tworzenia teorii państwa, będącego sztuczną świadomością, najbardziej uniwersalnym fizycznym wytworem ludzkim.

Idealistyczna odpowiedź na kartezjański problem zakłada, że świat materialny nie istnieje. Istnieje jedynie świat zdarzeń mentalnych. Nasze przeżycia, doświadczenia są takie, jakie są, nie istnieją żadne obiekty fizyczne będące ich przyczyną. Wszystko wokół nas wydaje się normalnie istnieć, chociaż w rzeczywistości nie ma nic poza naszymi przeżyciami. Ponieważ przedmioty zewnętrzne nie różnią się od naszych subiektywnych wrażeń zmysłowych, dlatego nie ma problemu z poznaniem przez nas naszych własnych subiektywnych wrażeń⁴⁷. Idealizm stwierdza, że materia jest po prostu konstrukcją umysłu i sama nie posiada niezależnego istnienia. Idealizm nie powinien więc mieć problemów związanych z wyjaśnieniem świadomości, ale z wyjaśnieniem tego, w jaki sposób obiekty fizyczne mogą być częścią rzeczywistości.

Według I. Kanta, istnieje metafizyka doświadczenia jako nauka o formach przedmiotów doświadczenia. Formy te podmiot narzuca w sposób ko-

⁴⁷ Taką odpowiedź na kartezjański dualizm proponuje G. Berkeley (por. jego formułę: „esse est percipi”). Berkeley neguje rozróżnienie na jakości pierwszorzędne i drugorzędne twierdząc, że obydwa rodzaje jakości są ideami istniejącymi w umyśle (por. Berkeley 1957: 27–30).

nieczny przedmiotom. Poznanie jest zawsze poznaniem zjawisk, poznaniem czegoś subiektywnego, a nie rzeczywistości samej w sobie. Zjawiska nie są jeszcze, w sensie Kantowskim, *doświadczeniami*. Doświadczenie powstaje dzięki pojęciowemu określeniu zjawiska. Jednak pojęcia (także czyste pojęcia intelektu) pozwalają się odnieść tylko do zjawisk, ich obszar zastosowania pokrywa się z obszarem możliwych doświadczeń. Używając języka Wittgensteina możemy dodać, że według Kanta, istnieją wprawdzie logiczne pojęcia, które nie odnoszą się do doświadczenia, ale analityczne zdania logiki nie mówią nic o świecie.

W krytyce paralogizmów psychologii racjonalnej Kant twierdzi, że jego poprzednicy traktowali posiadający idee umysł w ten sam sposób, w jaki jakaś rzecz posiada swoje cechy według tradycyjnej metafizyki. Podmiot („ja”) jest wtedy prostą, niezłożoną substancją, podobną do innych substancji w świecie. Kant argumentuje za tezą przeciwną⁴⁸. Aby przedstawić posiadane przez nas poznanie (zmysłowego świata), podmiot musi być charakteryzowany jako różny od przedmiotów świata, jako przypisujący prawa światu, jako niezaprzeczynowany i cechujący się spontanicznością. Musimy zacząć od omawiania relacji umysłu do przedmiotów i do świata, a nie oczekiwać, że dopiero mając odpowiednie fakty będziemy mogli tę relację uzasadnić.

W. James, podobnie do Kartezjusza, nie traktuje świadomości jako czegoś problematycznego. Jego zdaniem świat naszych przeżyć składa się z części obiektywnej i subiektywnej⁴⁹. Subiektywna część przeżyć jest wewnętrznym stanem, w którym zachodzi myślenie. Stan ten jest tym, co ogólnie określamy jako jaźń. „Ja” (*I*) i „mnie” (*me*) stanowią dwie części jaźni. „Ja” odnosi się do aktywnie poznającego aspektu jaźni, „mnie” odnosi się do poznawanego aspektu jaźni. Za teorią Jamesa opowiada się H. Bergson⁵⁰.

Od Kartezjusza do Jamesa subiektywność nie była postrzegana jako „trudna” do zrozumienia, chociaż dostrzegano zarówno istnienie perspektyw

⁴⁸ Według Kanta, subiektywna jedność świadomości mnie samego w różnych momentach czasowych nie wystarczy, aby ustanowić obiektywną tożsamość podmiotu lub duszy (por. Kant 1957: A 363–364).

⁴⁹ „Cały świat naszego doświadczenia składa się – i składał się zawsze – z dwóch części – przedmiotowej [*an objective – J.B.*] i podmiotowej [*a subjective part*]. [...] Część przedmiotowa – to suma tego wszystkiego, co w jakimś czasie możemy przemyśleć; podmiotowa zaś – to «stan» wewnętrzny, w którym całe myślenie odbywa się we własnej jaźni, która postawę tę przybiera” (James 2001: 383).

⁵⁰ „Zauważmy, w istocie, że subiektywnym nazywamy to, co zdaje nam się być w zupełności i całkowicie znane, obiektywnym, co jest nam znane w ten sposób, że wznagająca się ciągle wielość wrażeń mogłaby zastąpić pojęcie, które o nim mamy w danej chwili” (Bergson 1913: 59).

pierwso- i trzecioosobowej, jak i obecność zjawiskowej cechy stanów mentalnych. Sytuacja uległa zmianie wraz z pojawieniem się filozofii pozytywistycznej. Większość filozofów tej tradycji próbuje albo wyeliminować subiektywność, albo podać jej funkcjonalistyczne lub epifenomenalistyczne wyjaśnienie⁵¹. Anglojęzyczni filozofowie analityczni z drugiej połowy minionego stulecia (D.M. Armstrong, H. Feigl, J.J.C. Smart, W. Sellars) stali tym samym w opozycji do europejskiej filozofii kontynentalnej (konkretnie do fenomenologii E. Husserla, R. Ingardena czy do teorii pamięci H. Bergsona). We współczesnej filozofii analitycznej pytanie o subiektywność stanowi jeden z jej centralnych problemów. Przykładowo Searle zalicza subiektywność do własności strukturalnych świadomości, twierdząc, że to właśnie ona odpowiada za wszystkie filozoficzne kłopoty wiążące się ze świadomością⁵².

Chociaż liczni twórcy powstałych w ostatnim dziesięcioleciu teorii deklarują, że ich teorie wyjaśniają świadomość, to przy bliższej analizie okazuje się, że wiele z tych teorii ma braki merytoryczne (por. rozdziały III, IV i V niniejszej pracy). Obecna sytuacja, gdy chodzi o teorie świadomości, jest sytuacją „przedparadygmatyczną” (żadna ze wspomnianych teorii nie została zaakceptowana jako „paradygmat”) [por. Kuhn 1968: 93, 101]. To właśnie subiektywność blokuje nie tylko redukcję stanów umysłu do stanów mózgu, ale także redukcję potocznego widzenia świata do tego, co nazywamy jego naukowym obrazem. Problemy związane z ujęciem subiektywnego charakteru świadomości znajdują odzwierciedlenie w prowadzonych w ostatnich dziesięcioleciach psychologicznych i filozoficznych badaniach umysłu.

1.5. Podsumowanie

W rozdziale I, mającym charakter szeroko rozumianego wstępu do całej pracy, omówiliśmy podstawową terminologię używaną w filozofii świadomości oraz zakreslili obszar dalszych badań.

Rozróżnienie dwóch „pojęć” świadomości: świadomości zjawiskowej (fenomenalnej) i świadomości dostępu, nosi przede wszystkim charakter te-

⁵¹ Od tej linii odbiegają antropologiczne rozważania prowadzone przez K. Ajdukiewiczza, który uważa subiektywność za najistotniejszą cechę odróżniającą zjawiska psychiczne od fizycznych (por. Ajdukiewicz 1960: 319). Ajdukiewiczza koncepcję prywatności doznań omawia J. Woleński (2002: 8–10).

⁵² Por. Searle 1999: 180; por. s. 84, 134. Z jednej strony, Searle opowiada się za naturalizmem biologicznym (świadomość jest emergentną cechą procesów w mózgu), z drugiej podkreśla, że subiektywność jest kluczem do zrozumienia świadomości (por. Żegleń 2003: 251–254).

oretyczno-metodologiczny i nie oznacza dwóch od siebie empirycznie oddzielonych rodzajów świadomości.

Nawet krótkie i ograniczone przedstawienie psychologicznego i filozoficzno-metafizycznego spojrzenia na świadomość zjawiskową wskazuje, że spojrzenia te muszą być uwzględnione przy poszukiwaniu unifikującej teorii świadomości (por. rozdział VII, 7.6, 7.6.1). Przedstawiona w ostatnim rozdziale niniejszej pracy prototeoria świadomości zjawiskowej jest opartą na neurologii teorią filozoficzną, zakładającą różne płaszczyzny organizacji tego samego indywiduum i należąca do teorii emergencji.

Filozoficzne spojrzenie na świadomość umożliwia nam dwie rzeczy: (a) porównanie teoretycznych ujęć podstawowej cechy świadomości zjawiskowej – jej subiektywności. W niniejszej pracy podajemy nie tylko porównania wybranych, filozoficzno-materialistycznych teorii świadomości, ale także krytycznie te teorie oceniamy. Między innymi to stanowi o metateoretycznym charakterze naszego opracowania. Przedstawiona w rozdziale VII prototeoria również ten charakter posiada, chociażby przez to, że opiera się na naukowych badaniach tzw. neuronalnych korelatów świadomości. (b) Jeżeli stwierdzamy, że współczesne materialistyczne teorie świadomości nie potrafią wyjaśnić jej subiektywnych własności, to będziemy mogli w miarę szczegółowo pokazać, dlaczego tak jest.

Jedną z przyczyn nieadekwatności rozwiązań materialistycznych leży w tym, że rezygnują one z uwzględniania filozoficzno-psychologicznych opisów świadomości, rezygnując tym samym z pytania dotyczącego „trudnego” problemu świadomości i z próby jego teoretycznego ujęcia. Przedstawiona przez nas, unifikująca prototeoria świadomości taką próbę podejmuje.

W dalszych analizach świadomości zjawiskowej wychodzimy od jej całościowego (holistycznego) stanu *jak to jest być świadomym*, który obejmuje przeżycia podmiotu związane z postrzeganiem, ze stanami emocjonalnymi oraz poznawczymi. Przyjęcie takiego rozumienia świadomości tłumaczy zarazem, dlaczego tak istotna jest teza o jedności świadomości. Proponowane przez nas holistyczne rozumienie świadomości odróżniamy od ujęcia, które nazywamy „ściśle holistycznym”.

Przyjęcie założenia o holistycznie rozumianej świadomości ma na względzie przede wszystkim korzyści metodologiczne (por. I.6): stany świadomości zjawiskowej podmiotu w danym czasie t mogą być traktowane jako elementy całościowego stanu świadomości: *jak to jest być świadomym* w tymże czasie t . Całościowy stan świadomości – jak zobaczymy – ma złożoną strukturę. Tym samym (por. I.7), holistyczne rozumienie świadomości jest jedynie opisem pewnego stopnia organizacji, do którego wyjaśnienia potrzebujemy odpowiedniej teorii świadomości. Holistycznie rozumiana świadomość

pozwała także na nowe sformułowanie „trudnego” problemu świadomości: w psychologiczno-filozoficznym, a także w potocznym sensie świadomość jest holistyczna i od strony zjawiskowej odznacza się jednością. W przeciwieństwie do tego mózg – jak jest on rozumiany w teoriach neurologicznych – zdaje się nie odznaczać tego rodzaju jednością. Poprzez poszukiwanie neuronalnych korelatów staramy się odpowiedzieć na pytanie o rodzaj związku (identyczność, przyczynowość, paralelność, niezależność) pomiędzy stanami neuronalnymi [Φ_j], a cechującym się jednością zjawiskiem typu *jak to jest* być świadomym. Sama korelacja może być słaba. W 7.1 nadajemy jej silniejszą interpretację (jakiś [Φ_j] jest korelatem jakiegoś [Ψ_j], gdy [Φ_j] jest warunkiem wystarczającym dla [Ψ_j]) i opierając się na niej, tworzymy wspomnianą prototeorię świadomości.

ROZDZIAŁ DRUGI

FILOZOFICZNO-PSYCHOLOGICZNY OPIS SUBIEKTYWNOŚCI ŚWIADOMOŚCI

2.1. Wprowadzenie: cechy świadomości

Poniżej używamy terminów „filozoficzny opis” lub „filozoficzno-psychologiczny opis” w celu przedstawienia metodologiczno-holistycznie rozumianej świadomości i związanej z nią jedności stanów mentalnych typu: *jak to jest odczuwać X; jak to jest przeżywać Y*¹.

Najpierw zostaną omówione trzy cechy świadomości: subiektywność (2.1), jedność (2.2) i jej charakterystyki jakościowe (2.4). Pojęciowe oddzielenie owych cech nie oznacza, że w rzeczywistości nie są one ze sobą powiązane. Jedynie uzupełniająco wspominamy o innych cechach świadomości (2.4.1; 2.4.2).

Po części przyjmujemy charakterystyki świadomości zebrane przez R. Souśę, który twierdzi, że „wielkie zagadki”, jak „zagadka świadomości”, składają się z „małych sztuczek” (*little tricks*). Pisząc o problemie świadomości, nie mamy, jego zdaniem, do czynienia z jednym, lecz z wieloma różnorodnymi problemami. Jeżeli uda nam się je rozwiązać naturalistycznymi metodami, to możemy powiedzieć, że tym samym rozwiązaliśmy wszystkie

¹ Autorzy angloamerykańscy używają w tym miejscu wyrażenia „phenomenology of consciousness”, co jednak kojarzy się z „fenomenologicznym” (husserlowskim) rozumieniem świadomości. Dlatego rezygnujemy z używania tego zwrotu, zastępując go zwrotem „filozoficzno-psychologiczne ujęcie świadomości”.

ogólne „problemy świadomości” (por. Sousa 2002: 2). Sousa nie zamierza odpowiadać na pytanie: czym jest świadomość?, lecz, skupiając się na terminie „subiektywność”, podaje listę związanych z nią problemów. Nie interesuje go to, czy niektóre z tych problemów mogą zostać sprowadzone do innych, lecz twierdzi, że jeżeli będziemy umieli rozwiązać każdy z nich z osobna, to tym samym rozwiążemy problem subiektywności w tym sensie, że będziemy go mogli poddać analizom naukowym.

Proponowane przez Souś badania subiektywności są często spotykanym postępowaniem (podobnej metody, jak zobaczymy w rozdziale IV, używają D. Dennett i J. Searle). W naszych analizach subiektywności świadomości wychodzimy (w heurystycznym sensie) od proponowanej przez Souś metody i przedstawiamy najpierw kilka wybranych własności świadomości, wynikających z jej subiektywności. Tym samym precyzujemy terminy używane w dalszej części pracy (na przykład „punkt widzenia”, „subiektywne stany świadomości”).

W odróżnieniu od Souś uważamy, że owe własności podpadają pod założenie o całościowej i zjawiskowej jedności świadomości, które zostanie przedstawione w 2.2. Tym samym nie wychodzimy od elementarnych własności świadomości (które następnie należy złożyć razem), lecz od holistycznego rozumienia świadomości, której subiektywność jest wyrażona przez: *Jak to jest być świadomym*.

Przyjmujemy założenie, że można mówić o prototeoretycznym modelu holistycznego rozumienia świadomości wraz z jej poszczególnymi członkami, odnoszącymi się do *jak to jest* odczuwać dany stan? Model taki omawiamy w 2.3. Nie mamy żadnej gwarancji, iż tworzone przez Souś czy innych autorów „listy” aspektów i problemów subiektywności są zupełne. Wychodząc od całościowego ujęcia świadomości, staramy się tego problemu uniknąć.

2.1.1. Subiektywność ujęta ontologicznie i epistemologicznie

„Our conviction that we are somehow *directly acquainted* with special properties or features in our experience is one of the most powerful intuitions confronting anyone trying to develop a good theory of consciousness” (Dennett 1991: 359).

Badając subiektywność świadomości mamy na myśli jej stronę ontologiczną (jaka jest natura istnienia subiektywnej rzeczywistości) i epistemologiczną (jak możemy wiedzieć, że istnieje coś takiego jak subiektywność). Od strony ontologicznej możemy powiedzieć: „Moje” doznanie bólu jest „moim” w inny sposób aniżeli „moja” ręka jest „moją” ręką. Ból jest związany

inaczej z jego podmiotem aniżeli ręka². Ból nie istnieje gdzieś w obiektywnej czasoprzestrzeni. Trudno bowiem powiedzieć, że świadome przeżycia są w czasie i przestrzeni tak, jak są w niej przedmioty przestrzenne, ale także trudno powiedzieć, że są one gdzieś poza przestrzenią na wzór przedmiotów abstrakcyjnych, jakimi są chociażby liczby³.

Z ontologicznym rozumieniem subiektywności wiążą się problemy natury epistemologicznej (to znaczy dostępu do stanów mentalnych i przypisywania ich osobom trzecim). Muszę jakoś zdawać sobie sprawę z tego, że jestem świadomy. Mojemu doznaniu bólu przysługuje subiektywny sposób istnienia, który jest inaczej dostępny poznawczo mnie samemu (jako podmiotowi tego doznania), a inaczej jakiemuś trzeciemu obserwatorowi moich „bólowych” zachowań.

Do tak rozumianego ontologicznego i epistemologicznego wymiaru subiektywności odwołuje się J. Searle⁴. Należy zaznaczyć, że o ile – jego zdaniem – można mówić o epistemologicznej odrębności zjawisk umysłowych (w tym świadomości) i neuronalnych, to nie można mówić o ich ontologicznej odrębności.

Pisząc o subiektywności zjawisk świadomych Searle ma na myśli pewną kategorię ontologiczną, a nie tylko modalność epistemologiczną. Podając jako przykład doznanie bólu stwierdza, że owemu „obiektywnemu faktowi” przysługuje subiektywny sposób istnienia (por. Searle 2000: 568). Świadomość, według Searle’a, właśnie tym różni się od innych zjawisk biologicznych, że posiada subiektywną (albo pierwszoosobową) ontologię. Ontologiczna subiektywność świadomości nie powstrzymuje nas od przyjęcia jej epistemologicznej obiektywności (patrz niżej). Zatem zdaniem Searle’a, możemy zbudować obiektywną naukę o świadomości. Podobnie jak Kartezjusz,

² Por. Wittgenstein 2000: § 283. Searle twierdzi: „Ból musi być czymś doznaniem bólu i to w znacznie mocniejszym sensie niż ten, w którym, na przykład, noga musi być zawsze czyjąś nogą. Operacje przeszczepiania nóg są możliwe; nie można natomiast, w tym sensie, dokonać przeszczepu doznania bólu” (Searle 1999: 135). Tye zaznacza: „My pains, for example, are necessarily private to me. You could not feel any of my pains” (Tye 1996: 71).

³ Na różnicę pomiędzy „rozciągłością w przestrzeni” zjawisk fizjologicznych i brakiem takiej rozciągłości dla „objawów życia umysłowego” zwraca uwagę K. Twardowski (1997: 88-89).

⁴ Searle zakłada, że: „Świadome stany i procesy mentalne mają pewną swoistą własność, która nie przysługuje innym zjawiskom przyrody, własność subiektywności [...] mizernie efekty większości prac badawczych w dziedzinie filozofii umysłu i jałowość akademickiej psychologii w ciągu ostatnich pięćdziesięciu lat, [leżą] w uporczywej niezdolności do rozpoznania [...] że ontologia tego co mentalne, jest w sposób nieredukowalny ontologią z punktu widzenia pierwszej osoby” (Searle 1999: 134, 136).

Searle wiąże pierwszoosobową ontologię stanów mentalnych z epistemologią twierdząc, że pierwszoosobowy punkt widzenia jest „pierwotny” (por. Searle 1999: 40).

Subiektywność stanów mentalnych oznacza zatem, według Searle’a, dwie rzeczy. (a) Każdy z nas posiada bezpośredni dostęp do swoich własnych stanów mentalnych. (b) Posiadamy jedynie pośrednie empiryczne metody dochodzenia do takich samych stanów występujących u innych podmiotów. Subiektywne zjawiska nie są dostępne dla trzecioosobowych metod badań (por. Searle 1999: 109). Co do pytania o subiektywność od strony epistemologicznej, Searle stwierdza: „I podobnie jak ja pozostaję wobec moich stanów świadomych w pewnej szczególnej relacji, która różni się od mojej relacji do stanów świadomych innych ludzi, tak inni ludzie pozostają z kolei w pewnej relacji do swoich stanów świadomych, która różni się od tej relacji, w jakiej ja sam pozostaję wobec ich stanów świadomych” (Searle 1999: 135).

Równocześnie Searle zaznacza, że jego argumentacja nie jest argumentacją za „uprzywilejowanym dostępem” (*privileged access*), ponieważ nie ma żadnego uprzywilejowania ani żadnego uprzywilejowanego dostępu (por. Searle 1999: 330). Rozumiany po kartezjańsku uprzywilejowany dostęp jest: bezpośredni (nie jest on sprawą wnioskowania), niekorygowalny (nie można go odrzucić jako oczywistości) i możliwy dzięki introspekcji. Zdaniem Searle’a: „[...] fakt istnienia subiektywnej ontologii nie oznacza wcale, że nie możemy mylić się (*be mistaken about*) co do własnych stanów mentalnych”⁵.

Na zakończenie tego punktu należy zaznaczyć:

1. Poszukiwana ontologia tego, co mentalne, musi uwzględniać punkt widzenia pierwszej osoby i to, że każdy stan mentalny jest stanem jakiegoś podmiotu. Od strony epistemologicznej punkt widzenia pierwszej osoby różni się istotnie od punktu widzenia trzeciej osoby. Na przykład, pierwszoosobowo mogę wiedzieć, że jestem świadomy, chociaż wykazanie tego z punktu widzenia trzeciej osoby mogłoby się okazać niemożliwe⁶.

2. Różne interpretacje jakiegoś zdarzenia lub rzeczywistości (ze względu na epistemologiczną subiektywność) nie implikują, że owo zdarzenie nie zaszło lub że ono nie istnieje (ontologiczna subiektywność). Dwa podmioty

⁵ Searle 1999: 197; por. s. 201–202. Argumentacja Searle’a nie pokazuje jednak, dlaczego z bezpośredniości naszego dostępu nie wynika kartezjańska teza, że poznanie naszych własnych stanów jest niekorygowalne.

⁶ Zbliżone do takiego zjawiska przypadki obserwujemy na przykład w chorobie mięśni zwanej *locked-in syndrome* (SMS) [por. rozdział VII].

mogą obserwować to samo zdarzenie z dwóch perspektyw i różnie je opisywać, co nie oznacza, że samo zdarzenie zachodzi na dwa różne sposoby. Możemy powiedzieć tak: fakt, że nasze rozumienie natury jest formowane przez język, nie implikuje na sposób logiczny ontologii rzeczywistości. Z epistemologicznie rozumianej subiektywności nie wynika z logiczną koniecznością relatywizm ontologiczny.

3. Materializm domaga się, aby ontologia umysłu była na wskroś obiektywna. Obiektywność, chociaż założona jako kategoria ontologiczna, jest przez materialistów najczęściej charakteryzowana w terminach epistemologicznych: Zjawisko jest obiektywne, gdy jest intersubiektywnie dostępne (jak jednak korzystać z „intersubiektywności”, skoro jej fundament jest pierwszoosobowy?). Oznacza to zarazem, że odróżnienie perspektywy pierwszo- i trzecioosobowej nie ma tylko charakteru teoriopoznawczego, lecz pociąga za sobą odmienne kategorie ontologiczne.

2.1.2. Perspektywa pierwszo- i trzecioosobowa

Opisując świadomość – zarówno zjawiskową, jak i dostępu – używamy abstrakcyjnych zwrotów: „pierwszo-” lub „trzecioosobowy punkt widzenia”. Perspektywę pierwszoosobową rozumiemy jako indywidualną i subiektywną, a trzecioosobową jako publiczną i intersubiektywną.

II.1 Zwrot „pierwszoosobowy punkt widzenia” jest „podstawowym systemem pojęciowym”, odnoszącym się do subiektywności.

Poniższa, otwarta lista zawiera często spotykane znaczenia zwrotów: „pierwszoosobowy punkt widzenia” oraz „perspektywa pierwszoosobowa”.

1. Dyskurs o perspektywie może się odnosić: (a) do „perspektywicznych” faktów, (b) do „subiektywnych stanów”.

Ad (a) Perspektywiczne fakty są wymieniane dzięki ich subiektywnemu widzeniu. Według W. Lycana, jakiś fakt posiada jako część składową cechę perspektywiczną, gdy dotyczące go pojęcie dostępne jest z pierwszoosobowego punktu widzenia (a nie jest dostępne z trzecioosobowego punktu widzenia)⁷.

Ad (b) Niektóre stany subiektywne „odczuwają się w pewien sposób” [Ψ_s], inne stany subiektywne „odzwierciedlają pojęciowy punkt widzenia podmio-

⁷ Por. Lycan: 1987: 79. Takie pojęcie jest „[...] function from worlds to sets of individuals” i „[...] any such function is objectively describable” (s. 79). Lycan nie podaje jednak, czym jest „obiektywnie opisywalna” funkcja przebiegająca od świata do szeregu indywidualów.

tu” [Ω_s]⁸. Przyjmujemy, że istnieje zawsze jakieś *jak to jest być* w stanie [Ψ_s], choć nie zawsze musi istnieć jakieś *jak to jest być* w stanie pojęciowym [Ω_s].

2. Ciało podmiotu jest „ulokowane” w czasie oraz przestrzeni, a on sam traktuje przestrzeń jako wychodzącą od jego „tutaj”, a czas jako wychodzący od jego „teraz”. Ujmuje on świat jedynie skądś, nigdy zaś znikąd (nagłowskie „from nowhere”). Subiektywny punkt widzenia mogą mieć tylko indywiduala, które potrafią coś widzieć, słyszeć, odczuwać⁹.

3. Do epistemologicznej perspektywy sądów pierwszoosobowych odwołuje się L. Wittgenstein twierdząc, że wypowiedź „mam bóle” nie ma uzasadnienia w obserwacjach, lecz jest rodzajem prymitywnej reakcji. Reakcja ta ustala zarazem kryteria znaczenia wypowiedzi „mam bóle” (por. Wittgenstein 2000: §§ 398–411; Hallett 1977: 439–441). W potocznym użyciu, według Wittgensteina: „«Ja» nie nazywa żadnej osoby, «tu» – żadnego miejsca, «to» nie jest żadną nazwą”¹⁰.

Wypowiedź w pierwszej osobie („ja mam bóle”) nie jest opisem z perspektywy pierwszoosobowej. Wypowiadając „ja mam bóle” coś wyrażam, a moja tożsamość jako tego, który ma bóle, leży w tym, że to ja jestem podmiotem wypowiadającym „mam bóle”. Tak samo jak moja tożsamość leży w tym: „że jęczę z bólu”. Jest tak dlatego, gdyż „[...] tym, kto ból przejawia, jest cierpiąca osoba”¹¹.

4. Zdaniem Wittgensteina, subiektywne, pierwszoosobowe przeżycia dadzą się wyeliminować z określania *znaczenia* terminów mentalnych, a tym samym mogą przy takim określaniu w ogóle nie występować (por. Wittgenstein 2000: § 293). Możemy powiedzieć, że każdy z nas ma „pudełko” (świadomość), w którym znajduje się żuk (odczucia, subiektywne przeżycia), nikt jednak nie potrafi zajrzeć do pudełka kogoś drugiego, a jedynie do swojego, gdyż jest ono dostępne tylko z perspektywy pierwszoosobowej. Wówczas jest całkowicie obojętne, czy wszyscy mamy to samo w pudełku, czy coś całkiem innego. Owo coś w pudełku nie mogłoby zatem stanowić znaczenia nazwy „żuk” (nie mogłoby określać reguł używania tej nazwy). Stąd Wittgen-

⁸ Odróżnienie to odzwierciedla Kantowski podział na „odbiorczość wrażeń” i „samo-zrębność pojęć” (por. Kant 1957: B 74/A50).

⁹ Można to wyrazić słowami Searle’a: „Kolejną konsekwencją subiektywności jest to, że wszystkie świadome formy intencjonalności, dostarczające mi informacji o świecie, który jest ode mnie niezależny, reprezentują zawsze pewien konkretny punkt widzenia” (Searle 1999: 135; por. Sousa 2002: 3, 11).

¹⁰ Wittgenstein 2000: § 410; por. „Ja nie jest przedmiotem” (Wittgenstein 1995: 66).

¹¹ Wittgenstein 2000: § 302. Nie zajmujemy się tutaj kwestią, czy Wittgensteina należy zaliczać do ekspresywiistów, według których zdanie „boli mnie głowa” wyraża stan mentalny (podobnie jak krzyk), a nie opisuje ich (por. Garrett 1998: 112–121).

stein wnioskuję, że język nie może być czymś prywatnym. A jeżeli mój język nie jest prywatny, to także nie mogę tylko dla siebie wyrazić moich pozornie prywatnych przeżyć. Jeżeli myślenie zakłada istnienie języka, to musi się w nim realizować. Moje prywatne przeżycia nie są tylko dostępne dla mnie samego. To, co pozornie prywatne, jest samo przez się intersubiektywne (zawiera w sobie postępowanie według reguły). Tak rozumiana intersubiektywność jest założona w dyskursie dotyczącym obiektywności terminów psychologicznych¹².

5. Mówiąc o punkcie widzenia T. Naglowi chodzi z jednej strony o nasze pojęciowe rozumienie nas samych jako podmiotów, posiadających świadome przeżycia w rodzaju: *jak to jest być mną?* W tym sensie Nagel przyjmuje $[\Omega_s]$, odzwierciedlający pojęciowy punkt widzenia (por. wyżej 1.b): (i) *Jak to jest* dla podmiotu rozumieć świadomy stan C, w odróżnieniu od (ii) *jak to jest* dla podmiotu odczuwać C. Nagel przyjmuje (i) gdy mówi o tym, *jak to jest* dla jakiegoś gatunku, posiadać punkt widzenia.

Z drugiej strony Naglowi chodzi o punkt widzenia, zawierający w sobie związek z konkretnym miejscem („tutaj”) i czasem („teraz”). Właśnie dlatego, jego zdaniem, problem subiektywności jest „trudny”. „Powodem jest to, że każdy subiektywny fenomen jest związany z jednym jedynym (*single*) punktem widzenia i nieuchronnym wydaje się, że obiektywna teoria fizykalna ten punkt pominię” (Nagel 1997: 206).

Opisując lub wyjaśniając subiektywność musimy, według Nagla, podać sens terminu „punkt widzenia”, a to wydaje się niemożliwe. Subiektywność jest w sposób istotny związana z punktem widzenia, a każda obiektywna analiza musi taki punkt wyeliminować. Istnieje subiektywny charakter doznań (*jak to jest być*), który jest dostępny z konkretnego punktu widzenia, i dlatego: „Nie uchwytyje go żadna ze znanych ostatnio obmyślonych redukcyjnych analiz tego, co mentalne, wszystkie są one bowiem logicznie do pogodzenia z brakiem tego charakteru” (Nagel 1997: 204–205).

Nagel nie mówi, że jego twierdzenie dotyczy istotnie prywatnych przeżyć typu $[\Psi_s]$. Możemy sobie, według niego, wyobrazić, *jak to jest być* drugim człowiekiem, bo ten jest wystarczająco podobny do nas. Organizmy tego samego rodzaju posiadają (w różnym stopniu) wspólne, zjawiskowe $[\Psi_s]$ i behawioralne $[\Omega_s]$ stany, a tym samym i punkty widzenia. Ponieważ nie mamy tej samej perspektywy przeżyciowej ani nie spostrzegamy świata za pomocą

¹² Tak rozumie Wittgensteina T. Nagel: „Wittgenstein sądzi, że pojęcia psychologiczne spełniają ten warunek – rządzą nimi obiektywne reguły, a to ze względu na związki między ich użyciem w pierwszej i trzeciej osobie. Właśnie ten rodzaj obiektywności nadaje się do tego, co w swej istocie jest subiektywne” (Nagel 1997c: 47; por. Nagel 1997d: 249).

echolokacji, to ani nie potrafimy sobie wyobrazić, *jak to jest* być nietoperzem.

Rozważania o punkcie widzenia mają dla Nagla nie tylko charakter teoriopoznawczy (por. Nagel 1997b: 238), co pokazuje chociażby jego pytanie o sens ludzkiego działania, widzianego z perspektywy wewnętrznej i pozwalającego na uzasadnianie działań w kategoriach innych działań. Znaczenie wewnętrzne określamy przez odniesienie do perspektywy zewnętrznej, obejmującej cele i zaangażowania podmiotu. Podstawowym dylematem tego rodzaju eksternalizmu jest jednak: czy w pełnym przedstawieniu umysłu, podkreślającym dynamiczne interakcje z otoczeniem, istnieje jakieś wewnętrzne, zjawiskowe jądro? Nagel przyjmuje istnienie takiego wewnętrznego (internalistycznego) jądra. W badanym przez nas teoriopoznawczym kontekście Nagel podkreśla, że dyskurs o świadomości wymaga od nas uznania ogólnej myśli o perspektywie, bez zawężania tej świadomości do znanych, jej subiektywnych form¹³.

Nagel zgadza się z Wittgensteinem, że czysto pierwszoosobowych pojęć nie ma. Próba oddzielenia pierwszoosobowego użycia od trzeciosobowego jest filozoficzną iluzją w rodzaju, że podmiot mojej świadomości mógłby zostać zamieniony parę minut temu na jakiś inny, a wszystkie moje wspomnienia mogłyby zostać przetransferowane do nowego podmiotu, obecnego teraz w tym samym ciele. W dodatku cała ta zmiana zaszłaby bez jakiegokolwiek zauważalnej zmiany zewnętrznej (czy to fizycznej, czy psychologicznej). Zdaniem Nagla, czysta pierwszoosobowa idea „ja”, definiująca jakieś indywiduum, może mieć sens w tak rozumianym eksperymencie myślowym. Uzasadnione wydaje się jednak, że faktyczna idea „ja” utraciła swoje miejsce w tym eksperymencie¹⁴.

Mówiąc o pierwszoosobowym punkcie widzenia, o świadomości zjawiskowej, a tym samym o subiektywności, odwołujemy się w tej pracy do argumentacji przytoczonej przez Nagla. Czynimy tak z dwóch powodów:

¹³ „To samo jest z umysłem. Jeżeli chcemy zrozumieć szczególne rodzaje doświadczenia i nie korzystać przy tym z tego, że sami przeżyliśmy lub potrafimy sobie wyobrazić tego rodzaju doznania, to musimy przyjąć ogólną ideę perspektywy i nie ograniczać jej do form znanych nam z subiektywnego doświadczenia lub z innych źródeł” (Nagel 1997c: 32).

¹⁴ Por. Nagel 1998a: 341–342. Omawiane zagadnienie nawiązuje do Kanta, który twierdzi, że subiektywna jedność świadomości mnie samego w różnych momentach czasowych, nie jest wystarczająca, aby ustanowić obiektywną tożsamość podmiotu lub duszy (por. Kant 1957: A 363–364). Kant zaznacza, że transcendentálną jedność apercpcji nazywa się przeto „[...] przedmiotową i trzeba ją odróżnić od podmiotowej jedności świadomości” (B 139).

(i) naglowskie ujęcie, jak zobaczymy w rozdziale V, zachowuje filozoficzne rozumienie subiektywności, wskazując zarazem, dlaczego subiektywność stanowi „trudny” problem. (ii) Nagel zajmuje się subiektywną perspektywą jakiegoś rodzaju organizmów (np. nietoperza). Czyni to również wtedy, gdy mówi, że perspektywa ta nie może być zbadana metodami obiektywnymi.

W odróżnieniu od Nagla zakładamy holistyczne rozumienie świadomości. Nagel wychodzi od pojęcia identyczności jaźni (*self*) w czasie i uznaje tożsamość jaźni za niezależne pojęcie psychologiczne (por. Nagel 1997c: 42, 48–54). Jaźń leży u podstaw psychicznych ciągłości (tam, gdzie one istnieją), chociaż nie posiada koniecznych lub wystarczających warunków, specyfikalnych w terminach tych ciągłości¹⁵. Nagel odwołuje się jedynie sporadycznie do jedności świadomości (*unity of consciousness*) i do „subiektywnej jedności” (*subjective unity*) [Nagel 1997c: 51, 64], którą my rozumiemy jak holistyczną, synchroniczną jedność świadomości.

2.1.3. Perspektywa czasowa i przestrzenna

W inny sposób podchodzimy do rezultatów naszych działań, mających zająć w odległej przestrzeni aniżeli w odległym czasie. Czas jest asymetryczny (inaczej traktujemy to, co wydarzy się w przyszłości, od tego, co wydarzyło się w przeszłości). Przestrzeń nie ma takiego ukierunkowania. Asymetryczność czasu konstytuuje bardziej podstawowy rodzaj subiektywności aniżeli perspektywa przestrzenna (por. Sousa 2002: 3–4).

Zmiany, szczególnie w perspektywie czasowej, wpływają na relatywną wartość naszych celów. Na przykład, czasowo bliższe cele zdają się posiadać dla nas większe znaczenie aniżeli cele odległe. Wspomniane zmiany rodzą pytanie, która perspektywa jest wtedy faktycznie *moja*? Czy istnieje tak wiele jaźni, jak wiele istnieje perspektyw czasowych? Czasowa asymetria niesie ze sobą bycie w określonym czasie i wymaga uwzględnienia jaźni jako działającego podmiotu. Przestrzenność perspektywy pierwszoosobowej jest określona poprzez współrzędne ciała podmiotu i jego sposób poruszania się.

Spostrzeżeniowe treści obecne w podmiocie są treściami jakościowymi w tym sensie, że są one zawsze przedstawiane z jakiegoś punktu widzenia. Mieć wzrokowe przeżycie zarysu budynku, to posiadać je jako przeżycie budynku stojącego w pewnej przestrzennej relacji do podmiotu, i w pewnej

¹⁵ Polskie tłumaczenie tekstu Nagla nie jest w tym miejscu dokładne (s. 48). Dosłownie: „[...] the self is something that underlies the psychological continuities where they exist but has no necessary or sufficient conditions specifiable in terms of them” (Nagel, *The View from Nowhere*: 38).

przestrzennej relacji do otoczenia (do „tła”) tego budynku. Trudno poprowadzić linię oddzielającą centrum pola widzenia (budynek) od peryferii jego otoczenia. Mimo tego nie odnosimy zwykle wrażenia, że nasze spostrzegane wzrokowo otoczenie jest rozmyte na peryferiach tego pola (faktyczne rozmycie jest związane z nierównomiernym rozmieszczeniem receptorów na siatkówce oka). Chcąc ogarnąć leżącą przed nami przestrzeń, poruszamy oczami, „nadstawiamy uszu”, odwracamy głowę, poruszamy całym ciałem. Spostrzegające przeżywanie otaczającej przestrzeni nie jest czymś wycinkowym, lecz rozciągniętym w czasie zaangażowaniem podmiotu.

W dalszych rozważaniach rezygnujemy z aspektu subiektywności uwzględniającego czasowe cele podmiotu, lecz zachowujemy rozróżnienie przestrzeni spostrzegania na „centrum” i „tło”. Nie będzie nas przy tym interesować pytanie o diachroniczną jedność podmiotu (czyli jedność podmiotu obejmująca pamięć w okresie czasu od t_1 do t_2), lecz zajmiemy się synchroniczną jednością świadomości (łączenie różnych przeżyć w jednolitą całość w momencie t_1). Pytania o ten rodzaj jedności dotyczą przeżywania siebie samego oraz swoich stanów mentalnych. Wychodzimy od twierdzenia, że:

II.2 Strumień świadomych stanów musi się w sposób konieczny odznaczać jednością w czasie t .

Synchroniczną analizę jedności podmiotu należy uzupełnić o założenia postulujące jedność jego świadomości, chociażby o tak zwaną *unity thesis* postulowaną przez T. Bayne’a i D. Chalmersa (2003: 1–3). Przy jej uzasadnianiu Bayne i Chalmers odwołują się do metafizycznej tezy o holistycznie rozumianej świadomości.

2.1.4. Odczuwanie własnego ciała

Normalnie zdajemy sobie sprawę z przestrzennego rozmieszczenia naszych rąk, nóg, mamy radę z zamkniętymi oczami trafić palcem we własny nos. Doświadczeni kierowcy potrafią „poszerzyć” przestrzenne odczuwanie ciała, przenosząc je niejako na samochód, gdy na przykład muszą nim wykonać bardzo precyzyjne manewry. Osoby używające laski „odczuwają” dotykane nią przedmioty, jak gdyby dotykały ich ręką.

Świadomość przestrzeni zajmowanej przez ciało nie jest jednak czymś oczywistym. W przeprowadzonym przez V. S. Ramachandrana eksperymencie zwanym „iluzją fantomowego nosa” badana osoba siedzi z zasłoniętymi oczyma, a przed nią siedzi inna osoba, patrząca w tym samym kierunku. Stojący w pobliżu badanej osoby naukowiec bierze swoją lewą ręką lewy palec

wskazujący teje osoby i głaszcze nim nos siedzącej przed nią osoby (por. Ramachandran 1998: 1855). W tym samym czasie palcami swojej prawej ręki eksperymentator głaszcze w dokładnie zsynchronizowany sposób nos badanej osoby. Po kilku sekundach osoba ta odnosi wrażenie, że albo jej nos się przemieścił albo się znacznie wydłużył. Zdaniem Sousa:

„Szczególnie intrygujące w tej iluzji jest to, iż w rzeczywistości ów «zmysł» (*sense*), który prowadził twoją rękę wcale nie jest zmysłem, gdyż nie posiada on «organu». Co więcej, o ile tylko *podmiot* potrafi dokonać odpowiednich obserwacji, jest to oczywiście forma subiektywności. Nie możemy *mieć* fantomowej iluzji nosa kogoś innego. Ale nie jest to jedynie quale albo wiązka qualiów” (por. Sousa 2002: 6).

Na co dzień nie zastanawiamy się ani nad właśnie wymienionymi formami subiektywności, ani nad czasoprzestrzenią, w której żyjemy i poruszamy się, ani nad wymiarami czy miejscem naszego ciała w niej. Według S. Gallagera jest to możliwe dzięki posiadaniu przez nas odpowiedniego schematu ciała (*body schema*). Składa się nań cały system przedświadomych procesów, odgrywających dynamiczną rolę w zachowaniu postawy ciała i w jego przemieszczaniu się¹⁶. Schematowi ciała Gallagher przeciwstawia obraz ciała (*body image*), odpowiadający naszej świadomej, mentalnej reprezentacji własnego ciała. Naszą uwagę na wybrane części ciała skierowujemy dzięki wykorzystaniu obrazu ciała. Schemat ciała wywiera wprawdzie wpływ na przeżycia poznawcze, nie posiada jednak statusu świadomego przedstawienia. Można powiedzieć:

II.3 Podczas gdy schemat ciała funkcjonuje w sposób zintegrowany, automatyczny i całościowy, obraz ciała jest przedstawieniem ciała z perspektywy pierwszoosobowej.

W dalszej części pracy przyjmujemy, że **II.3** odnosi się do normalnego (zdrowego) podmiotu i pomija przypadki zaburzeń w jego obrazie ciała i w schemacie ciała. Wprowadzone przez Gallagera rozróżnienie jest bardziej opisem aniżeli faktycznym rozróżnieniem empirycznym. Trudno chociażby potocznie oddzielić funkcje sensoryczne od motorycznych, który to rozdział Gallagher zakłada.

¹⁶ Por. Gallagher i Cole 1995: 2–4. Według P. Duusa: „Tylko mała część impulsów dotyczących propriocepcji dochodzi do świadomości człowieka, tzn. dociera do kory mózgu” (Duus 1989: 23; por. s. 227).

Odczuwanie własnego ciała w sensie „obrazu ciała” jest dla podmiotu istotnym elementem jego rozumienia tego, *jak to jest* posiadać ciało, a tym samym jest ono istotnym elementem holistycznej świadomości. Można utracić zdolność do świadomego spostrzegania własnego ciała (czyli mieć zakłócony obraz ciała), a mimo tego być zdolnym do poruszania, jeżeli tylko posiada się nienaruszony schemat ciała (por. Sacks 1996: 67–80).

2.2. Subiektywność wyrazem jedności świadomości

Świadome przeżycia podmiotu ukazują się jako części jego odznaczającej się jednością świadomości. Siedząc przed komputerem widzę nie tylko kolorowy ekran, ale odczuwam dotyk klawiatury, słyszę szum pracującego wentylatora. Wszystko to przeżywam jako pojedynczą, cechującą się jednością całość¹⁷. Opisuując jedność świadomości, T. Bayne i D. Chalmers wychodzą z intuicyjnego założenia, że: „[...] istnieje jakiś podstawowy sens, w którym wszystkie przeżycia podmiotu – włączając przeżycia spostrzegania, cielesne, emocjonalne i poznawcze – mogą się cechować jednością” (Bayne i Chalmers 2003: 3; por. Searle 2000: 561–562).

Chodzi tutaj o jedność w tym sensie, że wszystkie te przeżycia są jakoś ujęte w pojedynczym, całościowym stanie świadomości. Możemy powiedzieć, że dwa świadome przeżycia cechują się jednością, gdy występują one w jednym, całościowym stanie świadomości. Bayne i Chalmers odróżniają cztery rodzaje jedności świadomości: jedność przedmiotowa (stany świadome odznaczają się jednością, gdy są skierowane na ten sam przedmiot), jedność przestrzenna (stany odznaczają się jednością, gdy przedstawiają przedmioty jako będące w tej samej części przestrzeni), jedność podmiotowa (stany świadome należą do tego samego podmiotu) i jedność obejmująca (*subsumptive unity*, stany cechują się jednością, gdy są ujęte w tym samym pojedynczym stanie świadomości). Przeżycia wzrokowe, słuchowe, dotykowe są ujęte w jednym, obejmującym je stanie spostrzegającej świadomości. Na tej podstawie można przyjąć, że:

II.4 Istnieje całościowy stan świadomości zjawiskowej podmiotu, obejmujący jego stany zjawiskowe (sposstrzegania, cielesne, emocjonalne, poznawcze, ...).

¹⁷ Mówiąc o otaczającej i przenikającej mnie w każdym momencie przeżywania całości, fenomenologowie używają wspomnianego już terminu „pole świadomości” (por. Piłat 1993: 11–13).

Wspomniany stan nie jest jedynie prostym złożeniem świadomych przeżyć (*A* i *B*), lecz sam jest świadomym stanem, „osobliwością stojącą za wielością”¹⁸. Bayne i Chalmers zaznaczają, że mówiąc o jedności obejmującej stany świadomości, nie mają na myśli jedności neurofizjologicznej, zgodnie z którą stany świadome muszą mieć związek z jakimś pojedynczym obszarem czy mechanizmem w mózgu.

Charakteryzując bliżej ten typ jedności, wspomniani autorzy odpowiadają na pytanie: Co to znaczy dla stanu mentalnego być stanem świadomym? Powołują się przy tym na wspomniane już przez nas rozróżnienie pomiędzy świadomością dostępu, a świadomością zjawiskową (por. także 4.3.3) i podają dwie definicje: funkcjonalnie rozumianą definicję dostępu do stanu mentalnego i definicję zjawiskowego stanu mentalnego:

(a) Stan mentalny *A* jest świadomy w sensie dostępu, jeżeli treść *A* jest dostępna dla sprawozdania lub racjonalnej kontroli. Świadomość dostępu jest definiowana w terminach przyczynowej, dynamicznej roli, jaką dany stan odgrywa.

(b) Stan mentalny *A* jest zjawiskowo świadomy, jeżeli jest coś takiego jak „bycie w stanie *A*”. Świadomość zjawiskowa – jak powiedzieliśmy wcześniej – jest definiowana w terminach mówiących o tym, jak dany stan się odczuwa, jak się go doznaje. Świadomość zjawiskowa implikuje, zdaniem Chalmersa, świadomość dostępu. Wykorzystując definicje zarówno świadomości dostępu, jak i zjawiskowej, Bayne i Chalmers wprowadzają dwie następne definicje: świadome stany *A* i *B* odznaczają się jednością na sposób dostępu, gdy koniunkcyjna treść *A* i *B* jest dostępna dla podmiotu; świadome stany *A* i *B* odznaczają się jednością zjawiskową, jeżeli *A* i *B* są przeżywane razem, jeżeli jest coś takiego, jak „posiadać razem *A* i *B*” (por. II.4).

Dwa stany odznaczają się jednością w sensie obejmującym i przeżyciowym, jeżeli podmiot może doświadczyć tego, *jak to jest być* równocześnie w obydwu tych stanach¹⁹. Można powiedzieć tak: całościowy zjawiskowy stan uchwytuje owo *jak to jest być* podmiotem w danym czasie *t*. Gdy podmiot jest w tym stanie, wtedy wszystkie stany zjawiskowe odznaczają się jednością i to – zdaniem wspomnianych autorów – na sposób konieczny (*total phenomenal unity thesis*).

¹⁸ „[...] this total state is not just a conjunction of conscious states. It is also a conscious state in its own right. If such a total conscious state exists, it can serve as the «singularity behind the multiplicity» – the single state of consciousness in which all of a subject’s states of consciousness are subsumed” (Bayne i Chalmers 2003: 4).

¹⁹ „That is, two states are [subsumptively] phenomenally unified when they have a conjoint phenomenology: a phenomenology of having both states at once that subsumes the phenomenology of the individual states” (Bayne i Chalmers 2003: 8).

W dalszych rozdziałach książki pod terminem „jedność świadomości” rozumiemy jedność odpowiadającą twierdzeniu II.4. Sugeruje ono, że zawsze istnieje jakiś pojedynczy stan świadomości, który obejmuje wszystkie stany zjawiskowo-przeżyciowe podmiotu w danym czasie t . Można powiedzieć tak: Gdy podmiot posiada taki stan, jego świadomość cechuje się jednością na sposób subiektywny. Chodzi w nim bowiem o *jak to jest* być w danym, cechującym się jednością stanie. Definicja subiektywnej jedności świadomości ma swoją podstawę w holistycznym rozumieniu tej ostatniej.

Przyjmujemy zatem, że świadomość posiada wewnętrzną, zjawiskową jedność, która jest skorelowana ze stanami w mózgu i jakoś od nich zależy. Nie przyjmujemy ani dualistycznego założenia, że jedność ta jest jednością jakiejś mentalnej substancji, ani założenia materialistycznego, że jedność ta jest jedynie wyrazem przyczynowego oddziaływania anatomicznej lub fizjologiczno-funkcjonalnej organizacji mózgu, albo że jedność zjawiskowa jest z tą organizacją identyczna. Nie przyjmujemy także założenia o ściśle holistycznej jedności, według której świadomość nie ma części albo elementów, por. 1.3.4). Precyzujemy tezę Chalmersa i Bayne (*total phenomenal unity thesis*) w tym sensie, że przyjmujemy jej słabsze holistyczne rozumienie, czyli jak jest ono traktowane w emergentystycznych teoriach świadomości. Przyjęcie tak sformułowanego założenia o jedności świadomości jest – jak zobaczymy – uzasadnione jego znaczeniem wyjaśniającym.

Pojęcie jedności odgrywa także centralną rolę w Kartezjusza filozofii świadomości. Zastanawiając się nad swoim umysłem jako nad rzeczą myślącą, Kartezjusz stwierdza, że nie potrafi w sobie odróżnić żadnych części, lecz wyraźnie pojmuje siebie jako kogoś jednego. Ponieważ więc umysł nie składa się z części, nie może składać się z materii (gdyż wszystko co materialne jest podzielne na części). Również I. Kant odwołuje się do pojęcia jedności świadomości przy analizowaniu jej subiektywności. Ani Kartezjusz, ani Kant nie operują kryterium subiektywności podobnym do naglowskiego *jak to jest być* świadomym.

Pytanie, które się nasuwa w związku z II.4, brzmi: Czy subiektywna jedność świadomości może ulec załamaniu? Gdy jedność świadomości ulegnie załamaniu, to nie ma niczego w sensie *jak to jest* być podmiotem w danym czasie, albo przynajmniej nie ma jakiegoś pojedynczego *jak to jest być*, obejmującego wszystkie jakościowe stany podmiotu. Trudno sobie wyobrazić możliwość całkowitego braku tak rozumianej subiektywnej jedności świadomości, co dodatkowo potwierdza zasadność przyjęcia założenia o subiektywnej jedności świadomości.

2.3. Model świadomości zjawiskowej

„Are dreamed pains real pains, or a sort of imagined pains?”
(Dennett 1991: 61).

„Model” – rozumiany w szerokim sensie tego słowa – jest mniej lub bardziej odpowiednim „substytutem” badanego na sposób naukowy obiektu lub procesu. Modele są używane chociażby dla symulowania przebiegu zjawisk niedostępnych dla empirycznych obserwacji. Najpowszechniejszymi modelami, do których odwołujemy się w naukach ścisłych, są modele matematyczne. Pomimo obecności modeli w nauce, nadal nie dysponujemy ani wspólnym słownictwem, ani jakąś jedną wspólną teorią wyjaśniającą relację modelu do badanego rzeczywistego obiektu.

Posiadanie odpowiedniego modelu świadomości zjawiskowej powinno się nie tylko okazać pomocne przy tworzeniu jej teorii, ale model ten mógłby pomóc zarazem naszej wyobraźni operować zbudowaną na jego podstawie teorią²⁰. Poszukiwany model świadomości zjawiskowej musi uwzględniać obecne w niej cechy jakościowe. Gdybyśmy poszukiwali modelu świadomości dostępu, to musiałby on uwzględniać potoczne cechy świadomości dotyczące jej przyczynowej roli oraz jej wpływu na obserwowalne zachowania. Interesujący nas model świadomości zjawiskowej należy uzupełnić o tak zwany „komentarz” objaśniający i interpretujący analogię, zachodzącą pomiędzy znanymi jakościowymi cechami wyjaśnianych obiektów, a obiektami modelu.

W filozofii analitycznej są znane różne modele świadomości (choćby funkcjonalne modele komputerowe czy modele neuronalne). W rozdziale IV zostanie omówiony Dennettowski model świadomości zwany Modelem Wielu Szkieł, jako przeciwstawny modelowi nazywanemu Kartezjańskim Teatrem. Model Wielu Szkieł niejako programowo nie obejmuje subiektywnych przeżyć, lecz odwzorowuje zachodzące w mózgu procesy przetwarzania informacji. Do modelu „teatru umysłu” (*theater of mind*) odwołuje się rozwinięta przez B. J. Baarsa teoria globalnej przestrzeni działania (*global workspace theory*, GWt). Globalna przestrzeń działania jest w tym modelu rozumiana na podobieństwo teatralnej sceny i jest traktowana jako forma biologicznej adaptacji spełniającej funkcje integracji i dystrybucji informacji²¹. Funkcje te są również funkcjami świadomości. Baars sugeruje, że to właśnie

²⁰ Przedstawione tutaj rozumienie modelu pochodzi od W. Sellarsa (por. Bremer 1997: 32–33).

²¹ Według B. J. Baarsa, „The stage receives sensory and abstract information, but only events in the spotlight shining on the stage are completely conscious. The actor in the spotlight frets and struts his hour upon the stage, directed by the playwright and di-

świadomość umożliwia kooperację pomiędzy różnymi sieciami neronalnymi, prowadząc przez to do rozwiązania takiego czy innego problemu. Całościowym zadaniem świadomości jest umożliwienie szerokiego dostępu, który z kolei może służyć koordynacji i kontroli. Niektóre rodzaje informacji mogą globalnie oddziaływać na całą sieć. Poszczególne neuronalne systemy w mózgu tworzą podstawę dla tak rozumianej globalnej przestrzeni. Tak rozumiana świadomość jest niejako „bramą” do mózgu, umożliwiającą kontrolę pojedynczych neuronów i całych ich populacji.

Zarówno model proponowany przez Dennetta jak i model proponowany przez Baarsa mają zasadniczo status „metafory opisującej świadomość” (por. Baars 1997a: 302), status „eksperymentu myślowego”, a nie status modeli używanych w teoriach nauk ścisłych. Podobny status ma proponowany przez nas model świadomości zjawiskowej.

W stosunku do modelu proponowanego przez Baarsa powstaje wątpliwość, czy bycie świadomym odczuwa się podobnie do siedzenia w ciemnym teatrze i patrzenia na zdarzenia rozgrywające się na jego scenie. Naszą świadomość w czasie t odczuwamy raczej jako bycie zanurzonymi w centrum świata, który jest dla nas jakoś cały obecny, chociaż naszą uwagę potrafimy skierować tylko na jego wybrane cechy. Otaczające nas zjawiska są uzewnętrznionymi zdarzeniami sensorycznymi podobnie jak nasz własny obraz ciała, a także te zdarzenia, które odczuwamy, iż zachodzą w naszym ciele albo w naszym umyśle. Stąd nasze poszukiwania innego modelu oraz innego systemu pojęciowego, aby uchwycić stronę zjawiskową (a w rezultacie całą świadomość).

Odwołując się do badań A. Revonsuo można powiedzieć, że charakterystyczne własności zjawiskowe ukazują się w marzeniach sennych występujących w fazie snu REM. Marzenia te mogą służyć jako model świadomości zjawiskowej²². Revonsuo pisze: „Marzenie senne jest w typowy sposób, niezwiązaną z jakimkolwiek wejściem i wyjściem, wyizolowaną płaszczy-

rector, against a background created by scene setters. These behind-the-scenes influences are context operators, unconscious systems that shape conscious events. The spotlight selects the most significant actors on stage, and once lit up, their messages are distributed to an audience consisting of all the unconscious routines and knowledge sources – the vast array of unconscious tools we use to adapt to the world” (Baars 1997a: 301). „GWT is based entirely on well-established empirical contrasts between pairs of conscious and nonconscious events” (s. 294).

²² Por. Revonsuo 1998: 206–208. Podobny, oparty na marzeniach sennych model „intrzyńszcznie” rozumianej świadomości został przedstawiony w: Llinas i Ribary 1998: 1844–1847. Neuronalnym korelatem świadomości we wspomnianym modelu jest zsynchronizowana aktywność oscylacyjna w systemie: wzgórze – kora mózgu (por. s. 1845).

zną zjawiskową (tzn. treści świadomości nie są we śnie ani modulowane przez faktyczną sensoryczną informację, ani nie są one w stanie spowodować motorycznych wyjść w sensie aktywnej interakcji: organizm–otoczenie)”²³.

Podstawowym pojęciem dla tej płaszczyzny organizacji jest doświadczenie *obecności*, doświadczenie *bycia zanurzonym*, w „spozrzanym świecie”, który jest światem „wyśnionym”, chociaż wydaje się być rzeczywistym. Mówi się wówczas o tak zwanej wirtualnej obecności albo o doświadczeniu znajdowania się „poza mózgiem”. Wyrażenia te odnoszą się do iluzji, że wydarzenia, których jesteśmy bezpośrednio świadomi, rozgrywają się w perceptualnym świecie poza mózgiem, chociaż – według Revonsuo – rozgrywają się one faktycznie w mózgu, do którego nie docierają sygnały percepcyjne. W komentarzu do omawianego modelu należy oczywiście zaznaczyć, że istnieją także istotne różnice pomiędzy snem a stanem przebudzenia (na przykład we śnie nie poznajemy tak samo, jak poznajemy na jawie).

W dalszych analizach odwołujemy się do podobnego modelu zjawiskowej świadomości, rozumianej jako subiektywny stan, obejmujący wspomniane już *jak to jest być* świadomym. Przyjmujemy, że w snach mamy do czynienia z holistycznie rozumianą świadomością. W odróżnieniu od Revonsuo, mówimy o korelatkach pomiędzy śpiącym mózgiem a świadomością zjawiskową i pytamy o zależność: mózg–świadomość.

II.5 Obserwacja mózgu w czasie fazy snu REM pokazuje, że sensoryczne wejścia i motoryczne wyjścia nie są *konieczne* dla pojawiania się cech zjawiskowych. Tak rozumiany mózg jest skorelowany z cechami zjawiskowymi dostarczając nam wglądu w przebieg procesów *wystarczających* dla ich powstania.

Marzenia senne wskazują na skorelowaną z mózgiem świadomość zjawiskową. Śniący podmiot jest świadomy, ale odcięty od spozrzanego świata zewnętrznego poprzez wejścia zmysłowe. Tak wyizolowana płaszczyzna zjawiskowa ukazuje się prawie zawsze jako całość w formie rzeczywistości doświadczonej. Podczas marzenia sennego pojawiają się sen-

²³ Revonsuo 1998: 206. Revonsuo chodzi o wejścia sensoryczne i motoryczne wyjścia, por. s. 207–208. Posiadanie snów jest tutaj rozumiane jako generowanie przez mózg płaszczyzny zjawiskowej. Naszym zdaniem nie chodzi o sny spowodowane np. przepelnionym pęcherzem lub napięciem seksualnym, opisane w pracach medycznych od S. Freuda poczynając.

soryczne przeżycia w wielu modalnościach (przeżycia słuchowe, wzrokowe, dotykowe, ...) ²⁴.

W snach zachowujemy podobny obraz ciała do tego, jaki posiadamy wtedy, gdy nie śpimy. Lokalizację własnego ciała, jego obraz i ruchy przeżywamy tak, jak gdyby zachodziły one w otaczającym nas na jawie świecie przestrzennym. Śniący podmiot odczuwa treści zjawiskowe w podobny sposób, w jaki odczuwa je na jawie.

Należy przypomnieć dwie rzeczy:

1. Sama metafora „wirtualnej rzeczywistości” nie jest wyjaśnieniem świadomości zjawiskowej, lecz co najwyżej uproszczonym opisem tej natury świadomości, o której nie wiemy, jak ją wyjaśnić chociażby w sensie: w jaki sposób i na ile mózg wywołuje świadomość? Niezależność tak przyjętego opisu jest zagwarantowana przez nieredukowalność uzyskanych przy jego pomocy wyjaśnień. Metafora „wirtualnej rzeczywistości” ujmuje świadomość jako charakterystyczny, tworzony w izolacji od spostrzeżeń zewnętrznych, model świata. Z jednej strony, cechy zjawiskowe wyśnionych „obiektów” tego modelu [Ψ_o] są analogiczne do zjawiskowych cech obiektów w rzeczywistości czasoprzestrzennej [Ψ_o]. Z drugiej strony, cechy te nie są powiązane z odpowiadającymi im normalnymi zachowaniami przebudzonego podmiotu (spostreżaniem, ruchami głowy, sekwencyjnym myśleniem, itd.).

2. Badanie mózgu „śniącego podmiotu” nie tylko uwzględnia subiektywność stanów świadomości, ale pozwala po części odizolować płaszczyznę zjawiskową świadomości od sensorycznych wejść i motorycznych wyjść (odpowiedzialnych na przykład za poruszenia głową na jawie). Nie potrafimy bezpośrednio „widzieć” wyśnionego przez drugą osobę świata. Niemniej jednak mamy wiele danych o zjawiskowych treściach snów, chociażby dzięki sprawozdaniom podmiotów oraz dzięki neuroobrazowaniom mózgów śpiących podmiotów. Proponowany w II.5 model uwzględnia subiektywność świadomości i skłania do zadawania pytań o możliwość bezpośredniego neuroobrazowania korelatów świadomości zjawiskowej. Tym samym można go włączyć do systematycznej, naukowej metodologii badań nad świadomością.

²⁴ Według Dennetta, tworzenie marzeń sennych podlega tym samym interakcjom z systemem kontroli i pamięcią, którym podlegają dane pochodzące z analiz spostrzegania (por. Dennett 1998: 155).

2.4. Inne przejawy świadomości

Mówiąc o subiektywności świadomości należy wspomnieć o związanej z nią problematyce samoświadomości oraz qualiów. Potocznie termin „qualia” (albo „jakości doznaniowe”) odnosi się do introspekcyjnie dostępnych, zjawiskowych cech naszego życia mentalnego, a tematyka z nimi związana stanowi centrum współczesnych dyskusji nad problemem subiektywności świadomości²⁵. Zdaniem D. Chalmersa, ontologiczne i epistemologiczne wyjaśnienia kwestii qualiów tworzą „trudny” problem świadomości (por. Chalmers 1996: 4).

2.4.1. Qualia

Opisy i teoretyczne wyjaśnianie subiektywności może być pojęciowo oddzielone od problemu qualiów, chociaż empirycznie problemy te są ze sobą ściśle powiązane.

Według D. Dennetta, klasycznie rozumiane qualia są: (1) niewyraźalne (*ineffable*), (2) wewnętrzne (*intrinsic*), (3) prywatne (*private*), (4) bezpośrednio uchwytnie w świadomości (*directly apprehensible in consciousness*) (por. Dennett 1988: 3–4). Qualia są przypisywane wrażeniom cielesnym (odczuwanie głodu, ciepła), emocjom, pasjom, odczuciom depresji, aktom spostrzegania (wzrokowego, słuchowego, ...), którym towarzyszą odczucia jakości barw, dźwięków. G. Strawson uzupełnia powyższą listę o przeżycie rozumienia zdania, przeżycie myślenia, przeżycie przypominania sobie. Oznacza to, że przynajmniej niektórym stanom intencjonalnym można przypisać cechy jakościowe²⁶.

Wrażenia kolorów, dźwięków, odczuwanie bólu są bardzo realne, a mimo wszystko trudno znaleźć w obiektywnie opisywanym świecie cechujące je jakości zjawiskowe. Co nastawiona na obiektywne opisy i wyjaśniania nauka może powiedzieć o kolorze niebieskim? Naukowiec wskazuje zazwyczaj na istnienie fali elektromagnetycznej określonej długości (emitowanej lub odbijanej przez różne obiekty i docierającej do siatkówki oka) i dodaje, że

²⁵ O podstawowych, różniących się między sobą ujęciach qualiów (por. Dennett 1988: passim; por. Tye 2003: 1–13; por. Dziarnowska 2001: 176–177; por. Northoff 1995: 270–289; por. Stubenberg 1998: passim).

²⁶ Pogląd Strawsona jest kontrowersyjny. Fenomenalny aspekt rozumienia różni się od lingwistycznego czy werbalnego obrazu, który ma fonologiczną i syntaktyczną strukturę. Wyobrażenia i wrażenia nie są zawsze dostępne dla podmiotu i zdają się nie być istotne dla procesu myślenia.

w świecie zewnętrznym nie ma czegoś takiego, jak potocznie rozumiany niebieski, lecz jest jedynie odpowiednie promieniowanie świetlne. Gdzie jednak owo niebieskie – mówiąc potocznie – jest, jeżeli nie ma go w świecie zewnętrznym? Naukowa odpowiedzią może być: gdy światło padnie na nasze oko, różne częstotliwości aktywują czopki i pręciki siatkówki, z których pobudzenie jest przekazywane za pośrednictwem nerwu wzrokowego z oka do ośrodków wzrokowych w mózgu wywołując powstanie odpowiedniego stanu neuronalnego, a następnie jakaś „przestrzeń zewnętrzna” zostaje wypełniona przedstawieniami koloru, które mogą być analizowane przez inne części mózgu celem wywołania behawioralnych zachowań. Intuicyjnie zgadzamy się oczywiście, że coś podobnego zachodzi w naszych mózgach, chociaż takie „naukowe spojrzenie” niesie ze sobą dalsze pytania natury epistemologicznej i ontologicznej.

Od strony epistemologicznej przyjmuje się czasem, że bycie świadomym jakiegoś wzrokowego przeżycia jest podobnym do widzenia wewnętrznego, niefizykalnego „obrazu” czy „danej zmysłowej” (np. spostrzegające „ja” jest podmiotem dla niefizykalnego przedstawienia niebieskiego kwadratu). Introspekcja dostarcza nam treści oraz wewnętrznych własności tego obrazu. Od strony ontologicznej qualia są najczęściej traktowane jako wewnętrzne, dostępne dla świadomości, nieintencjonalne cechy danych zmysłowych i innych niefizykalnych obiektów zjawiskowych, które są odpowiedzialne za przeżyciowy charakter tychże obiektów. Zakłada się przy tym, że qualia mogą się zmieniać bez zmiany treści intencjonalnej przeżycia. Są one mentalnymi odpowiednikami określonych własności obiektu (na przykład koloru).

W literaturze filozoficznej traktującej o qualiach pojawiają się zazwyczaj trzy tematy, które teraz krótko omówimy.

(1) Problematykę qualiów eksplikuje się najczęściej poprzez odwołanie się do *eksperymentów myślowych*. Jednym z nich jest antyredukcyjny eksperyment mówiący o możliwości istnienia wspomnianych już zombie. Pod dwoma względami zombie stanowią wyzwanie dla fizykalistycznego traktowania qualiów.

(a) Jeżeli zombie są metafizycznie możliwe, wtedy istnieje argument pokazujący, że stany zjawiskowe nie są identyczne z obiektywnymi stanami fizycznymi. Przyjmijmy, że obiektywny, fizyczny stan F może (w sensie metafizycznego „może”) w jakimś zombi wystąpić bez zjawiskowego stanu Z . Intuicyjnie zakładamy jednak, że Z nie może wystąpić bez odczuwania Z . Ból nie może być odczuwany bez bólu, chociaż sam alarm bólowy może się pojawić z powodów patologicznych (ból fantomowy). Tak więc, F może wystąpić bez odczuwania Z , a Z nie może. Tym samym, na mocy zasady Leib-

niza (że jakieś x jest identyczne z jakimś y , wtedy i tylko wtedy gdy x i y posiadają te same własności), Z nie jest identyczne z F .

(b) Jeżeli jakiś osobnik fizycznie identyczny ze mną, umieszczony w tych samych warunkach zewnętrznych (w przeszłych i teraźniejszych), może nie mieć wrażeń zjawiskowych, wtedy fakty dotyczące wrażeń i odczuć, fakty o qualiach nie są na sposób *konieczny* zdeterminowane przez obiektywne fakty mikrofizyczne. Fizykalista nie może jednak dopuścić do czegoś takiego, gdyż musi on przynajmniej wierzyć, że stany mikrofizyczne determinują wszystkie fakty, że świat, który byłby dokładnie podobny do naszego pod wszystkimi mikrofizycznymi względami, musiałby być koniecznie podobny do naszego pod wszystkimi innymi względami (z tymi samymi górami, jeziorami, istotami odczuwającymi)²⁷.

(2) Kolejnym tematem powracającym w dyskusji o qualiach jest ich „przejrzystość” oraz zależność od introspekcji. Powoływanie się na introspekcję ma uzasadnić twierdzenie, że qualia są wewnętrznymi, nieintencjonalnymi jakościami doznaniowymi. Zdaniem eksternalisty M. Tye’a, introspekcja nie ukazuje nam jednak takich jakości (por. Tye 1996: 30–31). Przypuśćmy, że patrzymy na białą ścianę, na której widzimy kwadratową, niebieską plamę. Zdajemy sobie przy tym bezpośrednio sprawę z niebieskości i z kwadratowości jako obecnych w świetle. Odwróćmy następnie uwagę od tego, co widzimy przed sobą „w świetle” i skierujmy ją na nasze wzrokowe przeżywanie. Starajmy się skupić naszą uwagę na wewnętrznych cechach przeżywania, które odróżniają je od innych przeżyć. Zadanie to wydaje się niemożliwe do wykonania – nasze zdawanie sobie sprawy zdaje się przesuwać z przeżycia niebieskości i kwadratowości na niebieskość i kwadratowość jako zrealizowane razem w zewnętrznym przedmiocie. Według Tye’a:

„Ktoś przeżywa coś niebieskiego i kwadratowego jako cechę jakiejś zewnętrznej powierzchni, a introspekcja zdaje się nie odkrywać jakichś dodatkowych, odmiennych cech przeżycia ponad te, które przeżywa on w iluzji. Wzrokowe przeżycie zatem jest transparentne albo przejrzyste, jakim zresztą w ogóle jest świadomość zjawiskowa”²⁸.

²⁷ Fizykalistycznej odpowiedzi na argument z zombi udziela B. Loar, wychodząc z założenia, że: „phenomenal qualities () to be identical with physical-functional properties of the sort envisaged by contemporary brain science” (Loar 1990: 83).

²⁸ Tye 1996: 30–31. Według T. Metzinger: „Das die meisten phänomenalen Zustände durchsichtig sind, bedeutet demnach, daß wir sie nicht als phänomenale Zustände erleben, sondern daß wir durch diese Zustände hindurch schauen und ihren Gehalt und Modus der direkten Gegenbenheit wahrnehmen” (Metzinger 1996: 25).

Gdy patrzymy na plamę na ścianie, jesteśmy świadomi pewnych własności w świecie. Gdy zwrócimy uwagę na ukazywanie się nam tych własności, jesteśmy świadomi tych samych własności razem z faktem, że nasz stan mentalny je przedstawia; żadne nowe własności tego przedstawienia koloru niebieskiego nie ukazują się. W tym sensie wzrokowe przeżywanie jest transparentne czy przejrzyste (por. Tye 2003: 8). Przypadek taki zachodzi także wtedy, gdy mamy halucynacje albo śnimy, to znaczy, gdy przed nami nie ma żadnej realnej plamy na ścianie. Mamy po prostu przeżycie pewnej kolorowej i posiadającej jakiś kształt plamy. W tym przypadku nasze przeżycie jest błędnym przedstawieniem (*misrepresentation*). A gdy skierujemy swoją uwagę na swoje przeżycie, będziemy znowu widzieli przez to przeżycie te same jakości.

(3) Dennett zaznacza, że klasycznie *qualiom* przypisywano wspomniane cztery cechy²⁹, i traktuje tak rozumiane *qualia* jako część kartezjańskiego spojrzenia na umysł. Pod hasłem „*quining qualia*”³⁰ krytykuje on tak zwanych miłośników *qualiów* (*qualiophiles*)³¹, do których zalicza S. Shoemakera, N. Blocka, J. Searle’a i T. Nagla (por. Dennett 1991: 442; Block 1990: 53–54, 71; Shoemaker 1996a: 161–162). Miłośników *qualiów* łączy nadrzędne dążenie do funkcjonalistyczno-naturalistycznego opisu przeżyć³², chociaż, zdaniem G. Northoffa, wiążą oni *qualia* z perspektywą pierwszej osoby³³. Dla Blocka i Shoemakera *qualia* nie istnieją jako swoiście subiektywne obiekty, możemy więc wiedzieć, czy różni od nas osobnicy mają *qualia*, czy też ich nie mają. Możemy się bowiem odwołać do perspektywy trzecioosobowej, w której to perspektywie *qualia* są relacjami funkcjonalnymi. Nie możemy jednak wiedzieć, jakie określone *qualia* owi osobnicy posiadają (indywidualne *qualia* są wewnętrznymi, neurofizjologicznymi faktami) [por. Horgan 1984a: 464; Lycan 1990: 114, 117, 124–27]. Gdy chodzi o *qualia*, to

²⁹ Por. Dennett 1988: 2–3; por. Dretske 1994: 49–51; por. Dennett 1979: 94.

³⁰ „The verb «to quine» is even more esoteric. It comes from, [...] a satirical dictionary of eponyms: «quine, v. to deny resolutely the existence or importance of something real or significant» (Dennett 1988: 1).

³¹ Pod *qualiophiles* Dennett rozumie tych, „którzy nadal wierzą w istnienie *qualiów*” (Dennett 1991: 386). Według J. Levine *qualiophile* jest terminem wprowadzonym przez Dennetta „[...] for one who thinks there is more to qualitative character than functional role” (Levine 1997: 404; por. Dennett 1988: 4–5).

³² Por. Block 1978: 293. Opisując *qualia*, Shoemaker rezygnuje z rozróżnienia pomiędzy tym, co fizyczne, i tym, co funkcjonalne (por. Shoemaker 1982: 372–73).

³³ Por. Northoff 1995: 277–278. Zdaniem Shoemakera, nawet jeżeli powiedziano wszystko na temat, jak-to-jest-widzieć-czerwona-plamę, to nadal pozostaje jakaś niewypowiedziana reszta (por. Shoemaker 1991: 515).

Dennett proponuje swoistą kombinację albo też „closed feedback loop” pomiędzy fenomenologią auto- i heteronomiczną (por. Dennett 1982: 174, 160–163). Swoje spojrzenie na qualia nazywa on „redukcyjnym”, gdyż sprowadza ono qualia do zespołów dyspozycji. Qualia nie są niczym więcej jak dyspozycyjnymi cechami poznania. Stąd szczególnym celem jego ataku jest idea, że qualia zawierają w sobie coś niewypowiedzianego.

Odwołując w tej pracy do qualiów, skupiamy się na ich *jak to jest* być aspekcie (*jak to jest* widzieć czerwone, *jak to jest* słyszeć ton C[#]). Oddzielamy (przynajmniej pojęciowo) istnienie qualiów jako cech neuronalnych, od istnienia qualiów jako cech zjawiskowych. Qualia rozumiane jako cechy neuronalne mogą istnieć bez ich przeżywania. Pozostaje pytanie, co czyni je przeżywanymi (bowiem nie wystarcza do tego posiadanie jedynie neuronalnych cech).

Istnieje coś, co nazywamy byciem świadomym organizmem. Ta subiektywna cecha odpowiada holistycznemu rozumieniu świadomości. Jakiś organizm jest świadomy, jeżeli jest coś takiego jak *jak to jest* nim być. Podobnie, dany stan mentalny jest świadomym, jeżeli istnieje jakieś jakościowe *jak to jest* być w tym stanie. W rozdziałach IV i V zostaną przedstawione teorie świadomości zbudowane przez Dennetta oraz Nagla i krótko omówione ich teorie qualiów. W swoim ataku na qualia Dennett neguje ich нефизыкалność i pozbawioną możliwości błędu daność, natomiast według Nagla, qualia są jedynymi determinantami wspomnianego jakościowego charakteru przeżyć.

2.4.2. Samoświadomość

Twierdzenie, że dyskurs o świadomości podmiotu musi się odwoływać do jego samoświadomości, jest negowane z dwóch powodów: (a) niektóre prymitywne istoty zdają się nie posiadać samoświadomości, (b) przyjmuje się, że samoświadomość jest wyspecjalizowaną zdolnością, a tego rodzaju zdolność nie musi towarzyszyć wszystkim stanom świadomym. Odpowiadając na te kwestie można wskazać na dwa odmienne rozumienia samoświadomości: (A) świadomość siebie jako siebie, (B) posiadanie metapsychologicznych myśli.

(A) Ludzkie przeżycia zakładają istnienie świadomego podmiotu, który ma także przeżycie bycia sobą w sensie jego pojęciowej świadomości siebie samego. Samoświadomość jako świadomość siebie samego nie zakłada jedynie posiadania świadomości jakichś stanów, w tym także stanów rodzaju *jak to jest być*, lecz również posiadanie świadomości „przedmiotu”, który te stany posiada. Zdaniem N. Blocka, samoświadomy podmiot posiada pojęcie siebie samego i potrafi go używać myśląc o sobie (por. Block 1996: 547). Oznacza to równocześnie zdolność do odpowiedniego działania zgodnego

z otrzymaną informacją o sobie samym, a ponadto zakłada to pojęciową wiedzę o byciu świadomym siebie jako siebie. Tego rodzaju świadomość jest świadomością siebie samego jako podmiotu przeżyć. Jako ludzie jesteśmy zdolni do odnoszenia się do samych siebie, co jest związane z używaniem zamimka pierwszoosobowego. Jest sprawą kontrowersyjną, czy zdolność taką posiadają inne istoty – naturalne lub sztuczne. U niektórych wyższych ssaków możemy zaobserwować, że rozpoznają się w lustrze i wykazują zainteresowanie zależnościami pomiędzy ich własnymi ruchami, a ruchami odbitych w lustrze obrazów.

Tak rozumiana samoświadomość nie jest ani (i) świadomością otaczających nas obiektów, ani (ii) świadomością naszych własnych przedstawień. (i) Świadomość obiektów jest świadomością, jaką mamy o otaczającym nas świecie. Do tego świata należy także nasze ciało. (ii) Świadomość własnych przedstawień, to świadomość naszych stanów psychicznych. Dzięki tym przedstawieniom jesteśmy świadomi otaczającego nas świata. Możemy porównywać ze sobą różne modalności, poprzez które jest nam dany ten czy inny przedmiot: *jak to jest* widzieć dany przedmiot można porównać z tym, *jak to jest* słyszeć ten sam przedmiot³⁴.

Ktoś posiadający świadomość siebie zdaje sobie sprawę nie tylko z siebie jako podmiotu, ale – używając terminologii Kanta – zdaje sobie z tego sprawę jako pojedynczy, wspólny podmiot pewnej liczby przedstawień i pojedynczy podmiot konkretnych działań³⁵. Jest to inny rodzaj świadomości aniżeli świadomość przedmiotów czy przedstawień. Gdy bowiem jesteśmy świadomi siebie jako podmiotu przeżyć, czy jako podmiotu działań, to nie jesteśmy świadomi jakiejś ilości przedmiotów. Gdy ktoś jest świadomy siebie samego jako podmiotu, to nie jest on świadomy siebie jako przedmiotu³⁶. Ra-

³⁴ Różnice pomiędzy percepcją wzrokową i słuchową omawia A. Klawiter (1999a: 329–336).

³⁵ Zgodnie z Kanta krytyką pierwszego paralogizmu czystej psychologii: „Tak wiele brak do tego, aby można było te własności [że jako istota myśląca trwam sam dla siebie, że nie ginę] wywnioskować wyłącznie z czystej kategorii substancji, że przeciwnie, musimy na podstawie doświadczenia założyć trwałość pewnego danego [nam] przedmiotu, jeżeli chcemy zastosować do niego pojęcie substancji, empirycznie użyteczne. Otóż nie przyjęliśmy żadnego doświadczenia za podstawę naszego twierdzenia, lecz wywnioskowaliśmy je jedynie z pojęcia odniesienia, jakie posiada wszelkie myślenie do Ja jako wspólnego podmiotu (*als das gemeinschaftliche Subjekt*), któremu przysługuje” (Kant 1957: A 350).

³⁶ „Gdyby uzyskanie czystego poznania rozumowego istoty myślącej w ogóle miało być możliwe, to owo Ja musiałoby być daną naoczną, która, jako że byłaby założona przy myśleniu w ogóle (przed wszelkim doświadczeniem), dostarczałaby, jako naoczność, zdań syntetycznych *a priori*. Jednakże owo Ja w równej mierze nie jest daną naoczną, jak pojęciem jakiegokolwiek przedmiotu” (Kant 1957: A 382).

czej jest on świadomy i wie, że jest świadomym jednej i tej samej rzeczy poprzez jakąś liczbę przedstawięń.

(B) Samoświadomość, to posiadanie metapsychologicznych myśli – od takiej definicji samoświadomości wychodzi R. J. Gennaro³⁷. Jako podstawę dla swoich analiz samoświadomości przyjmuje on D. Rosenthala „higher-order thoughts” (HOTs) teorię świadomości³⁸. Myślenie o własnych stanach mentalnych należy, według Gennaro, do definicji samoświadomości, a posiadanie wspomnianego rodzaju myśli wydaje się koniecznym i wystarczającym warunkiem mówienia o samoświadomości. Nic jednak nie wymaga, żeby posiadanie wspomnianych metapsychologicznych myśli było świadome. Są one nieświadomymi myślami i nie ma żadnej racji, aby negować istnienie nieświadomych myśli metapsychologicznych. HOTs są intencjonalnymi stanami reprezentującymi podmiot jako posiadający stany mentalne. Nakierowanie myśli wyższego rzędu (HOTs) na konkretny stan mentalny czyni go stanem świadomym. Stany mentalne nie będące celem HOTs nie są stanami świadomymi, same HOTs też nie są świadome. Stany mentalne mogą więc być świadome lub nieświadome. Gennaro pisze: „Jeżeli mam rację odnośnie do tego, co czyni stan mentalny świadomym, wówczas posiadanie stanów świadomych obejmuje samoświadomość” (Gennaro 1996: 17; por. s. 124–126).

Twierdzenie to należy także odnieść do stanów zjawiskowych. W tym sensie samoświadomość nie jest świadomością siebie jako siebie.

Można powiedzieć: świadome organizmy posiadają subiektywną perspektywę, która często nie jest im świadoma. Mając świadome przeżycia nie zdają sobie sprawy ze sposobu, w jaki przedmioty są przedstawiane ich świadomemu umysłowi. Ich pojęciowy aparat nie jest im (we wspomnianym sensie) świadomy, nie są go oni świadomi, lecz są świadomi jedynie przedmiotów.

Zgodnie z holistycznym rozumieniem świadomości, posiadanie przez nas stanów przeżyciowych nie wyklucza jakiegoś rodzaju samoświadomości w sensie (A) lub (B). To, co czyni te stany zjawiskowymi zależy od tego, czy należą one do ogólnego stanu *jak to jest* być świadomym.

³⁷ Por. Gennaro 1996: 16–21. Gennaro nawiązuje do J. Locke’a i jego tezy o spostrzeganiach w nas samych operacji dokonywanych przez nasz własny umysł (Locke 1955: II, kap. 1, par. 4).

³⁸ Według tej teorii stany mentalne są świadome, gdy reprezentujemy samych siebie jako będących w tych stanach, a sama ta reprezentacja jest możliwa dzięki myślom wyższego rzędu (HOTs) [por. Rosenthal 1996a: 426–435]. Stubenberg stwierdza krytycznie: „Resistance against the HO[T] account of state consciousness is grounded in a tacit commitment to the intrinsicness of consciousness. Given this view of consciousness, it is mysterious how consciousness arises in a state merely because some other mental state represents it” (Stubenberg 1998: 188; por. s. 188–199).

2.5. Podsumowanie

Nakreślone ramy pojęciowe oraz metodologiczne, w których badamy wspomniane w tym rozdziale aspekty subiektywności, zakładają realizm ontologiczny (istnienie rzeczywistości jest niezależne od postrzegania jej przez świadomy podmiot) i epistemologiczny (podmiot może tę rzeczywistość – zasadniczo – poznać). Subiektywność świadomości, jej jedność i cechy jakościowe, odróżniają ją istotnie od innych przedmiotów w świecie. Chociaż cechy te są empirycznie powiązane ze sobą i ze świadomością dostępną, można je traktować oddzielnie ze strony pojęciowej.

Intuicyjnie trudno sobie wyobrazić, jak jakiś stan mógłby być subiektywny bez odznaczania się holistyczną jednością. Gdyby, powiedzmy, nasza świadomość rozpadła się na dwie części, to raczej nie mówilibyśmy o pojedynczym podmiocie z dwoma stanami świadomości, lecz o dwóch podmiotach świadomości.

Jako model holistycznej świadomości zjawiskowej można przyjąć zjawiskową świadomość w fazie snu REM. Model ten jest ograniczony, pozwala jednak na: (a) jakiś stopień odizolowania świadomości zjawiskowej od docierających do podmiotu sensorycznych wejść i powodowanych przez niego motorycznych wyjść, (b) empiryczne porównanie podobieństw i stwierdzenie różnic pomiędzy „śpiącym” i „czuwającym” mózgiem.

Wyśniony świat nie pokazuje nam oczywiście, jak nasz mózg wywołuje świadomość, chociaż możemy obrazować co się w nim dzieje, gdy śnimy. Sny stwarzają iluzję, że przeżywane przez podmiot zdarzenia nie zachodzą w nim i jego mózgu, lecz w zewnętrznym świecie, że jest on obecny w innym świecie aniżeli jego aktualnie śpiące ciało.

Świadomość w sensie subiektywnych przeżyć okazuje się być istotną w opisie zależności umysł–mózg, której nauki o systemie nerwowym nie mogą i nie powinny ignorować. Satysfakcjonujące materialistyczne rozwiązanie problemu świadomości powinno by umieć wykazać, że subiektywność jest konieczną i wystarczającą cechą neuronalnej organizacji mózgu (to znaczy, że jest niemożliwe, aby mózg wywoływał świadomości w inny sposób, aniżeli obecnie przyjmujemy, że to czyni, por. 5.6.1).

Subiektywność świadomości nakłada mocne wymagania na teorię świadomości. Powstaje bowiem pytanie: Czy nasza świadomość jest subiektywna na sposób *konieczny*? Jeżeli tak faktycznie jest, to poszukiwana teoria świadomości powinna móc wyjaśnić tak rozumianą subiektywność. Przedstawione w tym rozdziale filozoficzno-psychologiczne rozumienie problemu subiektywności zakłada potoczne intuicje, które muszą być uwzględnione w konstruowanej, nowej teorii świadomości. Poszukiwana, nowa teoria świadomości powinna:

(a) Umieć określić, co jest wymagane ze strony mózgu dla powstania świadomości i jaka jest zależność pomiędzy świadomością a mózgiem.

(b) Pozwolić odróżnić procesy fizjologiczne bezpośrednio warunkujące świadomość od tych, które tego nie czynią. Teoria powinna także określić, jakie rodzaje stanów neurologicznych i jaki poziom strukturalny mózgu należy brać pod uwagę, gdy mówimy o holistycznej świadomości; jaka płaszczyzna mózgu jest właściwa dla mówienia o świadomości: płaszczyzna neuronów i synaps, czy płaszczyzna anatomicznych części mózgu?

(c) Pomóc zrozumieć, w jakim sensie zwierzęta i sztuczne inteligencje są (mogą być) subiektywnie świadome.

Do (a) i (b) wracamy w rozdziale VII pytając o neuronalne odpowiedniki zdarzeń i stanów świadomych. Naukowe poszukiwanie tego typu korelacji nie zależy, przynajmniej na razie, od tego, czy jest ono prowadzone z pozycji dualistycznych, materialistycznych czy, jak u C. McGinna, z pozycji traktujących świadomość jako coś „zagadkowego”. Niezależnie bowiem, które z tych metafizycznych założeń zostanie przyjęte, jesteśmy zainteresowani w określeniu (a w końcu w zidentyfikowaniu) chociażby tych procesów neurofizykalnych, które wystarczają do mówienia o świadomości albo o jej minimalnych przejawach. Materialista stara się zidentyfikować procesy neurologiczne ze stanami świadomości zjawiskowej³⁹, a dualista szuka czegoś dodatkowego, towarzyszącego świadomości.

Subiektywność można rozumieć w sensie (i) *jak to jest* być człowiekiem, oraz (ii) indywidualnego *jak to jest* być mną. Subiektywność rozumiana jak w (ii) jest indywidualnym wyrazem świadomości. Jej częściowego wyjaśnienia można szukać w genetyce, według której tylko kilka procent genów naszego genomu różni nas od genomu innych istot, a jednak wystarczy to, aby każdy z nas był inny. Czym jest zatem subiektywność w sensie (ii)? Z punktu widzenia przyrodniczego można ją potraktować jako wyraz odrębności osobniczej każdego z nas uwarunkowanej genetycznie i kulturowo. Odrębność ta wiąże się chociażby z pytaniem: dlaczego każdy z nas ma swój własny immunofenotyp?

³⁹ O tego rodzaju „przyczynowe” wyjaśnienie pyta Searle: „What exactly are the neurobiological correlates of conscious states (NCC), and which of those correlates are actually causally responsible for the production of consciousness?” (Searle 2000: 558).

ROZDZIAŁ TRZECI

MIEJSCE ŚWIADOMOŚCI W WYBRANYCH TEORIACH UMYSŁU

3. Wprowadzenie: subiektywna świadomość a mózg

„Najgłębszą przyczyną lęku przed świadomością jest to, że świadomości rzeczywiście przysługuje owa rażąca własność subiektywności” (Searle 1999: 84).

W latach rozkwitu behawioryzmu unikano używania terminu „świadomość” i odwoływano się, jeżeli w ogóle, raczej do „umysłu”. W pierwszej części obecnego rozdziału omawiamy powstałe w tym czasie teorie umysłu. Spotykana we współczesnych teoriach umysłu konfrontacja holistycznego rozumienia świadomości z jej neurofizjologicznym rozumieniem ukazuje wspomnianą już „lukę w wyjaśnianiu” (*explanatory gap*)¹: na obecnym poziomie rozwoju nauki nie potrafimy sobie chociażby przedstawić, jak neuronalne substraty naszego bólu mogłyby wyjaśnić, dlaczego odczuwam go w taki, a nie inny, subiektywny sposób.

Niektórzy filozofowie analityczni twierdzą, że luki tej nie da się przezwyceńczyć, gdyż istnieje odpowiadająca jej luka ontologiczna dotycząca zdarzeń w świecie. Przeżycia i odczucia mają nieredukowalnie subiektywne, niefizykalne jakości (por. Chalmers 1996: 127–130). Inni uważają, że pewne fizykalne własności czy stany same są nieredukowalnymi subiektywnymi jedno-

¹ Por. Levine 1983: 358–359. Tak rozumiana luka wyjaśniająca odnosi się głównie do niemożliwości redukcji jednego szeregu pojęć do innego (np. poprzez analityczne definicje) i nie musi oznaczać, że istnieje „luka ontologiczna” (że jakiś rodzaj cech lub faktów nie może być zredukowany do innego rodzaju).

stkami (por. Searle 1999: 135). Są tacy, którzy utrzymują, że luka wyjaśniająca może zostać przewyciężona, dzisiaj jednak brak nam pojęć, aby scalić perspektywę obiektywną i subiektywną². Kolejna materialistyczna wersja zakłada, że pojęcia dotyczące jakości zjawiskowych są po prostu pojęciami indeksykalnymi, z tym że zastosowanymi do stanów zjawiskowych, poznawanych przez introspekcję³. Pojęcia dotyczące cech zjawiskowych są pojęciami pierwszoosobowymi, różnymi od innych pojęć.

W obecnym rozdziale i w dwóch następnych odwołujemy się do wprowadzonego poniżej podziału teorii świadomego umysłu. Teorie dualistyczne jedynie wspominamy (por. 3.1). Szerzej zostaną omówione klasyczne już dzisiaj, monistyczno-materialistyczne teorie umysłu, których twórcy twierdzą, że wspomianej luki w wyjaśnianiu nie ma, bo świadomość można albo zredukować do obiektów neurologicznych, albo ją wyeliminować.

Cechą wspólną przedstawionych poniżej rozwiązań materialistycznych jest rezygnacja z perspektywy pierwszoosobowej oraz – z nielicznymi wyjątkami – eliminacja szczegółowych analiz neurologicznych.

Poszukiwane rozwiązanie problemu subiektywności świadomości musi wyjść od uwzględnienia argumentacji filozoficznej i neurologicznej. Znane z historii filozofii ujęcia subiektywności świadomości (por. 1.3) należy skonfrontować z neurofizjologicznymi badaniami mózgu. Te ostatnie, zdaniem neurologa B. J. Baarsa, wskazują na fundamentalną rolę świadomości w integracji funkcji mózgu⁴. Odpowiadając na pytanie o relację pomiędzy zdarzeniami w mózgu a subiektywnymi przeżyciami, i uwzględniając zależność mózgu od reszty ciała organizmu, można sformułować następujący tryemat:

- III.1** [1] Stany i zdarzenia w mózgu określają stany świadomości zjawiskowej.
 [2] Stany świadomości zjawiskowej określają stany oraz zdarzenia w mózgu i *vice versa*.
 [3] Istnieją jedynie stany i zdarzenia w mózgu.

Tezy [1]–[2]–[3] wskazują na lukę w wyjaśnianiu, a tym samym na centralny problem związany z naukowym wyjaśnieniem świadomości: na zależ-

² Na przykład C. McGinn przyjmuje, że może się okazać, iż qualia są faktycznie obiektami fizykalnymi, ale na razie nie posiadamy jasnej idei, czym one są (por. McGinn 1999: 31, 46, 60, 84).

³ Por. Lycan 1987: 54–58. Lycan przeprowadza analizy indeksykalnie rozumianego pojęcia „ja”.

⁴ Por. Baars 2003: 102–110; 2002: 47–50. Swoje ujęcie ścisłej relacji pomiędzy świadomością i mózgiem Baars nazywa „conscious access hypothesis” i skupia się głównie na przypadkach świadomości dostępu.

ność pomiędzy jej subiektywnym aspektem, a stanami fizycznymi mózgu. Teoria świadomości stoi przed zadaniem wyjaśnienia, jak owe subiektywne stany świadomości wbudować w obiektywny świat naukowo stwierdzalnych procesów w mózgu. Celem objęcia szerokiego spektrum teorii używamy w tezach [1], [2] terminu „określają” (bardziej ogólnym terminem jest „korelują”), a nie „przyczynują” albo „wywołują”. W tezach [1]–[2]–[3] chodzi o holistyczne bycie świadomym oraz o poszczególne stany świadomości zjawiskowej.

[1]–[2]–[3] pozwalają na uporządkowanie współczesnych analitycznych teorii umysłu i świadomości. Do grupy [1] zaliczamy klasyczne teorie dwu aspektów (*dual-aspect theory*) i współczesne teorie epifenomenalistyczne, nazywane także „słabym dualizmem” (D. Chalmers, F. Jackson). Epifenomenalizm dopuszcza tak zwaną *upwards* przyczynowość (mózg oddziałuje na świadomy umysł), negując tak zwaną *downward* przyczynowość (oddziaływanie umysłu na mózg). Używane w teoriach epifenomenalistycznych argumenty są związane z możliwością istnienia zombie (świadomość wprawdzie istnieje, ale nie posiada wpływu na zachowania, tzn. nie powoduje żadnej zmiany w fizycznej rzeczywistości). Sugeruje to możliwość oddzielenia świadomości od mózgu i jej eliminację. Do intuicji zombie odwołują się także D. Dennett i T. Nagel. Zdaniem tego ostatniego subiektywność nie może być uchwycona poprzez jakąkolwiek z obecnie znanych, redukcyjnych analiz świadomych zjawisk mentalnych, gdyż wszystkie one są również logicznie zgodne w przypadku jej braku. Pytaniem nadal pozostanie, czy możliwość istnienia zombie należy traktować jako element argumentacji.

Do grupy [2] należą teorie interakcjonistyczne i dualistyczne (R. Kartezjusz, G. Frege, J. Eccles), nazywane także „mocnym dualizmem”: subiektywne cechy świadomości są zasadniczo różne od aktywności mózgu. Dualista musi jednak również odpowiedzieć na pytania: (a) Jeżeli świat zawiera subiektywne elementy, to w jaki sposób współdziałają one z fizycznymi (czasoprzestrzennymi) obiektami?⁵ (b) Jakie (dzisiaj nieznanne) zasady powodują wyłanianie się (emergowanie) tego rodzaju subiektywnych elementów?

Grupę [3] stanowią teorie redukcjonistyczno-materialistyczne, a także funkcjonalistyczne i eliminatywistyczne (J. Smart, P. Churchland, D. Dennett). Odpowiadają one najmniej potocznemu rozumieniu świadomości,

⁵ Według M. Kinsbourne’a, „If consciousness has causal efficacy, as some theorists have held (Eccles, Sherrington, Sperry) than one must explain how it interacts with neural systems” (Kinsbourne 1997: 336–337). Wymaganie stawiane przez Kinsbourne’a jest bezpodstawne. W biologii opisano występowanie wielu zjawisk, a dopiero po latach wyjaśniono sposób ich zachodzenia.

a tym samym jej subiektywności⁶. Podstawowym założeniem materializmu redukcyjnego jest teza o przyczynowej zamkniętości obszaru fizycznego.

Niektórych materialistów trudno jednoznacznie przyporządkować. Próby połączenia perspektywy pierwszoosobowej z tezami materialistycznymi [1] i [3] podejmują między innymi: C. McGinn (w jego „naturalizmie”), J. Searle (w jego „biologicznym naturalizmie”) oraz T. Nagel (poszukujący tak zwanej obiektywnej fenomenologii). Z jednej strony, według McGinna, świadomość jest czymś „tajemniczym”, z drugiej jednak strony należy ją dostosować do porządku naturalnego⁷.

Prowadzona przez J. Searle’a krytyka materializmu wynika z jego rozumienia zjawisk mentalnych jako własności mózgu. Zjawiska mentalne są tak samo zjawiskami biologicznymi, jak procesy zachodzące w żołądku (czyli nie-[2])⁸. Proponowane przez niego „emergencyjne” rozumienie świadomości unika dualizmu i przypomina o eksperymentalnej ważności perspektywy pierwszoosobowej, której pomijanie wynika z obecnej w filozofii i w naukach szczegółowych tendencji do obiektywizowania (por. Searle 1999: 26–27; 1987: 145). Nagel jest zaliczany zarówno do anty-, jak i do materialistów⁹: do antymaterialistów na podstawie jego nieredukcjonizmu, do materialistów na podstawie jego „obiektywnej fenomenologii” (por. rozdział V).

Jako przedstawiciela rozwiązań pośrednich, gdy chodzi o problemy subiektywności, można wymienić S. Shoemakera, który łączy funkcjonalistyczne teorie umysłu z założeniem, że subiektywność oraz qualia nie są wyjaśnialne przez żadną taką teorię (Shoemaker 1991: 510–524; 1993: 906–909). Wszystkie stany funkcjonalne są realizowane przez stany fizyczne, a te mają własności jakościowe. Ponieważ ta sama funkcjonalna organi-

⁶ „Kartograficzne” zestawienie poszczególnych teorii świadomości podaje F. Varela, umieszczając je pomiędzy fenomenologią, funkcjonalizmem, redukcjonizmem i „misterianizmem” (por. Varela 2000: 339–341).

⁷ Por. McGinn 1999: 46–50. Trudności związane z teoretycznym ujęciem perspektywy pierwszoosobowej pochodzą, według McGinna, z dwóch źródeł: (1) z konstytutywnych ograniczeń ludzkiego poznania. Pojęcia o świadomości tworzymy bowiem poprzez introspektywną analizę swoich własnych, wewnętrznych stanów. (2) Naszym subiektywnym przeżyciom brak pojęciowego połączenia z mózgiem. Introspekcja nie przedstawia świadomości jako aspektu fizycznego mózgu. McGinn sugeruje, że „trudny” problem (w jaki sposób zdarzenia neuronalne wywołują świadome przeżycia) posiada wyjaśnienie, nie jest ono jednak dostępne naszemu poznaniu. Por. McGinn 1991: 58, 60, 68.

⁸ Searle stwierdza: „Świadomość jest emergentną czy też wyższego rzędu własnością mózgu, [...] w tym samym sensie, w jakim ciekły stan skupienia jest emergentną własnością wyższego rzędu cząsteczek H₂O” (Searle 1999: 32).

⁹ Na sposób antymaterialistyczny tezy Nagla interpretuje G. Madell, powołując się głównie na jego artykuł *Fizykalizm* (por. Madell 1981: 23, 71).

zacja może być zrealizowana na różne fizykalne sposoby, stąd opis funkcjonalny ujmuje wprawdzie wzajemne związki pomiędzy stanami, nie mówi jednak nic o tym, jak podmiot odczuwa te stany jako swoje stany wewnętrzne. Tym samym odwoływanie się do organizacji funkcjonalnej staje przed trudnością związaną z ujęciem odczuwania, jak i z ujęciem qualiów posiadających nieintencjonalne własności. Te ostatnie leżą poza wyjaśnianiem proponowanym przez teorie funkcjonalistyczne¹⁰. Dlatego, według Shoemakera, w wyjaśnianiu tego rodzaju własności zawsze pozostanie jakaś niewypowiedziana reszta.

3.1. Dualizm versus materialistyczny monizm

Tak jak z jednej strony tylko niektórych współczesnych filozofów analitycznych można nazwać behawiorystami¹¹, tak też, z drugiej, jedynie poglądy niektórych filozofów – analitycznych i nieanalitycznych – można zakwalifikować do słabego dualizmu (na przykład „naturalistyczny dualizm” D. Chalmersa). W ostatnich dziesięcioleciach obserwujemy radykalną zmianę w podejściu do dualizmu. Warto wspomnieć, że jeszcze dla G. Ryle’a dualizm był „oficjalną doktryną”. Dzisiejsze, sceptyczno-krytyczne nastawienie do dualizmu nie wynika bezpośrednio z prowadzonych badań naukowych, lecz przesłanek pośrednich. Należą do nich daleko idące, monistyczne ekstrapolacje i interpretacje wyników badań naukowych, które przypisują dualizmowi antynaukowość¹². Z tego względu dualizm jest odrzucany między innymi przez D. Dennetta: „Moim zdaniem, ta fundamentalnie antynaukowa postawa dualizmu, najbardziej go dyskwalifikuje powodując, dlaczego w tej książce przyjmuję pozornie dogmatyczną regułę, że dualizm musi być odrzucony *za wszelką cenę*” (Dennett 1991: 37).

¹⁰ Shoemaker stwierdza dosłownie: „Qualia are supposed to be nonintentional features of experiences that somehow underlie their intentional features” (Shoemaker 1991: 515).

¹¹ „Behaviorism is dead”, tym zdaniem G. Strawson rozpoczyna swoją książkę o świadomości (por. Strawson 1994: xi). Zbudowaną przez siebie teorię świadomości Strawson nazywa „znaturalizowanym kartezjanizmem”. Łączy w niej podstawowe cechy materializmu z tezą, że jedynymi odmiennymi od materialnych zjawiskami są świadome przeżycia.

¹² Zdaniem neopragmatyka R. Rorty’ego: „Postawa, dla której dualizm kartezjański jest dzisiaj tak niepopularny, nie wynika z odwoływania się do skutecznych metod współczesnej filozofii analitycznej, ale po prostu z tego, że czytamy w „Life” albo w „The Scientific American” o cerebryalnym umiejscowieniu emocji, myśli czy wrażeń zmysłowych, o ich wytwarzaniu przez wprowadzenie elektrod i tym podobnych urządzeń” (Rorty 1972: 218).

Klasyczny dualizm kartezjański jest krytykowany nie dlatego, że jego argumenty są uważane za metodologicznie wadliwe, lecz dlatego, iż przyjmuje istnienie niematerialnych obiektów, niepasujących do współczesnego, naturalistyczno-naukowego obrazu świata. Postulowane przez dualistów oddziaływanie obiektów niematerialnych na neuronalne połączenia w mózgu zachodzi – zdaniem monistycznych materialistów – bez udziału fizycznie rozumianej energii. Tego rodzaju oddziaływanie – nazywane często „przyczynowością skierowaną ku dołowi” albo „makrodeterminacją” – przeczy fizycznej zasadzie zachowania energii.

Pytając o możliwości obiektywnego opisu stanów subiektywnych T. Nagel stawia dualizmowi dwa zarzuty: (a) postulowanie istnienia dodatkowej, niematerialnej substancji, bez pokazania, w jaki sposób substancja ta potrafi spełniać funkcje podmiotu (być nosicielem podmiotowych cech), (b) brak wyjaśnienia (podobnie jak nie czyni tego fizyczna koncepcja rzeczywistości), w jaki sposób obiektywnie opisywalnej rzeczywistości mogą być przypisywane stany subiektywne (por. Nagel 1997c: 37–42).

Według innych materialistów (D. M. Armstrong, J. Smart), współczesna nauka przyjmuje, że ludzki podmiot nie jest niczym więcej jak tylko fizyczno-chemicznym mechanizmem. Prywatne dane zmysłowe nie istnieją, istnieją jedynie zdarzenia w mózgu i połączone z nimi zachowania. W naukowym obrazie świata znajdujemy tylko zespoły fizycznych elementów, co daje obraz świata niekompatybilny z dualizmem (przez „fizyczne” rozumie się tutaj najczęściej te obiekty i własności, których istnienie postuluje współczesna fizyka teoretyczna). Zakłada się, że podmiot i jego świadome funkcjonowanie są wyjaśnialne w terminach naturalistycznych, zaś sama świadomość jest produktem powstałym na drodze ewolucji z niższych form życia, które z kolei powstały z materii nieorganicznej. Stąd teza o emergencji, że jeżeli to, co mentalne, jest czymś ponad i poza tym, co opisywane fizycznie, to albo już jest ono obecne w świecie nieorganicznym, albo wyłonilo się (wyemergowało) z obiektów tegoż świata.

Materializm ontologiczny jest tezą, że nie tylko wszystkie mikro- i makroobiekty są fizyczne, ale także wszystkie ich własności są fizycznie opisywane. Jeżeli cechy te nie są natury fizycznej, jak to się może zdarzyć przy cechach funkcjonalnych lub logicznych, to są takimi cechami, których opis może zostać wydedukowany z opisów fizycznych w połączeniu z fizycznymi prawami. Pogląd, że istnieją нефизyczne własności mentalne, chociaż wszystkie obiekty są fizyczne, jest zaliczany do tak zwanego słabego dualizmu lub do kierunku zwanego materializmem cech (w przeciwieństwie do tego silny dualizm twierdzi, że istnieją cechy mentalne i obiekty mentalne; umysł i mózg mogą istnieć oddzielnie jako dwie substancje).

Należy zauważyć trzy ściśle ze sobą związane zagadnienia.

(1) Wielu materialistów, w tym także Dennett, wyraźnie neguje dualizm przechodząc szybko do mówienia o procesach neurofizjologicznych, które „generują” albo „powodują”, albo wręcz „są” subiektywnymi zjawiskami. Tego typu negacja ma jednak sens tylko z dualistycznego punktu widzenia.

(2) Materialiści muszą płacić wysoką cenę za propagowaną przez siebie eliminację subiektywnych stanów umysłu. Eliminacja taka jest związana chociażby z pytaniem o status wspomnianej już psychologii potocznej (*folk psychology*), do której z kolei odwołuje się filozoficzno-psychologiczny obraz świata i podmiotu. Czymś niewątpliwym w ramach psychologii potocznej jest to, że istnieje jakaś jaźń, źródło aktów wolitywnych, poznania, działań i subiektywnych przeżyć¹³. Także w wielu dziedzinach nauki potocznie rozumiana jaźń oraz jej stany są czymś niekwestionowalnym. Powstają pytania czy rozwój nauki zmusza do odrzucenia koncepcji potocznej jaźni oraz, czy jaźń da się przedstawić jako konstrukt teoretyczny?¹⁴

(3) Następny problem leży w samym rozumieniu psychologii potocznej: (i) czy jest to „teoria” o tym, co potocznie wiemy na temat jaźni? (ii) czy też chodzi o potoczną, psychologiczną praktykę, w którą każdy z nas jest codziennie zaangażowany? W (ii) opisujemy ludzi jako posiadających stany psychologiczne, wyjaśniamy i przewidujemy ich zachowania, odwołując się do tychże stanów¹⁵. Eliminatywni materialści rozumieją psychologię potoczną jako zawodną prototeorię, mającą wyjaśnić ludzkie zachowania. Dodatkowo wskazują oni na fakt, że istnieją całe obszary ludzkich zachowań, o których psychologia potoczna milczy, jak na przykład zachowania ludzi z poważnymi uszkodzeniami mózgu. Naukowo akceptowalna teoria – według eliminatywistów – powinna proponować empirycznie sprawdzalne prawa, pozwalające na przewidywanie zmian, które w określonych warunkach by ją potwierdziły (rzecz miałaby się podobnie jak w fizyce, gdzie znajomość równań opisujących prawo grawitacji, pozwala na obliczenie przyspieszenia ciała w polu grawitacyjnym).

Problem leży jednak w tym, że psychologia potoczna i rozwinięte na jej podłożu teorie filozoficzno-psychologiczne nie zakładają ścisłych, empirycz-

¹³ Por. Kałuża 1992: 78. Według H. Putnama, „common-sense” spojrzenie na świat jest tak samo uprawnione, jak spojrzenie naukowe (por. Putnam 1999a: 23).

¹⁴ Próbę takiego przedstawienia podaje W. Sellars (1963: 38–40). Stwierdzenia materialistów eliminatywnych dotyczących „psychologii ludowej” przedstawia S. Judycki (1997: 37–39).

¹⁵ Swoją krytyką tak zwanego mitu daności (*the myth of the given*) W. Sellars przyczynił do się powstania tezy, że nasze potoczne rozumienie stanów mentalnych można traktować jako prototeorię umysłu (por. Bremer 1997: 70–80).

nych praw, a testowalne w ich ramach korelacje są trudno falsyfikowalne za pomocą eksperymentów (na przykład to, że osoby w danych warunkach zachowują się generalnie w określony sposób, że działają według takich, a nie innych motywów, pragnień, przekonań). Należy dodać, że jeżeli ktoś odrzuca obecną w psychologii i filozofii kategorię przekonań, to *implicite* odrzuca także ideę prawdy, jak dalece ta jest oparta na zdaniach, które są o tyle prawdziwe lub fałszywe, o ile wyrażają prawdziwe lub fałszywe przekonania.

3.2. *Behawioryzm logiczny i nomologiczny*

Teorie behawiorystyczne nie przyjmują introspekcji jako źródła poznania stanów mentalnych. Jedynym rodzajem poznania, jaki posiada podmiot, mówiący o swoich własnych stanach mentalnych, są jego zewnętrznie obserwowalne zachowania. Poniżej w punkcie (A) przedstawimy krótko teorie logicznego i nomologicznego behawioryzmu, a w punkcie (B) ich stosunek do związanych z subiektywnością wypowiedzi pierwszoosobowych.

(A) Radykalną formą behawioryzmu jest „behawioryzm logiczny” („filozoficzny”) z jego tezą, że znaczenie terminów mentalnych może być zdefiniowane w terminach odnoszących się do: zachowań zewnętrznych i do dyspozycji do zachowań (por. Judycki 2002: 276–277). Istotą behawioryzmu logicznego jest semantyczna teza, że każde zdanie, w którym występują terminy mentalne, da się przetłumaczyć na równoważnościowe zdanie, w którym występują jedynie terminy fizykalne.

Przypisywać podmiotowi stany mentalne, to przypisywać mu określone dyspozycje do zachowań (por. Lowe 2000: 42–43). „Behawioralna dyspozycja” jest rozumiana jako tendencja lub skłonność podmiotu do zachowywania się w określony sposób, w określonych warunkach. Na przykład, przypisanie komuś bólu żołądka jest po prostu przypisaniem mu dyspozycji do skręcania się, sięgania po krople przeciwbólowe itp. Lista tego rodzaju zachowań jest otwarta. Nie ma bowiem ograniczeń co do sposobów, w jaki ktoś może okazywać ból żołądka. Podobne analizy prowadzi się, gdy chodzi o tak zwane postawy propozycjonalne (sądy, przekonania). Moje przekonanie, że pada deszcz, objawi się w tym, że wychodząc z domu zabiorę parasol. Także w tym wypadku lista zachowań jest otwarta.

Otwartość wspomnianych list stanowi jednak problem dla logicznego behawioryzmu: Kto, albo co, łączy i jednoczy wszystkie te zachowania? Może być to jedynie fakt, że są one zachowaniami, których ktoś racjonalnie oczekuje od podmiotu odczuwającego ból albo od podmiotu niemającego, że pada deszcz. Wydaje się, że abyśmy mogli generować kolejne pozycje występujące na takich listach, musimy już wiedzieć, co to znaczy przypisać ko-

muś ból lub przypisać mu mniemanie, że pada deszcz. Określone przekonania wyjaśnią jakieś działanie tylko w kontekście określonych pragnień. Dlatego nie można owych list używać do wyjaśnienia, co to znaczy przypisywać komuś ból lub przekonanie o rodzaju pogody.

W behawioryzmie logicznym identyfikacja stanów mentalnych ze zjawiskami fizycznymi (z zachowaniami) wynika z analizy znaczenia terminów mentalnych. Zamiast o „identyfikacji” mówi się także o tym, że zachowania „konstytuują” stany mentalne. Oczywiście, ból jest występującym powszechnie zjawiskiem. Gdy ktoś odczuwa ból – coś rzeczywiście się wydarza. Konsekwentnie nie może to być jakaś niezaktualizowana dyspozycja do zachowania. Co więcej, ból jest raczej przyczyną rzeczywistego czy domniemanego zachowania bólowego. Rwący ból odczuwany przez podmiot, który nigdy go nie manifestuje w jakimś konkretnym zachowaniu, nie jest identyczny z dyspozycją do odpowiedzi: „czuję rwący ból”, lecz jest raczej *przyczyną* tej dyspozycji i jako taki wyjaśnia, dlaczego w ogóle podmiot miał taką dyspozycję.

Ponadto, podmiot może mieć ból żołądka lub może mniemać, że pada deszcz i nie okazywać zachowań, wymienionych na wspomnianych listach. To, jak podmiot się zachowa w danych warunkach, nie zależy tylko od jego momentalnego odczucia lub przekonania, ale od wielu innych nastawień. Może on być przeciwnikiem zażywania lekarstw, może nie chcieć nosić parasola. Tym samym nie ma zachowania, które byłoby jedynie charakterystycznym zachowaniem kogoś, posiadającego określone stany wrazeniowe lub propozycjonalne (por. 3.3.5).

Behawioryzm logiczny nie potrafi adekwatnie przedstawić jakościowego charakteru świadomości. Sposób traktowania przez logicznych behawiorystów zachowań podmiotu jest głęboko antyintuicyjny; przypomina on traktowanie automatycznych zachowań kogoś, będącego pod wpływem środków odurzających. Podobny problem powstaje wtedy, gdy chodzi o jakości wrażeń, jakich doznaje podmiot, albo gdy chodzi o obiekty podmiotowego zdawania sobie sprawy (obrazy powidokowe, bycie-w-stanie-bólu).

Behawioryzm nomologiczny (naukowy) stwierdza, że jedyna oczywistość, na podstawie której świadomemu podmiotowi możemy przypisywać stany mentalne, to zachowania się podmiotów w danych warunkach. Behawioryzm logiczny zawiera w sobie behawioryzm nomologiczny, ale nie odwrotnie.

Ściśle mówiąc, behawioryzm nie jest teorią materialistyczną, lecz jest próbą odpowiedzi na dualistyczne problemy z interakcją przyczynową pomiędzy stanami mentalnymi i fizycznymi. Dwóm rodzajom stanów mentalnych (wrażeniom i postawom propozycjonalnym) odpowiadają dwa rodzaje stanów fizycznych: wolitywne działania i poruszenia (te mogą być wolityw-

ne lub nie). Odczuwając ból jestem, według behawiorysty, dysponowany do skręcania się, do krzyczenia. Tak więc mój ból jest redukowalny do moich zachowań. U podstaw tej redukcji leży założenie o istnieniu psychofizycznych praw natury.

(B) Potocznie rozumiane wypowiedzi pierwszoosobowe okazują się być problematyczne dla behawioryzmu. Dzieje się tak z dwóch względów. (i) Podczas gdy wypowiedzi trzecioosobowe odnoszą się do zachowań, wypowiedzi pierwszoosobowe zdają się posiadać bezpośredni związek ze zdarzeniami mentalnymi, które je wywołują. Trudno mówić, na przykład, o dwuznaczności wypowiedzianego przeze mnie zdania: „jestem głodny”. Dwuznaczność, o którą tutaj chodzi, może wystąpić, gdy powiem: „Piotr jest głodny”. (ii) Wypowiedzi pierwszoosobowe mają ponadto funkcję ekspresywną (tzn. coś wyrażają), a nie tylko opisową. W tym sensie ból może być traktowany zarówno jako zjawisko, zachodzące wewnątrz podmiotu, jak i zjawisko zewnętrzne (potrafiące wywołać społeczne reakcje) [por. Wittgenstein 2000: § 533].

O tym, co sami odczuwamy i co myślimy, nie wypowiadamy sądów na podstawie obserwacji swoich własnych zachowań (na przykład w lustrze). Mogą oczywiście istnieć przypadki, w których inne osoby znają lepiej to, co zachodzi w naszym umyśle, aniżeli my sami, a tym samym mogą zmuszać nas do przyjęcia trzecioosobowego punktu widzenia siebie samych. Najczęściej jednak, a szczególnie przy stanach nieintencjonalnych, jesteśmy od tych osób w lepszej sytuacji, gdy mamy powiedzieć, co odczuwamy lub myślimy. Ponadto potrafimy, jak wspomniałem wyżej, opisać nasze wrażenia bez konieczności odwoływania się do procesów, prowadzących do określenia, czym te wrażenia są, na podstawie obserwacji lub wniosku z bardziej podstawowych danych.

Behawiorystyczne opisy naszego wyrażania się o innych różnią się od behawiorystycznych opisów naszego wyrażania się o nas samych. Tym samym behawioryzm nomologiczny nie podaje adekwatnego przedstawienia rozumianego holistycznie: *jak to jest* być podmiotem, lecz stara się odpowiedzieć na pytanie: Z czym jest związane posiadanie umysłu? Nie interesuje go ani to, że posiadamy wewnętrzne stany mentalne, ani to, że stany te są subiektywnie przeżywane. Behawioryzm nie zajmuje się tym, że w bólu jest coś więcej aniżeli cielesne zachowania podmiotu, który doznaje bólu. Nie można odczuwać bólu bez zaistnienia bólu, niezależnie od tego, jakiego rodzaju zachowania będą obserwowane. Stąd twierdzenie:

III.2 Behawiorystyczne założenie, że wypowiedzi pierwszoosobowe można traktować jak wypowiedzi trzecioosobowe, jest upraszczające.

Stosowanie trzecioosobowych metod weryfikacyjnych do zdań pierwszoosobowych jest pozbawione sensu. Gdyby w naszym odczuwaniu bólu nie było nic więcej aniżeli zachowanie się w dany sposób (albo posiadanie dyspozycji do takiego zachowania się), moglibyśmy mieć do czynienia z absurdalną konsekwencją, że podmiot musiałaby czekać ze sprawozdaniami o swoich własnych stanach mentalnych aż do momentu, gdy sam się w odpowiedni sposób zachowa.

Ponadto może istnieć cała gama symulowanych, bólowych zachowań podmiotu, chociaż nie odczuwa on bólu. Skoro jednak mogą istnieć zachowania bez odpowiadających im stanów mentalnych (bólu), stąd stany mentalne nie muszą być identyczne z zachowaniami.

3.3. Teoria identyczności typów

Do teorii behawiorystycznych zaliczamy teorie identyczności typów oraz teorie identyczności zdarzeń, określane także jako teorie identyczności egzemplarzy. Według redukcyjnej teorii identyczności typów (*type-type identity theory*) każdy rodzaj (*type*) stanów mentalnych może być identyfikowany z jakimś rodzajem stanów fizycznych (na przykład ból – z pobudzeniem nerwowych włókien C). Według teorii identyczności zdarzeń (*token-token identity theory*), każde mentalne zdarzenie (*token*) może być identyfikowane z jakimś zdarzeniem fizycznym, chociaż ten sam typ stanu mentalnego może być identyczny z różnymi rodzajami stanów neurofizjologicznych (na przykład dany ból może być raz pobudzeniem włókien C_n , a raz innych włókien). Teoria identyczności zdarzeń jest także nazywana „nieredukcjonistycznym fizykalizmem”. W tym punkcie zajmiemy się teorią identyczności typów, teoria identyczności zdarzeń zostanie omówiona w 3.4.

Mówiąc o identyczności musimy odróżnić jej dwa pojęcia: bycie jedną i tą samą rzeczą (identyczność numeryczna) i identyczność w sensie bycia dokładnie podobnym (identyczność jakościowa). Na przykład, dwie kserokopie tego samego tekstu nie są numerycznie identyczne, a są takie jedynie jakościowo. Przypomnijmy, że zasada Leibniza (która łączy owe dwa rodzaje identyczności) brzmi: *a* jest identyczne z *b*, wtedy i tylko wtedy, gdy każda cecha, którą posiada *a*, jest także posiadana przez *b* i odwrotnie. Zasada ta nie mówi o tożsamości jakościowej. Normalnie bowiem zaliczamy rzeczy do jakościowo identycznych, chociaż nie mają wspólnych wszystkich swoich cech. Na przykład, dwie białe kartki papieru mogą być jakościowo identyczne, chociaż jedną z nich trzymam w prawej, a drugą w lewej ręce (co oznacza, że nie mają one tych samych cech relacyjnych). Z tej racji wprowadza

się dalsze rozróżnienie pomiędzy tym, jaką dana rzecz jest w sobie, a jaką jest w stosunku do innych rzeczy (czyli relacjonalnie).

Zgodnie z numeryczną identyecznością typów to, co mentalne, jest jedną i tą samą rzeczą z tym, co fizyczne (a nie dokładnie przypominające fizyczne). W tym sensie: umysł = mózg. Chodzi przy tym o empiryczną identyeczność, a nie identyeczność pochodzącą z definicji (tzn. terminy mentalne nie oznaczają tego samego, co terminy fizyczne, ale odnoszą się do tych samych rzeczy, do których odnoszą się terminy fizyczne). W przeciwieństwie do behawioryzmu logicznego terminy mentalne i fizyczne nie są synonimami.

W ramach teorii identyeczności typów rodzaj bólu jest (numerycznie) identyeczny z rodzajem stanów, zdarzeń czy procesów, zachodzących w centralnym systemie nerwowym. Ból jest nie tylko ściśle wewnętrznym zdarzeniem, lecz odgrywa przyczynową rolę w bólowym zachowaniu się. Nie mówi się tutaj, że identyeczność jest prawdą analityczną (jak w behawioryzmie logicznym), lecz że jest odkrywana empirycznie. Hipoteza identyeczności typów jest tym samym przygodną hipotezą *a posteriori*.

Przedstawiona w rozdziale VII niniejszej pracy metoda poszukiwania neuronalnych korelatów świadomości jest odbiciem teorii identyeczności typów. Zamiast pytać, jak to ma miejsce w teorii identyeczności typów, o ogólne psychofizyczne prawa, pytamy początkowo o możliwość uogólnienia słabszych relacji (korelacji), zachodzących pomiędzy tym, co mentalne, i tym, co fizyczne.

Pierwsza trudność z teorią identyeczności typów pojawia się wraz z (postulowanymi) inteligentnymi istotami, których „biochemia” opierałaby się – powiedzmy – na krzemie w taki sposób, jak nasza biochemia na związkach węgla. Według teorii identyeczności typów stany świadomości, mające swoją podstawę w innym podłożu materialnym, nie mogłyby być identyeczne ze stanami naszych mózgów. Wynika stąd twierdzenie: jeżeli ból jest identyeczny z aktywnością pewnych włókien w systemie nerwowym, to bez nich nie może być bólu.

Dlatego właśnie teoria identyeczności typów jest kwestionowana przez tezę o tak zwanej wielorakiej realizacji (*multiple realizability*): rodzaje stanów mentalnych mogą być realizowane na różnorodnych podłożach. Twierdzenie, że określonego rodzaju ból jest identyeczny ze ściśle określonym rodzajem aktywności neuronalnej, jest problematyczne. Bardziej przekonujące wydaje się twierdzenie, że:

III.3 Organizmy (lub systemy) z odmiennymi rodzajami struktur neuronalnych mogą być zdolne do odczuwania tego samego rodzaju wrażeń.

Problem wielorakiej realizacji zostanie zanalizowany dokładniej przy omawianiu funkcjonalizmu (por. 3.5).

Druga trudność z teorią identyczności typów pojawia się z pytaniem: Jak rozumieć zakładaną w niej identyczność? Przedstawiciele teorii identyczności typów (J. J. C. Smart, D. M. Armstrong i D. K. Lewis), próbują wyjaśnić, w jaki sposób identyczność stanów mentalnych i stanów fizycznych może być prawdą przygodną. W przeciwieństwie do nich T. Nagel, który poszukuje teorii świadomości uwzględniającej stany świadomości zjawiskowej, pyta:

„[...] w jaki sposób jakiś rodzaj identyczności umysł–ciało może być prawdą konieczną. Wymagałoby to nie tylko wyobrażenia pojęć, które mogą uchwycić to połączenie, ale także jakiegoś przedstawienia, w jakim stosunku nasze istniejące pojęcia, powinnyby stać do tych pojęć oraz do innych” (Nagel 1998a: 348).

Mówiąc o pierwszo- i trzecioosobowych warunkach aplikacji pojęć mentalnych, oraz o fizykalnych pojęciach używanych do opisu mózgu możemy mieć na myśli identyczność w sensie przygodnym (kontyngentnym) lub koniecznym (por. S. Kripkego krytykę teorii identyczności, 3.4.1).

Trzecią trudność z teorią identyczności typów można sformułować następująco: określenie stanów mentalnych w terminach dotyczących zachowań domaga się wyspecyfikowania, jak podmiot zachowałby się w różnych sytuacjach. „Odczuwać ból” znaczy na przykład: „krzyczeć, gdy jesteś w sytuacji K, L, M”, „wić się, gdy jesteś w sytuacji N, P”, „mówić «au», w sytuacji Q, R, T”. Aby uchwycić tak rozumianą całość tego, co mentalne, musimy mieć holistyczną teorię mentalnej przyczynowości.

Należy dodać, iż podanie wystarczająco ogólnej charakterystyki typów stanów mentalnych prowadzi do ich identyfikacji z abstrakcyjnie funkcjonalnymi rolami, odgrywanymi przez ich poszczególne egzemplarze (*tokens*) w mediowaniu pomiędzy *bodźcem* i *odpowiedzią*. Dochodzimy tym samym do funkcjonalistycznego traktowania stanów mentalnych. Tak powstały funkcjonalizm pozostaje obojętny w stosunku do wspomnianych problemów materializmu typów (problem zombie, problem empirycznej identyczności). Zachodzi to oczywiście przy założeniu, że przyjęty dla jakichś form funkcjonalizmu materializm (w stosunku do poszczególnego stanu mózgu czy zdarzenia) będzie nosicielem zjawiskowości mentalnej, zakładając, że taka zjawiskowość jest w tej formie w ogóle brana pod uwagę.

3.3.1. Materializm ontologiczny według Paula Churchlanda

Materialista P. M. Churchland ani nie neguje istnienia jakości przeżyciowych (qualiów), ani nie przypisuje im wewnętrznych własności. Można akceptować istnienie jakości przeżyciowych i zarazem negować to, że odgrywają one jakąś rolę w teorii identyczności typów, odniesionej do własności

mentalnych. Churchland sugeruje, że możemy być teoretykami identyczności, gdy chodzi o jakości przeżyciowe, i równocześnie być funkcjonalistami, gdy chodzi o inne stany mentalne. W tym sensie Churchland jest zarazem materialistą i funkcjonalistą: poznanie funkcjonowania umysłu, a tym samym i świadomości zjawiskowej jest do osiągnięcia poprzez poznanie funkcjonowania mózgu.

Churchland przyjmuje, że rodzaje subiektywnych przeżyć są czysto fizykalne i można je przypisać fizycznej strukturze mózgu. Odwołuje się przy tym do odpowiedniości pomiędzy strukturami fizycznymi oraz jakościowymi i mówi o tak zwanych „uporządkowanych parach” (por. Churchland 1985: 9–11). Jakości przeżyciowe istnieją jako neuronalne stany i struktury, a więc nie mogą być traktowane jako proste, lecz jako kompleksowe. Jakości te są określonymi, fizyczno-neurologicznymi cechami mózgu. Nie przedstawiają one sobą wewnętrznych, niefizykalnych cech, lecz są wewnętrzne (w sensie „fizykalnego świata wewnętrznego”) w przeciwieństwie do fizycznego „świata zewnętrznego” (por. Churchland 1989: 74–76). Fizyczny świat wewnętrzny odnosi się do struktur mózgu. Tym samym, jakości przeżyciowe są relacyjno-organizacyjnymi własnościami połączeń neuronowych i są sprowadzalne do cech opisywalnych fizykalnie.

Churchland traktuje spostrzeganie jakości przeżyciowych jako spostrzeganie świata. Za pomocą jakości przeżyciowych identyfikujemy i odróżniamy struktury fizyczne¹⁶. Poprzez introspekcję nie dochodzimy do poznania jakichś specyficznych, ontologicznych przedmiotów, lecz do struktur fizycznych. Do tych samych struktur dochodzimy za pomocą metod trzeciosobowych. Jeżeli potrafimy nauczyć się myśleć i rozpoznawać stany wewnętrzne wśród opisów stanów mentalnych, to tak samo możemy się nauczyć rozpoznawania i myślenia o nich w bardziej precyzyjnych opisach neuronalnych. Możemy się nauczyć wyrażać za pomocą języka neurofizjologii nasze przeżywanie jakości. A tym samym – poprzez introspekcję – możemy dojść do poznania, jaki szczególny proces dokonuje się w mózgu. Dlatego Churchland wskazuje na tak zwany argument z poznania (*knowledge-argument*): nie zdobywamy dodatkowej wiedzy o nowych przedmiotach, lecz posiadamy tylko nowe dojście do tego samego przedmiotu, czy też nową formę jego przedstawiania. Churchland nie określa jednak bliżej rozumienia introspekcji, którym operuje. Naszym zdaniem:

¹⁶ Por. Churchland 1989: 32; 1988: 39–41. Według Churchlanda: „If one’s pains [...] do not introspectively seem like elektrochemical states in a neural network, that may be only because our faculty of introspection [...], is not sufficiently penetrating to reveal such hidden details” (s. 15).

III.4 W ramach teorii materialistycznych każda próba rozumienia introspekcji musi się odwoływać do tego, co rozumiemy pod pojęciem dyspozycji.

Tym samym powstaje pytanie: Czy poprzez odwołanie się do dyspozycji uda się uchwycić aktualność analizowanego stanu świadomości. Jakości przeżyciowe nie posiadają charakterystyk swoistych dla naturalnych rodzajów, w szczególności nie są one redukowalne do możliwych, nieznanych „wewnętrznych struktur”, albo do „rzeczywistych treści”. Kolory nie są identyfikowalne z długością fal elektromagnetycznych, ponieważ teza dyspozycyjna definiuje kolory w terminach sensorycznego zjawiania się.

C. McGinn dodaje, że subiektywność jakości przeżyciowych implikuje ich relatywność (*relativity*, mogą się one ukazywać różnie różnym odbiorcom), a ich subiektywne zjawianie się ustala ich tożsamość¹⁷.

Na zakończenie należy zauważyć, że w przeciwieństwie do D. Dennetta, P. Churchland nie określa jakości przeżyciowych poprzez funkcjonalne stany posiadające zewnętrzne relacje, lecz poprzez wewnętrzne, fizycznie opisywane cechy mózgu. Na przykład różowość czyjegoś wzrokowego wrażenia może zostać opisana jako zsynchronizowany 95Hz/80Hz/80Hz akord w odpowiednim systemie kortykalnym¹⁸ (por. 7.5.2). Churchland nie neguje istnienia takich jakości, odmawia im jednak dostępności w pierwszej osobie i charakterystycznej homogeniczności (por. 2.4.1). Tym samym Churchland rezygnuje z poszukiwania teorii subiektywności, redukując jakości przeżyciowe do obiektów neurofizjologii (w tym sensie jest on eliminatywnym materialistą). Cała dyskusja o jakościach, a tym samym o subiektywności, jest, według Churchlanda, podobna do prowadzonych w ubiegłym stuleciu dyskusji o witalistycznej „sile życia”. Wówczas także nie potrafiono wyjaśnić przejścia z materii nieożywionej do ożywionej, dlatego postulowano istnienie siły życia. W końcu jednak – zdaniem Churchlanda – okryto, że zjawisko życia jest warunkowane odpowiednim składem chemicznym i połączeniami

¹⁷ „The subjectivity of secondary qualities implies their relativity, since the same objects could appear different to different perceivers and their subjective appearance fixes their identity” (McGinn 1983: 16; por. s. 12, 13).

¹⁸ Podobnie argumentują F. Crick i Ch. Koch. Świadomość wizualna ma swoją podstawę w zsynchronizowanej, czasowej oscylacji, zachodzącej w odpowiednich neuronach. Jeżeli komórki w mózgu, zaktywowane przez spostrzeganie konkretnego przedmiotu, „odpalają” w sposób zsynchronizowany, wtedy różne cechy tego przedmiotu (kształt, kolor, ruch, głębia) mogą zostać połączone w jedną całość, mimo że w całym tym procesie są zaangażowane różne komórki nerwowe (por. Crick i Koch 1990: 265–74).

strukturalnymi. Churchland powiedziałaby, że podobnie, przyjmowane dzisiaj istnienie subiektywności, zostanie kiedyś wyjaśnione przez odpowiednią neurofizykalną teorię.

3.3.2. Johna Smarta teza o ontologicznej neutralności

Według proponowanej przez J. J. C. Smarta odmiany teorii identyczności typów, istnieje ścisła, chociaż przygodna (empiryczna), identyczność pomiędzy zjawiskami mentalnymi i procesami w mózgu. Przez ścisłą identyczność Smart rozumie to, że zjawiska mentalne i procesy w mózgu są jedną i tą samą rzeczą w tym samym czasie. Krytycy tego założenia twierdzą, że Smart, zamiast o identyczności, mówi jedynie o korelacji, która może się okazać nieprawdziwa (dlatego, gdy mówimy o wrażeniu, nie mówimy jeszcze o procesie w mózgu). Odpowiadając na ten zarzut, Smart stwierdza, że nie możemy (w ścisłym tego słowa znaczeniu), mówić o korelatach, ponieważ „nie można korelować czegoś ze sobą samym” (Smart 1959: 142). Według Smarta, zarzut ten potwierdza jedynie to, że dyskurs o wrażeniach nie oznacza tego samego, co dyskurs o procesach w mózgu. Zarzut ten nie pokazuje natomiast, że chodzi tutaj o dwie odmienne rzeczy (o wrażenia i o procesy fizyczne). Z przygodności faktu „wrażenie *w* jest identyczne z procesem *p*, zachodzącym w mózgu” nie wynika, że wrażenia i procesy nie są identyczne.

W znanej argumentacji S. Kripke twierdzi (por. 3.4.1), że każda identyczność jest relacją konieczną, a tym samym nie można mówić o przygodnych (kontyngentnych) identycznościach. Każda rzecz jest koniecznie identyczna ze samą sobą. Jeżeli zatem przygodne jest to, że wrażenie *w* jest procesem *p*, to nie są one ściśle identyczne. Jeżeli zaś są one ściśle identyczne, to są takie w sposób konieczny. Czy więc Smarta teza o identyczności jest tezą empiryczną, czy tezą *a priori*? Jeżeli zadamy pytanie: czy wrażenia są procesami w mózgu, czy też procesami w innej części ciała?, to według Smarta jest to pytanie empiryczne. Gdy jednak rozważymy pytanie: czy jakaś forma materializmu jest poprawna?, albo: czy możemy przyjąć epifenomenalizm tego, co mentalne?, to Smart nie udziela nam empirycznej odpowiedzi. W rozstrzygnięciu pomiędzy materializmem i epifenomenalizmem nie można się powołać na żadną oczywistość. Jediną odpowiedzią zdaje się być wówczas to, że epifenomenalizm nie jest teorią naukową, gdyż brzytwą Ockhama („nie mnożyć bytów bez potrzeby”) eliminuje go jako teorię naukową. Tym samym można przyjąć, że Smart opowiada się za materializmem.

Teorie identyczności (do których należy teoria Smarta) stwierdzają empiryczną identyczność pomiędzy stanami mentalnymi i fizycznymi. Wynika stąd, że nasza metoda identyfikacji obiektu mentalnego musi być różna od

metody, za pomocą której identyfikujemy to, co fizyczne (kwestia ta jest bardzo ważna przy naukowym określaniu neuronalnych korelatów świadomości). Inaczej nie potrzebowalibyśmy odkrywać identyczności między nimi. Wydaje się jednak, że różne metody identyfikacji wymagają odmiennych (różnych) identyfikowalnych własności. Gdy zatem identyfikujemy stany mentalne poprzez ich wewnętrzne i nieredukowalne cechy, to dualizm cech zdaje się być nie do uniknięcia. Nie potrafimy, na przykład, wykazać, że wrażenia nie są procesami w mózgu, ale widzimy, że podmiot inaczej przeżywa wrażenia, inaczej zaś odpowiadające im w mózgu procesy. Muszą zatem istnieć jakieś cechy wrażeń, których nie da się zredukować do cech fizycznych. Mamy tutaj klasyczny dualizm własności: wrażenia muszą posiadać jakieś własności zjawiskowe, których brak procesom w mózgu.

Przyjmijmy, że nie posiadamy niezależnego dojścia do identyfikacji i do natury tego, co mentalne, albo że nie wiemy, że posiadamy takie dojście, a chcemy zidentyfikować to, co mentalne, w sposób pomijający pytanie, czym są jego wewnętrzne własności. Załóżmy, że mentalne przypisywania są tematycznie neutralne (*topic-neutral*) w tym sensie, że wychwytyują one rzeczy bez przypisywania im jakichś wewnętrznych własności (a tym samym bez wykluczania możliwości, że same własności fizyczne wystarczą do prawdziwości tych przypisywań).

(i) Mówiąc „jestem przekonany, że *p*”, mówimy „jest coś zachodzącego do mnie i składającego do zachowań odpowiadających prawdziwości *p*”.

(ii) Według Smarta, kolor jakiegoś przedmiotu, to zdolność tego przedmiotu do wywoływania określonych reakcji u normalnych podmiotów w normalnych warunkach. „To wygląda pomarańczowo” znaczy wtedy mniej więcej: „To wywołuje we mnie to, co zawsze wywołują we mnie pomarańczowe rzeczy”. Gdy stwierdzam, że mi się wydaje, iż na ścianie jest żółtopomarańczowa plama, to stwierdzam, że zachodzi we mnie coś, co jest podobne do tego, co zachodzi we mnie wtedy, gdy spoglądam na żółtopomarańczową plamę w normalnym oświetleniu, sam jestem zdrowy itd. Gdy potrafimy opisać w tematycznie neutralnych terminach, co stwierdzamy, gdy zdajemy sprawozdanie albo z występowania bólu, albo z wyobrażenia go sobie, to być może potrafimy podać adekwatną ocenę naszych przeżyć, bez odnoszenia się do mentalnych czy zjawiskowych cech, i zostawiając w spokoju zjawiskowe indywidua. Powiedzieć, że coś wygląda mi na pomarańczowe, to po prostu powiedzieć, że moje przeżycie jest podobne do przeżycia, które posiadam, gdy widzę coś pomarańczowego. Mój opis jest neutralny, gdy chodzi o to, czy jest to podobieństwo mentalne czy fizyczne. Ten rodzaj neutralności spotykamy u Smarta, w proponowanej przez niego wersji teorii identyczności:

„Sugeruję co następuje. Gdy jakaś osoba mówi, «widzę żółtopomarańczowy obraz powidokowy (*after-image*)», to mówi ona coś w rodzaju: «*zachodzi coś, co jest podobne do tego, co zachodzi, kiedy mam otwarte oczy, nie śpię, a przede mną jest dobrze oświetlona pomarańcza, to jest, kiedy faktycznie widzę pomarańczę*»” (Smart 1959: 149).

Po objaśnieniu, że słowa wyróżnione kursywą odnoszą się do mówienia neutralnego tematycznie, Smart podkreśla, że nie identyfikuje on jakiegos jednostkowego przedmiotu zjawiskowego (obrazu powidokowego) z procesem w mózgu, a jedynie twierdzi, iż przeżywanie posiadania obrazu powidokowego jest procesem w mózgu¹⁹.

Smarta teoria identyczności typów została zmodyfikowana przez D. M. Armstronga, dając początek kolejnej wersji teorii identyczności (tak zwanemu materializmowi centralnego stanu, *central-state materialism*). Wspomniana teoria składa się z dwóch części: (i) z analizy pojęć mentalnych (trzon tej analizy stanowi teza Smarta o neutralnej analizie pojęć mentalnych), (ii) z materializmu identyfikującego stany wewnętrzne z fizykochemicznymi stanami umysłu (ludzki organizm może być w pełni wyjaśniony w terminach dotyczących stanów centralnego systemu nerwowego). Tak rozumiany umysł jest przyczyną naszych zachowań (por. Armstrong 1968: 89–90). Według Armstronga, świadomość jest stanem kolejnego rodzaju, nie zaś cechą stanu, którego jesteśmy świadomi²⁰. Jest ona spostrzeganiem własnego stanu wewnętrznego. Jeżeli więc potrafimy podać przyczynową analizę spostrzegania, możemy podać taką samą analizę świadomości. Armstrong identyfikuje jakości drugorzędne (kolory, dźwięki) z kompleksowymi i dysjunktywnymi własnościami fizycznymi. Spostrzeganie pomarańczowej plamy jest stanem, powodującym rozróżnienie pomiędzy przedmiotami ze złożoną własnością pomarańczowości a tymi, które jej nie posiadają.

Holistyczna teoria stanów mentalnych m_1, m_2, \dots, m_n pochodzi od D. K. Lewisa. Bierze on pod uwagę wszystkie wypowiedzi o umyśle, które znamy, tworzy ich koniunkcję i zapisuje w postaci potocznej teorii tego, co mentalne (T): $T(m_1, \dots, m_n)$. Terminy dotyczące stanów mentalnych możemy zastąpić zmiennymi i poprzedzić kwantyfikatorem szczegółowym: $\exists x T(x_1, \dots, x_n)$, otrzymując tym samym zdanie Ramseya dla teorii T²¹.

¹⁹ Jest to przeżycie: „[...] which is reported in the introspective report. [...] and this experience is not a yellow-orange something” (Smart 1959: 150–151).

²⁰ „[...] consciousness is simply a further mental state, a state «directed» towards the original inner states” (Armstrong 1968: 94).

²¹ Zdanie Ramseya (*Ramsey sentence*) pozwala o czymś mówić, bez używania nazwy tego, o czym mówimy. Chcąc zbudować zdanie Ramseya dla terminu „ból” bierzemy pod uwagę wszystkie zdania teorii psychologicznej, mówiące o bólu, zastępujemy termin

Zdanie to mówi, że istnieją obiekty, posiadające role zdefiniowane przez teorię **T**. Zdanie Ramseya może być traktowane jako teoretyczna definicja terminów, dotyczących stanów mentalnych (m_n):

$$(a) (m_1, \dots, m_n) = (x_1, \dots, x_n) \text{ w } \mathbf{T}(x_1, \dots, x_n).$$

Formuła (a) oznacza, że potocznie rozumiane stany mentalne są obiektami posiadającymi przyczynowe role, przypisane stanom mentalnym przez potoczne myślenie o tym, co mentalne. Materialista, przyjmujący *central state materialism*, może powiedzieć, że istnieje szereg stanów fizycznych (f_1, \dots, f_n), które odpowiadają przyczynowym rolom w zdaniu Ramseya. Zgoda na ten stan rzeczy oznacza zarazem logiczną zgodę na przyjęcie identyczności: $(m_1, \dots, m_n) = (f_1, \dots, f_n)$. Na tej podstawie Lewis twierdzi, że teoria identyczności jest logicznie implikowana przez prawdziwość analizy przyczynowej oraz to że ludzkie zachowanie może być całkowicie wyjaśnione przez funkcjonowanie centralnego systemu nerwowego²².

Wracając do teorii Smarta i Armstronga należy zauważyć, że bez odpowiedzi pozostaje w nich pytanie o istnienie wspomnianych jednostkowych przedmiotów zjawiskowych (indywiduów zjawiskowych). Stąd nasze twierdzenie:

III.5 Jeżeli istnieją indywidua zjawiskowe, to teoria identyczności jest falsyfikowalna przez zasadę Leibniza.

„ból” przez zmienną x i kwantyfikujemy egzystencjalnie. Zamiast więc mówić, że ból ma takie a takie cechy, zdanie Ramseya mówi, że istnieje „coś”, co ma te cechy. Gdy uczynimy to dla wszystkich teoretycznych terminów danej teorii, otrzymamy zdanie Ramseya dostarczające nam „topic-neutral” struktury tej teorii (*topic-neutral*, ponieważ struktura ta jest neutralna, gdy chodzi o to, co owe terminy denotują). Powiedzmy, że mamy prostą „teorię” składającą się z terminów: „pragnienie” (x_1) i „stan zadowolenia” (x_2) oraz z dwuargumentowego predykatu (S), takiego, iż zachodzi: (a) „gdy ktoś jest x_1 i widzi wodę, pije ją i osiąga x_2 ” oraz „gdy ktoś jest x_1 i nie widzi wody, rozgląda się za nią i nadal pozostaje x_1 ”. Nasza teoria może być zapisana: $S(x_1, x_2)$. Wstawiając przed ten zapis kwantyfikatory szczegółowe otrzymamy zdanie Ramseya dla tej teorii: $(\exists x_1)(\exists x_2)[S(x_1, x_2)]$. Zdanie to mówi, że istnieją dwa „przedmioty” spełniające tę teorię. Dla interesującego nas przypadku psychologii potocznej, zdanie Ramseya stwierdza, że istnieje szereg obiektów x_1, \dots, x_n , ukazujących te same relacje, jakie ukazują stany, które nazywamy za pomocą terminów mentalnych.

²² „My argument is this: The definitive characteristic of any (sort of) experience as such is its causal role, its syndrome of most typical causes and effects. But we materialists believe that these causal roles which belong by analytic necessity to experiences belong in fact to certain physical states” (Lewis 1983: 100).

Indywidualia zjawiskowe mają bowiem cechy, jakich nie posiadają stany fizyczne w mózgu. Możemy oczywiście próbować pozbyć się owych indywidualiów odwołując się do tak zwanej teorii przysłówkowej. W teorii przysłówkowej jakaś zjawiskowa cecha przypisywana zjawiskowemu indywidualium, jest przypisana podmiotowemu sposobowi przeżywania lub odczuwania. Mówimy wtedy o tak zwanych przysłówkowych modyfikacjach bezobiektywnego odczuwania. Tego rodzaju przysłówkowe analizy proponuje na przykład W. Sellars. Zamiast wypowiedzi: „mam pomarańczowy obraz powidokowy” mówimy wtedy „odczuwam w pomarańczowy sposób”. Problemy z teorią przysłówkową pojawiają się wówczas, gdy chcemy w niej wyrazić kompleksowe wrażenia, na przykład: „wrażenie zielonej powierzchni, graniczącej z czerwoną powierzchnią”. Czy takie wrażenie ma być analizowane jako przypadek, w którym mamy dwa sposoby odczuwania (w sposób zielony i czerwony), czy też jako przypadek, w którym mamy jedno odczuwanie w całkiem nowy sposób (w zielono-czerwony sposób)?²³

Nie wchodząc w inne problemy przysłówkowych modyfikacji należy zauważyć, że adekwatny sposób ich traktowania jest dodatkowo ograniczony przez fenomenologię naszego potocznego przeżywania. Przeżywanie to skłania nas do uznania, że kolory są widziane, co z kolei oznacza uznanie zjawiskowego koloru. Gdyby obrazy powidokowe były czymś istniejącym jedynie intencjonalnie (tak jak pojęcia), wtedy prawdopodobnie żadne cechy zjawiskowe nie byłyby w nich egzemplifikowane (*instantiated*). Jednak takie założenie komplikuje sprawę, gdyż czemukolwiek, zajmującemu jakąś część pola wzrokowego, należy przyznać jakiś status ontologiczny. Zaś często podawane przykłady na intencjonalne istnienie (powiedzmy na istnienie złotej góry, chociaż nie obrazu złotej góry) nic o takim statusie nie mówią.

Tym samym wątpliwym okazuje się twierdzenie, że pojęcie *topic neutrality* prowadzi do materialistycznego rozwiązania problemu cech zjawiskowych. Jeżeli bowiem, według Smarta, przeżywam żółtopomarańczowy obraz powidokowy, to nie spostrzegam jakichś jego wewnętrznych własności. Potrafię jedynie zdać sprawozdanie z obecności czegoś wewnętrznego, co jest podobne do tego, co zachodzi we mnie, gdy widzę pomarańczowy przedmiot. Wtedy jednak nie napotykam zjawiskowych, kolorowych własności, gdyż przeżyciowe kolory musiałyby być jakąś wewnętrzną cechą obrazu powidokowego, widzianego żółtopomarańczowo. Bez tych cech trudno zaś zobaczyć, jak zachodzące w mózgu procesy mogą konstituować przeżycie [Ψ].

²³ Por. Sellars 1975: 311–314. Przysłówkową teorię, proponowaną przez W. Sellarsa, oraz jej krytykę ze strony F. Jacksona i M. Tye'a omawiamy w: Bremer 1997: 144–152. Sellarsa teorię identity omawiamy w: Bremer 1998: 173–178.

Procesów takich może być bowiem wiele $[\Phi_j - \Phi_n]$. Jak zatem – nie mając owych identyfikujących własności zjawiskowych – wychwycić ten proces, który jest owym przeżyciem?²⁴

Smart zaznacza, że stan wewnętrzny, uchodzący za przeżywanie obrazu powidokowego (*after-image*) może zostać uchwycony przez jego podobieństwo do stanu przeżywania spowodowanego patrzeniem na pomarańczę. Nie pisze on jednak nic o zdawaniu sobie przez podmiot sprawy z tego podobieństwa. Widocznie w przewidywanym przez niego postępowaniu nie ma miejsca na tego rodzaju zdawanie sobie sprawy. Może dlatego Smart nie mówi o zdawaniu sobie przez podmiot sprawy z podobieństwa, lecz skupia się na jego zdolności do relacjonowania faktu, że jedna rzecz jest podobna do drugiej.

Następna trudność, w proponowanym przez Smarta neutralnym rozwiązaniu, wynika z przypuszczenia, że świadomy podmiot jest w stanie całkowitej niewiedzy o swoich subiektywnych przeżyciach. Przyjęcie tego typu założenia napotyka dwie trudności.

(a) W przypadku posiadania jakiegoś obrazu powidokowego zdajemy sobie sprawę nie tylko z tego, że coś zachodzi, ale także ze specyficznych własności (kolor, kształt). Negacja tych własności stawia nas wobec pytania o wyjaśnienie dwóch spraw. Z jednej strony, *jak to jest* dla podmiotu przeżywać widzenie pomarańczy, gdy chodzi o kolory, z drugiej strony, jak podmiot może odróżnić posiadanie obrazu powidokowego od (powiedzmy) wyobraźni słuchowej. Nie wystarczy czysto neutralne poznanie czegoś „zachodzącego” lub czysto neutralne mówienie o czymś podobnym.

(b) Ponadto, jeżeli brakuje nam poznania wewnętrznej natury przeżywania obrazu powidokowego, to o takich samych brakach możemy mówić w stosunku do innych przeżyć sensorycznych (intensywne sny, halucynacje). A więc przynajmniej niektóre z nich muszą być uznane za subiektywnie nieodróżnialne od prawdziwego przeżywania. Jeżeli jednak halucynacyjne widzenie (powiedzmy) pomarańczy jest nieodróżnialne od prawdziwego widzenia, to wynika stąd, że nasze widzenie pomarańczowego koloru, gdy faktycznie widzimy pomarańczę, jest także tylko widzeniem tematycznie neutralnym. Można zatem zapytać, czy neutralne widzenie, w którym nie ma

²⁴ Prowadzone przez Smarta analizy powidokowego przeżywania można uściślić odwołując się do wprowadzonej przez Sellarsa „opisowej treści” (*descriptive content*). Podmiot, któremu wydaje się, że widzi żółtopomarańczową plamę, widzi coś, czemu odpowiada „opisowa treść”. „For, phenomenologically speaking, the descriptive core [*descriptive content* – *J.B.*] consists in the fact that something in some way red and triangular is in some way present to the perceiver other than as thought of” (Sellars 1975: 310).

miejsca dla poznania wewnętrznych barw, nie wykluczy także poznania jakości pierwszorzędnych, na przykład kształtu.

Na zakończenie chcemy zwrócić uwagę na generalną trudność w przyjęciu twierdzenia, że każdej własności mentalnej da się jednoznacznie przyporządkować neuronalne korelaty. Raczej przyjmujemy, że własności mentalne są realizowane w wieloraki sposób (szczególnie u gatunkowo różnych osobników). Jeżeli jakaś psychologiczna własność (np. „bycie zmęczonym”) może być w równym stopniu odniesiona do ssaków i do ptaków, i zarazem widać, że odpowiednie fizyczne korelaty tej własności są odmienne w obydwu wypadkach, to teoria identyczności zostaje postawiona pod znakiem zapytania. Z czymś podobnym możemy mieć do czynienia nawet w przypadku jednej i tej samej osoby, u której psychologiczno-neuronalne korelacje mogą ulec zasadniczym zmianom. Coś takiego zachodzi, gdy, na przykład, określone funkcje uszkodzonej tkanki zostają przejęte przez inne części mózgu (tak zwana „neuroplastyczność mózgu”, por. 6.1.5).

3.4. Teoria identyczności zdarzeń

Dla teorii identyczności typów problematycznym okazuje się być pytanie o fizyczną realizację stanów mentalnych. Pytaniem tym zajmuje się słabsza teoria identyczności, zakładająca tak zwaną *token-token* identyczność: każdy poszczególny egzemplarz lub zdarzenie mentalne jest numerycznie identyczne z poszczególnym egzemplarzem fizycznym lub zdarzeniem opisywalnym fizycznie. Identyczność zdarzeń dotyczy odnoszenia się (desygnacji) terminów, nie ich znaczenia. H. Feigl określa ten rodzaj identyczności następująco:

„Używając wprowadzonego przez Fregego rozróżnienia pomiędzy *sensem* (*Sinn*, «meaning», «sense», «intention») a *znaczeniem* (*Bedeutung*, «referent», «denotatum», «extension»), musimy powiedzieć, że terminy neurofizjologiczne i odpowiadające im terminy zjawiskowe, chociaż bardzo różniące się w sensie, a zatem także gdy chodzi o modele potwierdzania zdań je zawierających, mają jednak identyczne obiekty, do których się odnoszą (*referents*)” (Feigl 1960: 68).

Ponieważ terminy mentalne i fizyczne różnią się sensem, dlatego fakt, że mimo wszystko odnoszą się one do tego samego konkretnego obiektu, musi być potwierdzony odpowiednimi badaniami naukowymi. Analogicznie zresztą, jak za pomocą odpowiednich obserwacji i teorii astronomicznych odkryto, że terminy „gwiazda wieczorna” i „gwiazda poranna”, chociaż nie są synonimiczne, to odnoszą się do tego samego obiektu – do planety Wenus.

Problem z powyższą analogią leży jednak w tym, że terminy mentalne różnią się, gdy chodzi o znaczenie, tak radykalnie od terminów neurofizjolo-

gicznych, iż mówimy o, wspomnianej już kilka razy, „luce w wyjaśnianiu”. Tego rodzaju luki nie ma natomiast, gdy mówimy o terminach dotyczących „gwiazdy wieczornej” i „gwiazda porannej”. Termin „świadomość” posiada jako intuicyjny sens, rzeczywistość intencjonalną i zjawiskową. Dlatego twierdzenie, że przedmiot, do którego ten termin się odnosi, jest identyczny z czasoprzestrzennym przedmiotem, do którego odnoszą się terminy neurofizjologiczne, jest tak samo zadziwiające, jak twierdzenie, że „liczba” odnosi się do tego samego obiektu, co „świadomość”.

Według teorii identyczności zdarzeń każdy stan mentalny jest identyczny z jakimś stanem fizycznym. Stan mentalny lub zdarzenie, denotowane przez wyrażenie „mój ból w czasie t ”, jest identyczny ze stanem fizycznym lub zdarzeniem, denotowanym przez wyrażenie „moja stymulacja włókien C w czasie t ”. Tym samym stwierdza się (mówiąc ogólnie i odwołując się do tezy o wielorakiej realizacji), że dwa podmioty, egzemplifikujące dokładnie te same stany fizyczne, nie mogą się różnić, gdy chodzi o egzemplifikację stanów mentalnych. Dopuszczalne natomiast jest to, że podmioty egzemplifikujące dokładnie te same typy stanów mentalnych mogą się różnić co do egzemplifikowanych typów stanów fizycznych. Mówimy wtedy, że typy stanów mentalnych danego organizmu, w danym czasie *superwenują* na jego typach stanów fizycznych, zachodzących w tym samym czasie.

Podstawowa trudność z każdego rodzaju teorią identyczności ujawnia się wtedy, gdy przejdziemy do analizowania zdarzeń mentalnych, posiadających przeżyciową naturę. Ta ostatnia jest bowiem różna od natury zdarzeń fizycznych, z którymi, jak się zakłada, zdarzenia mentalne są identyczne. Mój ból w czasie t ukazuje mi się jako mający zjawiskowo-przeżyciowy charakter. Tenże charakter zdaje się pod każdym względem nie zgadzać z charakterem zdarzeń neurofizjologicznych, znanym i opisywanym w nauce. W jaki sposób te różne charaktery mogą być połączone w jednym pojedynczym zdarzeniu?

Jeżeli bowiem stan czy akt zdawania sobie sprawy jest jakąś wydarzającą się nieprzestrzenną rzeczywistością²⁵, wtedy wyrażenia „moje zdawanie sobie sprawy z x w czasie t ” i „moje neurony, będące w stanie s w czasie t ” nie mogą denotować tych samych zdarzeń. Pojedyncze zdarzenie nie może zarazem zajmować miejsca w obszarze przestrzeni i być nieprzestrzennym.

Pozostawmy na boku kwestię, czy jakiemuś zdarzeniu mentalnemu brak przestrzenności i skupmy się na jego jakościowym (treściowym) charakterze. Również to założenie nie wzmacnia wyjaśniającej roli przypisywanej teorii identyczności zdarzeń, co można wyrazić w twierdzeniu:

²⁵ W nieprzestrzenności świadomych przeżyć C. McGinn widzi poważną trudność dla materializmu (por. McGinn 1999: 105–118).

III.6 W odróżnieniu od tego, że bycie białym jakiegoś przedmiotu wyklucza jego bycie czarnym (w tym samym czasie t), jakościowy charakter bólowego zdarzenia nie wyklucza jego charakteru neuronalnego.

Wspomniany brak wzajemnego wykluczania się jest jednak wszystkim, co można powiedzieć dla podtrzymania twierdzenia o pojęciowej zrozumiałości i naukowej adekwatności teorii identyczności zdarzeń.

Gdy bowiem skupimy się na zdarzeniu bólu „jak ukazuje się on” w introspektywnym zdawaniu sobie z niego sprawy, to okazuje się, że brak nam sposobu połączenia tego zdarzenia z jakąś własnością zdarzenia fizycznego. Tak więc, chociaż urzeczywistnienie obydwu własności (mentalnej i fizycznej) w pojedynczym zdarzeniu wydaje się być logicznie albo metafizycznie możliwe (por. III.6), to założenie, że są one urzeczywistniane w różnych zdarzeniach, jest także możliwe do utrzymania. Innymi słowy: teoria identyczności zdarzeń stwierdza, że każde zdarzenie mentalne jest numerycznie identyczne ze zdarzeniem fizycznym, co implikuje, że jedno i to samo zdarzenie musi urzeczywistniać dwie tak różne własności, choć brak nam zrozumienia, jak one obydwie mogą opisywać pojedyncze zdarzenie.

Dodatkowe problemy z teorią identyczności zdarzeń pojawiają się wtedy, gdy przyjmiemy tezy J. J. C. Smarta i D. M. Armstronga o przygodnym rodzaju związku pomiędzy bólem i związanym z nim zdarzeniem neuronalnym. Z jednej strony, jakościowa własność $[c\Psi]$ mojego bólu w czasie t jest istotna dla tego bólu (por. 3.4.1. [2]). Ból musi mieć ową własność, aby być tym bólem, chociaż zachodzi to najwidoczniej przy stymulacji moich włókien C w czasie t , czyli jest jakimś neurologicznym stanem systemu $[\Phi_j - \Phi_n]$. Z drugiej strony, istotna dla stymulacji moich włókien C nie jest jakościowa własność $[c\Psi]$, lecz ta własność, jaką bólowi przypisuje teoria naukowa. Musimy bowiem – odwołując się do twierdzeń S. Kripkego – zapytać o możliwe światy, w których ból mógłby występować bez własności jakościowej (por. 3.4.1). A jeżeli tak jest, to wtedy twierdzenie, że „zarówno własność bólu $[c\Psi]$, jak i cecha stymulacji moich włókien C $[c\Phi]$ są urzeczywistniane w tym samym, pojedynczym zdarzeniu”, może być twierdzeniem tylko przygodnie prawdziwym (to znaczy może także nie być prawdziwym). Wydaje się wprawdzie, że teoria identyczności obejmuje twierdzenie o przygodności, ale tak nie jest. Jeżeli bowiem, jak się w owej teorii utrzymuje, ból i stymulacja włókien są jednym i tym samym zdarzeniem, wtedy tożsamość ta *musi* zachodzić w każdym możliwym świecie, w którym obydwa występują, a to sprzeciwia się tezie o przygodności. I o ile trudno pokazać, że teza o przygodności jest błędna, to tym trudniej pokazać, że dany ból i stymulacja włókien C , są identyczne.

W stosunku do teorii identyczności zdarzeń należy wspomnieć jeszcze o dodatkowych trudnościach.

(1) Klasyczny eksternalizm odrzuciłby tezę o identyczności zdarzeń (a tym samym także o identyczności typów). Według niego stany mózgu podmiotu pozostałyby te same (chciaż jego otoczenie zmieniłoby się na „bliźniaczą Ziemię”), tak długo, jak długo jego wewnętrzne cechy nie uległyby zmianie. Jednakże jego intencjonalne stany mentalne mogłyby być różne. Tym samym, na mocy zasady Leibniza, stany mentalne byłyby różne od stanów mózgu.

(2) Teoria identyczności zdarzeń nie uchwytuje podstawowej intuicji fizykalizmu. Można sobie bowiem wyobrazić sytuację, w której dwa jakościowo (ale nie numerycznie) identyczne, fizyczne zdarzenia są każde numerycznie identyczne z bardzo odmiennymi mentalnymi zdarzeniami (na przykład ja i mój zombie możemy posiadać jakościowo identyczne ewolucyjne i społeczne biografie, a różnić się radykalnie pod względem mentalnym). Teoria identyczności zdarzeń nie wyraża adekwatnie fizykalizmu, gdyż nie mówi nic o sposobach, w jakich to, co mentalne, jest generowane przez to, co fizyczne. Jeden ze sposobów, w jaki to, co fizyczne, generuje to, co mentalne, jest następujący. Ponieważ relacja identyczności jest relacją przechodnią i symetryczną, to jeżeli mentalne zdarzenie $[\Psi_i]$ byłoby identyczne z fizycznym zdarzeniem $[\Phi_i]$ i fizyczne zdarzenie $[\Phi_i]$ byłoby identyczne z mentalnym zdarzeniem $[\Psi_j]$, wtedy $[\Psi_i]$ i $[\Psi_j]$ będą ze sobą identyczne. Jeżeli jednak $[\Psi_i]$ i $[\Psi_j]$ byłyby różnymi zdarzeniami mentalnymi, to nie mogą być identyczne ze sobą, a tym samym nie mogą obydwie być identyczne z $[\Phi_i]$. Jeżeli $[\Psi_i]$ jest identyczne z $[\Phi_i]$, to $[\Psi_i]$ musi być identyczne z jakimś $[\Phi_j]$. Dwa mentalnie różne zdarzenia muszą być identyczne z dwoma numerycznie różnymi fizycznymi zdarzeniami. Nie oznacza to jednak automatycznie, jak zdaje się tego oczekiwać (choćby tylko intuicyjnie) fizykalizm, że wspomniane zdarzenia fizyczne są jakościowo różne.

3.4.1. Saula Kripkego krytyka teorii identyczności

Wychodząc z założenia, że bóle są odczuwane jako pełne bólu, S. Kripke formułuje dualistyczny zarzut przeciwko tezie o identyczności typów oraz tezie o identyczności zdarzeń (Kripke 1988: 144; por. Rosenthal 1996: 353–354): twierdzenia o identyczności przyjmują tezę o przygodnej (kontyngentnej) identyczności, którą należy odrzucić²⁶. Wypowiedzi stwierdzające identyczność muszą być konieczne prawdziwe.

²⁶ Por. Horgan 1984: 464–467. Zdanie konieczne prawdziwe nie może być fałszywe w żadnych możliwych warunkach. Zdanie kontyngentnie prawdziwe może w pewnych warunkach być fałszywe.

Kripke odwołuje się do pojęcia „twardego” (*rigid*) desygatora („ściśle wyrażenie oznaczające”), czyli wyrażenia desygnującego to samo indywiduum w każdym możliwym świecie, w którym owo indywiduum istnieje (albo we wszystkich możliwych światach, w których owo wyrażenie coś desygnuje)²⁷. W tym sensie, na przykład, zdanie „ciepło = ruch molekuł” może być *konieczne i a posteriori*.

Kripke mówi o ontologicznej konieczności, a nie tylko o epistemologicznej i dopuszcza poprawność zdania, iż jakiś inny fizyczny proces może być odpowiedzialny za odczucie ciepła. Powiedzmy, że jakąś przygodną cechą ciepła jest, iż wywołuje w nas pewne odczucie. W końcu czymś przygodnym jest to, że na naszej planecie w ogóle żyją odczuwające istoty. Nie wiemy zatem *a priori*, jakie fizyczne zjawisko (*fz*), opisane w podstawowych pojęciach teorii fizycznej, jest tym zjawiskiem, które wywołuje owe odczucia. Nie wiedzieliśmy tego, ale w końcu odkryliśmy, że zjawiskiem tym jest ruch molekuł. Tym samym odkryliśmy tożsamość ukazującą nam istotną cechę tego zjawiska. „Odkryliśmy zjawisko, które we wszystkich możliwych światach będzie ruchem cząsteczek – które nie mogłoby zawieść w tym, że jest ruchem cząsteczek, ponieważ to jest to, czym *jest* owo zjawisko” (Kripke 1988: 133).

Z drugiej strony cecha, z którą uprzednio identyfikowaliśmy to zjawisko, cecha wywoływania w nas danych wrażeń, nie jest cechą konieczną, lecz przygodną. To samo zjawisko mogłoby istnieć bez bycia odczuwanym jako ciepło (na przykład ze względu na odpowiednie różnice w naszej strukturze nerwowej)²⁸.

Używając pojęcia twardego desygatora Kripke wykazuje, że każde zdanie mówiące o identyczności, którego terminy są twardymi desygatorami, jest *koniecznie* prawdziwe. Tak więc jeżeli „mój ból w czasie *t*” i „moje pobudzenie włókien *C* w czasie *t*” są desygatorami, wtedy zdanie stwierdzające:

²⁷ Terminy, takie jak „H₂O”, „ciepło”, są twardymi desygatorami, odnoszą się bowiem dokładnie do tych samych rzeczy we wszystkich możliwych światach. W tym sensie zdania: „Woda = H₂O” „Ciepło = ruch molekuł” są koniecznie prawdziwe.

²⁸ Aby „ustalić desygnację” (*to fix the reference*) danego terminu *T*, możemy użyć definicyjnego opisu: „*x* taki, że *Fx*”. Tego rodzaju umocnienie kontrastuje z definiowaniem *T* jako równoważnego w znaczeniu, to znaczy jako skrótu definicyjnego opisu; w pierwszym przypadku, ale nie w drugim, zdanie „jeżeli *T* istnieje, wtedy *T* jest *x* takie, że *Fx*” będzie raczej przygodnie, a nie koniecznie prawdziwe. Według S. Shoemakera, nazwy potoczne są twardymi desygatorami, zaś wiele definicyjnych opisów nimi nie jest. Gdy jakiś definicyjny opis jest używany do wprowadzenia nazwy (a tym samym do wprowadzenia twardego desygatora), jest on używany raczej do „umocnienia jego desygnacji aniżeli do „definiowania” albo do podania jego „znaczenia” (por. Shoemaker 1984: 194).

(A) „Mój ból jest identyczny z pobudzeniem moich włókien *C* w czasie *t*”

jest koniecznie prawdziwe. Nie jest to bowiem twierdzenie o dwóch różnych przedmiotach, które jedynie w aktualnym świecie są identyczne, a pomiędzy którymi nie ma identyczności w jakimś innym możliwym świecie. Wyrażenia odnoszące się desygnują raczej pojedynczą rzecz, która musi pozostać identyczna ze samą sobą w każdym możliwym świecie, w którym ona występuje.

Pojęcie konieczności nie jest przez Kripkego – w odróżnieniu od I. Kanta – używane jako równe znaczeniowo z pojęciem „analityczne” albo „aprioryczne” (por. Kant 1957: A1, B2). Nie każda konieczność jest koniecznością *a priori*, istnieje także metafizyczna konieczność *a posteriori*. Czymś koniecznym jest to, co nie może być inaczej, co jest prawdziwe we wszystkich możliwych światach.

Lecz tak samo, jak niekwestionowalne wydaje się powyższe zdanie (A), tak niekwestionowalne wydaje się to, że potrafimy sobie przedstawić możliwe światy, w których bóle podmiotu nie są stanami mózgu. Tego rodzaju przedstawienie pokazuje, że połączenie pomiędzy bólem a stanami mózgu jest jedynie przygodne. Istotą bólu jest odczuwanie bólu, a naszą intuicją jest to, że jest ono jedynie przygodnie związane z mózgiem. Tym samym twierdzenie o koniecznej identyczności pomiędzy tym, co mentalne i fizyczne, jest niemożliwe do pogodzenia ze wspomnianą, skądinąd oczywistą, intuicją. Znana jest teza, że nasza intuicja o przygodnym połączeniu jest kompatybilna z fizykalistyczną tezą o identyczności zdarzeń²⁹. Teza ta nie stwierdza, że każdy przedstawialny ból jest identyczny ze stanem mózgu. Przyjmuje ona słabsze założenie, że zdarzenia występowania bólu w rzeczywistym świecie są stanami mózgu.

Teza o identyczności zdarzeń przyjmuje, że mój ból w czasie *t* jest identyczny ze stymulacją moich włókien *C* w czasie *t*, a identyczność ta zachodzi w każdym możliwym świecie. Jasne wydaje się jednak, że potrafimy sobie wyobrazić, z jednej strony, możliwy świat, w którym pobudzenie moich włókien *C* w czasie *t* nie jest odczuwane jako ból i tym samym przestaje być bólem. Z drugiej strony, możemy sobie wyobrazić świat, w którym mój ból występuje w czasie *t* chociaż moje włókna *C* nie są stymulowane (powiedzmy, że jest logicznie możliwe uznanie twierdzenia, jakoby fizyczne mózgi nie istniały). Podobne intuicje mamy w stosunku do myśli. Świadome przeżywanie myślenia o czymś, przeżywanie bezpośredniości myśli zdaje się

²⁹ Argument taki podaje McGinn, akceptujący zarazem Kripkego tezę o konieczności identyczności (por. McGinn 1977: 78–79).

w tym świecie tylko przygodnie być połączone z mózgiem. Nie wyklucza się jednak możliwości, że nie jest tak w innym świecie. Ponadto trudno zrozumieć sugestię, że odczuwanie jest przygodnie identyczne z pobudzeniem włókien *C* u ludzi, a nie u innych podmiotów.

Przedstawioną argumentację Kripkego można streścić następująco: Odczucie bólu jest nie tylko bolesne, ale jest czymś *istotnie* bolesnym. Pobudzenie włókien *C*, o którym możemy twierdzić, iż identyfikujemy z nim ból, może być *przygodnie* bolesnym. Dlatego, na mocy zasady Leibniza, pobudzenie włókien nie może być identyczne z odczuciem bólu. Odwołując się do twardej desygnatorów, stoimy przed dwiema możliwościami: (1) wyrażenie „mój ból w czasie *t*” desygnuje twardo, (2) wyrażenie „mój ból w czasie *t*” nie desygnuje twardo (por. Lund 1994: 100). Rozpatrzmy te dwie możliwości.

(1) Utożsamiliśmy ból z tym, co Kripke nazywa jego bezpośrednią, zjawiskową jakością (ból jest istotnie bólem i odczuwa się go jako ból). Ponieważ własność ta identyfikuje ból w każdym możliwym świecie, w którym on istnieje, taka identyfikacja dostarcza twardej interpretacji wyrażenia [a] „mój ból w czasie *t*”. Następnie przyjęliśmy, że [b] „pobudzenie moich włókien *C* w *t*” jest także twardym desygnatorem. Założenie to okazało się sprzeczne z naszymi potocznymi intuicjami. Musimy więc zobaczyć, czy wprowadzając jeden lub drugi desygnator ([a] lub [b]) jako nietwardy, identyczność okaże się możliwa do zaakceptowania.

Próba skonstruowania desygnatora [b] „pobudzenie moich włókien *C* w *t*” jako desygnatora nietwardego jest, z materialistycznego punktu widzenia, mało interesująca. Jeżeli włókna *C* są czymś naturalnym, czymś, czego wewnętrzne lub istotowe cechy znamy dzięki badaniom naukowym, to cechy te muszą wystarczyć do uchwycenia ich funkcjonowania we wszystkich możliwych światach. Nieprawdopodobną wydaje się myśl, że jakościowe cechy bólu charakteryzowałyby wewnętrznie włókna *C*, podczas gdy nauka ukazywałaby tylko ich neutralne, relacyjne, strukturalne, trzecioosobowe cechy. Nieprawdopodobną wydaje się myśl, że (ograniczone do perspektywy pierwszej osoby) cechy mogłyby być istotnymi cechami materii (konkretnie materii systemu nerwowego). Jeżeli podstawową cechą materii jest jej dostępność z punktu widzenia trzeciej osoby, to jej wewnętrzne lub istotowe cechy nie mogą być ograniczone do punktu widzenia pierwszej osoby. Ponadto, jeżeli cechy zjawiskowe byłyby dla materii istotne, to charakteryzowałyby one ją wszędzie, a nie tylko w materii mózgu. Konsekwencją takiego spojrzenia jest panpsychizm. Aby go uniknąć, musielibyśmy przyjąć, że materia mózgu różni się od pozostałej materii.

(2) „Mój ból w czasie *t*” nie jest twardym desygnatorem, nie desygnuje tego samego indywiduum we wszystkich możliwych światach. Istnienie ta-

kich światów implikowałoby, że moja stymulacja włókien *C* w czasie *t* nie jest istotnie bólem, co wydaje się nie wzbudzać kontrowersji. Takiemu pobudzeniu mogłoby brakować jakościowego charakteru bólu, a przez to mogłoby ono być czymś, co nie boli. Tym samym jednak mój ból w czasie *t* mógłby mi dolegać, ale i nie dolegać. Kripke uważa:

(I) „[...] jeśli coś jest bólem, jest nim z istoty, i absurdalne wydaje się założenie, że ból mógłby być jakimś innym zjawiskiem niż to, jakim jest. To samo dotyczy terminu «pobudzenie włókien *C*», pod warunkiem, że «włókna *C*» – to ściśle wyrażenie oznaczające, jak chcę tu przyjmować” (Kripke 1988: 148).

Tak więc, według Kripkego, jeżeli jakieś pojedyncze wystąpienie bólu byłoby identyczne z pojedynczym stanem mózgu, to byłaby to identyczność konieczna. Jeżeli jednak nie może być ona konieczna, to nie może być w ogóle identycznością.

Wracając do (I) można powiedzieć, że teoretyk identyczności musiałby znaleźć sposób negowania tezy, iż jakościowy charakter bólu jest dla bólu istotny. Perspektywy dla takiej negacji nie wyglądają jednak zbyt optymistycznie, gdy zauważymy jak trudno jest się oprzeć intuicji, że bóle jawią się doznającym ich podmiotom bezpośrednio i transparentnie. Czymś niezaprzeczalnym wydaje się to, że bóle mają jakościowy charakter i ukazują go niezależnie od tego, czym byłaby ich fizyczna podstawa. W przypadku bólu trudno zrozumiałe zdaje się być również twierdzenie, że jakość ich przeżywania nie jest podmiotowi bezpośrednio znana. Miałoby tak być dlatego, ponieważ epistemologiczny dostęp podmiotu do tej jakości jest zniekształcony poprzez jego konceptualizację tej jakości albo poprzez jego sposób odnoszenia się do niej.

Podobnie do Kripkego argumentuje T. Nagel, pytając o wyjaśniającą rolę teorii identyczności. Jego zdaniem, hipoteza, że określony stan mózgu ma mieć z konieczności pewien subiektywny charakter, jest niezrozumiała bez dalszych wyjaśnień. Wyjaśnienia takiego nie dostarczają nam teorie identyczności, traktujące relację umysł–mózg jako przygodną. Być może, według Nagla, istnieją inne wyjaśnienia, których jeszcze nie poznaliśmy (por. Nagel 1997: 215). W dniu dzisiejszym brak nam jednak pojęć, aby wyjaśnić, jak cechy subiektywne i neurologiczne mogą być cechami jednego i tego samego konkretnego przedmiotu (por. 5.5.1). Proponowane przez Nagla wyjaśnienie przesuwania rozwiązania zależności umysł–mózg w nieokreśloną przyszłość. Naszym zdaniem, dziś poszukiwanego rozwiązania należy szukać w ramach teorii emergencji (por. 3.6), która nie jest jednak teorią ściśle naukową (w sensie dzisiejszego rozumienia teorii naukowych).

J. Levine porównuje wspomniane już wyrażenia (por. Levine 1983: 357): (i) „ciepło = ruch molekuł”, (ii) „ból = pobudzenie włókien C”. Istotna różnica pomiędzy nimi leży w tym, że identyczność (i) w przeciwieństwie do (ii), jest wyjaśnialna dzięki teoriom termodynamiki. Na podstawie znajomości praw fizyki możemy zrozumieć, dlaczego ruch molekuł spełnia dokładnie tę przyczynową rolę, która jest przypisywana ciepłu, gdyż ruch molekuł jest ciepłem. Makrofizykalne pojęcie ciepła może być całkowicie określone przez ową przyczynową rolę. Innymi słowy: przyjmuje się, że makroskopowa cecha ciepła jest – gdy chodzi o wyjaśnianie – redukowalna do ruchu molekuł. W tym sensie cecha ta nie jest cechą emergentną. Wyrażenie (ii) pozostawia coś istotnego niewyjaśnionym, gdyż pojęcie bólu nie wyczerpuje się w jego przyczynowej roli (do cech bólu należy także jego jakościowy charakter, „ból jest pełen bólu”).

III.7 Nie wiemy, w jaki sposób wrażenia mogą być redukowane (gdzie chodzi o wyjaśnianie) do stanów neurofizykalnych i funkcjonalnych. Nie wydaje się bowiem możliwe, za pomocą ogólnie obowiązujących, znanych z neurologii, praw natury wyprowadzenie tego, że systemy neuronalne posiadają wszystkie cechy charakterystyczne dla stanów przeżyciowych.

Wydaje się, że nie istnieje nic, co przemawiałoby za tym, że pobudzenie włókien C pasuje lepiej do przeżyć związanych z bólem aniżeli do jakichś innych przeżyć.

Podsumowując możemy powiedzieć, że argumentacja Kripkego opiera się na założeniu: zdania o identyczności stanów mózgu i wrażeń muszą być konieczne prawdziwe. Zdania te mają jednak jedynie status zdań przygodnych. Pozostaje pytanie, czy epistemologiczna możliwość fałszywości wypowiedzi o identyczności implikuje możliwość ontologiczną. Twierdzenia o pobudzaniu włókien C należą do wypowiedzi o subiektywności stanów mentalnych, mogą one być jednak jedynie traktowane jako „wolne miejsca” w adekwatnej, neuropsychologicznej analizie bólu.

W dalszych rozdziałach przyjmujemy, że na obecnym etapie badań pierwszym krokiem do takiej analizy jest określenie neuronalnych korelatów świadomości. Gdy chodzi o podmioty ludzkie, to stwierdzamy w tym kroku jedynie istnienie korelacji albo pomiędzy specyficznymi stanami mentalnymi i stanami fizycznymi, albo pomiędzy zdarzeniami mentalnymi i zdarzeniami fizycznymi. Nie wyjaśniamy w nim natomiast chociażby tego, w jaki sposób stany fizyczne wywołują stany mentalne. Dla uzyskania takiego wyjaśnienia potrzebujemy odpowiedniej teorii świadomości.

3.4.2. Teorie superwencji

Tezę o superwencji zdarzeń mentalnych można rozumieć jako jeden z przypadków tezy o identyczności zdarzeń (egzemplarzy)³⁰. Według D. Davidsona, pojedyncze mentalne zdarzenia $[\Psi_i]$, $[\Psi_m]$ są identyczne z pojedynczymi fizycznymi zdarzeniami $[\Phi_j]$, $[\Phi_n]$ ³¹. Nie ma natomiast odpowiedniości pomiędzy typami zdarzeń mentalnych i typami zdarzeń fizycznych. Davidson neguje istnienie praw psychofizycznych. Pojęcia tego, co mentalne, mają bowiem zawsze jakieś racjonalne warunki ich aplikacji, podczas gdy stany fizyczne takich warunków nie mają. Według Davidsona „zdarzenia” są ograniczone do tego, co zachodzi w czasie i przestrzeni³². W tym sensie zdarzenia są także niepowtarzalnymi indywiduami.

Jedną z metod opisu relacji pomiędzy cechami mentalnymi i cechami fizycznymi jest odwołanie się do jakiejś wersji teorii superwencji. Teorie superwencji – mówiąc ogólnie – posiadają dwie podstawowe charakterystyki.

(i) Cechy superwencyjne (mentalne, $[c\Psi]$) nie muszą być identyczne z cechami podstawowymi (fizycznymi $[c\Phi]$), chociaż są od nich zależne.

(ii) Szereg cech lub faktów mentalnych $[\Psi_i-\Psi_m]$ superwencjuje na szeregu cech lub faktów fizycznych $[\Phi_j-\Phi_n]$, jeżeli nie ma zmian w szeregu $[\Psi_i-\Psi_m]$, bez zmian w szeregu $[\Phi_j-\Phi_n]$. Cechy subwencyjne powodują (na sposób konieczny) występowanie cech superwencyjnych $[c\Psi]$. Powołując się na prace J. Kima, można przyjąć następujące twierdzenie: cechy mentalne $[c\Psi]$ superwencjują (są nadbudowane) na cechach fizycznych $[c\Phi]$ w tym sensie, że jeżeli coś urzeczywistnia mentalną cechę $[c\Psi]$ w czasie t , to istnieje podstawowa, fizyczna cecha $[c\Phi]$, taka, że przedmiot posiada $[c\Phi]$ w czasie t , i w sposób konieczny coś posiadającego $[c\Phi]$ w czasie t , posiada $[c\Psi]$ w czasie t .

³⁰ Tezy o superwencji (wraz z podaniem odpowiedniej literatury) omówiono krótko w: Bremer 2001: 122–127. Formalne zapisy relacji zachodzących w poszczególnych typach superwencji oraz krytykę też J. Kima podaje A. Wykowska (2002: 27–55; por. Poczobut 2000: 31–32).

³¹ W swojej teorii umysłu Davidson zakłada holizm zdarzeń mentalnych (por. Bremer 2001: 120). Po linii holizmu Davidsona i przeciwko behawioryzmowi argumentuje także S. L. Hurley: „When we conceive of intentional agents at the personal level, we make sense of what they do holistically” (Hurley 1998: 5, por. s. 424–428).

³² Davidson przyznaje, odpowiadając na krytykę W. V. O. Quine’a dotyczącą kryterium identyfikacji zdarzeń, że proponowane przez Quine’a kryterium jest „neater, and better” (LePore i McLaughlin 1985: 175). Davidson kwestionuje tym samym swoje przy czynowe kryterium identyfikacji zdarzeń (zdarzenia są identyczne, gdy pokrywają się we wszystkich swoich przyczynach i skutkach) i przyjmuje kryterium Quine’a: dwa zdarzenia są identyczne, gdy zachodzą w tym samym czasie na tym samym miejscu.

Jakie relacje zachodzą pomiędzy: (a) różnymi cechami mentalnymi oraz (b) cechami mentalnymi i fizycznymi?

(a) Powiedzmy, że jakiś przypadek mentalnej cechy $[c\Psi]$ wywołuje powstanie innej cechy $[c\Psi_j]$. Pytaniem jest, w jaki sposób powstaje cecha $[c\Psi_j]$. Aby na to pytanie odpowiedzieć, przechodzimy do kwestii (b).

(b) Cecha $[c\Psi]$ ma swoją superweniencyjną podstawę w fizycznej cesze $[c\Phi]$, a cecha $[c\Psi_j]$ ma swoją podstawę w cesze $[c\Phi_j]$. Cecha $[c\Psi_j]$ jest urzeczywistniona, ponieważ urzeczywistniona jest cecha $[c\Phi_j]$. Można powiedzieć tak: jak długo występuje $[c\Phi_j]$, musi występować $[c\Psi_j]$. Dzieje się tak bez względu na to, czy występowanie $[c\Psi_j]$ jest lub nie jest poprzedzone występowaniem $[c\Psi]$. W tym sensie występowanie $[c\Phi_j]$ jest warunkiem koniecznym i wystarczającym wystąpienia $[c\Psi_j]$.

Wracając do (a) możemy powiedzieć, że aby cecha $[c\Psi]$ była w przyczynowej relacji do cechy $[c\Psi_j]$, cecha $[c\Psi]$ musiałaby być przyczyną cechy $[c\Phi_j]$. Jednak uwzględniając przyczynową zamkniętość obszaru tego, co fizyczne $[\Phi]$ należy powiedzieć, że $[c\Phi]$ jest wystarczającą przyczyną dla $[c\Phi_j]$. Tym samym, jeżeli $[c\Psi]$ powodowałaby (w sensie przyczynowym) $[c\Phi_j]$, to albo naruszałoby to zasadę przyczynowej zamkniętości obszaru tego, co fizyczne, albo mielibyśmy do czynienia z przyczynową naddeterminacją (nałożenie dwóch różnego rodzaju przyczyn: $[c\Psi]$ i $[c\Phi]$) w stosunku do $[c\Phi_j]$. Uwzględniając to, co wyżej powiedziano stwierdza się, że pomiędzy $[c\Psi]$ a $[c\Psi_j]$ brak przyczynowego połączenia, co jest sprzeczne z (a).

Relacja superweniencji $[c\Psi]$ względem $[c\Phi]$ jest z jednej strony relacją nieredukcjonistyczną, a z drugiej strony musi ona przedstawiać wzajemną zależność (determinację), oraz jednostronną współzmiennność (czyli kowariancję). Relacja ta nie może być zbyt mocna (aby zachować tezę o nieredukcjonizmie), ani zbyt słaba (aby zachować tezę o zależności pomiędzy $[c\Psi]$ oraz $[c\Phi]$). Stąd tezy Kima o psychofizycznej superweniencji (por. Kim 1995: 183):

(A) Każdy wewnętrzny (*internal*) stan psychiczny jakiegos organizmu jest superweniencyjny na jakimś synchronicznym wewnętrznym stanie fizycznym. Równocześnie Kim przyjmuje tezę o silnej kowariancji (*strong covariance*, którą możemy także nazwać tezą o „silnej superweniencji”)³³:

(B) Dla dowolnych obiektów x i y i dowolnych światów w_i i w_j zachodzi: jeżeli x jest w świecie w_i , B-nierozróżnialne od y w świecie w_j , (to znaczy, x ma w w_i dokładnie te B-cechy, które y ma w w_j), wtedy x w w_i jest A-nierozróżnialne od y w świecie w_j .

Tezy (A) i (B) możemy odnieść do dwóch podmiotów, x i y . Otrzymamy wtedy:

³³ Kim odróżnia cztery typy wymienionej relacji kowariancji (por. Kim 1990: 10).

(C) jeżeli podmioty te są nieodróżnialne co do ich wewnętrznych stanów fizycznych, to dzielą one także te same stany mentalne. Teza (C) ma następstwa, gdy chodzi o metodologiczne określenie neuronalnych korelatów świadomości, zwłaszcza gdy chodzi o rozumienie czasowej zależności „zsynchronizowania”. Do tezy tej wrócimy w rozdziale siódmym.

Przez lata przyjmowano, że psychofizyczna superweniencja mogłaby gwarantować relację zależności stanów mentalnych od stanów fizycznych. Jednak nawet pojęcie silnej superweniencji nie implikuje stosunku zależności, lecz jedynie twierdzenie o silnej kowariancji³⁴. Dlatego zasada psychofizycznej superweniencji nie gwarantuje fizykalistycznego ujęcia stanów mentalnych. Co więcej, zasadę tę można pogodzić z epifenomenalizmem, okazjonalizmem i z dualizmem³⁵. Dodatkową cechą teorii superweniencji jest, że właściwie nie stawiają one sobie zadań wyjaśniających w sensie wyjaśniania rozumianego jak w teoriach redukcyjnych³⁶.

3.5. Funkcjonalizm

W teoriach funkcjonalistycznych stany mentalne podmiotu są określane poprzez trzy relacje: (a) ich odniesienie do zewnętrznych bodźców (*inputs*), (b) ich przyczynowe oddziaływania na inne stany mentalne, oraz (c) przez behawioralne reakcje ze strony podmiotu (*outputs*). Same twierdzenia (a)–(b) nie implikują jeszcze fizykalizmu, ponieważ nie mówią one nic o naturze stanów mentalnych (teoretycznie mogą to być także stany kartezjańskiej „substancji myślącej”).

Dlatego (a)–(b) są – podobnie do teorii superweniencji – uzupełniane o twierdzenie dotyczące fizycznej realizacji stanów mentalnych: jeżeli jakiś system x (powiedzmy, mózg) posiada w momencie czasu t mentalną cechę M , to mózg jest przedmiotem materialnym i posiada M dzięki posiadaniu w t fizycznej cechy F , która urzeczywistnia cechę M .

³⁴ Analizując pojęcie „silnej superweniencji” (SS) Grimes stwierdza: „For even in its strongest form, (SS), this concept fails to serve as a general form of dependency” (Grimes 1988: 159; por. Bremer 2001: 124). Kim stwierdza: „I am trying to keep the idea of covariance and that of dependence separate” (Kim 1990: 10).

³⁵ Por. Poczobut 2000: 36–37. Dla zagwarantowania chociażby minimalnego fizykalizmu, należałoby zasadę psychofizycznej superweniencji uzupełnić o tezę, że wspomniana korelacja ma swoją podstawę w fizycznej strukturze systemu.

³⁶ To, że relacja superweniencji nie stanowi metafizycznie głębokiej relacji, dostarczającej wyjaśniania np. problemu umysł–ciało, nie musi być jej wadą (por. Poczobut 2000: 36–38, 40–42).

Funkcjonalista odrzuca ontologiczną identyfikację (redukcję) typów stanów mentalnych oraz cech mentalnych z typami behawioralnych dyspozycji lub z cechami fizycznymi. Dąży on jednak do zdefiniowania stanów mentalnych przez pośrednie odwoływanie się do zachowań. Stany mentalne są definiowane w terminach ich abstrakcyjnych, funkcjonalnych ról, które są konstytuowane przez ich przyczynowe relacje typu (a), (b), (c)³⁷. Stany mentalne są tym samym definiowane w terminach przyczynowych ról, odgrywanych w determinowaniu zachowania się podmiotu w określonych warunkach. Typy stanów mentalnych są typami stanów funkcjonalnych, które uzyskują indywidualne cechy przez ich przyczynowe role albo cechy relacyjne, nie zaś przez strukturę czy kompozycję medium, w którym są one egzemplifikowane.

Zdaniem funkcyjnalisty, zachowanie podmiotu w konkretnych warunkach nigdy nie może zostać wyjaśnione tylko w terminach posiadania przez niego stanów mentalnych danego typu. Adekwatne wyjaśnienie wymaga uwzględnienia przyczynowych interakcji, mogących zachodzić pomiędzy różnymi stanami tegoż podmiotu. Funkcjonalista uznaje, że przypisywanie stanów mentalnych podmiotowi, odwołujące się do jego zachowań, jest otwarte dla różnorodnych interpretacji. Możemy jedynie prawdopodobnie określić, że podmiot posiada dane stany mentalne.

Oryginalność funkcyjnalistycznego ujęcia zdarzeń mentalnych ukazuje się przy przeciwstawieniu go ujęciom behawiorystyczno-nomologicznym i ujęciom proponowanym w teorii identyczności³⁸.

Teoria identyczności typów jest konfrontowana z tezą o wielorakiej realizacji stanów mentalnych: stany mentalne są stanami, które mogą być egzemplifikowane w różnych mediach. Stany mentalne nie są tym samym ograniczone do systemów biologicznych (mózgów). Także o innych systemach (na przykład o komputerach) można powiedzieć, że posiadają one coś w ro-

³⁷ Lowe zaznacza: „Another term for «causal role» in this sense is «functional role»: hence the name «functionalism»” (Lowe 2000: 45). Lowe odróżnia trzy rodzaje przyczynowych relacji, które mogą być związane z przyczynową rolą określonych typów stanów mentalnych. (a) Otoczenie podmiotu może być przyczyną posiadania przez niego określonych typów stanów mentalnych (na przykład padające na oko światło powoduje wrażenia wizualne). (b) Określone typy stanów mentalnych podmiotu mogą oddziaływać przyczynowo na inne jego stany mentalne (wrażenie wzrokowe powoduje powstanie przekonania o czymś). (c) Stany mentalne mogą powodować zachowania się podmiotu (np. zamknięcie oczu). Stany mentalne są faktycznie istniejącymi stanami podmiotu i można się do nich odwoływać w przyczynowym wyjaśnianiu jego zachowań.

³⁸ Porównanie funkcyjnalizmu z innymi teoriami materialistycznymi można znaleźć w: Block 1995: 191–193.

dzaju analogicznych stanów mentalnych. Tym samym pomocną metaforą dla wyjaśnienia różnicy pomiędzy zdarzeniem cielesnym a przeżywanym zdarzeniem mentalnym okazuje się rozróżnienie: *hardware–software* (por. Maloney 2001: 333–334).

Behawioryzm definiuje zdarzenia mentalne w ich relacji do sensorycznych wejść i behawioralnych wyjść (pod tę definicję podpadają jednak również systemy, którym raczej nie chcemy przypisywać stanów mentalnych). Funkcjonalizm uwzględni nie tylko relacje wyjścia i wejścia dotyczące danego stanu mentalnego, lecz także jego relacje przyczynowe z innymi stanami wewnętrznymi.

Funkcjonalista nie musi się wypowiadać na temat identyczności stanów mentalnych podmiotu ze stanami jego mózgu. Dzięki funkcjonalnej identyfikacji typów stanów mentalnych z typami stanów funkcjonalnych może on podać definicyjne cechy stanów mentalnych oraz może dokonywać uogólnień dotyczących natury tego, co mentalne, których ani teoria identyczności typów, ani teoria identyczności zdarzeń nie potrafią dostarczyć. Chociaż ogólne cechy tego, co mentalne, nie są (po prostu) fizyczne, mogą one uchodzić za funkcjonalne cechy fizycznie opisywanych systemów (a może i fizycznych systemów).

Zdaniem S. Shoemakera, czysto behawioralne definicje stanów mentalnych są obarczone swoistą cyrkularnością (por. Shoemaker 1984: 261–263). Ktoś przekonany, że pada deszcz, może być dysponowanym do zabrania parasola. Tak jest jednak tylko wtedy, gdy chce się uchronić przed deszczem. Ktoś, chcący się uchronić przed deszczem, może być dysponowanym do zabrania parasola, jednak tylko wtedy, gdy jest przekonany, że pada deszcz. Celem podania behawioralnej definicji któregoś z tych stanów mentalnych (przekonania, chcenia), musimy więc wymienić inny stan. Tym samym pojawia się kwestia możliwości sformułowania niecyrkularnej, czysto dyspozycyjnej definicji obydwu tych stanów, albo czysto behawioralnej definicji jednego z nich. Aby wyjść z tego koła, Shoemaker proponuje założyć, że istnieje teoria T , której wszystkie terminy są psychologicznymi predykatami. Φ_1 jest zmienną zastępującą „jest przekonany, że pada”, a Φ_2 jest zmienną zastępującą „chce mieć suche ubranie”, zaś zdanie Ramseya o T może być napisane jako: $\exists\Phi_1\dots\exists\Phi_n: T(\Phi_1\dots\Phi_n)$. Wtedy, według Shoemakera:

- (1) „ x jest przekonany, że pada $\equiv \exists\Phi_1\dots\exists\Phi_n: [T(\Phi_1\dots\Phi_n) \& x$ posiada $\Phi_1]$ ”,
- (2) „ x chce mieć suche ubranie $\equiv \exists\Phi_1\dots\exists\Phi_n [T(\Phi_1\dots\Phi_n) \& x$ posiada $\Phi_2]$ ”.

Funkcjonalna definicja cechy „bycia przekonany, że pada” będzie identyfikowała tę cechę z cechą wyrażoną przez predykat po prawej stronie znaku równoważności w (1). Podobnie z funkcjonalną definicją (2) „chcieć mieć suche ubranie”. Predykaty te kwantyfikują cechy lub stany mentalne,

ale nie wyszczególniają jakiegoś specyficznego stanu, tak więc nie ma cyrkularności w definiowaniu przekonania w terminach (1) i pragnienia w terminach (2).

Identyfikując stany mentalne z abstrakcyjnymi funkcjonalnymi rolami, funkcjonalista potrafi zaakceptować intuicję S. Kripkego, że zdania o tożsamości stanów mózgu i wrażeń mają status zdań przygodnych. Do wyjaśnienia intuicji, że ból może istnieć przy braku stymulacji włókien C, można przywołać tezę o wielorakiej realizacji bólu. Pewnej uwagi wymaga jednak pytanie o granice odwoływania się do tej intuicji. Większość funkcjonalistów deklaruje się jako materialistyczni moniści, ale ponieważ stany mentalne nie mogą być redukowane do wywołujących je neurologicznych lub komputerowych mechanizmów, to są one czymś więcej aniżeli tylko stanami fizycznymi.

N. Block odróżnia dwa podstawowe typy funkcjonalizmu: „funkcjonalizm maszynowy”³⁹ i „funkcjonalizm teleologiczny”. Według przedstawicieli tego ostatniego, stan fizyczny jakiegoś organizmu zalicza się do stanów urzeczywistniających dany opis funkcjonalny tylko wtedy, gdy organizm ten posiada faktyczną organiczną integralność, a stan ten odgrywa swoją funkcjonalną rolę, właściwą dla tego organizmu w teleologicznym sensie terminów „dla” oraz „funkcja”. Dany stan musi wywoływać to, co wywołuje, jako realizację swojego biologicznego celu.

Do tego typu funkcjonalizmu możemy zaliczyć funkcjonalizm reprezentowany przez D. Dennetta, który zostanie omówiony w 4.3.4. Dennett traktuje zachowania behawioralne jako manifestacje posiadanych przez podmiot zdolności. Zdolności te są wyjaśnialne przez odwołanie się do podmiotu jako do systemu wzajemnie połączonych komponentów. Dennett dekomponuje stany mentalne w hierarchię przypominającą wielką wspólnotę, zawierającą kooperujące ze sobą jednostki, podjednostki itd., aż do osiągnięcia poziomu neuronalnego. Każda z tych jednostek jest identyfikowana przez odniesienie do funkcji, którą wykonuje.

Jeżeli przyjmiemy, że zakładana przez funkcjonalizm globalna nieredukowalność cech mentalnych oznacza, że nie istnieją ogólne (globalne) psychofizyczne prawa łączące stany mentalne z fizycznymi, to nie oznacza to jeszcze, że cechy mentalne (rozumiane jako funkcjonalne), są emergentne (por. 3.6.2).

Funkcjonalizm musi się liczyć z konsekwencjami wynikającymi z ekstermalizmu (por. 1.3.5). Z jednej strony funkcjonalista może zaakceptować ujęcie, że stan wewnętrzny jest stanem przekonania dzięki posiadaniu odpowie-

³⁹ Możemy go nazwać „funkcjonalizmem maszyny Turinga” i zaliczyć do niego funkcjonalizm reprezentowany przez H. Putnama (por. Bremer 2001: 107–108).

dniej wewnętrznej funkcjonalnej roli. Treść takiego wewnętrznego stanu może zależeć od jego relacji z otoczeniem zewnętrznym. Z drugiej strony funkcjonalista może tak rozszerzyć swoją teorię, że stanie się ona kompatybilna z eksternalizmem. Funkcjonalista twierdzi, że stan mentalny jest zdefiniowany poprzez jego funkcjonalne role (w relacji do wejść, wyjść i innych stanów mentalnych). W potocznych sformułowaniach funkcjonalizmu wejścia i wyjścia są zwykle traktowane jako obiekty czy stany należące do ciała podmiotu, a nie do jego zewnętrznego otoczenia. Zgodnie z tezami „aktywnego eksternalizmu” ograniczenie to może zostać zniesione.

Kończąc ogólną charakterystykę funkcjonalizmu, przejdziemy w następnych punktach do omówienia teoretycznego rozumienia stanów przeżyciowych, proponowanego przez jego przedstawicieli.

3.5.1. Funkcjonalistyczna analiza „trudnego” problemu

Funkcjonalista W. Robinson przedstawia jedność naszych stanów mentalnych, bez odwoływania się do substancjalnej jaźni lub aktów zdawania sobie sprawy. Neguje on zarazem twierdzenie, że umysł może być całkowicie wyjaśniony w terminach struktur fizycznych. Jego teoria dopuszcza dualistyczne wyjaśnianie wrażeń i fizykalistyczne wyjaśnianie intencjonalności. Robinson nie twierdzi, że „trudny” problem (HP) faktycznie istnieje, interesuje go jedynie to, dlaczego problem ten się pojawia (por. Robinson 1996: 17–19). Parafrazując tezy Robinsona można zadać pytanie:

(HP) Dlaczego podmiot S powinien mieć świadome przeżycie typu $[\Psi_i]$, zawsze wtedy, gdy zachodzi w nim neuronalne zdarzenie typu $[z\Phi_j]$,

(HP) ani nie zakłada, że *tylko* neuronalne zdarzenia $[z\Phi_j]$ stoją w regularnym związku ze świadomym przeżyciem $[\Psi_i]$, ani nie określa relacji pomiędzy świadomym przeżyciem $[\Psi_i]$, a zdarzeniami neuronalnymi. Możemy więc mówić zarówno o relacji przyczynowej, o relacji identyczności, albo o emergencji, jako o relacjach specyfikujących „trudny” problem, albo o neuronalnych korelatach świadomości, jako specyfikujących „trudny” problem (por. rozdział VII).

Robinson argumentuje podobnie do Nagla, twierdząc, „[...] że nie potrafimy dać odpowiedzi na (HP) w ramach naszego obecnego systemu pojęć, teorii i podstawowych zasad”⁴⁰.

⁴⁰ Robinson nie akceptuje pesymistycznej tezy McGinna, że nigdy nie będziemy mogli odpowiedzieć na pytanie (HP) [por. Robinson 1996: 15, 24].

Dwa fakty prowadzą nas do konkluzji, iż nie umiemy odpowiedzieć na pytanie (HP).

(F1) Wyjaśnianie regularności pomiędzy wystąpieniami dwóch rodzajów faktów zachodzi przez wyprowadzenie odpowiedniości pomiędzy strukturami tych wystąpień.

Twierdzenie (F1) jest zbliżone do twierdzeń podanych przez D. Chalmersa, a dotyczących funkcji i struktury stanów mentalnych. Twierdzenia te nie zostały jednak przez niego wyczerpująco zanalizowane (por. Chalmers 1995: r. III–V). Robinson uzupełnia (F1) o twierdzenie:

(F2) Jeżeli regularność, zawierająca dany rodzaj faktów ma być wyjaśnialna, to rodzaj ten musi być wyrażalny jako jakaś struktura.

Do (F1) i (F2) zostają dodane dwa twierdzenia, pomocne przy wyjaśnianiu „trudności” trudnego problemu:

(F3) Gdziekolwiek możemy znaleźć jakąś strukturę, tam możemy oczekiwać znalezienia jej wyjaśnienia. Na przykład ból możemy opisywać poprzez rytmiczny wzrost i spadek jego intensywności. Taki opis dostarcza jakiegoś rodzaju struktury, którą możemy łączyć z aktywnością niektórych neuronów w mózgu, odpowiadających za wzrost i spadek intensywności bólu (por. np. Coghill 1999: 1937).

(F4) Pośród wszystkich cech przeżyć świadomych, istnieje zawsze przynajmniej jedna, która nie posiada strukturalnego sposobu wyrażania.

Tezy (F2) i (F4) [wzięte razem] stoją w opozycji, implikując tym samym brak możliwości rozwiązania „trudnego” problemu. Robinson pyta: dlaczego jednak powinniśmy przyjąć (F4)? Przykład z pulsującym, rwącym bólem pokazuje, że niektóre cechy świadomego przeżycia mogą faktycznie zostać wyjaśnione. Niemniej jednak przykład ten zawiera *jakościową* cechę: rytmicznie wzrastający i malejący ból. Cecha ta nie posiada, w sposób konieczny, wewnętrznej struktury. Bóle są zróżnicowane, różnią się intensywnością czy inną jakością zależnie od tego czy innego miejsca w ciele⁴¹. Wydaje się jednak, że tego rodzaju zróżnicowanie nie jest *konieczne*, gdyż intensywność i inne jakości bólu *mogą* być jednolite bez względu na jakiś cielesny region. Ponadto, typowa dla bólu zmienność nie jest tym, co konstytuuje jakieś odczucie jako bolesne.

Dodatkowo tezy (F2) i (F4) odnoszą się do świadomego przeżywania w ogóle, a nie tylko do bólu. Na przykład obrazy powidokowe (powidoki, *after image*) *mogą* być jednolitego koloru, doświadczenia słuchowe *mogą*

⁴¹ Neurolog F. Strian podaje przykładowe określenia bólu: „[...] palący, kłujący, rwący, napinający, elektryzujący, raniący, kruszący [...]” (Strian 1996: 44).

mieć czasowo stałą tonację⁴². W żadnym z tych przypadków nie istnieje szczególny rodzaj zmienności, który konstryuuje jakość. Faktycznie, nie miałoby to sensu. Wymienione cechy są treściami, które się zmieniają wtedy, gdy mamy zmienne świadome przeżycia.

Ktoś może zaoponować, że barwy powidokowych obrazów posiadają swoistą strukturę, mianowicie podział na odcień i nasycenie. Robinson potwierdza, że przy pulsowaniu bólu, gdy raz rozpoznamy tę strukturę, możemy mieć nadzieję na uzyskanie szukanego wyjaśnienia. Podobnie w pełni zrozumiałe jest (choć nie może być zagwarantowanym *a priori*), poszukiwanie jakiejś neuronalnej cechy (neuronalnego korelatu [$c\Phi_1$]), która by dostarczała jakiegoś słabego nasycenia koloru poprzez różne odcienie, i jakiejś innej neuronalnej cechy [$c\Phi_2$], która dostarczałaby danego odcienia poprzez różne stopnie nasycenia. Jeżeli potrafilibyśmy znaleźć takie cechy, to moglibyśmy się odwołać do ich obecności, aby wyjaśnić, dlaczego, przy jakiejś szczególnej okazji, podmiot posiada obraz powidokowy o określonym odcieniu i stopniu nasycenia.

Tego rodzaju wyjaśnienie zakłada jednak swoistą prawidłowość pomiędzy, na przykład, jakąś neuronalną cechą i odcieniem obrazu powidokowego. Wyjaśnienie, *dlaczego* taka prawidłowość powinna tutaj zachodzić, odsyła nas z powrotem do kwestii „trudnego” problemu. Mamy bowiem (jakiś szczególny) odcień obrazu jako nasz cel wyjaśniania, a nie posiadamy sposobu wyrażenia go jako struktury.

Robinson traktuje funkcjonalistyczne ujęcia problemu umysł–ciało jako próbę naświetlenia problemu wyjaśniania w ogóle. Polega ona na podaniu jakiegoś (domniemanego) wyrażenia cech świadomego przeżycia w strukturalnej formie, to znaczy jako struktury przyczynowych relacji. Robinson rozróżnia dwie wersje funkcjonalizmu.

(V1) Bóle są świadomymi przeżyciami, spowodowanymi przez określoną przyczynę (na przykład przez zranienie), która normalnie wywołuje określone skutki (jęk, branie pigułek, i tym podobne).

W wersji (V1) mamy zaproponowany pewien sposób identyfikacji tego przeżycia, które spośród wielu innych możliwych świadomych przeżyć jest nazywane *bolesnym*. Nie dostarcza ona jednak strukturalnego wyrażenia sa-

⁴² Rozważania Robinsona dotyczące powidoków (rozumiemy, że chodzi tutaj o wywołanie barwy dopełniającej dla każdego uczucia „barwnego”) nie są istotnym argumentem w kwestii stany neuronalne – stany mentalne. Obrazy powidokowe nie dotyczą tylko siatkówki (gdyż występują w różnych modalnościach percepcyjnych), jednak zawsze są następstwem bezwładności procesów chemicznych i fizycznych w samych receptorach, tzn. zmęczenia pewnych obszarów siatkówki (por. Zeugner 1965: 108-117).

mych przeżyć. Tym samym zakłada ona taką regularność (pomiędzy okolicznościami i cechami świadomych przeżyć), której nie potrafi wyjaśnić. Ze względu na to ograniczenie Robinson analizuje silniejszą wersję funkcjonalizmu:

(V2) Ból jest jakimkolwiek zdarzeniem wywołanym zazwyczaj w określony sposób, powodującym określone skutki (jęk, branie pigułek, i tym podobne) (por. Robinson 1996: 18).

Aby nie przejść z powrotem do wersji (V1), wersja (V2) musi być uzupełniona o „obojętnie, czy to zdarzenie odczuwa się jako coś lub nie”. Funkcjonalizm w wersji (V2) zawiera ujęcie relacjonalnej struktury, znaczącej dla bólu. Struktura ta nie jest jednak strukturą cechy [cΨ] świadomego przeżywania (w tym przypadku bolesności), a jedynie strukturą relacji, w które to przeżywanie wchodzi. Potrafimy, w zasadzie, zidentyfikować neuronalne zdarzenie [Φ_j], które posiada przyczynowe relacje wymienione w (V2), a to zdarzenie może mieć strukturę (na przykład jakiś porządek zmiany potencjałów w różnych neuronach). Ale nie jest to struktura cechy bycia bolesnym, dlatego wyjaśnienie tej struktury nie będzie wyjaśnieniem bolesności. To znaczy, wyjaśnienia tej struktury nie będą wymagane do rozwiązania „trudnego” problemu. Zwolennicy (V2) mogą wprawdzie negować istnienie jakiejś cechy bolesności [cΨ], którą należy przedstawić, jednak taka negacja oznacza po prostu odrzucenie „trudnego” problemu. Mogą oni także zaakceptować tę cechę i próbować ją przedstawić w inny sposób. Jednak w każdym z tych przypadków wersja (V2) nie podaje rozwiązania omawianego problemu.

Pytając, dlaczego powinniśmy przyjąć (F4), Robinson zaznacza, iż żadna z odmian funkcjonalizmu nie może być zaakceptowana jako dostarczająca tego rodzaju struktury, o której (F4) mówi, że ją wyklucza. Inne dwa przypadki, które rozważyliśmy, posiadają następujący charakter: dostarczają jakiejś struktury, która zakładając (F3), stwarza nam nadzieję, że może nam umożliwić jakieś wyjaśnienie – implikują one jednak, że istnieją elementy w tej strukturze. Elementy te ponownie wprowadzają „trudny” problem. Dlatego, zdaniem Robinsona, jak długo nie będziemy mieli jakiegoś innego rodzaju wyjaśniania, to fakt, że każde wyjaśnienie jednej cechy wprowadza na nowo „trudny” problem dla innej cechy przekonuje o tym, że teza (F4) jest faktycznie poprawna.

Wynika stąd, że „trudny” problem jest dlatego taki złożony, gdyż wyjaśnienia regularności mogą ujawniać tylko te cechy, które są wyrażane jako mające jakąś strukturę, a wszystkie świadome przeżycia mają cechy, które nie mają strukturalnego wyrażenia (por. Robinson 1996: 19). Stąd wspomniany wniosek Robinsona, że w ramach naszego obecnego systemu pojęć nie możemy rozwiązać „trudnego” problemu.

Odnosząc do (HP) uwagi T. Nagla możemy powiedzieć, że poszukiwany system pojęciowy (*conceptual framework*) [por. Nagel 1998a: 347] musi być wynikiem myślenia w świetle danych o podmiocie, który chcemy zrozumieć oraz myślenia w świetle wyszukiwania pojęć, które takiemu myśleniu lepiej odpowiadają aniżeli nasze obecne pojęcia. Do podobnego celu zmierzają w filozofii umysłu teorie fizykalistyczno-funkcjonalistyczne. Zdaniem Nagla, teorie te nie odniosły spodziewanego sukcesu. Funkcjonalizm próbuje tak reinterpretować pojęcia tego, co mentalne, aby je uczynić odpowiednimi elementami systemu pojęć nauk szczegółowych. Zamiast tego potrzebujemy jednak analiz, które rozpoczną się od pojęciowej niezrozumiałości w jej obecnej formie, od połączenia ujęcia subiektywnego z obiektywnym. Próbę tego rodzaju analiz podajemy w rozdziale VII.

Według Robinsona, trudności, które mamy z „trudnym” problemem, pochodzą od natury przypisywanej świadomym przeżyciom oraz z wymagań wyjaśniania jej. Można przyjąć, że natura świadomych przeżyć nie ulegnie zmianie. „Trudny” problem pozostanie zatem tak długo problemem, jak długo nasz system pojęciowy zachowa swoje wymagania dotyczące wyjaśniania. Aby rozwiązać „trudny” problem, musimy znaleźć inne aniżeli dotychczasowe sposoby wyjaśniania.

Na zakończenie Robinson odpowiada na zarzut, że proponowane przez niego przedstawienie jest uproszczone, i nie wymienia subiektywności świadomych przeżyć (mimo iż ciągle używa on słowa „świadomość”). Robinson zaznacza, iż odrzuca relacyjny rodzaj patrzenia na świadomość zjawiskową. Nie mówi więc o „podmiocie”, w relacji do którego stoją świadome przeżycia. Odwołujące się do wspomnianej relacji teorie utrzymują, że dla podmiotu X bycie świadomym podmiotem, to bycie w jakiejś relacji R do jakiejś innej rzeczy Y . Teorie takie różnią się w zależności od tego, jak jest rozumiane owo R oraz różnią się w opisach tego, jakim rodzajem rzeczy musi być Y ⁴³. Konsekwencją odrzucenia teorii relacyjnej jest to, że świadome przeżycia nie są traktowane jako świadome dzięki relacji do czegoś innego. Są one świadome wewnętrznie (intrynsecznie). Gdy chodzi o ujęcie relacji świadomości do podmiotu, to można powiedzieć, że Robinson przyjmuje holistyczną teorię świadomości. Nie przyjmuje jej natomiast, gdy chodzi o świadomość w sensie całościowego stanu przeżyciowego.

⁴³ Robinson uważa, że D. M. Rosenthal jako relację R proponuje „bycie myślą o”, gdzie Y jest treścią myśli (por. Robinson 1996: 20). Do obrońców teorii relacyjnej Robinson zalicza: D. M. Rosenthala, P. Carruthersa, G. E. Moore’a i D. Armstronga.

3.5.2. Relacyjne wyjaśnianie stanów jakościowych

Reasumując prowadzone przez Robinsona analizy można powiedzieć: Jeżeli funkcjonalista uznaje istnienie jakiejś rzeczywistości mentalnej, to potrafi on jedynie podać relacyjne wyjaśnienie jakościowych stanów mentalnych. Wyjaśnienie takie odbiega od tego, w jaki sposób stany te identyfikujemy z perspektywy pierwszej osoby. Przyjmijmy, że ból występuje tak, jak jego występowanie jest rozumiane w funkcjonalizmie (to znaczy występuje tylko wtedy, gdy zostaje spełniona określona funkcjonalna czy przyczynowa rola). Wówczas jednak należy sprawdzić, czy taka rola wystarczy do identyfikowania tego, czego ja sam doświadczam jako bólu. Jeżeli mogę sobie zdawać sprawę ze wszystkich znaczących wejść, to wiele wyjść jest zachowaniem wynikającym z tych wejść, zachowaniem leżącym w przyszłości w stosunku do czasu, w którym dokonuję identyfikacji. A tym samym, moje zdawanie sobie sprawy z bólu, nie może po prostu być funkcjonalistycznym zdawaniem sobie sprawy z wejść i wyjść.

Ból identyfikujemy dzięki jego zjawiskowo-przeżyciowemu charakterowi. Nie bierzemy przy tym pod uwagę, przynajmniej na pierwszym miejscu, jego przyczynowej roli. Co więcej, rola ta nie jest tym, o czym zazwyczaj myślimy, gdy mówimy o bólu.

III.8 Specyficzny charakter bólu, jako bólu, jako coś różnego od naukowego wyjaśniania bólu albo przewidywania jego wystąpienia, nie odnosi się do przyczynowej roli bólu, lecz do przeżywania w sensie: *jak to jest*, gdy boli.

To, co wydaje się oczywiste w przypadku bólu, zdaje się być bardziej oczywiste, gdy chodzi o identyfikację wrażeń kolorów. Ich związek z wejściami i behawioralnymi wyjściami wydaje się być jeszcze słabszy.

W tym sensie należy rozumieć argumentację J. Searle'a (1999: 49–95), zestawiającego siedem „falszywych” twierdzeń, tworzących „podstawy” współczesnego materializmu. Należy do nich funkcjonalistyczna teza, że stany mentalne są definiowalne przez relacje przyczynowe, w jakich się one znajdują (por. Searle 1999: 40–41). Teza ta jest, według Searle'a, jedynie aplikacją standardowych, obiektywnych metod naukowych do badań nad świadomością. Nie uwzględnia ona istnienia jakiegoś „ja”, które posiada stany mentalne, nie uwzględnia tego, że ontologia tego, co mentalne, jest ontologią z punktu widzenia pierwszej osoby (w tym sensie możemy powiedzieć, że Searle'a ujęcie świadomości jest relacyjne). Searle podkreśla, że w prowadzonych przez niego analizach świadomości pierwotnym jest pierwszoosobowy punkt widzenia.

Funkcjonalizm odwraca się od subiektywności, przedefiniowując ontologię w kategoriach epistemologii oraz przyczynowości i twierdzi: stany mentalne są definiowane przez relacje przyczynowe. Z obiektywnego punktu widzenia trzeciej osoby jedyną odpowiedzią na epistemologiczne pytanie: skąd możemy wiedzieć o stanach mentalnych u innych ludzi?, jest odpowiedź: z obserwacji zachowań.

Ludzka świadomość jest dla Searle'a zjawiskiem subiektywnym, do którego nie ma dostępu metodami redukcjonistycznymi. Subiektywność świadomości uniemożliwia jej redukcję do czegoś fizycznego. Searle dodaje jednak:

„Nieredukowalność świadomości nie ma żadnych niefortunnnych konsekwencji dla nauki. Nadto, mówiąc o nieredukowalności świadomości, mówię tylko o jej nieredukowalności na gruncie standardowych modeli redukcji” (Searle 1999: 170).

Searle nie wyklucza, że poszukiwana, wyjaśniająca redukcja może zostać przeprowadzona w przyszłości, a tym samym program funkcjonalistyczny może być zrealizowany.

Ponadto, jeżeli zasadnicze tezy funkcjonalizmu są poprawne, wtedy dwa podmioty (czy ten sam podmiot w różnych przedziałach czasu) nie mogą mieć różnych stanów mentalnych, jeżeli nie mają one różnych stanów funkcjonalnych. Różnica mentalna zachodzi tylko przy różnicy funkcjonalnej. W związku z tym rodzą się dwa pytania:

(a) Czy może zatem zachodzić różnica jakościowa bez jakiejś różnicy funkcjonalnej? Musimy przyznać, że nie wiemy czy faktycznie istnieją tego rodzaju empiryczne przypadki, dlatego jesteśmy po części skazani na analizie „eksperymentów myślowych”.

(b) Pytaniem jest: czy funkcjonalizm potrafi uwzględnić jakościowy, jak i subiektywny, charakter stanów mentalnych? Czy w funkcjonalistycznych analizach można znaleźć miejsce dla owego *jak to jest być* w danym stanie mentalnym? To właśnie odczuwanie w rodzaju *jak to jest być* powoduje, że zjawiskowe stany mentalne są tymi a nie innymi stanami mentalnymi. Trudność funkcjonalizmu z ujęciem zjawiskowego charakteru stanu mentalnego wynika z abstrakcyjności funkcjonalnych ról. Można sobie bowiem (na przykład) wyobrazić różne znaczenia „qualiów”, odgrywające tę samą funkcjonalną rolę. Dopuszczenie takiej możliwości oznacza jednak uznanie, że w opisach, odwołujących się do funkcjonalnych ról, nie został uchwycony jakościowy aspekt tego, co mentalne.

Do eksperymentów myślowych z odwróconym spektrum (*inverted spectrum*) i z nieobecnymi qualiami (*absent qualia*) odwołują się dwa znane za-

rzuty przeciwko funkcjonalistycznej teorii qualiów (por. Chalmers 1996a: 369). Eksperyment myślowy z odwróconym spektrum pochodzi od J. Locke'a, zdaniem którego, ta sama rzecz miałaby w umyśle różnych ludzi wywoływać rozmaite idee. Na przykład idea, jaką za pośrednictwem oczu fiołek wywołuje w umyśle jednego podmiotu, byłaby tożsama z ideą, wywołaną przez nagietek w umyśle innego, i *vice versa*⁴⁴. Eksperymenty myślowe z „nieobecnyymi”, „znikającymi” i „tańczącymi” qualiami zawdzięczamy N. Blockowi, J. Searle'owi i D. Chalmersowi⁴⁵. Jeżeli funkcjonalizm nie potrafi ująć przypadków opisanych w tych eksperymentach, to trudno uznać, że proponuje on adekwatną teorię świadomości zjawiskowej (por. pytanie [b]).

Wskazując na twierdzenia Searle'a należy dodać, iż trudno zrozumieć, jak funkcjonalizm mógłby podać adekwatne ujęcie qualiów. Chcąc je opisać lub wyjaśnić, musi się on odwołać do cech funkcjonalnych lub przyczynowych. A te są cechami relacyjnymi. Czym jednak mogłaby być czysto relacyjna charakterystyka homogenicznego koloru? Operujący z perspektywy trzeciej osoby funkcjonalista, niejako sam generuje problem qualiów. A to dlatego, że jakiś przeżyciowy stan mentalny posiada nie tylko cechy relacyjne, ale także wewnętrzne. Ktoś może powiedzieć, że zjawiskowy kolor jest cechą zjawiskowego indywiduum. Sama negacja istnienia takich indywiduów nie będzie jednak wystarczająca dla uniknięcia problemu qualiów. Funkcjonalne analizy, odwołujące się do zdarzeń, również generują te same problemy.

Chociaż nie potrafimy przedstawić sobie wewnętrznych cech stanu mentalnego jako innych cech, aniżeli one są, możemy sobie wyobrazić, jak wchodzi on w całkiem odmienne relacje z innymi stanami i obiektami. Możemy powołać się na eksperymenty myślowe, w których przedstawiamy sobie świat, w którym dany stan zatrzymuje swoje wewnętrzne cechy, a jego relacje się zmieniają. Ponieważ funkcjonalizm nie dostarcza opisu wewnętrznych cech danego stanu, nie dostarcza on także opisu, czym ten stan jest.

3.5.3. Teoria zmysłu wewnętrznego

Funkcjonalista W. G. Lycan proponuje materialistyczną teorię umysłu, (*inner sense theory*) w której ramach subiektywny charakter tego, co mental-

⁴⁴ Por. Locke 1955: II, rozdz. 32, par. 15. A. Byrne i D. R. Hilbert negują hipotezę, że przypadki z odwróconym spektrum są empirycznie możliwe (por. Byrne i Hilbert 1997: 268).

⁴⁵ Por. Block 1978: 278–279; por. Searle 1980: 420–421; 1999: 39, 69; por. Chalmers 1996a: 370–373; 1996: 253–270. Argumentację Chalmersa przedstawia W. Dziarnowska (2001: 179–185) oraz A. Stephan (1999: 193–194).

ne, jest rozumiany jako coś obiektywnego. W książce *Consciousness and Experience* Lycan broni funkcjonalizmu, przyjmując:

(a) Założenie Locke'a o „zmyśle wewnętrznym” (*internal sense*) albo „refleksji”⁴⁶, według której świadomość jest rodzajem wewnętrznego monitoringu (obrazowania).

(b) Założenie, że cechy mentalne są sprowadzalne do cech przedstawialnych lub funkcjonalnych. Lycan twierdzi: „[...] umysł nie posiada specjalnych cech, które nie mogłyby zostać wyczerpane poprzez jego cechy przedstawialne (*exhausted by its representational properties*) oraz poprzez ich ewentualną kombinację z funkcjonalną organizacją jego komponentów” (Lycan 1987: 11).

Założenia (a) i (b) odnoszą się do cechującej się jednością, funkcjonalistycznej teorii umysłu i świadomości. Założenia te są od siebie niezależne (to, że do świadomości należy monitorowanie, nie zawiera w sobie tezy, że monitorowany stan sam jest przedstawialny [*representational*]). Mentalność jakiegoś stanu może istnieć w jego cechach przedstawialnych, mimo braku świadomości lub zdawania sobie sprawy ze stanów wyższego stopnia. Chcąc osłabić, opierające się na subiektywności świadomości, zarzuty przeciwko funkcjonalizmowi, Lycan odwołuje się do założenia (a). Jeżeli funkcjonalistyczna teoria (uwzględniająca (a) oraz (b)), okaże się skuteczna, to przyczyni się ona do uznania materialistycznej tezy, że ontologia umysłu jest czymś na wskroś obiektywnym.

Lycan łączy terminologię odnoszącą się do subiektywności z analizami prowadzonymi w ramach obiektywnej (materialistycznej) ontologii. Nie traktuje przy tym subiektywności jako kategorii ontologicznej, a tym samym nie musi przyjmować i uzasadniać istnienia subiektywnych faktów czy zdarzeń. Ontologia, według Lycana, zajmuje się tylko obiektywnymi zjawiskami. Lycan dopuszcza istnienie swoistego rodzaju obliczalnej informacji (*computational information*), dostępnej tylko z perspektywy pierwszoosobowej⁴⁷.

Przyjęcie wspomnianego rodzaju informacji jest związane z Lycana rozumieniem introspekcji, która jest procesem, dzięki któremu mózg skanuje

⁴⁶ Lycan 1987: 13. Według Locke'a wszystkie nasze przedstawienia pochodzą albo od zmysłów zewnętrznych, albo z refleksji (por. Locke 1955: II, kap. 1 par. 5).

⁴⁷ Lycan odróżnia *coarse-factual information* (warunki prawdy i desygnację) i *computational information* (zawiera funkcjonalne role mentalnych przedstawień). Dwa przedstawienia mogą wyrażać tę samą *coarse-factual information* (dany mężczyzna Piotr) przy różnej *computational information* (terminy: „Piotr”, „mąż Kasi” różnią się funkcjonalnie). Por. Hershfield 1998: 230.

swoje własne stany pierwszego stopnia (*first-order brain states*). Obrazowanie jest przeprowadzane przez mechanizm wytwarzający stany mózgowe drugiego stopnia, wyrażające informację o stanach pierwszego stopnia. Stan pierwszego stopnia jest świadomy na podstawie bycia obrazowanym, i ukazuje się w pewien sposób swojemu podmiotowi, ale samo obrazowanie nie jest już niczym obrazowane. Introspekcja jest więc formą wyższego stopnia mentalnego przedstawiania⁴⁸. Świadomość – według Lycana – jest funkcjonującym wewnętrznym mechanizmem uwagi, skierowanym na niższego rzędu psychologiczne stany i zdarzenia.

Świadomość pojawia się w różnych stopniach intensywności⁴⁹. Na przykład komputery mogą mieć niski stopień świadomości. Lycan nie podaje jednak, jak go zmierzyć. Ponadto Lycan twierdzi, że w jakimś pojedynczym ciele może istnieć nieokreślona liczba „świadomości” (*consciousnesses*), ale nie wiadomo, czy odnosi się to do jakichś jednostek chorobowych czy też nie.

Lycan dopatruje się subiektywności stanów mentalnych w dwóch cechach procesu obrazowania. (1) Prymitywne mentalne elementy (z których złożone są introspektywne przedstawienia drugiego stopnia) nie są wyrażalne w języku potocznym. (2) Inni mogą myśleć o naszych stanach pierwszego stopnia i odnosić się do nich, a tym samym wyrażać tę samą prostą faktyczną informację (*coarse-factual information*), jaką wyrażają nasze własne introspektywne przedstawienia drugiego stopnia.

Lycan uważa, że opisując subiektywność z trzecioosobowego punktu widzenia, szukamy jedynie lepszych opisów przeżywania zdarzeń jako całości. Według niego: „Subiektywność jest immanentna dla przeżywanego zdarzenia, lecz nie ogranicza nas ona w obiektywnym opisie całego zdarzenia” (Lycan 1987: 56).

⁴⁸ Według założenia Lycana o *hegemony of representation*: „[...] the mind has no special properties that are not exhausted by its representational properties. It would follow that once representation is (eventually) understood, then not only consciousness in our present sense but subjectivity, qualia, «what it is like», and every other aspect of the mental will be explicable in terms of representation without the positing of any other ingredient not already well understood from the naturalistic point of view” (Lycan 1997: 755).

⁴⁹ Podobne stwierdzenie przyjmuje Searle: „Świadomość może przyjmować różne stopnie intensywności nawet w czasie czuwania” (Searle 1999: 120). Lycan mówi, że jego faktyczny brak zgodności z Searle’em nie dotyczy stopni świadomości, lecz „[...] the question of whether a creature or device could have a much lower degree of consciousness than is ordinarily enjoyed by human beings and still qualify as being conscious at all” (Lycan 1997: 770).

Obliczalna informacja, wyrażona przez introspekcyjne pojęcie (w którym zawarte jest odniesienie do subiektywności tego, co mentalne), zawiera rodzaj prezentacji tego, do czego się odnosi. To, do czego się ona odnosi, jest pierwszostopniowym stanem mózgu, a rodzaj prezentacji jest określoną funkcjonalną rolą. Obydwa są obiektywnymi zjawiskami. Dlatego obliczalna informacja, wyrażona przez introspekcyjne pojęcie, jest dostępna intersubiektywnie. Powstaje pytanie, czy Lycan nie myli tutaj dwóch znaczeń pojęcia „informacja”: informacji cybernetycznej (rozumianej jako uporządkowanie systemu) i potocznej (rozumianej jako treść aktu świadomości). Ponadto należy zapytać, czy istnieją informacje „niedostępne” i „nieobliczalne” (skoro często przyjmuje się, że obliczalność i dostępność informacji wynika z jej definicji)⁵⁰.

Podstawowe założenie, z którego wychodzi Lycan, brzmi: jak przeddefiniować termin „subiektywność”, aby ten nie odnosił się do czegoś subiektywnego ale obiektywnego? Odpowiadając, Lycan przedstawia ściśle intencjonalną teorię umysłu. Zakres tego, co mentalne, jest wyczerpany przez cechy przedstawialne, a stany, które są świadome, są nimi na podstawie bycia przedstawialnymi (*being represented*). Naszym zdaniem natomiast:

III.9 Pod znakiem zapytania stoi funkcjonalistyczna teza, że opisując subiektywność z trzecioosobowego punktu widzenia, ani jej nie eliminujemy, ani jej nie zastępujemy uproszczonym, obiektywnym opisem.

Przedstawiona przez Lycana teoria umysłu jest w tym sensie teorią „intencjonalną”, że zakres tego, co mentalne, jest wyczerpany przez cechy przedstawialne (*representational features*), zaś stany, które są świadome, są nimi na podstawie bycia przedstawialnymi (*being represented*).

Dodatkowe, nieco uogólnione zarzuty wobec proponowanej przez Lycana teorii świadomości brzmią następująco.

Po pierwsze, Lycan rozwija Locke’a teorię zmysłu wewnętrznego (*inner sense theory*), twierdząc między innymi, że umożliwia ona najlepsze rozwiązanie problemu subiektywności, jak go rozumieją T. Nagel lub F. Jackson. Gdy chodzi o subiektywność, jego teoria implikuje rozróżnienie pomiędzy zjawianiem się a rzeczywistością (*appearance/reality distinction for subjec-*

⁵⁰ Powołując się na J. Ecclesa i R. Penrose’a W. Marciszewski stwierdza, że na poziomie kwantowym mózg ma zdolność do pewnych procesów, które nie mają charakteru algorytmicznego (obliczalnego). „Jeśli takie procesy rzeczywiście istnieją, to wśród nich mogłyby się znaleźć intuicje intelektualne” (Marciszewski 2004: 15). Por. także uwagi Marciszewskiego na temat „funkcji nieobliczalnych” (Marciszewski 2002: 13).

tivity). Między innymi dlatego jest ona – jak zobaczymy – krytykowana przez eliminatystę D. Dennetta⁵¹. Lycan zgadza się, że powyższe rozróżnienie jest dla problemu świadomego zdawania sobie sprawy raczej osobliwe i że dramatyczne przypadki występowania różnicy: zjawisko–rzeczywistość, są raczej rzadkie i patologiczne. Jego zdaniem, nie jest to jednak zarzut, który negowałby jego *inner sense theory*.

Po drugie, Lycan nie wyjaśnia: (i) w jaki sposób zdawanie sobie sprawy ze stanów mentalnych różni się od zdawania sobie sprawy z istnienia jakiegoś przedmiotu. Nie wystarczy podać, że różnica ta leży w tym, że stany mentalne są mentalne w odróżnieniu od przedmiotu zewnętrznego. (ii) Dlaczego świadomość (albo po prostu przedstawienie) jakiegoś stanu fizycznego wystarcza, aby same te stany uczynić świadomymi? Co jest specyficznym w jakimś rodzaju stanów fizycznych, że świadomość ich czyni je świadomymi? Jesteśmy świadomi fizycznego stanu naszego żołądka, ale nie rozróżniamy pomiędzy „świadomym żołądkiem” a „nieświadomym żołądkiem”. Odpowiedź Lycana, że stany mentalne, w odróżnieniu od stanów żołądka, są mentalne, jest mało zadowalająca.

Oraz po trzecie, niejasny pozostaje, jak wspomnieliśmy, podany przez Lycana opis „stopni świadomości” i relacji zachodzącej pomiędzy świadomymi podsystemami podmiotu, a świadomym podmiotem. Jeżeli bowiem tylko to, co spełnia funkcję monitorowania dla podmiotu, może dostarczyć czegoś dla świadomości, to jeżeli dane podsystemy mogą być świadome, to należałoby im przyznać jakąś podmiotowość. Nie widzimy nic problematycznego w mówieniu o stopniach-rodzajach-odmianach świadomości (nie muszą to być jedynie metafory), chociaż ilościowo trudno takie stopnie ująć.

3.6. *Teorie emergencji*

Problemy teorii redukcjonistycznych z wyjaśnianiem różnorodności zjawisk mentalnych spowodowały ponowny zwrot ku teoriom emergencji. Pierwszy okres rozwoju tych teorii przypada na lata dwudzieste minionego stulecia. Zbudowane wówczas teorie emergencji miały stanowić przeciwagę zarówno do „mechanistycznych”, jak i „witalistycznych” rodza-

⁵¹ Dennett zalicza teorię Lycana do negowanego przez siebie Kartezjańskiego Teatru, ponieważ zachowuje ona „reality/appearance distinction” oraz tworzy osobliwą kategorię „objectively subjective – the way things actually, objectively seem to you” (Dennett 1991: 132). Wysuwamy przeciwko Lycanowi zarzut Dennett opiera na swoim „operacjonalizmie pierwszej osoby” (por. Lycan 1997: 758).

jów wyjaśniania zjawiska świadomości i zjawiska życia⁵². W następnych latach filozofowie i biolodzy przestali się tymi teoriami zajmować. Swój ponowny rozwój teorie emergencji przeżywają od lat sześćdziesiątych minionego stulecia. O wzroście zainteresowania teoriami emergencji świadczy fakt, że dzisiaj spotykamy je nie tylko w filozofii umysłu, ale także w kognitywistyce, w kosmologiach ewolucyjnych oraz w teoriach samoorganizacji.

A. Stephan przypisuje teoriom emergencji następujące cechy (por. Stephan 1999: 14–65):

(1) Teorie emergencji przyjmują tezę o ontologicznym nieredukcyjnym naturalizmie oraz wykluczają istnienie obiektów „ponadnaturalnych” (*élan vital, entelechie*). Takie obiekty nie są eksplanacyjnie niezbędne.

(2) Liczne teorie emergencji dopuszczają pojawianie się w ewolucyjnym rozwoju przyrody czegoś zupełnie nowego (np. pojawienie się systemów obdarzonych życiem). Doprecyzowania wymaga określenie tego, do czego odnosi się słowo „nowego” (do obiektów, do cech, czy też do praw).

(3) Istnieją „systemowe (kolektywne) cechy”, których nie posiada żadna część systemu. Faktycznie „nowe”, gdy chodzi o cechy, mogą być tylko cechy systemowe.

(4) Występujące w naturze przedmioty, przynależą – w zależności od posiadanych charakterystycznych cech – do jednego z wielu obszarów (tzw. poziomów istnienia). Obszary te są związane z kolejnością ich ewolucyjnego powstania i tworzą hierarchiczny porządek. Najczęściej wyróżnia się obszary przedmiotów fizycznych, biologicznych i psychicznych.

(5) Według Laplace’a, świat jest wtedy i tylko wtedy zdeterminowany, gdy jego zdarzenia są (w zasadzie) przewidywalne. Emergentyści osłabiają związek pomiędzy „zdeterminowaniem” i „przewidywaniem”, dopuszczając możliwość istnienia zdeterminowanych procesów, w których jednak postuluje się (z zasady) ograniczoną zdolność poznawczą podmiotów. Gdy emergentysta mówi o zdeterminowaniu (albo zamiennie o „superweniencji”), rozumie pod tym terminem:

(a) Diachroniczne zdeterminowanie – w tych samych warunkach powstają te same struktury i systemy. Nie może być tak, że w dwóch światach, z tymi samymi warunkami początkowymi i z tymi samymi prawami, zostaną utworzone odmienne, nowe, struktury.

⁵² Por. Poczobut 2002: 403–406; por. Strawiński 1997: 171–194; por. Bremer 2001: 127–133. Zarówno teorie materialistyczne, jak i witalistyczne, zajmowały się pytaniem: Czy wszystkie rzeczy są materialne, czy też istnieją takie, które nie są redukowalne do materialnych? (por. Stephan 1999: 7–10).

(b) Synchroniczne zdeterminowanie – tak samo zbudowane systemy urzeczywistniają te same dyspozycje i cechy. To znaczy, że w jakimś systemie cechy i dyspozycje do zachowań zależą nomologicznie od jego mikrostruktury, czyli od cech części, które tworzą ten system oraz od uporządkowania tych części. Nie może istnieć różnica w cechach systemu bez równoczesnej różnicy w cechach elementów tego systemu. Teza o synchronicznym zdeterminowaniu zawiera tezę o mereologicznej superwencji (cechy systemu superwenują na cechach i relacjach jego części) [por. Stephan 1999: 26].

(6) Do centralnych charakterystyk teorii emergencji należą założenia: (A) o niededukowalności (nieredukowalności) oraz (B) o zasadniczej nieprzewidywalności systemowych cech. Z reguły założenia te są ze sobą połączone. Systemowe cechy, które są nieredukowalne, są także (z zasady) nieprzewidywalne.

(A) Przyjmijmy, że mamy jakąś systemową cechę, o której zakładamy, że zależy ona, zgodnie z tezą o synchronicznym zdeterminowaniu, od mikrostruktury posiadającego ją systemu *S*. Cecha ta jest wtedy nieredukowalna (albo według emergentysty C. D. Broada „emergentna”), gdy nie można jej wydedukować z uporządkowania części systemu *S* oraz z „wyizolowanych” cech tych części (tzn. cech, które dane części posiadają w izolacji od *S*)⁵³. Do nieredukowalnych cech Broad zalicza przeżyciowe stany świadomości.

We właściwym sensie tego słowa, dedukowalne nie są jednak cechy, lecz prawa. Tym samym jakaś systemowa cecha jest dokładnie wtedy nieredukowalna, gdy prawo, korelujące ją z mikrostrukturą systemu *S*, nie jest prawem, które można wydedukować z praw określających cechy i sposoby zachowania poszczególnych części systemu *S*. Tego rodzaju korelujące prawa Broad nazywa „prawami transfizykalnymi” (*trans-physical laws*) [por. Beckermann 2000: 51]. Prawa te są w sposób konieczny emergentne, ponieważ korelują one nieredukowalne jakości przeżyciowe oraz cechy mentalne, z jednej strony, z fizjologicznymi cechami – z drugiej.

(B) Według Broada, nieredukowalność systemowej cechy jest istotną własnością jej emergencji. Tej synchronicznej własności odpowiada w ewo-

⁵³ Podobną argumentację wprowadza Searle, zdaniem którego „świadomość jest przyczynowo emergentną własnością pewnych systemów neuronów” (Searle 1999: 155) i ich zachowania. Jej istnienie można wyjaśnić przyczynowymi interakcjami zachodzącymi pomiędzy elementami mózgu na poziomie mikroskopowym. Świadomości nie da się jednak „wylizyc” na podstawie czysto fizycznej struktury wyizolowanych neuronów, gdyż właśnie wzajemne oddziaływanie neuronów wywołuje stany świadomości. Świadomość jest więc przyczynowo redukowalna do procesów mózgowych, ale nawet idealna nauka o mózgu nie doprowadzi do ontologicznej redukcji świadomości (tzn. że sama świadomość [a nie jej „przedmioty”] nie sprowadza się do neuronów).

lucyjnych teoriach – zbudowanych przez emergentystów S. Aleksandra i C. L. Morgana – diachroniczna własność oznaczająca zasadniczą „nieprzewidywalność” wspomnianego już, nowego rodzaju cech⁵⁴. Własność systemowej nieprzewidywalności można wyrazić poprzez: Systemowa cecha E , która została po raz pierwszy urzeczywistniona w świecie \dot{S} przez system S o strukturze $[o_1 - o_n]$ w momencie czasu t , jest wtedy (z zasady) nieprzewidywalna, gdy pojawienie się E w czasie t w \dot{S} nie jest przewidywalne z panujących w \dot{S} stanów oraz z obowiązujących w nim praw, przed czasem t .

Mówiąc ogólnie: jakaś nowego rodzaju struktura jest (z zasady) nieprzewidywalna przed swoim pierwszym wystąpieniem, gdy:

(i) zachowanie części składowych systemu, odpowiedzialnych za utworzenie tej struktury, jest nieredukowalne (czyli nie może być wyprowadzone z zachowań tych komponentów w innych systemach),

(ii) zachowanie wspomnianych części systemu nie wynika z ich zachowania się w innych systemach, powstałych przed momentem czasu t .

(iii) świat, w którym owa nowa struktura powstaje, jest niezdeterminowany. Wcześniej była mowa o tym, że nieprzewidywalność nie wyklucza się z determinacją; pierwsza kategoria dotyczy poziomu epistemologicznego, druga zaś ontycznego.

(iv) procesy, prowadzące do powstania owej struktury przebiegają chaotycznie.

(7) Z teoriami emergencji łączy się twierdzenie, że w systemach posiadających emergentne cechy występuje tak zwana przyczynowość skierowana ku dołowi (por. Stephan 2002: 77–78). Neuropsycholog R. Sperry przyjmuje, że emergentne struktury wtedy oddziałują „przyczynowo ku dołowi”, gdy w ten sposób wpływają na zachowania części systemu, że zachowania te nie dadzą się sprowadzić do zachowań wspomnianych części w mniej kompleksowych systemach⁵⁵. Istotne pytanie brzmi: Czy teorie emergencji potrafią przyznać cechom mentalnym przyczynową rolę i równocześnie zachować te-

⁵⁴ Opisując „nieprzewidywalność” różnych cech systemowych Broad ma na myśli ich „nieredukowalność” (por. Stephan 1999: 39–48).

⁵⁵ Sperry traktuje świadomość jako zjawisko pojawiające się na pewnym stopniu ewolucyjnego rozwoju i odznaczające się przyczynowością (tzw. makrodeterminacją). Zjawiska zachodzące w świadomości nie tylko nie są redukowalne do procesów neuronalnych, ale mogą na nie przyczynowo wpływać. „The subjective mental phenomena are conceived to influence and to govern the flow of nerve traffic by virtue of their encompassing emergent properties. Individual nerve impulses and other excitatory components of a cerebral activity are simply carried along or shunted [...]” (Sperry 1969: 532). Sperry nie rozróżnia pomiędzy zdawaniem sobie sprawy a świadomością. Omawiając makrodeterminację skupia się na funkcjonalnej roli świadomych zjawisk, co oznacza, że zajmuje go raczej zdawanie sobie sprawy aniżeli świadomość zjawiskowa.

zę o przyczynowej zamkniętości obszaru fizycznego? Przypisując cechom mentalnym „skierowaną ku dołowi” przyczynową rolę, emergentyści muszą zrezygnować z tezy o przyczynowej zamkniętości obszaru fizycznego. Rezygnując z możliwości „skierowanej ku dołowi” przyczynowości, emergentyści stoi przed problemem właściwego ujęcia przyczynowej roli tego, co mentalne. Jedyną alternatywą wydaje się wtedy epifenomenalizm (cechy mentalne wprawdzie istnieją, ale nie odgrywają przyczynowej roli). Problem mentalnego oddziaływania przyczynowego zdaje się stawiać teorie emergentne wobec dylematu: jeżeli odrzuci się epifenomenalizm i „skierowaną ku dołowi” przyczynowość, wówczas cechy emergentne nie mogą być ujęte jako rzeczywiste⁵⁶ (por. 3.6.1, (2)).

Inne rozwiązania polegają na przyjęciu, że emergencja zachodzi już na poziomie fizycznym. Wówczas wszystkie poziomy wyższego rzędu (chemiczny, biologiczny, psychologiczny) będą zarazem emergentne i fizyczne.

3.6.1. Podstawowe rodzaje emergencji

Nie istnieje jeden, lecz wiele wariantów emergencji, z których poniżej omówimy trzy: emergencję słabą, synchroniczną i diachroniczną (por. Stephan 1999: 66–72; 2002: 78–92).

1. Słabą emergencję charakteryzują:

a. Fizyczny monizm – wszystkie obiekty naszego świata składają się jedynie z elementów fizycznych. Także cechy czy dyspozycje zaliczane do emergentnych są realizowane przez systemy, składające się wyłącznie z elementów fizycznych.

b. Przyjęcie założenia o istnieniu cech systemowych (albo zbiorowych) – żadna część składowa systemu nie posiada tego rodzaju cech, posiada je jedynie cały system. Cechy kompleksowych systemów można podzielić na dwie grupy: na cechy, które posiada system, a także jego części (np. ciężar, prędkość), oraz na cechy systemowe (np. oddychanie, posiadanie wrażeń).

c. Synchroniczna determinowalność – wewnętrzne cechy i zachowania jakiegось systemu zależą na sposób nomologiczny od jego mikrostruktury (czyli od cech jego części i ich uprządkowania). Nie istnieje różnica pomiędzy wewnętrznymi cechami systemowymi bez różnicy pomiędzy cechami części tego systemu⁵⁷.

⁵⁶ Zdaniem S. Aleksandra, bycie realną cechą oznacza posiadanie przyczynowej roli (por. Stephan 1999: 212).

⁵⁷ Stephan odróżnia: (i) tezę o „mereological supervenience” („that a system’s intrinsic properties supervene on its parts”), od (ii) tezy o synchronicznej determinacji. Teza

Teoria słabej emergencji odrzuca substancjalny dualizm i jest do pogodzenia z teoriami redukcyjnego fizykalizmu (np. teorią identyczności typów), który, przy wyjaśnianiu cech systemowych, uwzględnia każdorazową strukturę danego systemu. Teorie słabej emergencji nie implikują „przyczynowości skierowanej ku dołowi”.

2. Teoria synchronicznej emergencji powstaje przez uzupełnienie teorii o słabej emergencji o twierdzenie dotyczące cech systemowych⁵⁸. Cechy te można podzielić na redukowalne i nieredukowalne. Cechy są nieredukowalne, gdy: (a) nie można ich analizować w sposób behawioralny albo gdy (b) zachowanie części systemu, na którym one superwenują, jest nieredukowalne. W obydwu przypadkach cechy systemowe nie są wyprowadzalne z uporządkowania cech części systemu (czyli z jego mikrostruktury) i z cech, które te części posiadają w izolacji od tego systemu albo w prostszych systemach. Nawet najbardziej całościowa wiedza o tak rozumianych cechach elementów systemu nie pozwala na wydedukowanie emergentnych cech systemu (które są na sposób naturalny zależne od mikrostruktury systemu). Ten typ emergentyzmu znajdujemy w pracach C. D. Broada.

Synchroniczna emergencja w miarę jasno określa, co to znaczy, że jakaś systemowa cecha jest nieredukowalna (niededukowalna). Stephan zaznacza, że mamy tutaj do czynienia z dwoma niezależnymi rodzajami nieredukowalności, mającymi różne konsekwencje. Jeden z nich – odniesiony do cech mentalnych – zdaje się implikować „przyczynowość skierowaną ku dołowi”, drugi epifenomenalizm tychże cech. Owymi rodzajami nieredukowalności cech systemowych są: (a) Cecha systemowa P systemu S jest nieredukowalna, gdy nie jest tak, że (w zasadzie) z zachowania części systemu S wynika, że system posiada cechę P . (b) Cecha systemowa P systemu S jest nieredukowalna, gdy nie jest tak, że (w zasadzie), z zachowania części systemu w systemie z złożenia prostszym aniżeli S wynika, jak części te będą się zachowywały w S .

Z (a) i (b) można wyprowadzić zmodyfikowaną wersję tezy o nieredukowalności cech systemowych (tzn. cech synchronicznie emergentnych, których przykładem mogą być qualia):

(IR) Jakaś systemowa cecha jest nieredukowalna, jeżeli (a) nie można jej analizować ani mikro-, ani makrobehawioralnie, lub (b) specyficzne za-

(i) nie implikuje zależności cech systemu od cech jego części, lecz jedynie ich kowariancję (por. Stephan 2002: 80).

⁵⁸ Analizowane przez R. Poczobuta ontologiczne odmiany tezy o emergencji własności można zaliczyć (ze względu na nieredukowalność) do tezy o synchronicznej emergencji (por. Poczobut 2002: 414–417).

chowanie elementów systemu, na których ta systemowa cecha superwenuje, nie wynika z zachowania elementów rozpatrywanego w izolacji od tego systemu lub w złożeniach od niego prostszych.

Jeżeli cecha systemowa jest nieredukowalna z powodu (b), możemy mówić o „przyczynowości skierowanej ku dołowi”. Istnieje bowiem wtedy jakiś „skierowany w dół” przyczynowy wpływ wywierany przez sam system (albo przez jego strukturę), na zachowania elementów systemu. Tego rodzaju przyczynowość nie będzie naruszała przyczynowej zamkniętości obszaru fizycznego⁵⁹.

W przeciwieństwie do tego, występowanie cech, których nie można analizować, nie implikuje jakiegś „przyczynowości skierowanej ku dołowi”. Systemy, które posiadają tego rodzaju cechy zależne na sposób nomologiczny (naturalny) od ich mikrostruktur, nie muszą być ukonstytuowane w sposób wskazujący na nieredukowalność tych cech. Raczej wątpliwe jest, w jaki sposób niepodatne na analizy cechy mogłyby odgrywać przyczynową rolę. Ponieważ nie są one analizowalne na sposób behawioralny, trudno zrozumieć, jak mogłyby one odgrywać przyczynową rolę, co jednak nie musi oznaczać, że są one czymś epifenomenalnym.

Synchronicznie rozumiana emergencja przeżywa renesans we współczesnych debatach dotyczących nieredukcyjnych koncepcji umysłu i nie można jej pogodzić z teoriami redukcyjnego materializmu. Tego typu teorię emergencji (IR, [a]) przyjmujemy w dalszych analizach, mówiąc o neuronalnym podłożu świadomości w danym momencie czasu t .

3. Teorie diachronicznej emergencji były rozwijane zwłaszcza na początku XX wieku. Uwzględniają one wymiar czasowy i uzupełniają teorię o słabej emergencji twierdzeniem o nieprzewidywalności (*unpredictability*) autentycznie nowych cech (*novel properties*). We wszechświecie dochodzi ciągle do egzemplifikacji czegoś autentycznie nowego. W trakcie ewolucji powstają nowe układy części, tworzą się nowe struktury, które konstytuują nowe obiekty z nowymi cechami i nowymi sposobami zachowań. Samo dodanie do teorii o słabej emergencji twierdzenia o nowych cechach jeszcze jej nie wzmacnia (tego typu dodanie jest kompatybilne z redukcyjnym materializmem). Dopiero uzupełnienie tej teorii o twierdzenie dotyczące zasadniczej nieprzewidywalności nowych systemowych cech prowadzi do silniejszej teorii, czyli do diachronicznej emergencji.

⁵⁹ Zdaniem Stephana: „Within the physical domain, we would just have to accept additional types of causal influences besides the already known basal types of mutual interactions” (Stephan 2002: 89).

Rozróżnia się dwa warianty wspomnianej nieprzewidywalności cech, co ma swoją podstawę w: (a) nieredukowalności cech, (b) w zasadniczej nieprzewidywalności struktury systemu (powstanie niektórych nowych struktur, a tym samym urzeczywistnianie nowych cech jest zasadniczo nieprzewidywalne).

Jeżeli nowe cechy są diachronicznie emergentne, to są one także synchronicznie emergentne. Diachroniczny emergentyzm typu (a) różni się jedynie nieznacznie od synchronicznego emergentyzmu. Emergentyzm strukturalny typu (b) jest całkowicie niezależny od synchronicznego emergentyzmu i powstaje ze słabego emergentyzmu poprzez dodanie twierdzenia o zasadniczej nieprzewidywalności nowych struktur systemu.

3.6.2. Wrażenia jako cechy emergentne

W obecnej i w dalszych częściach niniejszego opracowania zajmiemy się synchronicznym emergentyzmem i proponowanym przez niego ujęciem problematyki subiektywności. Jeżeli zbudowane przez T. Nagla, C. McGinna, N. Blocka, F. Jacksona, J. Levine'a i D. Davidsona nieredukcjonistyczne („nieredukcjonistyczne” w sensie teorii identity typów) teorie mają odnieść wyjaśniający sukces, to implikują one niejako założenia spotykane w teorii emergencji⁶⁰. Emergentyści przyjmują jako podstawową tezę, że subiektywna świadomość oraz jakości przeżyciowe [$c\Psi$] są (z zasady) nieredukowalne do obiektów, stanów, zjawisk i cech fizyko-fizjologicznych [$c\Phi$].

Parafrazując eksperyment myślowy C. D. Broada można powiedzieć: nawet matematyczny archanioł (*mathematical archangel*) nie byłby w stanie przepowiedzieć (wydedukować) subiektywności świadomości z fizyko-fizjologicznych danych dotyczących mózgu⁶¹. Argument Broada przyjmuje postać:

(a) Archanioł zna chemiczną strukturę i neuronalną budowę naszego systemu nerwowego oraz wszystkie odnoszące się do tego systemu prawa.

(b) Na podstawie tej wiedzy nie potrafi jednak wydedukować cech stanów przeżyciowych (niemożność ta może służyć jako przykład „luki w wyjaśnianiu”).

⁶⁰ Argumenty przedstawione w eksperymentach myślowych proponowanych przez Nagla i Jacksona (por. rozdział V) mają wykazać, że teorie fizykalistyczne nie dostarczają adekwatnego wyjaśnienia subiektywnej świadomości. Pozostanie pytanie, czy argumenty te mają konsekwencje ontologiczne.

⁶¹ Por. Broad 1925: 70. Eksperyment myślowy z archaniołem przypomina Jacksona eksperyment z Mary. W odróżnieniu od Mary, archanioł jest pozbawiony nie tylko „wiedzy o stanach zjawiskowych”, ale także „wiedzy o stanach propozycjonalnych” podmiotu.

(c) Nawet, gdy zdarzenia przeżyciowe są zdeterminowane przez zdarzenia neurologiczne i są przez nie fizycznie urzeczywistnione, to nie są one redukowalne do zdarzeń fizycznych od strony wyjaśniającej.

(d) Stąd wniosek: zdarzenia przeżyciowe są emergentne.

Znając naukowe prawa, obowiązujące w neurofizjologii, archanioł wie jedynie, że w odpowiednich komórkach mózgu zachodzą elektrochemiczne zmiany. Z praw tych nie wynika, że zmianom tym towarzyszą określone, subiektywne wrażenia. Prawo stwierdzające, że takie a takie zmiany w systemie nerwowym podmiotu prowadzą do tych a nie innych wrażeń, nie jest dedukowalne z praw natury, dotyczących poziomu fizycznego. Jest to emergentne prawo natury, które łączy mikrostrukturę mózgu podmiotu z tymi jego cechami, których nie da się wydedukować z praw obowiązujących na poziomie fizycznym. Znając emergentne prawa natury, Archanioł mógłby wydedukować stany subiektywne. Umożliwia to redukcję *via* „prawa pomostowe” w sensie E. Nagla (por. niżej).

Prawa emergencji są prawami wertykalnymi. Do przewidzenia własności emergentnych nie wystarczą tylko „prawa horyzontalne” niższych rzędów. Paradoksalnie: prawa emergencji są prawami redukcji w sensie E. Nagla⁶².

Opisując zależność: świadomość–umysł, filozofowie analityczni powołują się często na jej podobieństwo do opisów „problemów” związanych ze rozumieniem zjawiska życia. Broad wskazuje na większe trudności, gdy chodzi o redukcję, w kwestiach dotyczących „problemów umysłu” aniżeli w kwestiach dotyczących „problemów życia”. Stanów mentalnych, w przeciwieństwie do poruszania się (jako jednej z charakterystycznych cech istot żywych), nie da się analizować w terminach behawioralnych. Dla pytań dotyczących życia istnieje tylko jedna oczywistość: określone sposoby zachowania się danego systemu. Do stanów świadomych mamy dwa sposoby dojścia: sposób pierwszo- i trzecioosobowy. Podejście trzecioosobowe zakłada obserwację zachowań. W podejściu pierwszoosobowym nie przypisujemy sobie stanów umysłu na podstawie obserwacji swoich zachowań. Przebiegającą według praw redukcję licznych procesów mentalnych dodatkowo komplikuje ich aspekt jakościowy (por. Broad 1925: 612–621; por. Beckermann 2000: 47–50).

Z jednej strony należy stwierdzić, że samo odwoływanie się do cech, dotyczących sposobów zachowania się nie jest *wystarczające* dla uchwycenia cech zjawiskowych (fenomenalnych). Z drugiej strony, takie stwierdzenie nie

⁶² Według U. Żegleń może istnieć emergencja bez emergentnych praw (por. Żegleń 2003: 297). Nie ma emergentnych praw ze względu na wieloraką realizację stanów umysłowych i ze względu na ich holistyczną naturę.

prowadzi do uzasadnienia, że cechy zjawiskowe mają *koniecznie* emergentny charakter. Broad uważa, że identyfikacja cech jakościowych (które jego zdaniem są „nieprzeustrzenne”) z procesami neurologicznymi jest kategoriyczną pomyłką. Nie wystarcza bowiem umieć wyjaśnić różnicę pomiędzy widzeniem czerwonego a widzeniem zielonego, lecz pytanie brzmi: czy zupełny wgląd w stany neuronalne pozwala – w ogóle – na dedukcję zdań o stanach jakościowych?

Dlaczego przeżycia jakościowe opierają się redukcjom wyjaśniającym (*explanatory reduction*)? Powołując się na J. Levine, Stephan odpowiada (por. Stephan 1999: 194–196): redukcja ma charakter wyjaśniający, gdy podczas redukcji jakiegoś obiektu lub cechy potrafimy ukazać mechanizm, dzięki któremu zostaje urzeczywistniona przyczynowa rola, konstytutywna dla tego obiektu lub cechy.

Redukcja wyjaśniająca przebiega dwustopniowo. Stopień pierwszy obejmuje quasi-aprioryczne wypracowanie pojęcia cechy, która ma być redukowana poprzez zidentyfikowanie przyczynowej roli, dla której szukamy leżących u jej podstaw mechanizmów. Stopień drugi obejmuje empiryczne badania tego, czym owe podstawowe mechanizmy są.

Jeżeli twierdzimy, że redukcja wyjaśniająca – powiedzmy redukcja subiektywności świadomości – jest niemożliwa, to zasadniczo nie mamy na myśli tego, że niewykonalne jest zadanie występujące na stopniu drugim (stąd poszukiwania neuronalnych korelatów świadomości). Najwidoczniej problem pojawia się na stopniu pierwszym i dotyczy indywiduacji cech jakościowych za pomocą ich ról przyczynowych. Na tyle, na ile istnieje jakiś element w naszym pojęciu jakościowego charakteru przeżycia [Ψ_i], który to element nie jest do uchwycenia przez cechy jego przyczynowej roli, na tyle element ten będzie się wymykał wyjaśniającym związkom obecnym w fizykalistycznej redukcji. O ile Nagel, Kripke, Jackson i McGinn chcą poprzez swoje myślowe argumenty pokazać nieadekwatność pozycji materialistycznych, o tyle Levine zaznacza, że wszelkiego rodzaju argumenty myślowe nie mają konsekwencji ontologicznych.

Uwzględnając to, co dotychczas powiedziano na temat emergentyzmu, możemy zdefiniować emergentyzm subiektywności za pomocą twierdzenia odnoszącego się do jej wyjaśniającej nieredukowalności:

III.10 Redukcja cechy emergentnej *E* ma tylko wtedy charakter wyjaśniający, gdy urzeczywistniająca ją podstawa spełnia dokładnie tę rolę, która jest konstytutywna dla *E*. Cechy zjawiskowe nie są w pełni wyjaśnialne przez wywołujące je przyczyny.

Przyjmując rozróżnienie pomiędzy „logiczną” i „ontologiczną” koniecznością możemy powiedzieć: nieredukowalność jest tylko na płaszczyźnie tego, co logiczne (epistemologiczne). Na razie nie mówimy nic o płaszczyźnie ontologicznej. Można próbować ominąć III.10 przez wskazanie na to, jak cechy, które nie są całkowicie do ujęcia przez ich przyczynowe role, mimo wszystko mogą zostać wyjaśniająco zredukowane. Odnośnie do twierdzenia III.10 należy zauważyć co następuje:

(1) Akceptacja III.10 oznacza zajęcie emergentystycznego stanowiska. Samodzielność cech emergentnych jest zagwarantowana poprzez ich nieredukowalność eksplanacyjną (wyjaśniającą), nie zaś poprzez ich przyczynową nieredukowalność. Nieredukowalność można rozumieć od strony eksplanacyjnej (wyjaśniającej), a niekoniecznie od strony ontologicznej (przyczynowej). Wtedy staje się możliwe, aby cechy emergentne (pomimo ich nieredukowalności eksplanacyjnej), związać wystarczająco silnie ontologicznie z płaszczyzną fizykalną, aby tym samym zapewnić relację superwencji.

W odniesieniu do umysłu, teoria emergencji wymaga, opartej na odpowiednich prawach, korelacji pomiędzy neuronalną mikro-płaszczyzną i emergentną makro-płaszczyzną. Tym samym, istnieje możliwość (albo istnienie takiej możliwości jest przynajmniej sprawą otwartą) redukcji za pomocą praw psychofizycznych. Pomimo tego pozostaje wystarczająco dużo miejsca dla nieredukowalności w specyficznym sensie (gdyż prawa te nie są do wyrowadzenia z naszych wiadomości o mikro-płaszczyźnie, tzn. z praw stosowanych w neurologii). Prawa te należy przyjąć jako postulaty. Tym samym mamy zagwarantowaną nieredukowalność eksplanacyjną. Mająca swoją podstawę w prawach natury korelacja mikro-płaszczyzny i makro-płaszczyzny może być następnie wyjaśniania w ramach teorii superwencji.

(3) Wspomniany w 3.6 (7) problem mentalnego przyczynowego oddziaływania zmusza emergentystów albo do przyjęcia epifenomenalizmu, albo do zaakceptowania „przyczynowości skierowanej ku dołowi”. Z jednej strony możemy przyjąć, że emergentyzm nie jest epifenomenalizmem, z drugiej strony możemy zanegować możliwość „oddziaływania ku dołowi”. Rozumiejąc bowiem i używając pojęcia nieredukcyjności cech w sensie eksplanacyjnym (III.10), emergentysta zyskuje możliwość interpretacji tezy o „przyczynowości skierowanej ku dołowi”, w sensie, w jakim rozumie ją redukcjonista (jako przyczynowość superwencyjną). Jest to jednak jedynie bardzo słabe rozumienie przyczynowości.

⁶³ Por. Broad 1925: 64–65. Broad mówi wprawdzie o „emergent character of the behavior of chemical compounds”, ale jego analizy można odnieść do emergentnych wrzeń.

(4) Jeżeli stwierdzamy, że jakaś cecha lub odpowiadający jej predykat jest nieredukowalna, musimy odpowiedzieć na pytanie: w jakich warunkach byłaby ona redukowalna? Zarówno odpowiadający na to pytanie funkcjoniści, jak i redukcjoniści materialiści powołują się na standardowy, logiczno-empiryczny schemat redukcji interteoretycznej sformułowany przez E. Nagla⁶⁴. Według Nagla: „Redukcja [...] jest wyjaśnieniem jednej teorii lub zbioru praw eksperymentalnych, ustalonych w pewnej dziedzinie badań, przez teorię zwykle, chociaż nie koniecznie, sformułowaną dla dziedziny innej” (Nagel 1970: 295–296).

Zredukować można jedną teorię do drugiej, a nie np. jakiś szereg cech do innego szeregu cech. Ogólny schemat redukcji interteoretycznej wygląda następująco. Teoria T_2 redukuje teorię T_1 wtedy, gdy prawa T_1 są wyprowadzalne z praw T_2 . Schemat ten opiera się na dwóch koniecznych warunkach formalnych:

1. „Trzeba przyjąć jakieś postulaty określające odpowiednie relacje między denotacją terminu A, a własnościami denotowanymi przez terminy teoretyczne istniejące w nauce pierwotnej. Charakter takich założeń trzeba będzie zbadać, bez przesądzenia z góry wyniku (warunek ten nazwiemy dla wygody «warunkiem spójności»”).
2. „Ze względu na owe dodatkowe założenia, wszystkie prawa nauki wtórnej, łącznie z zawierającymi termin A, muszą wynikać logicznie z postulatów teoretycznych i związanych z nimi definicji koordynujących nauk pierwotnej. Nazwijmy warunek ten «warunkiem wynikania logicznego»” (Nagel 1970: 308).

„Odpowiednie relacje” są często nazywane „zasadami korespondencji” albo „prawami pomostowymi (łączącymi)”. Warunki 1. i 2. możemy sformułować następująco. W teorii redukowanej T_1 wszystkie pojęcia muszą zostać połączone z pojęciami teorii redukującej T_2 za pomocą „zasad korespondencji” oraz że za pomocą tych zasad, wszystkie prawa T_1 muszą być wyprowadzalne z T_2 .

Zasady korespondencji mogą być definicjami, w których wyrażenia T_1 są definiowane poprzez wyrażenia T_2 . Mogą one także być empirycznie potwierdzonymi prawami korelacji, które twierdzą, że określone przez wyrażenia z T_1 albo T_2 cechy albo stany są nomologicznie ze sobą skorelowane. Empiryczne prawa korelują tutaj poszczególne mentalne predykaty z kontekstowymi, fizykalnymi predykatami. Jednak funkcjonalistyczne założenie o wielorakiej realizacji jakiejś cechy mentalnej przeczy możliwości istnienia tego rodzaju empirycznych, ogólnych praw korelacji (a tym samym

⁶⁴ Por. Nagel 1970: 294–342; Judycki 1997: 36–39. Ten model redukcji odrzuca J. Kim (2002: r. IV).

przeciwko redukcji stanów mentalnych do fizycznych). Jeżeli funkcjoniści dopuszczają istnienie jakichś lokalnych praw korelacji, to mogą oni być traktowani jako redukcjoniści.

W przeciwieństwie do tego, synchroniczni emergentyści ani nie wykluczają, ani nie przyjmują, że jakaś cecha mentalna jest ogólnie lub lokalnie redukowalna do jej fizycznego odpowiednika. Synchroniczno-emergencyjne założenie o nieredukowalności stwierdza, że jakaś cecha systemowa – powiedzmy cecha psychiczna – jest wtedy emergencyjna (nieredukowalna wyjaśniająco), gdy dotyczące ją stwierdzenia nie mogą być wyprowadzone (*via* prawa dotyczące fizycznych części systemu) z twierdzeń dotyczących cech i relacji zachodzących pomiędzy fizycznymi częściami systemu. Jakaś systemowa cecha jest wtedy redukowalna eksplanacyjnie, gdy takie wyprowadzenie jest możliwe.

Przy omawianiu zagadnień emergencyjności nie wchodziliśmy w głębsze analizy różnych rodzajów redukcji. Dopiero mając przed sobą podział różnych redukcjonizmów zrozumielibyśmy specyfikę danej wersji emergencyjności. Już z naszych analiz widać, że terminy „redukcja” i „emergencyjność” są ze sobą w różnoraki sposób skorelowane. My ograniczyliśmy się jedynie do krótkiego omówienia trzech rodzajów emergencyjności, z których w dalszych analizach będzie nas interesowała emergencyjność synchroniczna.

W rozdziale VII zostanie przedstawiona prototeoria świadomości, odwołująca się do zaproponowanego przez Nagla modelu interteoretycznej redukcji. Wspomniana prototeoria stwierdza emergencyjność na poziomie teorii oraz na poziomie stanów. Chodzi w niej o emergentny rodzaj zależności pomiędzy filozoficzno-psychologiczną „teorią” subiektywnej świadomości (MkT), a neurologiczną teorią świadomości (MiT), oraz o emergentny rodzaj zależności pomiędzy stanami zjawiskowymi a neurologicznymi.

3.7. Podsumowanie

Podstawowym zarzutem wobec behawioryzmu jest to, że nie uwzględnia on całościowego powiązania (holizmu) zjawisk mentalnych, tzn. faktu, że przekonania i pragnienia, spostrzegania i zamiary są powiązane ze sobą i z działaniem podmiotu. Posiadane przekonania wyjaśniają jakieś działania podmiotu w kontekście określonych pragnień, a określone pragnienia wyjaśniają dane działania jedynie w kontekście określonych przekonań. Klasyczny behawioryzm, chcąc wyjaśnić spostrzeganie i działanie, odwołuje się do prostego modelu: wejście–wyjście. Model ten projektuje izomorficznie płaszczyznę podmiotową (spozstrzeganie–działanie) na płaszczyznę subpodmiotową (rozdzielenie: wejście–wyjście). Projekcja taka wydaje się nie

do uniknięcia ze względu na behawiorystyczne założenie o jednokierunkowym łańcuchu przyczynowym oraz ze względu na czysto instrumentalne relacje pomiędzy spostrzeganiem i działaniem. Model wejście – wyjście skłania niejako do skupienia się na relacjach pomiędzy przyczynami i spostrzeżeniami oraz pomiędzy działaniami i skutkami.

Gdzie jednak przeciąć łańcuch przyczyn, prowadzących do przeżyć spostrzeżeniowych, aby określić, co determinuje ich jakościową treść? Gdzie przeciąć łańcuch skutków, w celu znalezienia tego, co determinuje wspomnianą treść? Czy tego rodzaju granica ma miejsce w „głowie” spostrzegającego podmiotu czy w jego otoczeniu? Co spostrzegamy bezpośrednio, a co jedynie pośrednio? Behawioryzm nomologiczny zakłada, że dla normalnej osoby różne rodzaje zachowań są skojarzone ze szczególnymi, spostrzeżeniowymi wejściami. Trudno jednak dostrzec jakieś rodzaje zachowań podmiotu, odpowiednie do jego otoczenia, bez dostrzeżenia chociażby związku tych zachowań z zamiarami podmiotu.

Teoria identyczności typów cechuje się prostotą. Stany mentalne są fizycznymi stanami mózgu, ból jest identyczny z pobudzeniem neuronów i niczym więcej. Relacja pomiędzy bólem a pobudzeniem neuronów jest identycznością typów. Świadome przeżycie jest sprawą fizycznych procesów zachodzących w mózgu. Zaletą teorii identyczności jest jej oszczędność ontologiczna. Jeżeli jednak ból jest procesem neurologicznym, to wtedy to, co nie jest takim procesem, nie jest także bólem. Stąd rodzi się teza o identyczności zdarzeń i funkcjonalistyczna teza o wielorakiej realizacji.

Czy cechy przeżyciowe da się ująć w obiektywnych terminach? Widać wyraźnie, że mówienie o stanach mózgu, będących (jako zdarzenia) identycznymi ze zdarzeniami lub ze stanami przeżyciowymi, nie wystarczy do równoczesnego, przyczynowego uchwycenia cech przeżyciowych. Nasze dzisiejsze odnoszenie się do stanów systemów neuronalnych [$\Phi_j - \Phi_n$] nie wystarczy nawet do tego, aby uchwycić odpowiednie cechy funkcjonalne (cechy funkcjonalne nie są definiowalne w języku fizyki lub neurofizjologii), które uchodzą za trzecioosobowe. Stąd potrzeba szukania dziś bardziej rozbudowanej, filozoficzno-neurologicznej metodologii, pozwalającej na precyzyjniejsze określenie neuronalnych korelatów świadomości zjawiskowej.

Przedstawiona w 3.4, 3.5 i 3.6 możliwość „autonomii” tego, co przeżyciowe (zjawiskowe), jest antyredukjonistyczna. Nieredukowalność subiektywności nie wynika z jakiegoś ogólnego antymaterializmu, ale z anomalności stanu zjawiskowego w stosunku do stanu fizyczno-funkcjonalnego. Wspomniana „autonomia” odpowiada z, jednej strony, antyredukjonistycznym tezom C. D. Broada i T. Nagła, z drugiej – tłumaczy eliminatywizm D. Dennetta (por. rozdział IV), który tej autonomii nie bierze pod uwagę.

Reprezentowanemu przez W. G. Lycana funkcjonalizmowi oraz reprezentowanemu przez P. Churchlanda materializmowi nie udaje się wykluczyć możliwości, że subiektywność świadomości jest opisywalna i wyjaśnialna jedynie z perspektywy pierwszoosobowej. Posiadanie tej perspektywy pozwala nam odnosić się do cech zjawiskowych jako takich. Pozostaje pytanie, czy inne okoliczności – na przykład przyczynowe role – pozwolą na niezależne odnoszenie się do tychże cech i na ich całkowite pojęciowe ujęcie. Należy jednak dodać, że nawet gdy nie posiadamy obiektywnych charakterystyk dla stanów subiektywnych (bo stany te są w sposób konieczny związane z perspektywą pierwszoosobową), nie musi to automatycznie oznaczać antymaterializmu. Materializm nieredukcyjny (w przeciwieństwie do redukcyjnego) twierdzi bowiem, że skupiając się na wyjaśnianiu, nie musimy przyjmować, iż wszystkie cechy są cechami fizycznymi.

Powyższe twierdzenie stoi w opozycji do funkcjonalizmu sugerującego, że nie tylko indywidualne zdarzenia mentalne, ale i subiektywne cechy mentalne, są identyfikowalne z jakimiś funkcjonalnymi rodzajami cech. To oznacza, że subiektywne cechy mentalne nie są jedynie emergentne w stosunku do cech neuronalnych. Lycan twierdzi, że funkcjonalistyczne ujęcie stanu subiektywnego nie oznacza, że wiemy, *jak to jest być* w tym stanie albo *jak to jest go posiadać*. W tym sensie, Lycan i Nagel, przyjmują jakiś rodzaj wewnętrznie zjawiskowych cech, które prowadzą do wspomnianego konfliktu z redukjonistycznym materializmem. Można powiedzieć, że cechy te odpowiadają cechom postulowanym przez anomalny monizm, z jego nieredukcyjnym traktowaniem cech mentalnych.

Oдноśnie do teorii emergencji pojawia się zasadnicze pytanie: do jakiego typu stanowisk należy je zaliczyć? Czy jest to stanowisko materialistyczne, czy też stanowisko należące do teorii dwu aspektów (por. 5.4.1, 5.4.2)? Wszystko zależy od tego, jak rozumiemy materializm. Jeżeli przyjmie się, że pozycje materialistyczne zakładają redukcję ontologiczną i redukcję wyjaśniającą (czyli że materializm nie może pozostawić niczego niewyjaśnionego), to pozycja emergentystyczna do tak rozumianego materializmu nie należy.

Zestawienie teorii materialistycznych oraz ich krytyczna analiza prowadzą do stwierdzenia, że dziś trudno odpowiedzieć na pytanie o ontologiczną redukcję stanów subiektywnych do stanów neurologicznych, dlatego nie zajmowaliśmy się tym pytaniem.

Gdy natomiast chodzi o redukcję wyjaśniającą, materializm pomija subiektywność świadomości. Teorią, która dziś w adekwatny sposób uwzględnia wyniki badań neurologicznych i holistyczne opisy świadomości zjawiskowej, jest jakiś rodzaj teorii synchronicznej emergencji. Teoria ta zakłada, że redukcja świadomości zjawiskowej jest tylko wtedy wyjaśniająca, gdy jej

podstawa realizacji spełnia dokładnie tę przyczynową rolę, która jest konstytutywna dla świadomości zjawiskowej. Świadomości zjawiskowej nie da się jednak całkowicie uchwycić poprzez cechy jej przyczynowej roli (III.10). Podstawowa trudność z akceptacją teorii emergencji leży w stawianym przed nią wymaganiem bliższego określenia wspomnianej *downward causation*.

Dlatego proponowane w rozdziale VII prototeoretyczne wyjaśnienie problemu subiektywności mówi o emergencyjnej zależności pomiędzy dwiema teoriami świadomości: teorią filozoficzno-psychologiczną (jej zarys przedstawiono w rozdziale II) i teoriami naukowymi, zakładającymi wyniki badań neurologicznych (por. rozdział VII). Potwierdzenie tego rodzaju emergencyjnej zależności ma swoją podstawę w wynikach uzyskiwanych dzięki współczesnym filozoficzno-naukowym metodom poszukiwania neuronalnych korelatów świadomości. Ponadto przyjmujemy emergencyjną zależność pomiędzy stanami mózgu a stanami świadomości.

Należy dodać, że niemożliwość wyjaśniającej redukcji subiektywnej świadomości do cech funkcjonalnych (albo obliczeniowych), nie prowadzi jeszcze sama z siebie do dualizmu cech mentalnych i fizycznych.

ROZDZIAŁ CZWARTY

DANIEL DENNETT – ŚWIADOMOŚĆ AKTYWNOŚCIĄ MÓZGU

4. Wprowadzenie: jaka teoria świadomości?

W obecnym rozdziale zostaną przedstawione analizy świadomości zjawiskowej prowadzone przez D. Dennetta. Wychodzi on z trzecioosobowego punktu widzenia i skupia się przede wszystkim na stanach intencjonalnych¹. Według niego, podmioty ludzkie są operującymi językiem, wyposażonymi w wejścia oraz wyjścia systemami intencjonalnymi. Posiadają one cechy, odziedziczone na drodze ewolucji po organizmach o prostszych systemach nerwowych. Właściwego punktu widzenia do opisanego i wyjaśnienia świadomości dostarczają nam neurofizjologia i nauki o sztucznej inteligencji.

W poszukiwaniu teorii świadomości Dennett nie wychodzi od definicji typu: (*x*) (*x* jest świadomym przeżyciem wtedy i tylko wtedy gdy, ...) [por. Dennett 1991: 459]. Podstawowe elementy swojej teorii świadomości Dennett podaje w tak zwanym miniaturowym szkicu (*thumbnail sketch*):

„Nie istnieje pojedynczy (*single*), jednoznacznie określony «strumień świadomości» (*stream of consciousness*) ponieważ nie istnieje żadna Centralna Kwatera Główna (*central Headquarters*), nie istnieje Teatr Kartezjański (*Cartesian Theater*), gdzie «wszystko schodzi się razem», aby móc zostać przeczytanym przez jakiegoś Centralnego Dawcę Znaczenia (*Central Meaner*). Zamiast takiego pojedynczego strumienia (nie wiadomo jak szerokiego) istnieje wielość kana-

¹ Wstępne omówienie argumentacji Dennetta można znaleźć w: Czarniecki 2002: *passim*; Rey 1994: 265–266. Dennett jest uczniem G. Ryle’a.

łów, w których specjalne obwody (*specialist circuits*) próbują, w paralelnych kotłowniach, realizować swoje rozmaite zadania, tworząc stopniowo Wiele Szkiców (*Multiple Drafts*). Wiele z tych fragmentarycznych szkiców «narracji» odgrywa krótkotrwałe role w modulowaniu rozgrywającej się właśnie aktywności. Niektóre zostają jednak w szybkim następstwie przekazane do następnych funkcjonalnych ról poprzez dynamikę mentalnej maszyny w mózgu. Seryjność tej maszyny (jej «von Neumannowski» charakter) nie jest strukturalnie utrwalona, lecz jest raczej wynikiem potoku chwilowych zespołów pomiędzy tymi wyspecjalizowanymi dynamikami.

Podstawowe specjalizacje są częścią naszego zwierzęcego dziedzictwa. Nie zostały one wytworzone do wykonywania właściwie ludzkich czynności, takich jak czytanie i pisanie, lecz do takich jak ukrywanie się, unikanie drapieżnych zwierząt, rozpoznawanie twarzy, chwytanie, rzucanie, zbieranie jagód, oraz do wykonywania innych istotnych zadań. Są one często czysto pragmatycznie zaangażowane w nowe role, do których mniej lub bardziej pasują dzięki wrodzonym talentom. Jeśli wynik nie jest chaosem, to tylko dlatego, że tendencje nałożone na te wszystkie formy aktywności są same w sobie produktami kompozycji (planu, zamiaru, strategii, wzorca). Część tego wzorca jest wrodzona, wspólna z innymi zwierzętami. Ale jest ona spotęgowana, a czasami nawet zdominowana w skali ważności poprzez mikronawyki, myśli rozwinięte w danym indywiduum, częściowo jako osobiste, charakterystyczne dla danego osobnika wyniki samopoznania, a częściowo jako zastane wpływy kultury. Tysiące przekazów, najczęściej zrodzonych w języku, lecz także poprzez bezsłowne «obrazy» i inne struktury danych, zajmują miejsce w jakimś indywidualnym mózgu, formując jego tendencje i tym samym przemieniają go w umysł².

Swoją teorię świadomości Dennett rozumie jako „realistyczną” i materialistyczną. Świadome przeżycia są realnymi zdarzeniami zachodzącymi w realnym czasie i w realnej przestrzeni mózgu. Są one czasowo i przestrzennie mierzalne³.

Do badań nad świadomością filozofia może wnieść analizy pojęciowe. Dennett przeprowadza takie analizy zarówno gdy chodzi o język potoczny, jak i o język naukowy⁴. Analiza pojęciowa ukazuje pomieszania językowe blokujące w przeszłości rozwój nauki o świadomości. Ponadto zgadza się on

² Dennett 1991: 253–254. Graficzny schemat wspomnianego „szkicu” można znaleźć w: Dennett 1998: 155.

³ Por. Dennett i Kinsbourne 1992a: 235.

⁴ Na przykład w *Time and the Observer* D. Dennett i M. Kinsbourne pokazują, jak eksperymenty naukowe mogą pomóc w rozwiązaniu problemu odróżniania modeli świadomości. Neurofizjolog mógłby wprawdzie takie rozróżnienie odkryć, jednak przeprowadzanie eksperymentów naukowych jest jedną sprawą, zaś analiza pojęć, do których te eksperymenty się odwołują – inną.

ze stwierdzeniem, że filozofowie patrzą na całość rzeczywistości, co jednak nie czyni ich uprzywilejowanymi, lecz uwrażliwia na potrzebę syntez jednoczących fakty pochodzące z różnych dyscyplin wiedzy. Pojęciowe syntezy mają pokazać kierunki przyszłego rozwoju nauki o świadomości. Przykładem tego rodzaju syntezy jest wspomniany w powyższym cytacie model umysłu zwany Modelem Wielu Szkiców (*Multiple Drafts Model*)⁵. Dennett zdaje sobie sprawę, że nie jest to model ściśle naukowy, ma jednak nadzieję, że będzie on stymulował poszukiwania adekwatniejszych teorii umysłu i świadomości (por. Dennett 1991: xi).

4.1. Heterofenomenologia – perspektywa trzecioosobowa

D. Dennett odwołuje się do „neutralnego” opisu, który nazywa opisem „heterofenomenologicznym”. Stanowi on istotę jego naturalistycznego rozumienia umysłu⁶ i związanej z tym rozumieniem tezą, że badanie świadomości należy prowadzić z perspektywy trzeciej osoby. Jak zobaczymy, metodologiczna neutralność heterofenomenologii obejmuje, z jednej strony, debatę na temat subiektywnych i obiektywnych podejść do klasycznej fenomenologii, zaś – z drugiej fizyczną czy нефизyczną rzeczywistość przedmiotów zjawiskowych (por. Dennett 1991: 95). Na pierwszoosobowe wypowiedzi podmiotu nie należy patrzeć jak na wypowiedzi o wewnętrznych zjawiskach poznawanych poprzez introspekcję, lecz podchodzić do nich z podobnym nastawieniem, z jakim podchodzi się do pisanych trzecioosobowo tekstów literackich.

Proponowana metoda składa się z dwóch części: (1) badany podmiot sporządza tekst składający się z opisu jego przeżyć (tekst ten może zostać nagrany), (2) tak otrzymany tekst jest następnie interpretowany. Zdaniem Dennetta, „Musimy przejść poza tekst, musimy go interpretować jako nagranie *czynności mowy* (*as a record of speech acts*), nie tylko jako wypowiedzi czy odczytywanie, ale jako stwierdzenia, pytania, obietnice, komentarze, prośby o objaśnienie, głośne rozważania, samoupomnienia” (Dennett 1991: 76).

Wydający dźwięki podmiot musi być traktowany jako ktoś racjonalnie działający, mający przekonania, pragnienia oraz inne stany mentalne wskazu-

⁵ Zwrot ten P. Czarnecki tłumaczy jako „model szkiców złożonych” (por. Czarnecki 2002: passim).

⁶ Według U. Żegleń, naturalizm Dennetta jest „[...] połączeniem pewnej wersji naturalizmu biologicznego [...] z funkcjonalizmem” (Żegleń 2003: 81). Żegleń zalicza Dennetta koncepcję umysłu do stanowisk „niematerialistycznych”, co jest kwestią wysoce dyskusyjną.

jące na intencjonalność, jako ktoś, kogo działania mogą być wyjaśnione (albo przewidziane) na podstawie treści tychże stanów mentalnych.

Tak interpretowany tekst jest heterofenomenologicznym opisem świata dokonany z perspektywy trzecioosobowej. Nie pytamy w nim o faktyczne istnienie obiektów, o których mówi podmiot. Metafizyczny status jego świata jest, według Dennetta, taki sam jak status świata wiktoriańskiego Pickwicka, a postawa naukowca wobec sporządzonego przez podmiot tekstu jest podobna do postawy czytelnika w stosunku do bohaterów powieści literackiej⁷. Ową metodologiczną neutralność Dennett określa mianem heterofenomenologicznego „minimalizmu metafizycznego”.

Należy zauważyć dwie rzeczy.

(1) Heterofenomenolog nie musi pasywnie akceptować treści subiektywnych sprawozdań odnoszących się do treści świadomości. Z jednej strony introspektywne sprawozdania wyrażają (*express*) określoną świadomą treść, z drugiej strony mogą być sprawozdaniami (*reports*) dotyczącymi innej treści (por. Korb 1993: 35–36). Treść introspekcyjnych relacji opisuje świat mentalny, który jest mniej lub bardziej koherentny i posiada związki z obiektami neurofizjologicznymi. Poza tym relacjonowana przez podmiot treść jest podobna lub niepodobna do introspekcyjnych relacji innych ludzi. Heterofenomenologiczne sprawozdania podmiotu o jego stanach mentalnych dotyczą także pośrednio przedmiotów świata zewnętrznego, dotyczą obiektów konkretnych i abstrakcyjnych. Tym samym heterofenomenologiczny tekst zawiera, oprócz sądów o stanach wewnętrznych, także weryfikowalne sądy o świecie.

(2) Podobnie fikcyjne są desygnaty niektórych terminów dyskursu naukowego, powiedzmy „środek ciężkości”. Nie istnieje nic, z czego składa się desygnat terminu „środek ciężkości”. Takie „przedmioty”, jak środek ciężkości, nie są ani czymś konkretnym, ani „[...] beczczynnymi fantazjami, lecz ciężko pracującymi fikcjami teoretyków”⁸.

⁷ „Mr. Pickwick is a fictional object, and so are the objects described, named, mentioned by the heterophenomenologist” (Dennett 1991: 95). Dennett zaznacza, że ludzie często nieświadomie tworzą różnego rodzaju fikcje, co powoduje, iż to, jak coś przedstawiają, może być dokładnie tym, jak im się coś wydaje. Heterofenomenolog poważnie traktuje, co ludzie mówią o tym – używając nagłowskiego sformułowania – jak to jest dla nich „rozwiązać problem, podjąć decyzję, rozpoznać dany obiekt” (s. 94).

⁸ Por. Dennett 1991: 95–96. Dodajmy, że znaczenie – w teorii Dennetta – nie jest czymś intrzymsycznym, ale jest konstruowane i podlega ewolucyjnemu procesowi selekcji, który kieruje tym, co jest aktualnie myślane i mówione. Zdaniem B. Voorheesa: „For Dennett, it is not a case of the Emperor having no clothes. It is rather that the clothes have no Emperor”. Albo używając samej teorii Dennetta – jak uzupełnia Voorhees – „[...] one could say that the clothes are the Emperor, a new twist on the old saw that the clothes make the man” (Voorhees 2000: 56).

Heterofenomenolog zajmuje się tekstami, z których wyodrębniono myślące i odczuwające podmioty. Heterofenomenologiczny świat podmiotu składa się z różnego rodzaju obrazów, zdarzeń, dźwięków, smaków, odczuć, co do których podmiot jest przekonany, że istnieją one w jego strumieniu świadomości. Strumień ten należy do przestrzeni zjawiskowej, którą można porównać do logicznej przestrzeni pojęć lub do przestrzeni opowieści literackich. W pewnym sensie wszystkie heterofenomenologiczne sprawozdania przekazują przekonania podmiotu, który jest ich świadomy⁹. Aby uniknąć przypisywania podmiotowi kartezjańskiej perspektywy pierwszoosobowej Dennett wprowadza heterofenomenologiczną perspektywę podmiotu w trzeciej osobie.

4.1.1. Postulowana jaźń

„The way to answer these «first-person point of view» stumpers is to ignore the first-person point of view and examine what can be learned from the third-person point of view” (Dennett 1991: 336).

Dennett umieszcza swoją teorię świadomości pomiędzy behawioryzmem i introspekcjonizmem. Behawioryzm nomologiczny ignoruje pierwszoosobowe wypowiedzi o odczuciach i o innych wewnętrznych stanach mentalnych, konstruując teorię umysłu typu: bodziec–reakcja. Introspekcjonizm przyznaje tymże wypowiedziom specjalny status, wynikający z uprzywilejowanego dostępu podmiotu do jego własnych stanów mentalnych. Z jednej strony, Dennett krytykuje behawioryzm, gdyż intencjonalność wprowadzonych przez niego podmiotów może być stwierdzana tylko wewnętrznie (behawioryzm nie umie chociażby wykazać, że podmiot za każdym razem opisuje taki sam przedmiot, gdyż ten jest dostępny tylko introspekcyjnie). Z drugiej strony Dennett odrzuca świadomość jako fenomen dostępny introspekcyjnie, a tym samym odrzuca opartą na introspekcji Husserlowską fenomenologię. Ta ostatnia zakłada, że introspekcja dostarcza danych niedostępnych dla badań zewnętrznych¹⁰. Dennett chce określić świat pojęciowy podmiotu z zewnątrz i odwołuje się do heterofenomenologii: Co podmioty mówią o ich własnych, wewnętrznych, subiektywnych przeżyciach jest przyjmowane

⁹ „Ludzie są niewątpliwie przekonani, że posiadają mentalne obrazy, bóle, przeżycia percepcyjne i wszystko pozostałe, i te fakty – fakty o tym, w co ludzie wierzą i co przekazują, gdy wyrażają swoje przekonania – są zjawiskami, które musi wyjaśnić jakaś naukowa teoria umysłu. Organizujemy nasze dane odnośnie do tych zjawisk w fikcje teoretyków, w «intencjonalne obiekty» w heterofenomenologicznych światach” (Dennett 1991: 98).

¹⁰ W podobny sposób argumentuje K. Twardowski (por. 1997: 89–90).

jako oczywiste (jako oczywistość, jak podmioty coś odczuwają, a nie jako bezpośrednia oczywistość, jakim coś faktycznie jest).

Gdy potocznie mówimy o odczuwaniu czy byciu o czymś przekonanym, to zakładamy istnienie podmiotu, który coś odczuwa, który jest o czymś przekonany. Brzmi to trywialnie, ale trywialnym nie jest, gdyż nie w każdym działającym systemie taki podmiot zdaje się być obecny. Dodatkowym pytaniem jest: jak ów ktoś musi być zbudowany, aby – jak się zdaje – mógł odczuwać ból, aby mógł postrzeżyć? Dennett podkreśla:

„Z punktu widzenia trzeciej osoby postulowałem podmiot, heterofenomenologiczny podmiot, swego rodzaju fikcyjny «do wszystkich, których to może dotyczyć», któremu faktycznie, my stojący z boku, poprawnie przypisalibyśmy przekonanie, że przeżył jakieś oddziaływujące poruszenie” (Dennett 1991: 128).

Heterofenomenolog rezygnuje z poszukiwań jakiegoś „realnego” podmiotu, gdyż ten ma dla niego status postulowanej „teoretycznej fikcji”¹¹. Można powiedzieć w ten sposób: heterofenomenologiczna metoda jest najpierw teoretycznym uproszczeniem. Przy bliższym spojrzeniu pokazuje się jednak, że aby można było obserwatorowi postawić jakieś dodatkowe zadanie, musi on odpowiedzieć na pytanie, jak wszystkie jego obserwacje są do pogodzenia z jego „ja”. Patrząc heterofenomenologicznie, to właśnie „ja” jestem podmiotem wszystkich pochodzących ode mnie opowiadań. Zdaniem Dennetta:

(A) „Jakaś jaźń, według mojej teorii, nie jest znanym matematycznym punktem, ale abstrakcją zdefiniowaną przez miriady określeń i interpretacji (łącznie z samookreśleniami i samointerpretacjami), które utworzyły biografię żyjącego ciała, a dla którego owa jaźń jest Środkiem Narratywnej Ciężkości. Jako takie odgrywa ono szczególnie ważną rolę w aktualnej kognitywnej ekonomii żyjącego ciała [...]” (Dennett 1991: 426–427).

Zdaniem Dennetta, nasze opowiadania są czasami tworzone przez nas, a czasami one tworzą nas. Nasza ludzka świadomość i nasza narratywna sobość są produktami tych opowiadań, nie ich źródłem (por. Dennett 1991: 418). Wprawdzie chętnie wierzymy w istnienie w nas rzeczywistego „ja”, ale takie „ja” nie jest, według Dennetta, wcale czymś oczywistym. Za naszymi przekonaniem i życzeniami nie musi stać jakieś realne ja, lecz musi w nich

¹¹ „The subject’s heterophenomenological world will be a stable intersubjectively confirmable theoretical posit, having the same metaphysical status as, say, Sherlock Holmes’s London or the world according to Garp” (Dennett 1991: 81). Dennett dodaje jednak: „If we choose to interpret at all, we have no choice but to posit a person whose communicative acts we are interpreting. This is not quite equivalent to positing an *inner* system that is the Boss of the body [...]” (s. 228).

jedynie panować minimalny porządek, zapewniający nam dalsze przeżycie. Zatrzaszczyła się o to ewolucja: biologiczne „ja” jest każdemu z nas dane. U zwierząt owo „ja” jest wszystkim (por. Dennett 1991: 413). Człowiek poza tym, że posiada mózg, snuje swoistą „pajęczynę dyskursów” (*web of discourses*) ze słów oraz czynów. I chociaż nie wie co robi, po prostu snuje pajęczynę, która go zarazem podtrzymuje¹².

„Pajęczyna dyskursu” jest takim samym biologicznym wytworem, jakim jest pajęczyna pająka. Ludzie rozwinęli język, owo symboliczne medium, dzięki któremu mogą wyeliminować zarówno ograniczenia indeksykalne typu „tutaj”, „teraz” jak i ograniczenia konkretnej przedmiotowości. Używając terminologii komputerowej Dennett porównuje język do *software*, zainstalowanego na *hardware*, którym jest odpowiednio przez ewolucję wyposażony mózg. To dzięki językowi mogą zostać opisane i przekazane wydarzenia zaszłe w innych miejscach i w innym czasie. Język umożliwia operowanie pojęciami indeksykalnymi, do których należy termin „ja”. Dennett zauważa: „Podobnie jak biologiczna jaźń, owa psychologiczna albo narratywna jaźń jest teraz innego rodzaju abstrakcją, a nie rzeczą w mózgu [...]” (Dennett 1991: 418).

Skoro tylko dysponujemy językiem, zaczynamy mówić o nas samych w taki sposób, jak gdyby w nas był ktoś, powodujący, że obserwator zewnętrzny i my sami jesteśmy przekonani o istnieniu jakiegoś podmiotu. W przeciwieństwie do tego, ja sam (według Dennetta) nie jestem niczym więcej jak tylko Środkiem Narratywnej Ciężkości i równocześnie swoim własnym słuchaczem. Chociaż zdajemy się posiadać zdolności do utworzenia własnego „ja”, to nie jest czymś oczywistym, że je faktycznie utworzymy. Sam bowiem *hardware* (mózg) do takiego tworzenia nie wystarczy, lecz musi na nim zostać zainstalowane odpowiednie memowe (*memes*) *software*. *Hardware* jest produktem ewolucji biologicznej, memowe *software* jest produktem ewolucji kulturowej. Memy są kulturowymi, replikującymi się jednostkami (do tematu „memes” wracamy w 4.4.). Systemy zainstalowane w naszych mózgach są, według Dennetta, wynikiem selekcji naturalnej oraz naszych indywidualnych historii życia.

Jaźń jest konstruktem utworzonym z wielu szkieł, a stosowana przez jaźń taktyka samozachowania, samokontroli i samookreślenia polega na opowiadaniu opowiadań (w tym także na tworzeniu i kontrolowaniu opowiadania o nas samych). Opowiadania te nie są tworzone przez pojedynczego ko-

¹² „Unlike a spider, an individual human doesn't just exude its web; more like a beaver, it works hard to gather the materials out of which it builds its protective fortress” (Dennett 1991: 416).

gość. Dennett neguje istnienie jakiegoś Centralnego Dawcy Znaczenia. Narratywny strumień obejmuje nasze jaźnie oraz wypowiedziane i niewypowiedziane odniesienia (*references*) do innych ludzi. Narratywne strumienie wpływają z danego osobnika tak, jak gdyby wypływały z jakiegoś pojedynczego źródła. W efekcie, jako słuchający, postulujemy istnienie jednolitego podmiotu, który wypowiada dane słowa.

Podobnie jak środek ciężkości jakiegoś przedmiotu ma swoje miejsce w fizyce, tak Środek Narratywnej Ciężkości ma swoje miejsce w naukach o człowieku (por. Dennett 1992: 1; 1991: 418). Jego istnienie zakładamy nie tylko u innych, ale także w samych sobie. Środek Narratywnej Ciężkości posiada jedynie te cechy, w jakie wyposażała go teoria, w ramach której jest on postulowany. Wynika stąd, iż jeżeli jakiś materialista pyta: „gdzie jest ta jaźń?” i szuka jej w mózgu to popełnia on – używając określenia Ryle’a – „pomyłkę kategorialną”. Jedyną własnością środka ciężkości jest jego nieprecyzyjnie zdefiniowana lokalizacja przestrzenno-czasowa. Bardziej precyzyjnych danych o lokalizacji jaźni może dostarczyć postęp badań neurofizjologicznych. Nie należy się jednak spodziewać, że kiedyś będziemy mogli powiedzieć: Tamta komórka, na prawo w środku hipokampa – to jest jaźń. Według Dennetta, nie tylko pacjenci z rozszczepionymi osobowościami, ale wszyscy jesteśmy „nowelistami”¹³. Wszyscy opowiadamy samym sobie opowiadania nie troszcząc się zbytnio o ich każdorazową prawdziwość.

Na potwierdzenie swojej teorii jaźni Dennett przytacza opisy badań prowadzonych przez neurofizjologa M. Gazzanigę nad pacjentami z rozdzielonymi półkulami mózgowymi (*split-brain subjects*) [por. Dennett 1991: 260]. Gazzaniga neguje, że pacjenci ci zachowali swoją przedoperacyjną jedność jaźni, i dodaje, że taka jedność jest iluzją. Nasz umysł nie odznacza się jednością, lecz jest raczej złożoną wiązką częściowo autonomicznych systemów. Na pytanie: czy operacja rozdzielenia półkul dzieli jaźń na dwie części? Dennett odpowiada, że po takiej operacji pacjenci nie zachowywali się (w normalnych warunkach), jak gdyby mieli dwie osobowości¹⁴.

Sprawozdaniom podmiotu możemy próbować przypisać jakiś rzeczywisty desygnat, ale czyniąc to musimy znaleźć pewien rodzaj korelacji pomiędzy mózgiem i heterofenomenologicznie ukonstytuowanym światem świa-

¹³ Dennett stwierdza: „[...] it does seem that we are all virtuoso novelists, who find ourselves engaged in all sorts of behavior, more or less unified, but sometimes disunified, and we always put the best «faces» on it we can. We try to make all of our material cohere into a single good story. And that story is our autobiography” (Dennett 1992: 6–7).

¹⁴ Opinie naukowców co do „wielości osobowości” (u tych pacjentów) są podzielone (por. Sperry 1990: 380–381; Nagel 1997a: 185–201).

domości. Właściwym desygnatem jakiegoś pojedynczego terminu używanego przez podmiot jest, dla Dennetta, takie zdarzenie w mózgu, które jak najbardziej odpowiada heterofenomenologicznemu obiektowi, o którym mowa w sprawozdaniu (tzn. obiektowi, który nosi tę nazwę w „świecie” ukonstytuowanym przez heterofenomenologiczny tekst).

4.1.2. Ocena heterofenomenologii

Proponowana przez Dennetta metoda jest zarówno deskryptywna, jak i normatywna (w sensie: co powinno się czynić, aby osiągnąć obiektywne rezultaty badań świadomości). Czynione w jej ramach poszukiwania odpowiedników stanów mentalnych w realnym świecie są ograniczone do obiektów, które mogą być dostępne z perspektywy trzeciej osoby. Dennett odrzuca „rzeczywistość” subiektywnych doświadczeń, chociaż równocześnie uznaje, że jedyny dostęp do ich treści mamy poprzez wypowiedzi podmiotu. Dostępnymi z perspektywy trzecioosobowej odpowiednikami subiektywnych stanów (o których mówią sprawozdania w pierwszej osobie) mogą być wzorce należących do podmiotu neuronalnych aktywności systemu [$\Phi_j - \Phi_n$].

Podmiotowa subiektywność jest formą iluzji podobną do tej, jaką posiada użytkownik komputera co do „ja” swojej maszyny¹⁵. W przypadku naszego własnego „ja” Dennett neguje istnienie „użytkownika” czy „obserwatora”, który ulega takiej iluzji. Niemniej jednak,

IV.1 Z filozoficzno-psychologicznego punktu widzenia trudny do zaakceptowania jest pogląd, że dennettowski podmiot, wychodząc od heterofenomenologicznej analizy siebie, dochodzi do wniosku, że w ogóle nie istnieje.

Dla Dennetta jaźń jest iluzją bez kogoś, kto ją posiada. Uwzględniając jednak rozumienie świadomości obecne w psychologii potocznej należy zauważyć:

(1) W psychologii potocznej traktujemy świadome zjawiska jako pochodzące z dokładnie określonych źródeł, co nie stwarza problemu, gdy chodzi o przedmioty wielkości ludzkiego ciała. Złudzeniu możemy ulec dopiero wtedy, gdy z potocznymi kategoriami („podmiot”, „świadome odczucie”) przejdziemy do naukowego obrazu świata i w mózgu zaczniemy szukać ja-

¹⁵ M. J. Farah zarzuca Dennettowi, że jego „[...] analysis would apply equally well to the nonconscious representations of an appropriately programmed PC” (Farah 1992: 209).

kiegoś osobliwego anatomicznego regionu, z którego pochodzi świadomość. W tym kontekście nasuwają się następujące pytania.

(2) Nasza jaźń oszukiwałaby samą siebie, gdy chodzi o przekonania (albo z Dennettowskiego punktu widzenia o dyspozycję do wierzenia, myślenia i mówienia), że jest ona czymś innym aniżeli zbiorem narratywnych dyspozycji. Najczęściej oszukiwanie dotyczy drugiej osoby, która nie wie czegoś, co jest mi znane. W przypadku samooszukiwania się mamy – choć nie zawsze – do czynienia z czymś wewnątrznie nieracjonalnym¹⁶.

Pozostaje nadal pytanie, jak to się dzieje, że owa jaźń (por. IV.1) jest świadoma? Czy świadomość jest czymś poza Środkiem Narratywnej Ciężkości? Prowadzone przez Dennetta funkcjonalistyczne analizy są ograniczone do metod trzecioosobowych i nie potrafią na te dwa pytania odpowiedzieć. Stąd między innymi konieczność zbudowania nauki o świadomości zjawiskowej, zawierającej przeżycia widziane z perspektywy pierwszej osoby. Potrzeba przy tym dostrzec konieczność swoistej komplementarności opisów z perspektywy pierwszej i trzeciej osoby. Mamy bezpośrednie doświadczenie świadomości zjawiskowej (w pierwszej osobie), ponieważ istnieje coś takiego, co pyta *jak to jest być* nami samymi. T. Nagel i J. Searle – co zresztą sam Dennett zauważa – podkreślają że heterofenomenologiczne studia nad świadomością pozostawiają nietknięty jej rzeczywisty problem – mianowicie owo *jak to jest być* (por. Dennett 1991: 96).

Zdaniem Dennetta, świadomość nie jest konstytuowana przez leżącą u podstaw Wielu Szkiców jaźń. Świadome przeżycie jest „sprawozdaniem”, które wypowiadamy, albo „sądem”, który czynimy, albo „przekonaniem”, które utrzymujemy o naszych przypuszczalnych przeżyciach. Tego rodzaju ujęcie świadomości niesie z sobą modułową koncepcję umysłu (należy przyjąć, że wiele modułów pracuje poza kontrolą świadomości) oraz sprowadzenie świadomości do tego, co intencjonalne.

Nie wynika z tego jednak, naszym zdaniem, że postawy propozycjonalne są „instrumentalną”, pragmatycznie korzystną fikcją pochodzącą z potocznej teorii umysłu. Dlatego przyjmujemy twierdzenie:

IV.2 Mentalny akt, w którym postulujemy cechujący się jednością podmiot, wymaga wewnętrznej idei jedności. Pytanie brzmi: czy taka idea jedno-

¹⁶ Przypadki tego rodzaju opisuje A. R. Mele (1997: 96–98). „Analyzing self-deception is a difficult task: providing a plausible set of sufficient conditions for self-deception is less demanding. Not all cases of self-deception need involve the acquisition of a new belief. Sometimes we may be self-deceived in retaining a belief that we were not self-deceived in acquiring” (s. 105).

ści może powstać bez odwołania się do cechującej się jednością, holistycznej świadomości?

Odpowiadając na IV.2 Dennett, naszym zdaniem, musiałby twierdzić, że wszystkie przypisywania jedności, także jedności nas samych, są po prostu sprawą językowych konwencji, bez znaczeniowej treści. Używając metaforycznego języka Dennetta możemy powiedzieć, że cechująca się jednością świadomość byłaby swoistym „wirtualnym kapitanem” (Dennett 1991: 228), małą grupą spokrewnionych pakietów informacyjnych, która powstała dla uzyskania czasowej dominacji w zabiegach prowadzących do uzyskania kontroli nad takimi poznawczymi (intencjonalnymi) aktywnościami, jak samomonitorowanie czy samowypowiadanie się. Aktywności te zachodzą w ogromnym zbiorze mózgowych mikropołączeń. Tego typu krótkotrwałe zjawiska uważamy następnie, zdaniem Dennetta, za coś więcej, aniżeli są, gdyż każde z nich jest w tym momencie „mną”.

Dennett traktuje cechującą się jednością świadomość jako rezultat określonych samoorganizujących się aktywności w mózgu. Zakłada on, że mając daną naturę mózgu jako ogromny zbiór neuronów otrzymujących od siebie i wysyłających do siebie sygnały, nasze poznanie nie może przyjąć innej formy aniżeli formę chaotycznie konkurujących ze sobą elementów treści, z których świadome są te, które w tej konkurencji wygrywają.

Należy podkreślić, że tak rozumiana jedność świadomości odnosi się nadal do stanów intencjonalnych. Dennett nie mówi nam nic o holistycznej jedności świadomości przeżyciowej, traktując ją jako coś epifenomenalnego.

4.2. Kartezjański Teatr

Dennett podkreśla, że rozumienie „umysłu” jako miejsca psychologicznych „zdarzeń mentalnych” nie musi być logicznie powiązane z dualizmem. Błędna idea „umysłu” jako czegoś niecielesnego może wystąpić w połączeniu z materialistycznymi teoriami identityczności typów. Wystarczy, że umysł zostaje „umiejscowiony” w mózgu, a zjawiska mentalne są traktowane jako identityczne ze zdarzeniami neurofizjologicznymi. Tego rodzaju umysłowo-mózgowe zdarzenia zostają następnie wkomponowane w „fizyczne” łańcuchy przyczyn. Pogląd taki, który według Dennetta reprezentują niektórzy neurofizjologowie, neguje dualizm przy zachowaniu wyobrażenia o jakimś centralnym (ale materialnym) miejscu w mózgu¹⁷.

¹⁷ Zdaniem B. Baarsa: „More broadly, Dennett and Kinsbourne maintain that there is no single place in the brain where «it all comes together», as suggested by Damasio, Crick and Koch, and others” (Baars 1998: 3; por. Kinsbourne 1997: 340).

Zdaniem Dennetta, tego rodzaju wyposażenie mózgu w cechy mentalne, prowadzi materialistów do błędnego rozumienia natury tych stanów. Dualistyczne mentalne obiekty, stany, procesy, stają się dla materialistów „naturalnymi” (najczęściej w sensie „nietajemniczymi”) obiektami lub stanami mózgu. Ponadto materialiści zakładają, że podstawowe cechy mentalne zostaną po takiej transformacji jakoś zachowane (na przykład takie cechy, jak to, że *ktos* coś obserwuje, że istnieje jakieś pojedyncze, indywidualne *coś*). Tego rodzaju pogląd Dennett nazywa kartezyjskim materializmem. Wyjaśniając zjawisko świadomości w ujęciu materialistycznym Dennett krytykuje zarówno kartezyjski dualizm, jak i rodzący się z niego kartezyjski materializm.

W krytykowanych przez Dennetta materialistycznych teoriach z dziedziny kognitywistyki przyjmuje się (jego zdaniem), że:

(a) istnieje swoiste miejsce w głowie, gdzie wszystko się schodzi „it all comes together”: wrażenia powierzchni, koloru i ruchu danego makroobiektu oraz jego ewentualne rozpoznanie;

(b) czasowe własności pojawiających się w owym miejscu zdarzeń determinują czasowe własności (następstwa, równoczesności, ...) subiektywnego „strumienia świadomości”¹⁸. Miejsce to Dennett nazywa Kartezyjskim Teatrem: „Kartezyjski Teatr jest metaforycznym obrazem tego, jak świadome przeżycie musi być usadowione w mózgu” (Dennett 1991: 107).

Kartezyjusz był przekonany, że mózg posiada centrum: szyszynkę (*corpus pineale*), będącą nie tylko drogą do świadomego umysłu¹⁹, ale także miejscem, w którym oddziałuje on na ciało (tzw. interakcjonizm kartezyjski). Takie rozumienie umysłu prowadzi, według Dennetta, do naruszenia zasady zachowania energii. Ponadto, neurologowie mówią nam (na przykład A. Damasio), że nic w szyszynce ani prawdopodobnie w całym mózgu nie jest siedzibą świadomości²⁰. Siedzibą taką nie są również ani kora śródmózgowa (*ento-*

¹⁸ Por. Dennett i Kinsbourne 1992a: 184. G. Rey zaznacza, że używana przez niektórych materialistów metafora „wewnętrznego teatru”, „specjalnego punktu widzenia”, nie może być rozumiana dosłownie (por. Rey 1994: 263).

¹⁹ Dennett i Kinsbourne piszą dosłownie: „For a person to be conscious of something, traffic from the senses had to arrive at this station [*pineal gland – J.B.*], where thereupon caused a special – indeed magical – transaction to occur between the person’s material brain and immaterial mind. When the conscious mind then decided on a course of bodily action, it sent a message back «down» to the body via the pineal gland. The pineal gland, then, is like a theater in which information is displayed for perusal by the mind” (Dennett i Kinsbourne 1992a: 185).

²⁰ Krytykując Kartezyjski Teatr A. Damasio stwierdza: „[...] there is no neuroanatomical structure in the cerebral cortex to which signals from all the sensory modalities that may be represented in our experience can converge, spatially and temporally. The ento-

rhinal cortex), ani hipokamp (*hippocampus*), którym takie zadania przypisywano. Pacjenci, którzy mieli uszkodzone powyższe organy, nie wykazywali zaburzeń świadomości w sensie, jaki interesuje Dennetta.

Dennett zaznacza, że wielu współczesnych materialistów zapomina, iż jeżeli odrzucili kartezjańską *res cogitans*, to w ich teorii nie ma miejsca dla jakiegś „centralnej drogi”, czy dla jakiegś funkcjonalnego centralnego miejsca w mózgu. Podstawowym błędem Kartezjańskiego Teatru – jako modelu umysłu – jest zdaniem Dennetta przyjęcie założenia, że mózg posiada jakiś wewnętrzny punkt, swoisty „deeper headquarter” (por. wyżej, 4), którego obecność jest koniecznym i wystarczającym warunkiem wystąpienia świadomego przeżycia (por. Dennett 1991: 106).

Odrzucając Kartezjański Teatr Dennett neguje cały szereg określeń związanych z rozumieniem świadomości i podmiotu. Kilka z tych określeń wymienia G. Rey²¹:

1. Istnieje „wewnętrzny teatr”, w którym wiele lokalnych spostrzeżeń jest ponownie re-reprezentowanych oraz zostają wypełniane luki pomiędzy tymi spostrzeżeniami (por. Dennett 1991: 113).

2. Umysł jest swoistym „miejszem subiektywności” (*locus of subjectivity*), treści obecne w teatrze determinują znane nam *jak to jest* być danym osobnikiem (por. Dennett i Kinsbourne 1992a: 183). Owo *jak to jest* być danym osobnikiem, jest częściowo determinowane przez to, co jest dostępne dla obserwacji albo dla przeżycia²².

3. Treści występujące w teatrze są zdeterminowane i wyraźne, nawet w przypadkach, w których introspektywnie wydają się one niewyraźne. Czas recepcji danej treści w świadomości jest czasem świadomego przeżywania tej treści (por. Dennett 1991: 107).

4. Obecne w owym teatrze treści tworzą pojedyncze opowiadanie („*single narrative*”) (Dennett 1991: 136). Teatr jest miejscem, w którym schodzą się wszystkie psychologiczne zdarzenia, potrzebne chociażby do podejmowania decyzji.

5. Potoczne intuicje sugerują, że nasz strumień świadomości składa się ze zdarzeń zachodzących w sekwencjach oraz że w jakimś momencie każdy element sekwencji może albo zostać uznany za świadome zjawisko, albo nie.

rhinal cortex and the hippocampus might be candidates for that sort of «integrative» role but they do not pass the necessary anatomical tests” (Damasio 1992: 208).

²¹ Por. Rey 1994: 262–263. O metodologicznych założeniach Dennetta teorii świadomości por. Dennett 1991: 40–41.

²² Por. Dennett i Kinsbourne 1992a: 183. Dennett: „Whatever there is a conscious mind, there is a point of view” (Dennett 1991: 101).

Rey powątpiewa zarówno w słuszność kartezyjskiego rozumienia umysłu, jak i w to, że Dennett podał wystarczające podstawy do odrzucenia tego rozumienia. W następnych punktach omówimy naturalistyczne wersje Kartezyjskiego Teatru, które są do pogodzenia z danymi empirycznymi, na które powołuje się Dennett w swojej negacji kartezyjskiego modelu umysłu.

4.3. *Wiele Szkiców jako model subiektywnych przeżyć*

„Dennett thinks there are no such things as qualia, subjective experiences, first-person phenomena, or any of the rest of it” (Searle 1998: 99).

Koncyptowana w ramach Modelu Wielu Szkiców świadomość nie jest ujednoliconym procesem, lecz czymś rozłożonym, czymś co przypomina edytowanie szkiców jakiegoś opowiadania. Świadomość jest mniej lub bardziej ciągłym (*continuing*), zmiennym strumieniem stale edytowanych, re-edytowanych, przechodzących z jednego w drugi szkiców (*drafts*)²³. Czasami inicjują one wypowiedzi, czasami zostają zachowane w pamięci, czasami zostają po prostu zapomniane. Docierająca do systemu nerwowego informacja podlega ciągłej „obróbce redakcyjnej”. Określenie „momentu publikacji”, a tym samym nazwanie jednego ze szkiców „tekstem kanonicznym”, jest sprawą arbitralną. Dennett zaznacza:

„Zgodnie z Modelem Wielu Szkiców, każda różnorodność aktów spostrzegania – właściwie wszystkie różnorodności myśli lub aktywności mentalnej – zachodzą w mózgu przez równoległe, wielośladowe procesy interpretacji i przetwarzania danych sensorycznych. Informacja wchodząca do systemu nerwowego podlega stalej «rewizji edytorskiej»” (Dennett 1991: 111).

Wspomniane procesy interpretacji i rozpracowania danych sensorycznych zachodzą w setkach milisekund. W tym czasie może dojść do różnego rodzaju uzupełnień, do tworzenia połączeń, do przetwarzania treści. Ustalenie cech czy ich odróżnienie dokonuje się tylko raz. Istnieją określone części mózgu odpowiedzialne za rozpoznawanie lub odróżnianie danych cech, a pochodząca od tych części informacja nie jest przesyłana do jakiejś nadrzędnej, „interpretacyjnej” części mózgu. Jeżeli jakaś obserwacja została dokonana,

²³ Dennett wskazuje na analogię z tekstami rosyłanymi drogą elektroniczną. „With the advent of word-processing and desktop publishing and electronic mail, it now often happens that several different drafts of an article are simultaneously in circulation, with the author readily making revisions in response to comments received by electronic mail” (Dennett 1991: 125).

to treść uzyskanej informacji nie musi być przesłana gdzieś indziej, aby została raz jeszcze przedstawiona jakiemuś innemu „Interpretatorowi”. Nie istnieje coś takiego, jak ponowne przedstawienie właśnie określonej cechy na scenie Kartezjańskiego Teatru. Dennett dodaje:

„Szkice przeżycia mogą być szybko przetwarzane, a żaden z nich nie jest bardziej poprawny aniżeli inny. Każdy odzwierciedla sytuację w czasie, w którym został dokonany” (Dennett i Kinsbourne 1992a: 185).

Zmiany w konkretnych częściach mózgu zaangażowanych w rozpoznawanie lub odróżnianie cech zachodzą w określonym czasie i przestrzeni, a ich zajście nie oznacza jeszcze pojawienia się świadomości ich treści.

Innymi słowy: następujące po sobie świadome zdarzenia rozpadają się na krótkie, mierzone w milisekundach, nieuprządkowane zdarzenia i to one tworzą świadomość. Nie ma jakiegoś pojedynczego, centralnego miejsca w mózgu (rozumianego jako miejsce materialne), gdzie zapadają świadome decyzje, gdzie zachodzą świadome przeżycia, gdzie subiektywnemu osądowi są prezentowane wszystkie rozróżnienia poszczególnych modalności przeżyć. Według M. Kinsbourne, do zadań neurofizjologa należy określenie właśnie tego szeregu stanów mózgu, którym towarzyszą stany subiektywne (por. Kinsbourne 1997: 340).

Zdaniem J. Glicksohna, Model Wielu Szkiców posiada swoją długą historię w psychologii kognitywnej i rozwojowej, gdzie spostrzeganie jest traktowane jako proces identyfikowania, poznawania, sądzenia, a wrażenia są traktowane jako minimyśli. Każda świadomość czegoś jest świadomością, że coś. Model dennettowski – według Glicksohna – można umieścić w ramach szerszej teoretycznej debaty dotyczącej chociażby pojęcia regresji (*regression*) w psychologii rozwojowej²⁴. Tym samym jest on podobny do modelu zbudowanego przez psychologa-kognitywistę H. Wenera. Ten ostatni sugeruje, wychodząc od obserwacji różnorodnych zachowań podmiotu, że mniej rozwinięte systemy zachowania zostają wprawdzie z czasem podporządkowane bardziej rozwiniętym systemom i są przez nie regulowane, ale nie znikają. Podobnie ma się sprawa z dennettowskimi wcześniejszymi „szkicami”, które w normalnych sytuacjach są jedynie trudniej dostępne, a które z reguły ujawniają

²⁴ „By regression I mean a shift to the mode of cognition that is characteristic of an earlier stage of functioning, [...] the problem is whether regression or reversion can be accommodated by the developmental theory [...] Are the former drafts retained after progressive development, and if so, what is their relation to the more advanced drafts that are presently dominant?” (Glicksohn 1995: 807).

niąją się dopiero w sytuacjach stresowych, w których system musi odpowiedzieć na nowe wyzwania²⁵.

Dennett i Kinsbourne zgadzają się z tezami Glicksohna. Odrzucają jednak jego założenie, że późniejsze „szkice” muszą być „bardziej rozwinięte”. Tego rodzaju założenie „flirtuje”, ich zdaniem, z wszechobecną mitologią dochodzenia przez podmiot do coraz wyższej świadomości, rozwijając się po drodze (por. Dennett i Kinsbourne 1995: 811).

4.3.1. Orwellowski i Stalinowski Model świadomego umysłu

Uzasadniając wprowadzenie Modelu Wielu Szkiców Dennett wskazuje na jego plusy przy wyjaśnianiu relatywnie prostego zjawiska – „zjawiska pozornego ruchu” (tak zwany *fenomen phi*)²⁶. Dwa leżące blisko siebie świetlne punkty – czerwony i zielony – zapalają się na zmianę i gasną. Obserwującemu je podmiotowi wydaje się wówczas, że pojedynczy punkt porusza się tam i z powrotem, zmieniając w połowie drogi swój kolor. Pytanie brzmi: W jaki sposób podmiot, zanim zabłyśnie druga plamka, potrafi wypełnić kolorem miejsce wzdłuż drogi biegnącej od pierwszego błysku do drugiego? (por. Dennett 1991: 114). Jeżeli odrzucimy zdolność podmiotu do jasnovidzenia, to możemy przyjąć, że jego mózg potrafi jakoś „przedatować” ten błysk w przeszłości.

W ramach modelu Kartezjańskiego Teatru istnieją, zdaniem Dennetta, dwa rozwiązania *zjawiska phi*: tak zwane modele „orwellowski” i „stalinowski”. Nazwy te odzwierciedlają praktyki okłamywania społeczeństw przez systemy totalitarne, które to praktyki można porównać z metodami tworzenia iluzji przez umysł. W mózgu może funkcjonować Model Stalinowski, w którym uzupełnia się luki pomiędzy poszczególnymi wrażeniami wzrokowymi, tak że stają się one ciągłe. Zanim wrażenia te zostaną uświadomione, są w pewnym sensie sfalszowane przez wprowadzenie „falsyfikacji” potrzeb-

²⁵ Glicksohn powołuje się na Dennetta i Kinsbourne’a: „[...] at any one point in time there are multiple «drafts» of narrative fragments at various stages of «editing» in various places in the brain. Probing this stream at different intervals produces different effects, elicits different narrative accounts from the subject [...] If one probes «to early» one may gather data on how early a particular discrimination is achieved in the stream, but at the cost of disrupting the normal progression of the stream” (Dennett i Kinsbourne 1992a: 185).

²⁶ Pozorny ruch można zobaczyć na stronie: <<http://www.yorku.ca/eye/balls/htm>>. Zjawisko to leży u podstaw naszego przeżywania ruchu w oglądanych obrazach telewizyjnych. „Motion pictures and television depend on creating apparent motion by presenting a rapid succession of «still» pictures (Dennett 1991: 114).

nych do uzyskania ciągłości. W mózgu może także funkcjonować model orwellowski, działający po uświadomieniu sobie wrażeń wzrokowych²⁷. Dennett zaznacza, że funkcjonowanie obydwu modeli jest zgodne z tym, co podmiot myśli, wypowiada i pamięta, oraz że obydwa posługują się swoistymi „falsyfikacjami”.

Mówiąc ogólnie, w modelu stalinowskim spostrzeżenie pierwszej plamki jest opóźnione (przesunięte), w modelu orwellowskim spostrzeżenie pierwszej plamki jest zachowane lub zapamiętane (por. Dennett 1991: 127). W modelu stalinowskim spostrzeżenia czegoś blokują naszą pamięć, ale możemy edytować nasze spostrzeżenia, w modelu orwellowskim możemy wymazać wyniki naszego początkowego spostrzeżenia poprzez przepisanie na nowo naszej pamięci. „Stalinowscy” teoretycy twierdzą, że edytujemy nasze spostrzeżenia, „orwellowscy” utrzymują, że przepisujemy na nowo naszą pamięć.

„Jak stwierdzić, która z tych dwóch hipotez jest poprawna?” – pyta Dennett. Jedna postuluje stalinowskie „wypełnienie” (*filling in*) na przedprzeżyciowej (*pre-experiential*) drodze, druga postuluje orwellowską „rewizję pamięci” na poprzęziociowej (*post-experiential*) drodze, i obydwie są do pogodzenia z tym, co dany podmiot mówi, myśli albo co sobie przypomina. Należy zauważyć, że wspomniana niezdolność do odgraniczenia tych dwu możliwości nie dotyczy jedynie zewnętrznych obserwatorów, którym może brakować jakichś prywatnych danych, do których podmiot ma „uprzywilejowany dostęp”. Według Dennetta niezdolność do odróżnienia dotyczy także samego spostrzegającego podmiotu: „Ty, jako podmiot podczas eksperymentu z fenomenem phi, nie potrafisz z twojej własnej pierwszoosobowej perspektywy odkryć nic w doświadczeniu, co stawiałoby jedną teorię nad drugą; przeżycie byłoby «odczuwane tak samo» w obydwu przedstawieniach” (Dennett 1991: 123).

Dennett podkreśla, że gdy chodzi o mózg, oba omawiane modele odwołują się do tej samej teorii. Oba modele zgadzają się co do tego, że błędna treść wchodzi (w mózgu) w przyczynowy łańcuch. Nie zgadzają się jednak

²⁷ Orwellowska „revision” odnosi się do sytuacji, w których wydaje się nam, że przypominamy sobie zdarzenia, które nigdy nie miały miejsca. „We might call such post-experiential contaminations or revisions of memory”. Stalinowski „show trial” oznacza „[...] to stage show trials, carefully scripted presentations of false testimony and bogus confessions, complete with simulated evidence. We may call this ploy Stalinesque” (Dennett i Kinsbourne 1992a: 191.). Stalinowski rodzaj iluzji jest tworzony poprzez zmianę samego przeżycia. Orwellowski model nie domaga się pokazowego procesu, a jedynie odpowiedniego spreparowania książek do nauki historii (musi być w nich napisane, iż proces był, chociaż go nie było).

co do tego, gdzie owo wejście ma zachodzić: przed świadomym przeżyciem czy po przeżyciu. Obydwa modele dają takie samo przedstawienie wyników niewerbalnych, z tą różnicą, że jeden podaje, że są one (te wyniki) rezultatem nieświadomie określonych (*discriminated*) treści, podczas gdy drugi podaje, że są one rezultatem świadomie określonych, ale zapomnianych treści. Obydwa modele odnoszą się do danych subiektywnych typu *jak to jest*: „[...] wyjaśniają dane subiektywne – cokolwiek jest dostępne «z perspektywy pierwszej osoby» – ponieważ zgadzają się one co do tego, jak coś powinno być odczuwane przez podmioty” (Dennett i Kinsbourne 1992a: 193).

Podmioty (zdaniem Dennetta) są niezdolne do wyrażenia różnicy pomiędzy „źle poczętymi” (*misbegotten*) przeżyciami a bezpośrednio „źle zapamiętanymi” (*misremembered*) przeżyciami (do interpretacji powyższego fragmentu wracam w 4.4.). W przykładzie z podejściem stalinowskim i orwellowskim Dennett stawia sobie za cel pokazanie, że w badaniach nad świadomością mamy do czynienia ze swoiście nierozwiązywalnym problem deskryptywnym.

4.3.2. Model Wielu Szkieców

Pomiędzy modelami orwellowskim i stalinowskim istnieje, według Dennetta, jedynie różnica werbalna. Nie istnieje jakaś rzeczywista linia oddzielająca jeden model od drugiego. (A) „Te dwie teorie przekazują dokładnie to samo opowiadanie za wyjątkiem tego, gdzie umieszczają mityczny Wielki Rozdział (*Great Divide*), punkt w czasie (a tym samym miejsce w przestrzeni), którego *precyzyjna* pozycja w niczym nie pomaga podmiotom w jego lokalizacji i którego umiejscowienie jest również neutralne w stosunku do wszystkich innych cech ich teorii” (Dennett 1991: 125).

Mówiąc o różnicy nie powodującej żadnej różnicy należy dodać, że Dennett przyjmuje istnienie w nas fizycznie opisywalnych różnic, których jednak nie jesteśmy świadomi. Dennett stwierdza w typowo „weryfikacjonistyczny” sposób, że ze względu na czasoprzestrzenne zatarcie (*spatiotemporal smearing*) się w mózgu punktu widzenia obserwatora, ani wszelka dostępna oczywistość ani wszelkie dowody, nie dostarczają podstaw do odgraniczenia orwellowskiego modelu świadomych przeżyć od modelu stalinowskiego. W tym sensie żadna różnica nie istnieje²⁸. Dennett i Kinsbourne

²⁸ W. G. Lycan nazywa to stwierdzenie „argumentem weryfikacjonistycznym”. Bez dalszego uzasadnienia Lycan nie zgadza się z tezą Dennetta i Kinsbourne’a [D&K] mówiącą o jedynie werbalnej granicy: „[...] I am completely unpersuaded by D&K’s claim that the Orwellian and Stalinesque accounts of (e.g.) color phi differ «only [...] verbal[ly]»” (Lycan 1992: 217).

stwierdzają, że oprócz cech dotyczących różnych synchronizacyjnych połączeń i skutków wywieranych na pamięć, na mówienie i na inne zachowania, nie istnieje żadna funkcjonalna czy neurologiczna cecha *K* umożliwiająca odróżnienie obydwu modeli (por. Dennett i Kinsbourne 1992a: 236).

W Modelu Wielu Szkiców nie mamy ani orwellowskiego „przepisywania wspomnień”, ani stalinowskiego „edytowania spostrzeżeń”. Kolejność wspomnień zostaje ustanowiona już w procesie ich formowania, dlatego nie ma orwellowskiej kolejności fałszowania wspomnień. Doświadczenie podmiotu może być jedynie subiektywne, dlatego nie ma fałszywego edytowania spostrzeżeń. Impuls wzrokowy powoduje szereg zdarzeń w mózgu, które realizują różnorodne, coraz bardziej wyspecjalizowane funkcje. W mózgu nie ma jednak „linii”, którą szeregi te muszą przekroczyć, aby przeżycie wzrokowe stało się świadome. Strumień kolejnych i równoległe przebiegających zdarzeń w mózgu (por. 4. *thumbnail sketch*) powoduje różne skutki i zachowania podmiotu. Uświadamianie sobie czegoś jest jednym z takich skutków, czego rezultatem jest strumień świadomych zdarzeń, które nie muszą zachodzić w tym samym porządku czasowym, w jakim zachodzą zdarzenia obiektywne. Subiektywne następstwo zdarzeń może być inne aniżeli następstwo obiektywne. Czas przeżywany subiektywnie nie musi się pokrywać z obiektywnym upływem czasu²⁹.

Z przyjęcia Modelu Wielu Szkiców wynika, według Dennetta i Kinsbourne’a, że gdyby ktoś chciał uznać jakiś czasowy moment przetwarzania zachodzącego w mózgu za moment już świadomy, to będzie to dowolna (arbitralna) decyzja (por. Dennett i Kinsbourne 1992a: 194). Możemy zawsze „pociągnąć linię” w zachodzącym w mózgu strumieniu przetwarzań. Nie ma jednak funkcjonalnej różnicy, która mogłaby nas skłaniać do zdeklarowania wszystkich wcześniejszych stanów i poprawek (*revisions*) jako postaw nieświadomych czy przedświadomych, zaś wszystkich kolejno następujących poprawek treści jako poprzezyciowego zakłócenia pamięci (*memory-contamination*). Dennett mówi tym samym o braku jasnej linii oddzielającej świadomość od nieświadomości.

Powiedzmy, że przyjmiemy: Bycie świadomym jest jakąś wewnętrzną cechą stanów mentalnych (na przykład jesteśmy bezpośrednio świadomi danego stanu, to znaczy, że zgodnie z Kartezjuszem, nic nie mediuje pomiędzy stanem mentalnym a naszą jego świadomością). Wtedy czymś naturalnym byłoby ustalić moment czasowy, w którym dany stan mentalny się pojawia.

²⁹ „The temporal order of subjective events is a product of the brain’s interpretational processes, not a direct reflection of events making up those processes” (Dennett i Kinsbourne 1992a: 183).

Można to uczynić odwołując się do momentu, w którym stanie się on świadomy. W takim modelu (w przeciwieństwie do modelu dennettowskiego) nie ma miejsca dla mózgu, aby reinterpretować albo rewidować sposób, w jaki ja jestem świadomy moich stanów mentalnych, ponieważ moje bycie świadomym tych stanów jest ich wewnętrzną cechą. Wtedy jest także sens wprowadzić wspomnianą linię podziału, której istnienie Dennett neguje.

W tym miejscu można jednak zapytać Dennetta (por. wyżej, [A]), czy rozplynięcie się granicy jest faktycznie dobrym argumentem przeciwko jej istnieniu? Można powiedzieć tak: dopóki nie znajdziemy „komórek holistycznej świadomości”, owych „specyficznych neuronów” [$\Phi_j - \Phi_n$] pobudzanych wtedy i tylko wtedy, gdy podmiot jest świadomy, tak długo pozostaną niejasności co do „gdzie i kiedy” zachodzenia świadomych procesów. Jeżeli jednak procesy świadome superwenują na procesach fizycznych, wtedy mimo wszystko jest jakieś „gdzie i kiedy” elementów świadomości: wszystkie procesy fizyczne rozgrywają się bowiem w czasie i przestrzeni.

4.4 Jedność świadomości

Wychodząc z założenia, że od strony przeżyciowej świadomość cechuje się jednością, a strumień świadomości jest czymś pojedynczym można zadać dwa pytania. Pierwsze brzmi: czy odrzucenie przez Dennetta anatomicznej wersji Kartezjańskiego Teatru oznacza również odrzucenie tezy o jedności świadomości? Odpowiadając na to pytanie J. Aronson zaznacza, że może istnieć funkcjonalna jedność świadomości, choć materialne podłoże świadomości może być anatomicznie poroździelane i operować na różnych poziomach oraz w różnych momentach czasu (por. Aronson 1992: 202). Można być na przykład przekonanym, że subiektywne własności zdarzeń nie są zdeterminowane gdzieś w mózgu przez obiektywne własności zdarzeń i równocześnie być przekonanym, że „wszystko gdzieś się razem schodzi” – mówiąc funkcjonalnie. Zdaniem Aronsona, można być przekonanym, że treści są ułożone czasowo (*stamped*), że towarzyszą im ich czasowe przedstawienia, nie rezygnując przy tym z myśli, że świadomość cechuje się jednością.

Drugie pytanie brzmi: czy z Modelu Wielu Szkiców nie wynika, że w każdym podmiocie mogą istnieć różnorodne świadome jaźnie (jedna niezależna od drugiej)? Należy stwierdzić, że dzisiejsze badania naukowe nie dostarczają żadnej logicznej racji, dla której mielibyśmy zaakceptować ten wniosek. Poza tym przeżycia poszczególnych podmiotów nie tylko nie popierają tego wniosku, ale co więcej, wniosek ten może być rozumiany jako swoista wersja Kartezjańskiego Teatru: zamiast jednego teatru istnieje kilka teatrów.

Nawiązując do powyższych pytań, przedstawimy w kolejnych dwóch punktach dwie krytyki Modelu Wielu Szkiców. W pierwszej z nich zarzuca mu się, że nie ujmuje potocznie rozumianej jedności świadomości (por. 4.4.1), w drugiej – że nie uwzględnia występowania świadomości zjawiskowej (4.4.2).

4.4.1 Jedność świadomości w ramach Funkcjonalnego Teatru

Wychodząc od proponowanego przez J. Aronsona modelu Funkcjonalnego Teatru możemy zarazem przejść dennettowski Model Wielu Szkiców oraz twierdzenia psychologii potocznej dotyczące jedności świadomości. Proponowany tym samym model świadomości nie wymaga dla niej pojedynczego *locum*, dopuszcza jednak pojedynczy strumień świadomości, który cechuje się jednością, mimo że jest stale przetwarzany.

Pytanie o wyjaśnienie potocznej koherencji i jedność świadomych przeżyć jest dla nas pytaniem o holistyczną jedność świadomości. Zgodnie z Aronsonem przyjmujemy, że:

IV.3 Obserwowaną potocznie jedność świadomości łatwiej jest wyjaśnić poprzez przyjęcie istnienia funkcjonalnego teatru. Jeżeli Dennett chce wyjaśnić świadomość rozumianą jako rodzaj cechującego się jednością procesu, to musi umieć ponadto wyjaśnić, w jaki sposób postulowane przez niego równoległe i przeciwstawne procesy schodzą się razem.

Samo mówienie o *time stamping*, jak to czynią Dennett i Kinsbourne, nie jest wystarczającym wyjaśnieniem jedności świadomości. Jeżeli zaś Dennett i Kinsbourne negują, że świadomość cechuje się jednością, to nie wyjaśnili świadomości, tylko ją po prostu wyeliminowali.

Poszukiwana teoria świadomości musi się zająć pytaniami o relację pomiędzy holistyczną jednością świadomości, jednością przeżycia, a poszczególnymi komponentami tego przeżycia – identyfikowalnymi kolorami, dźwiękami. Nasz świat potoczny jawi się nam jako swoiste *continuum*, chociaż procesy spostrzegania, na których obraz tego świata jest zbudowany, są pełne nieciągłości. To, w jaki sposób różnorodne sensoryczne cechy spostrzegalnych obiektów są analizowane przez mózg, tak że są one komponowane w cechujący się jednością przedmiot spostrzegania, nazywamy „problemem scalania” (*binding problem*)³⁰: w jaki sposób, na przykład, przedsta-

³⁰ „The problem of binding together in a single unified percept the different features of an object, which are represented in different locations in the brain” (Rees 2002: 266).

wienia kształtu obiektu i przedstawienia jego koloru zostają połączone razem w przedstawienie obiektu posiadającego kolor? Jak aktywność wewnątrz autonomicznie odseparowanego modułowego systemu, odpowiadająca różnym atrybutom bodźca wzrokowego (kolor, ruch, postać, głębia), działa w świadomym wzrokowym spostrzeganiu cechującego się jednością obiektu albo sceny? Problem scalania należy do centralnych problemów świadomości.

Jedną z odpowiedzi może być, że tego rodzaju wiązanie jest możliwe dzięki ciągłości samoświadomości. Jednak według Dennetta samoświadomość nie jest jakimś *continuum*, pomimo iż dostarcza ona wrażenia, że sama takim *continuum* jest.

Z jednej strony Dennett jest kognitywistą, z drugiej strony jest on – gdy chodzi o świadome przeżycia – neurofizjologicznym realistą³¹. Fikcyjne jaźnie i przeżycia są konstruowane na podstawie sądów. Kto jednak wypowiada owe sądy, które konstruują jaźń i przeżyciowy świat? Dennett odpowiada: nikt. To, co faktycznie istnieje, to różne zdarzenia ustalenia treści (*content-fixation*), zachodzące w mózgu w różnych miejscach i w różnych czasach (por. Dennett 1991: 365). Dennett dodaje, że nikt (rozumiany w sensie potocznego podmiotu) nie dokonuje czynności mowy (*speech acts*), i dlatego są one raczej podobne do czynności mowy: posiadają treść, przekazują informacje. Niektóre z tych ustaleń treści prowadzą do wypowiedzania zdań w języku (publicznie albo tylko wewnętrznie). W ten sposób powstaje heterofenomenologiczny tekst. W czasie jego interpretacji rodzi się iluzja, że istnieje jakiś autor tego tekstu.

Powtórzmy. Zdaniem Dennetta, co faktycznie istnieje, to różne neurologiczne zdarzenia prowadzące do ustalania treści, nie istnieje natomiast coś takiego, jak holistycznie rozumiana świadomość. Jednak to, co w dennettowskim modelu zalicza się do treści świadomości, nie jest dokładnie określone. Taka nieokreśloność może być korzystna, zwłaszcza gdy uzmysłowimy sobie fakt, że bez teorii świadomości nie mamy „jasnej i klarownej” idei, czym jest treść świadomości. To, co aktualnie przeżywamy, jest produktem wielu interpretacyjnych procesów edytorskich. Dennett nie mówi nam, co rozumie przez „ustalenia treści” świadomości. Jeżeli bowiem wypowiada się on w terminach procesów neuronalnych, wtedy odnosi się do „problemu scalania”.

Sposoby scalania cech podstawowych w jednoznaczną (i niejednoznaczną) reprezentację bodźca omawia A. Wróbel (2000a: s. 462–471).

³¹ „Świadome przeżycia są realnymi zdarzeniami zachodzącymi w realnym czasie i w realnej przestrzeni mózgu, i dlatego są one uchwytnie czasowo i przestrzennie wewnątrz odpowiednich granic dokładności dla realnych zjawisk ich rodzaju” (Dennett i Kinsbourne 1992a: 235).

Jeżeli zaś mówi o rzekomych przeżyciach pierwszoosobowych skorelowanych z neuronalną dynamiką, wtedy odnosi się do przeżywalnej jedności świadomości. Żaden z tych problemów nie zostaje rozwiązany przez odwołanie się do „ustalenia treści”. Co więcej, mowa o ustaleniu treści na nowo prowadzi do pytania o istnienie zombi, w których takiego ustalenia brak.

Podobnie do J. Aronsona argumentuje neurolog A. Damasio, którego zdaniem odrzucenie tezy o jednym jedynym biologicznym Kartezjańskim Teatrze nie musi oznaczać odrzucenia tezy o doznającej przeżyć jaźni. Istnieją bowiem neuronalne systemy, na podstawie których konstruujemy fałszywą intuicję, że w mózgu jest jedno miejsce, gdzie pojawiają się przeżycia³². Poszukiwany model świadomości – według Damasio – powinien uwzględnić sposób, w jaki owe zdeintegrowane, powstające w mózgu fragmenty, wytwarzają poczucie zintegrowanej jaźni. Dennettowskiemu modelowi (zdanem Damasio) brakuje dwóch koniecznych funkcji, bez których trudno zrozumieć, jak może powstać subiektywna świadomość iluzorycznych fenomenów. Pierwszą funkcją jest zdolność do stałego uaktualniania (*sustained updating*) istotnych kroków poznania, odnosząca się nie tylko do przeszłości, ale także do bezpośrednio zamierzonych czynności oraz do dalszych planów podmiotu. Drugą funkcją jest ciągłe monitorowanie somatycznych stanów podmiotu przeżyć. Z jednej strony, według Damasio, „[...] pozostawiony sobie mechanizm Wielu Szkiców będzie wytwarzał pozbawioną jaźni, beczcielną świadomość”.

Z drugiej strony Damasio obstaje przy tym, że zadowalający model świadomości powinien: „[...] brać pod uwagę to, w jaki sposób owe zdeintegrowane fragmenty operują wytwarzając zintegrowaną jaźń”³³.

³² Por. Damasio 1992: 208. Za występowaniem odznaczającego się jednością strumienia świadomości opowiada się także B. Baars. „However, recent single-cell studies by Sheinberg and Logothetis [...] suggest strong convergence of conscious visual object information in inferotemporal cortex and the superior temporal sulcus in the macaque. Ninety percent of visual neurons in these areas respond differentially to the conscious but not the unconscious visual flow in a binocular rivalry task. Since this area integrates many visual features into object representations, it may indeed be a place where conscious visual information comes together” (Baars 1998: 3).

³³ Damasio 1992: 208. Według Damasio, w niektórych przypadkach jest możliwa świadomość pozbawiona jaźni, czego przykładem może być anozognozja [uszkodzenie proto-ja (tzn. systemu wzajemnie połączonych neuronalnych wzorów), co rzutuje na funkcjonowanie Ja-autobiograficznego (opartego o działanie autobiograficznej pamięci), por. Wiener 2000: 6–12]. Sugestie Damasio dotyczą pytania o dochodzenie do cechującego się jednością sensu jaźni. W *Time and the Observer* Dennett nie zajmuje się tym pytaniem. Po linii Damasio argumentuje M. Velmans pytając Dennetta o zagadnienie integracji świadomości (por. Velmans 1992: 230).

W dalszej argumentacji wychodzimy od twierdzenia zakładającego jedność świadomości. Twierdzenia tego nie da się uzasadnić przez odwołanie się do funkcjonalnego teatru, niosącego ze sobą wszystkie trudności funkcjonalizmu. Zamiast tego przyjmujemy emergencyjną zależność pomiędzy: (a) stanami neurologicznymi a stanami zjawiskowymi, oraz (b) neurologiczną a filozoficzno-psychologiczną teorią świadomości (por. 7.6.1). Wstępnym krokiem do ustalenia i uzasadnienia tak rozumianych zależności jest określenie metody poszukiwania neuronalnych korelatów świadomości.

4.4.2. Wielość szkiców a świadomość zjawiskowa

Postawiona przez Dennetta teza o nieodróżnialności modelu orwellowskiego od stalinowskiego jest, według N. Blocka, prawdziwa tylko przy założeniu, że odróżnimy świadomość zjawiskową od świadomości dostępu. Block jest funkcjonalistą (wewnętrzne przedstawienia w naszych głowach odgrywają określone funkcje przy podejmowaniu decyzji, przy zastanawianiu się itp.) i naukowym realistą (przynajmniej niektóre teoretyczne terminy denotują istniejące obiekty, których nie możemy bezpośrednio obserwować).

Zdaniem Blocka „świadomość” jest terminem odnoszącym się do różnych rodzajów cech mentalnych. Block zajmuje się bliżej wspomnianymi już dwoma typami świadomości: świadomością zjawiskową i świadomością dostępu³⁴. Świadomość zjawiskowa nie jest opisywalna w terminach funkcjonalnych czy intencjonalnych, to znaczy nie można jej wyjaśniająco zredukować do świadomości dostępu³⁵. Trudno jest nam wyjaśnić lub opisać, *jak to jest* doświadczać bólu, podczas gdy równocześnie jakoś „wiemy”, *jak to jest* go doświadczać. Cechą świadomości dostępu jest (w przypadku istot używających języka) możliwość relacjonowania o swoich stanach. Obydwa typy świadomości przenikają się, czego mamy chociażby przykład, gdy odczuwamy ból (potrafimy zdać sprawę z tego przeżycia wskazując na bolące miejsce, opisując ból itd.).

Dennett stwierdza, że – ponieważ spostrzeżenie jakiegoś zdarzenia jest rozsiane po przestrzeni mózgu, a tym samym jest rozsiane w czasie – zaety-

³⁴ Podział na tego rodzaju świadomości jest kwestionowany przez D. M. Rosenthala. „There can be no doubt that Block’s two notions correspond to a robust distinction in commonsense between two distinct types of mental occurrence, a distinction to which any satisfactory theory must do justice. It is far less obvious, however, exactly how this commonsense distinction should be understood” (Rosenthal 1998: 2).

³⁵ Świadomość dostępu Block definiuje jako: „[...] a state with a certain content is access-conscious if, in virtue of one’s having the state, a representation which has that content is (1) poised to be used freely as a premise in reasoning, according to the capabilities of the reasoner, (2) poised to be used freely for control of action” (Block 1994: 26).

kietowanie jednego z tych stanów jako świadomy jest sprawą arbitralną. Odwołując się do powyższego twierdzenia Block rozumie swoją argumentację za świadomością zjawiskową jako odpowiedź na pytanie: „co jednak będzie, gdy niektóre z mózgowych przedstawień jakiegoś zdarzenia tworzących ową przestrzenno-czasową objętość, są świadome w sensie zjawiskowym, podczas gdy inne nie są?” (Block 1992: 205).

Odpowiadając na to pytanie przyjmujemy krytyczną w stosunku do Dennetta hipotezę:

IV.4 Przypisanie świadomości zjawiskowej do jednego lub drugiego przedstawienia w mózgu nie jest czymś arbitralnym. Istnienie świadomości zjawiskowej musi być odzwierciedlane na poziomie neuronalnym.

Może istnieją stwierdzalne empiryczne różnice pomiędzy tymi przedstawieniami mózgu, które są przeżyciowo świadome i tymi, które takimi nie są? Nie musimy przy tym wychodzić z założenia, że w mózgu istnieje jedno osobliwe miejsce, w którym występuje świadomość zjawiskowa. Być może przeżyciowo świadome zdarzenia zachodzą w różnych regionach mózgu (por. rozdział VII). Twierdzenie IV.4 nie opowiada się za przyjęciem Kartezjańskiego Materializmu, a jedynie za istnieniem świadomości zjawiskowej.

Należy dodać, że, zdaniem Blocka, w przeciwieństwie do Dennetta nasze dojście do świadomości zjawiskowej pochodzi z naszej z nią znajomości (*acquaintance with it*). Argumentacja Blocka zakłada tym samym możliwość rozróżnienia pomiędzy świadomością zjawiskową a niezjawiskową³⁶.

Dennett i Kinsbourne powtarzają, że nic, co może być spostrzegalne zarówno dla: (a) obserwatora wewnętrznego, jak i dla (b) obserwatora zewnętrznego nie potrafi doprowadzić do odróżnienia opcji stalinowskiej od orwellowskiej. Jeżeli jednak świadomość zjawiskowa – jak tego chce Block – istnieje, to zarówno rozumowania Dennetta i Kinsbourne’a dotyczące obserwatora wewnętrznego, jak i zewnętrznego są błędne.

(a) Block zaczyna od zagadnień związanych z obserwatorem wewnętrznym. W modelu orwellowskim mamy zjawiskowe przeżycie A, w modelu

³⁶ Z podobnego założenia przy komentowaniu tez Dennetta i Kinsbourne’a wychodzi R. van Gulick porównując swój Global Integrative Phenomenal State (GIPS) Model z Dennettowym Modelem Multiple Drafts (MD). Van Gulick zaznacza: „But unlike the MD model, it [GIPS – J.B.] takes seriously the distinction between phenomenal and non-phenomenal representation [...]. On the GIPS model, questions about a person’s state of consciousness at a given moment are questions about the content of her phenomenal representations” (Van Gulick 1992: 229).

stalinowskim nie mamy takiego przeżycia. Różnica pomiędzy tego rodzaju przeżyciem a jego brakiem jest różnicą wewnętrzną (jeżeli w ogóle jakąś jest). Mówiąc o czymś rozróżnialnym wewnątrznie (*discernible inside*) Dennett i Kinsbourne zdają się mieć na myśli coś rozróżnialnego w sądach wypowiedzianych przez podmiot. Według Blocka założenie, że to, co subiektywne, wyczerpuje się w sądach, jest przejściem obok pytania o świadomość zjawiskową, gdyż ta nie jest tylko sprawą sądów. Chociaż, według Blocka, wielu filozofów jest przekonanych o istnieniu świadomości zjawiskowej, niektórzy twierdzą, że można ją utożsamić z jakimś stanem funkcjonalnym albo intencjonalnym (por. Block 1994: 26). Dennett i Kinsbourne twierdzą, że obydwa modele: orwellowski i stalinowski, „[...] wyjaśniają dane subiektywne – cokolwiek jest dostępne «z perspektywy pierwszej osoby» – ponieważ zgadzają się one co do tego, jak coś powinno być odczuwane przez podmioty”³⁷.

Według Blocka, pomiędzy obydwooma modelami nie ma zgodności co do tego, jak jakaś rzecz jest odczuwana przez podmiot. Model orwellowski akceptuje zjawiskowo-świadome przeżycie A. Model stalinowski odrzuca zaś takie przeżycie. Block konkluduje, że obydwa modele zgadzają się jedynie co do sądów wydawanych przez podmiot, nie zgadzają się natomiast, gdy chodzi o to, co podmiot odczuwa.

Dennett i Kinsbourne chcą wykazać, iż odwołując się do introspekcji jako źródła poznania nie potrafimy rozróżnić pomiędzy stalinowskim a orwellowskim modelami naszych własnych przeżyć. Block nie kwestionuje tego założenia, ponieważ jest przekonany, że nasz uprzywilejowany dostęp, taki jaki jest, obejmuje nasze przeżycia, ale nie czyni nieomylną naszej pamięci o nich. Jeżeli chcemy rozstrzygnąć pomiędzy owymi modelami, musimy się odwołać do badań naukowych, to znaczy do obserwatora zewnętrznego.

(b) Jeżeli obecność albo nieobecność świadomości zjawiskowej jest realną własnością, to odpowiednio przemyślane eksperymenty będą mogły wykazać poszukiwaną różnicę. Block zauważa, że modele stalinowski i orwellowski mogą się zgodzić co do tego, gdzie błędna treść wchodzi w łańcuch przyczynowy. Nie zgadzają się one jedynie co do tego, czy treść ta jest prze-

³⁷ Pytaniem jest: Czy owe dwa – jak się wydaje odmienne – modele pozostają odmienne, gdy w grę wchodzi nie sekundy, lecz milisekundy. Dennett odpowiada: „Nie. W takim przypadku rozróżnienie pomiędzy przetwarzaniem spotrzezań a przetwarzaniem pamięci, które funkcjonuje skutecznie w innych skalach, nie ma na dalszą metę sensu. Weszliśmy bowiem w zamgloną przestrzeń, w której podmiotowy punkt widzenia jest zamazany przestrzennie i czasowo, a pytanie o model orwellowski lub stalinowski traci swój sens” (Dennett 1991: 119; por. Rockwell 1996: 468).

zyciowo świadoma czy też nie. Znana konkluzja Dennetta i Kinsbourne'a – gdy chodzi o różnicę pomiędzy tymi modelami – brzmi: „Jest to różnica, która nie powoduje żadnej różnicy”. O braku różnicy możemy mówić, według Blocka, tylko wtedy, gdy założymy, że nie istnieją rzeczywiste fakty świadczące o świadomości zjawiskowej. Jednak cała dennettowska argumentacja ani takiego założenia nie zawiera, ani go nie uzasadnia.

4.5. Kulturowa ewolucja świadomości – darwinizm społeczny

„All three media – genetic evolution, phenotypic plasticity, and memetic evolution – have contributed to the design of human consciousness, each in turn, and at increasing rates of speed” (Dennett 1991: 208).

Block wykazuje, że ani świadomość zjawiskowa, ani świadomość dostępu nie są konstruktem kulturowym. Dennett broni tezy przeciwnej i zauważa, że wprowadzony przez Blocka podział świadomości prowadzi do błędnego rozumienia całości świadomości, która jest konstruktem kulturowym. Block twierdzi natomiast, że świadomość dostępu jest podstawową biologiczną cechą osoby, dlatego nie może być takim konstruktem. Blok nie twierdzi jednak, że świadomość ta jest już obecna w chwili urodzenia się podmiotu.

Block pyta, czy świadomość zjawiskowa (zbliżona do tego, co Dennett rozumie pod terminem „świadomość”) jest konstrukcją kulturową. W odpowiedzi zaznacza, że musimy odróżniać pomiędzy poglądem, że kultura oddziałuje na świadomość zjawiskową, a poglądem, że świadomość zjawiskowa jako całość jest konstrukcją kulturową. Kultura ma także wpływ na ludzkie stopy, a mimo tego stopy nie są konstrukcją kulturową.

Wychodząc z wprowadzonego przez Blocka rozróżnienia na dwa rodzaje świadomości, stwierdzamy, gdy chodzi o genezę świadomości zjawiskowej:

IV.5 Powstanie świadomości zjawiskowej zawdzięczamy jakościowym przeżyciom, a nie temu, co o tych przeżyciach mówi nam kultura.

Dennettowski Model Wielu Szkiców jest teorią świadomości dostępu. Nasze pojęcie konkurencji mózgowej (*cerebral celebrity*) jest konstrukcją kulturową, ale sama konkurencja w mózgu nią nie jest³⁸. Block konkluduje: ponie-

³⁸ Block cytuje Denneta: „Consciousness is cerebral celebrity – nothing more and nothing less. Those contents are conscious that persevere, that monopolize resources long enough to achieve certain typical and «symptomatic» effects – on memory, on the control of behavior, and so fort” (Block 1995a: 238).

waż ani świadomość dostępu, ani świadomość zjawiskowa (osobno wzięte) nie są konstruktami kulturowymi, tym samym nie jest nimi świadomość jako taka.

Dennett zgadza się z Blockiem, że świadomość jest czymś złożonym, nie przyjmuje jednak jego podziału na dwa typy świadomości. Na świadomość należy patrzeć w całości, wtedy widać, że jest ona konstruktem kulturowym.

Ze zbudowanego przez Dennetta szkicu ewolucyjnego powstania świadomości wynika, że nie ma istotnych różnic pomiędzy mózgiem ludzi pierwotnych a naszym mózgiem. Rozwój gatunku *Homo sapiens* w ostatnich kilkunastu tysiącleciach zawdzięczamy plastyczności mózgu, który wytworzył coś podobnego do *software* udoskonalając tym samym swoje podstawowe zdolności. Nasi przodkowie nie byli świadomi w takim sensie, w jakim my jesteśmy (por. Dennett 1992: 5; 1991: 190). Dennett opowiada historię przypominającą opowiadanie W. Sellarsa o „naszych Ryle’ańskich przodkach” (por. Sellars 1963b: 178–180; Bremer 1997: 72–74), którzy w różnych okolicznościach zadawali pytania, nie kierując ich jednak do konkretnej osoby. Kiedyś jeden z nich zadał pytanie, ale nie było nikogo, kto by je usłyszał, więc sam sobie na nie odpowiedział, ustanawiając tym samym (i nie zauważając tego) komunikację pomiędzy częściami swojego mózgu, pomiędzy którymi z jakichś biologicznych względów istniał problem wzajemnego dostępu. Wystarczyło to, według Dennetta, do ustanowienia ewolucyjnego uzupełnienia podmiotu, które może być przekazywane kulturowo. Potocznie zachowanie takie nazywamy „mówieniem do samego siebie”.

Nasi przodkowie nauczyli się paru dobrych sztuczek, a jedną z nich jest wspomniana poznawcza autostymulacja: zadanie pytania i odpowiedź na nie (por. Dennett 1991: 196). W ramach tejże autostymulacji postuluje się istnienie sprzężenia zwrotnego, w którym różne formy naszego ludzkiego wyrażania się są zaangażowane w stymulowanie sensorycznych powierzchni. Zachodzi to w taki sposób, iż tworzą się wzorce neuronalnych aktywacji, skojarzonych z podwyższoną poznawczą organizacją. Tego rodzaju autostymulacja zaczęła się następnie stabilizować w ramach ludzkiej populacji. Część owych autostymulacji realizuje się poprzez używanie języka.

Pojęciową podstawę dla tak rozumianej autostymulacji Dennett znajduje w rozwiniętej przez zoologa R. Dawkinsa idei „memes”³⁹. Przykładami memów – jako jednostek kulturowego przekazu – są melodie, slogany, sposoby lepienia garnków, wynalezienie koła czy kalendarza. Zdaniem Dawkinsa:

³⁹ Dennett cytuje z tekstów Dawkinsa (1976: 206; 1982: 110–111) i pisze: „Once our brains have built the entrance and exit pathways for the vehicles of language, they swiftly become parasitized (and I mean that literally, as we shall see) by entities that have evolved to thrive in just such a niche: memes” (Dennett 1991: 200).

„Tak jak geny rozprzestrzeniają się same w sadzawce genów przez przekazywanie z ciała do ciała przez spermę lub jaja, tak też memy rozprzestrzeniają się same w sadzawce memów przeskakując z mózgu do mózgu poprzez proces, który, w szerszym sensie, może być nazwany naśladownictwem” (za: Dennett 1991: 202).

Memy są strukturami językowymi, które „pasożytują” na ludzkich mózgach. Są one postulowanymi, atomowymi jednostkami kultury, podobnie jak geny są atomowymi jednostkami biologicznego dziedzictwa. Geny są niewidzialne dla oka, są przekazywane poprzez organizmy, w których wywołują określone efekty (efekty „fenotypowe”). Memy są także niewidzialne i są przekazywane przez malowane obrazy, książki, opowiadania. Memy, podobnie jak geny, są potencjalnie nieśmiertelne. Podobnie jak w przypadku genów, istnienie memów zależy od łańcucha fizycznych mediów (na przykład przechowanie Platońskich memów poprzez serie kolejnych kopii jego dzieł). Przyporządkowanie memów do struktur językowych zapobiega traktowaniu ich jako coś urzeczowionego⁴⁰.

Występowanie świadomości jest dla Dennetta wynikiem procesu ewolucji kulturowej. Pomijając jego „luźne” opisy, należy dodać, że ewolucja nie wyjaśnia, jak świadomość się rozwinęła, może za wyjątkiem tego, że posiada określoną funkcję gdy chodzi o przeżycie organizmu i dlatego była faworyzowana przez naturalną selekcję. W przeciwnym razie świadomość nie rozwinęłaby się, aby pełnić tak centralną rolę w naszym życiu.

4.6. Komputerowo-funkcjonalistyczna struktura świadomego umysłu

Odwołanie się do idei memów służy Dennettowi do określenia umysłu i świadomości: ludzki umysł jest sztucznym tworem powstałym wtedy, gdy memy tak zrestrukturyzowały ludzki mózg, iż stał się on odpowiedniejszym miejscem dla ich pobytu.

„[...] ludzka świadomość jest sama jednym wielkim kompleksem memów (albo bardziej dokładnie kompleksem efektów działania memów w mózgach), który może być najlepiej rozumiany jako operacja wirtualnej maszyny «von Neumannesque» zrealizowanej w równoległej architekturze mózgu, który nie był przewidziany do takich czynności” (Dennett 1991: 207, 210).

W powyższym cytacie dochodzi do głosu analogia pomiędzy umysłem a komputerem. Przyjmijmy, że J. von Neumann wykorzystał abstrakcyjne

⁴⁰ Wheelwell krytykuje ideę memów jako próbę wprowadzenia niefizycznych, atomicznych komponentów do badanych przedmiotów fizycznych (por. Wheelwell 1998: 362, 370).

idee A. Turinga do zbudowania cyfrowego komputera, nazywanego maszyną von Neumanna. Dennett utrzymuje, że świadomość jest wirtualną symulacją sekwencyjnej maszyny, działającej na rzeczywistej maszynie równoległej. Wirtualna maszyna von Neumanna (stąd nazwa *von Neumannesque*) wywołuje strumień świadomości, który Dennett nazywa także „maszyną Joyce’a”⁴¹ i odwołuje się do niej przy wyjaśnianiu zjawiska świadomości. Maszyna Joyce’a tworzy pozorne zjawisko pojedynczego strumienia świadomości, jako występujące poza wieloma szkicami.

Zamiast jednego pojedynczego strumienia świadomości mamy w Modele Wielu Szkiców różne strumienie odpowiedzialne za poszczególne funkcje świadomości. Wszystkie one składają się na sztuczne urządzenie, powstałe na drodze wspólnego działania różnych modułów funkcjonalnych. Niektóre z tych modułów dzielimy razem ze zwierzętami, inne (dzięki zdolnościom językowym oraz rozprzestrzenianiu i przekazywaniu memów), są charakterystyczne tylko dla nas ludzi. Założeniem, z którego wychodzi Dennett jest określenie celu używania świadomości, stąd jego badanie mechanizmów prowadzących do tego celu.

Przykład takiego mechanizmu Dennett znajduje w programie badawczym modelowania zjawisk psychologicznych, rozwiniętym przez neurokognitywistę D. Marra (por. Dennett 1991: 276–277). Marr proponuje trzy stopnie analizy. Stopień najwyższy i najbardziej abstrakcyjny tworzy komputacyjna analiza problemu przetwarzania informacji, stopniem środkowym jest algorytmiczna analiza procesów umożliwiających przetwarzanie informacji, stopniem najniższym (fizykalnym) jest sieć neuronalna, gwarantująca materialne podłoże procesu algorytmicznego i komputacyjnego. Podane przez Marra stopnie mogą być także używane do analiz prostszych systemów aniżeli umysł. Dennett podkreśla jednak, że skupienie się na procesie przetwarzania informacji, jako na celu działania świadomości, jest obarczone ryzykiem uproszczenia (por. Dennett 1991: 277). Niektóre cechy świadomości posiadają wielorakie funkcje, niektóre funkcje są jedynie znikomo określane przez istniejące cechy świadomości.

Maszyna Joyce’a jest odpowiedzialna za ustalenie połączeń. Większość prostych połączeń może być ustalona rutynowo (nieświadomie). Natomiast

⁴¹ Chodzi o maszynę (*Joycean machine*) zdolną do generowania kompleksowego tekstu o „strumieniu świadomości” według podobnego wzoru, jakiego J. Joyce używa w *Ulyssesie*. Umysł przedstawiony w *Ulyssesie*, jest wynikiem równoczesnego występowania świadomości, pamięci, podświadomości, przypadkowych skojarzeń, do których dochodzą lęki i oczekiwania. Jedna wielka chaotyczna wędrówka zjawisk przewija się przez całe życie podmiotu, w jego śnie i na jawie.

znajdywanie trudnych lub dla podmiotu nieprzyjemnych powiązań wymaga „skupienia się na nich”, stosowania mnemotechniki czy innych technik (na przykład zapisywania). Są to strategie pozwalające używać naszych procesów percepcyjnych w nowy sposób. Dzięki tym strategiom powstaje myślenie hipotetyczne, zdolność do realizacji długoterminowych (choć może chwilowo nieprzyjemnych) projektów w oczekiwaniu na przyszłe korzyści, poszerzenie pamięci długotrwałej o wspomnienia operacji przeprowadzonych w umyśle, tak iż możemy sami sobie wiele rzeczy wyjaśnić. Antycypowana lub przypominana treść – która nie musi być koniecznie i zawsze świadoma (Dennett 1991: 278) – niesie z sobą kontekst, który z kolei czyni tę treść sensowną dla podmiotu. Podmiot może tę treść „świadomie uchwycić” i nadać jej znaczenie. Oznacza to, że możliwa jest świadoma i nieświadoma aktywacja różnych treści, a tym samym różnych połączeń w mózgu.

Zbudowana przez Dennetta koncepcja umysłu jest funkcjonalistyczna, co ze względu na tezę o wielorakiej realizacji oznacza, że w sensie analogicznym posiadanie umysłu możemy przypisać także komputerom i robotom. Podmiot jest przy tym traktowany jako zespół podsystemów, dlatego możemy Dennetta teorię nazwać teorią stopnia subpodmiotowego⁴². Suma skoordynowanych zachowań tych subsystemów dostarcza wyjaśnienia zachowania podmiotu. Tak rozumianą koncepcję umysłu widać w Dennetta ujęciu qualiów. Rozstrzygające dla tego, co nazywa on qualiami nie jest, zdaniem O. Flanagan, fizjologiczna płaszczyzna, lecz pewna płaszczyzna subosobowa i funkcjonalna, którą można usytuować ponad fizykalno-fizjologicznym a pod osobowym sposobem patrzenia⁴³. Przeżywany ból i struktury w mózgu nie mogą zostać ze sobą utożsamione. Nie dlatego, że ból ma jakieś ontologicznie nieredukowalne cechy, ale dlatego, że w mózgu nie ma żadnego centrum bólu. Ból zachodzi w ekstrynsycznych relacjach pomiędzy wejściem, stanem wewnętrznym a wyjściem określonego układu funkcjonalnego. Ból nie jest jakąś wewnętrzną, nieredukowalną jakością, lecz scharakteryzowanym przez wewnętrzne relacje stanem, który transformuje określoną informację do systemu nerwowego.

Przyjmując funkcjonalistyczną teorię umysłu Dennett stara się tym samym utworzyć podmiotową jaźń z podsystemów („functional areas”) dostęp-

⁴² Dennett mówi o „sub-personal cognitive psychology” (por. Dennett 1987: 235).

⁴³ O. Flanagan odnosząc się do proponowanego przez Dennetta ujęcia qualiów zauważa: „Such concepts are neither physical nor phenomenological, since they are not typed by appeals to the neural substrata nor by the way things seem” (Flanagan 1992: 127). Sam Dennett odróżnia jedynie osobową (intencjonalną) i subosobową (*level of brains and events in the nervous system*) płaszczyznę wyjaśniania (Dennett 1969: 93, 179; por. s. 148, 156).

nych jedynie od strony funkcjonalnej. Oznacza to wyróżnienie podsystemów funkcjonalnych (introspekcja, przechowywanie informacji, analiza danych spostrzeżeniowych, ...). Podział ten prowadzi Dennetta do wyodrębnienia uwagi jako osobnej kategorii⁴⁴, którą jego teoria świadomości musi umieć uwzględnić.

Następującą trudność z funkcjonalistycznie rozumianą teorią Dennetta dostrzega M. Antony: Świadome przeżycia nie odgrywają żadnej roli w funkcjonalnej organizacji mózgu⁴⁵. Nie stoją one bowiem w przyczynowej relacji do stanów i procesów, które są opisywane funkcjonalnie. Gdyby stały, to byłyby one same funkcjonalnymi stanami lub procesami. Wynika stąd, że świadome przeżycia nie mogą być dostępne introspekcji i nie można o nich zdawać sprawy. Nie mogą one być przyczynowane przez sensoryczne wejścia oraz nie mogą odgrywać przyczynowej roli w powstawaniu zachowań i myśli. W rzeczy samej – zdaniem Antony’ego – trudno jest zrozumieć, jak ktoś mógłby myśleć o swoich świadomych przeżyciach, jeżeli jego własne myśli nie byłyby przyczynowo związane z tymi przeżyciami. Stąd Antony wyciąga wniosek, że jedyną interpretacją tez Dennetta i Kinsbourne’a jest interpretacja eliminatywna (świadomość nie istnieje).

Dennett i Kinsbourne odrzucają podaną przez Antony’ego interpretację ich teorii konkludując (por. Dennett i Kinsbourne 1992a: 238): świadome przeżycia nie odgrywają roli w funkcjonalnej organizacji mózgu poprzez to, że spełniają kryterium bycia świadomym. To, że dane zdarzenie stanie się świadomym, nie oznacza, że zostało ono wyposażone w jakies przyczynowe możliwości, których przedtem nie miało. Teza Antony’ego, że świadome przeżycie nie może być identyfikowane ze stanami funkcjonalnymi jest prawdziwa tylko w tym sensie, że świadome przeżycie nie może być identyfikowane jako typ z jakimś szczególnym, funkcjonalnym typem. Dlatego, zdaniem Dennetta i Kinsbourne’a, ich teza nie jest „eliminatywna”. Zdarzenia konstytuujące świadome przeżywanie odgrywają funkcjonalne role, które mogą zachodzić w milisekundach, jednak kiedy i gdzie odgrywają one owe role jest niezależne od tego, kiedy czy też gdzie zdarzenia te wejdą do kręgu świadomych zdarzeń.

⁴⁴ „For clarity I restrict attention to six of the functional areas to which a theory of consciousness must do justice” (Dennett 1998: 156).

⁴⁵ Por. Antony 1992: 202. Do podobnych konkluzji dochodzi M. Velmans.

4.7. Stany zdawania sobie sprawy

Omawiając dennettowski model świadomości D. M. Rosenthal wskazuje na podstawową tezę kartezjanizmu: Bycie świadomym jest tym, co nazywamy „byciem mentalnym” jakiegoś stanu lub jakiegoś przedstawienia (por. Rosenthal 1992: 220). Tym samym kartezjańskie bycie świadomym stanów mentalnych jest bezpośrednią i wewnętrzną cechą tychże stanów, co z kolei ma konsekwencje dla ich subiektywnego czasowego następstwa. Przypuśćmy, że najpierw widzę zdarzenie *a*, a później zdarzenie *b*. Jeżeli bycie świadomym jest wewnętrzną cechą każdego stanu mentalnego, obydwa przypadki widzenia będą świadomymi stanami dokładnie w tym momencie, w którym występują. Równocześnie z widzeniem jakiegoś zdarzenia będę świadomym, że je widzę. Będę świadomym widzenia *a*, a później będę świadomym widzenia *b*. Rosenthal zauważa ponadto, że liczne stany mentalne nie są świadome, co implikuje możliwość rozróżnienia pomiędzy byciem świadomym, a tym, co mentalne w ogóle.

Według Dennetta skierowanie uwagi może być nieświadome, gdy mówimy o nim przykładowo w programach komputerowych. Zadania wykonywane przez komputer na zbiorach danych mogą być traktowane jako „skupianie” na nich uwagi. Uświadamiane nie są także prowadzone przez podmiot analizy danych spostrzegawczych biorące udział w rozpoznawaniu danego przedmiotu (mamy jedynie świadomy dostęp do wyników tych analiz, a nie do samych danych spostrzeżeńowych). Zdaniem Dennetta, aby wyjaśnić problem świadomości, musimy odwołać się do świadomej uwagi. Gdy zgodnie z Rosenthalem odrzucimy Kartezjańską ideę, że wszystkie zjawiska mentalne są bezpośrednio obecne w naszej przestrzeni mentalnej, to pojawia się inne, równie trudne pytanie, na które wskazuje T. Rockwell: jeżeli bowiem naszemu mózgowi udało się w końcu utworzyć kształty, kolory, dźwięki, to w jaki sposób jest on w stanie zdać sobie sprawę (*to become aware of*) z niektórych z tych mentalnych zjawisk i ignorować wszystkie inne? (por. Rockwell 1996: 465). Dlaczego zdajemy sobie sprawę tylko z niektórych z nich i w jaki sposób „przeskakujemy” z jednego świadomego zjawiska na drugie?

Udzielając na te pytania odpowiedzi Rockwell odwołuje się do Modelu Wielu Szkieł oraz do Rosenthala teorii *higher order thoughts* (HOTs). Przyjmując, że myśli wyższego rzędu nie są w konieczny sposób językowe lub symboliczne, możemy je traktować jako paradygmatyczne akty uwagi (*acts of attention*). Są one mentalnymi aktami specjalnego rodzaju, umożliwiającymi zdawanie sobie sprawy. Według Rockwella niekwestionowanymi zdają się być dwie sprawy. (a) W danym momencie spostrzegam wzrokowo wiele rzeczy,

z których nie zdaję sobie sprawy, ponieważ skupiam się na innych rzeczach obecnych w moim polu wzrokowym, jak również przesuwam moje zdawanie sobie sprawy z tego, co widzę na to, co słyszę, albo co czuję. (b) Procesy spostrzegania zachodzą w nas ciągle, chociaż nie zdajemy sobie z tego sprawy⁴⁶.

Powołując się na wyniki badań neurofizjologicznych Rockwell przyjmuje, że mózg jest równoległym procesorem wytwarzającym dużo więcej zjawisk mentalnych (spostrzeżenia, wrażenia, uczucia) aniżeli te, z których w danym momencie możemy zdać sobie sprawę. Czymkolwiek w naszych mózgach jest owo „coś”, co wyłącza (blokuje) te zjawiska, które są czasowo poza zdawaniem sobie z nich sprawy, spełnia to funkcję, która domaga się wyjaśnienia. Dennett poświęca wprawdzie wiele miejsca tradycyjnym pytaniom o to, jak zjawiska mentalne są kreowane przez mózg, ale nigdzie (według Rockwella) nie zajmuje się pytaniem o „zdawanie sobie sprawy” (*awareness*) jako o coś różnego od „spostregania” (*perception*). Rockwell dodaje: „Co więcej, mogę powiedzieć, MDM [*multiple drafts model*] nie ma istotnego znaczenia dla pytania o to, jak mentalne zjawiska są konstytuowane, chociaż przekazuje on [MDM] wiele ważnych rzeczy o zdawaniu sobie sprawy i o samoświadomości” (Rockwell 1996: 467).

Biorąc pod uwagę rozumowania Rockwella należy stwierdzić, że jedna z podstawowych tez Dennetta dotycząca różnicy pomiędzy modelem orwellowskim i stalinowskim jest wynikiem pomieszania językowego.

IV.6 Teza Dennetta: „Jeżeli podmiot uznaje jakiś moment procesów zachodzących w mózgu za moment świadomości, to jego decyzja będzie arbitralna [w sensie „subiektywna” – *J.B.*]” (Dennett 1991: 194) – wynika z pomieszania w słowie „świadomość” zdawania sobie sprawy i innych zjawisk mentalnych.

Zawarta w (IV.6) teza Dennetta jest prawdziwa gdy chodzi o zdawanie sobie sprawy, a nie gdy chodzi o takie zjawiska mentalne, jak spostrzeganie

⁴⁶ Rockwell cytuje Dennetta: „Are you constantly conscious of the clock ticking? If it suddenly stops, you notice this, and you can say right away what it is that has stopped; he ticks «you weren't conscious of» up to the moment they stopped and «would never have been conscious of» if they hadn't stopped are now clearly in your consciousness. [...] But how could you so clearly remember hearing something you hadn't been conscious of in the first place?” (Rockwell 1995: 466, cytowano z: Dennett 1991: 137–138). Dyskutując przypadki „odruchowego” prowadzenia samochodu, Dennett przyjmuje jako heterofenomenologiczny fakt, że spostrzeganie zdaje się być kontynuowane bez naszego zdawania sobie z niego sprawy oraz że zdawanie sobie sprawy może się przesuwać z jednego aktu spostrzegania na drugi.

czy świadomość zjawiskowa. My przyjmujemy, że zjawiska mentalne są: (a) emergentnymi elementami jakichś zdarzeń w mózgu, które wymagają określonego czasu do swojej realizacji, (b) zjawiskami posiadającymi określoną treść (por. 4.2.2).

Na przykładzie zabawy w „polowanie na naparstek” (*hunt the thimble*) Dennett omawia różnicę pomiędzy spostrzeganiem i zdawaniem sobie sprawy (por. Dennett 1991: 333–335). Dzieci wychodzą z pokoju, w tym czasie ktoś kładzie naparstek w jakimś miejscu. Dzieci wracają. Dziecko, które spostrzega naparstek, siada spokojnie i stara się nie patrzeć w jego stronę, podobnie następane dziecko. W końcu zostaje jedno, powiedzmy Betsy, które pomimo iż się intensywnie rozgląda, nie widzi naparstka. Naparstek jest blisko Betsy, lecz ona go nie dostrzega.

Wyjaśnienie, jakie Dennett proponuje dla braku spostrzegania w przedstawionej grze, różni się od Rockwellowego przeciwstawienia zdawania sobie sprawy i zjawisk mentalnych. Dennett przyjmuje następujące rozróżnienia:

(a) bycie w centrum przeżywania (*foreground of experience*)

(b) bycie w tle przeżywania (*background of experience*)

Elementy (a) i (b) są odmienne od:

(c) bycia tylko w tle spostrzegalnego otoczenia, bez bycia przeżywanym.

Dennett traktuje (a) i (b) jako dwa różne aspekty świadomości. Aby spostrzeganie było w centrum mojego przeżywania, musi istnieć specjalna relacja intencjonalności albo relacja „bycia o czymś”, zachodząca pomiędzy spostrzeganiem, a moimi innymi stanami mentalnymi. Zanim Betsy nie zobaczy naparstka w tym właśnie sensie, żaden jej stan spostrzegania nie dotyczy jeszcze naparstka. Ten rodzaj intencjonalnej relacji Dennett nazywa: „[...] połączeniami celowymi, charakteryzującymi się sterowaniem zwrotnym, korektą błędów, dopasowaniem wyników”⁴⁷.

Połączenia te są prerekwizytami dla tego sensu znajomości poznawczej, o który Dennettowi chodzi. Gdy raz zobaczyłem coś w tym sensie, mogę „coś z tym zrobić”. Podobnie dzieje się z Betsy, która nie widzi naparstka, zanim nie zidentyfikuje go jako naparstek.

Gdy Betsy widzi naparstek w założonym przez Dennetta sensie, to zna, gdy wypowie (głośno lub cicho) sąd „ah, naparstek”, to zdaje sobie sprawę z obecności naparstka. Tego rodzaju wyjaśnienie przyznaje rację fenomenologii „zdawania sobie sprawy” i odpowiada Modelowi Wielu Szkieletów. Nie potrzeba tutaj domagać się istnienia wyspecjalizowanego modułu zdawania sobie sprawy albo istnienia Centralnego Dawcy Znaczenia.

⁴⁷ „[...] feedback-guided, error-corrected, gain-adjusted, purposeful links” (Dennett 1991: 335).

Dennett unika pytania: czy identyfikacja naparstka przez Betsy następuje przed czy po uświadomieniu go sobie (*becoming conscious*) przez nią?⁴⁸. Odwołując się do wprowadzonego przez N. Blocka rozróżnienia możemy przyjąć, że Betsy uświadamia sobie naparstek bez zdawania sobie sprawy z jego pojawiania się. Chcąc jednak rzeczywiście na to pytanie odpowiedzieć, musimy się odwołać do badań naukowych świadomości zjawiskowej. Ponadto Dennett nie pyta: czy ktoś może widzieć i być świadomym obiektów zanim zostanie mu powiedziane, czym one są? Niektóre zwierzęta szybciej zauważyłyby naparstek aniżeli Betsy. W odróżnieniu od nich władająca językiem Betsy nie tylko zauważa naparstek, ale i wie, czym on jest (tzn. wie, że rzeczy, które wyglądają jak właśnie *te*, są naparstkami). Nie jest jednak tak, że akt intencjonalnego uświadamiania sobie tworzy także przeżycie, którego doznaje Betsy.

4.8. *Neuronalne podłoże subiektywności*

Zdaniem T. Rockwella, podmiot dlatego „widzi” na przykład ślepa płamkę na peryferiach swojego pola widzenia, ponieważ wzrokowy proces w mózgu jest w sprzężeniu zwrotnym z innymi procesami, które z kolei umożliwiają wyższy porządek myśli o ślepej płamce. Jakieś połączenie stanów mózgu, które dokonuje tego rodzaju poszukiwań opartych na zdobytym doświadczeniu, będzie zdolne do utworzenia myśli wyższego rodzaju (wspomnianych już HOTs). Nie ma przy tym podstaw, aby zakładać, że tylko jeden organ w mózgu realizuje tę funkcję; może być ona realizowana przez różne części mózgu w różnych czasach albo w tym samym czasie, ponieważ procesy poszukiwania nie muszą być połączone z tym samym Centralnym Dawcą Znaczenia.

„Ja”, z którym sami się identyfikujemy, jest największym, najbardziej aktywnym i najbardziej koherentnym szeregiem neuronalnych połączeń w mózgu. Ludzie z silnym „ja” łączą świat wokół siebie z zorganizowanym szeregiem celów i projektów określających kim oni są. Są oni „ześrodkowani”, ponieważ posiadają silny Środek Narratywnej Ciężkości. Ich historie życia są spójne. Ale na przykład w stresie czy chorobie wiele neuronalnych koalicji przeszukuje „świat” w różnorodny, często niespójny sposób. W takich momentach cechująca się jednością jaźń jest osłabiona (por. 7.6). Należy do-

⁴⁸ „Is her [Betsy, J.B.] identification of the thimble a subsequent effect of her becoming conscious of it, or a priori cause of her becoming conscious of it? This is the question – Orwellian or Stalinesque? – that the Multiple Drafts model teaches us not to ask” (Dennett 1991: 335).

dać, że jaźń (jak ją rozumie Dennett) nie jest przy tym jakimś nieuniknionym kartezjańskim surowym faktem, lecz „wytworem” istniejącym dzięki stałemu wsparciu ze strony mózgu.

Można wątpić w oczywistość tego, że obserwując coś „obiektywnie”, stawiam to we wzajemną relację ze „szkicami” definiującymi moje opowiadanie. Musimy jednak pamiętać, że obiektywność jest celem, do którego zdążamy. Cel ten wyjaśnia, dlaczego nasze zdawanie sobie sprawy jest skupione na tym, na czym jest, i dlaczego widzimy to, co widzimy. Obserwując zmianę koloru – zjawisko *phi* – jesteśmy w stanie mentalnym odpowiadającym byciu obserwatorem. Tego rodzaju nastawienie potwierdza założenie Dennetta:

„Gdy jakaś część świata zacznie w ten sposób tworzyć splot opowiadań (*a skein of narratives*), to owa część świata jest obserwatorem. Czyli jak to jest dla kogoś być obserwatorem w świecie [...]” (Dennett 1991: 137).

Obserwator zaczyna komponować splot opowiadań, snuć opowiadanie o sobie, ponieważ jest on nie tylko obserwatorem, ale zarazem i aktorem. Świadome podmioty są świadome nie tylko dlatego, gdyż znają obecny moment czasu, ale ponieważ rozumieją ów moment w kategoriach stanów przeszłych⁴⁹. Co dzieje się, gdy jakiś obserwator staje się czegoś świadomy w skali zjawisk mentalnych, które jego mózg stale syntetyzuje? Oznacza to, że napotkał on spostrzegane mentalne zjawisko (takie jak spostrzeżenie naparstka) wspólnie z jakimś rodzajem zwrotnie skierowanego połączenia, ustanawiającego relację pomiędzy owym mentalnym zjawiskiem i jakąś czynnością. Obserwacja nie musi być dokonywana przez pojedynczy „moduł świadomości”; może istnieć wielość takich połączeń (*links*). Nie ma racji przemawiającej za przyjęciem tezy, że połączenia te będą formowały jakiś naturalny rodzaj. Ponadto, jeżeli największa neuronowa „koalicja” jest wszystkim, czym jest jaźń, to nie musi istnieć jakiś Centralny Dawca Znaczenia, do którego wszystkie owe połączenia z powrotem dochodzą.

Opis tego, w jaki sposób coś jest w centrum (*foreground*) przeżycia, zakłada, że musi również istnieć tło (*background*) przeżycia. Dennett uznaje taki opis (por. 4.7), rozróżniając pomiędzy: „bezwiednymi reakcjami” a „tłem świa-

⁴⁹ Dennett potwierdza to, gdy pisze, że poziom świadomości umożliwiający cierpienie może istnieć tylko u istot zatroskanych o przyszłość. „Suffering is not a matter of being visited by some ineffable but intrinsically awful state, but of having one’s life hopes, life plans, life projects blighted by circumstances imposed on one’s desires, thwarting one’s intentions – whatever they are” (Dennett 1991: 449).

domych przeżyć⁵⁰. Przedmioty w tle przeżycia mogą być w jakimś momencie wychwytywane przez akt intencjonalnego śledzenia, i uczynione przedmiotami świadomości. Przedmioty te posiadają swoisty rodzaj bezpośredniej dostępności, na przykład dzięki odpowiedniemu treningowi, jakiemu został poddany podmiot. Trening ten może być prosty albo kompleksowy. Zawodowy stroiciel, chociaż nie wsłuchiwał się w dźwięki płynące z rozstrojonego pianina, może nam po pewnym czasie powiedzieć, które struny były rozstrojone. Słyszał on dźwięk pianina, chociaż nie zdawał sobie z tego sprawy. Stroiciel nie śledzący intencjonalnie muzycznego tła jest go, mimo to, świadomy.

Dennett podkreśla, że nie ma ostrej granicy pomiędzy byciem świadomym a byciem w stanie mentalnym. To, że Dennett mówi, iż Betsy jest świadoma napastrka przed i po tym jak go spostrzegła, zaciera niepotrzebnie różnicę pomiędzy „byciem w” i „byciem poza” centrum świadomości. Naszym zdaniem:

IV.7 Dennett musi umieć wyjaśnić zjawisko przesunięcia elementów z tła świadomości do jej centrum, a także odpowiedzieć na pytanie: jaka jest różnica pomiędzy tłem i centrum (zakładając, że z bycia w obydwu z nich można zdać sprawozdanie)?

Jeżeli istnieje tego rodzaju różnica, to musiałby on odrzucić wiele elementów ze swojego Modelu Wielu Szkiców, który to model można nazwać „operacjonalizmem w pierwszej osobie”. Tak rozumiany operacjonalizm: (a) neguje możliwość uświadomienia sobie danego elementu bez wystąpienia intencjonalnego zdawania sobie sprawy o jego pojawieniu się, (b) mówi w zasadzie, że akt intencjonalnego sondowania (próbowania, badania) konstrytuje wszystko, co jest obecne w świadomości, oraz że przeżycie nie ma tła, tylko pierwszy plan⁵¹.

Gdy Dennett twierdzi, że moment świadomości jest *arbitralny* to oznacza to, że jest on *subiektywny* (por. IV.6). Akty dociekań i intencjonalnego śledzenia są tym, co spośród zjawisk mentalnych tworzy strumień świadomości, powodując, że nie istnieją ustalone fakty mówiące o tym strumieniu niezależ-

⁵⁰ „Presumably it’s not sufficient for reflected light from the object merely to enter your eyes, but what further notice must your brain take of it – for the object to pass from the ranks of the merely unconsciously responded to into the background of conscious experience?” (Dennett 1991: 336).

⁵¹ Dennett broni tego rozumowania w przykładzie z setkami obrazów A. Warhola przedstawiającymi Marylin Monroe, które to obrazy, mimo że widzimy tylko ich część, rozpoznajemy jako portrety aktorki (por. Dennett 1991: 354–356).

ne od tychże dociekań. Ale zarówno zjawiska mentalne, jak i obecne w nich dociekające akty są kierowane przez podmiot, który jest obserwatorem i aktorem. Są one w ten sposób subiektywne, ale nie są one arbitralne, chyba że podmiot kieruje się w życiu tylko chwilowymi nastrojami, bez odwołania się do Środka Narratywnej Ciężkości.

Możemy powiedzieć tak: dennettowski stroiciel słyszy dane tony, ponieważ bycie stroicielem uczynił jednym ze swoich projektów życiowych. Tenże projekt jak i inne pomogły mu utworzyć różnego rodzaju stopnie zjawisk mentalnych, mających swoje podłoże w odpowiedniej neuroplastyczności mózgu (por. 6.1.5). Stopnie te z kolei stawiają granice temu, co może zostać odkryte przez pojedynczy akt intencjonalnego śledzenia.

4.9. Podsumowanie

M. V. Antony i J. Aronson nazywają zbudowaną przez Dennetta teorię świadomości „eliminatywną” (por. Antony 1992: 201) [to samo odnosi się do jego ujęcia subiektywności]. W. G. Lycan zarzuca Dennettowi „weryfikacjonizm”, a R. Van Gulick – „antyrealizm”. Według N. Blocka, Model Wielu Szkiców jest do zaakceptowania, gdy przyjmie się przemilczany przez Dennetta irrealizm dotyczący świadomości zjawiskowej. Lycan dodaje, że Model Wielu Szkiców potwierdza jego (Lycana) własną Locke’owską teorię wewnętrznego sensu (*inner sense theory*)⁵².

Chcąc naukowo wyjaśnić świadomość Dennett tworzy całościową teorię, opartą na instrumentalistycznym modelu wyjaśniania, związanych z jego rozumieniem teorii naukowych. Nie tyle ważnym w tym modelu jest wyjaśnienie zachowań (w sensie odpowiedzi na pytanie: co to jest świadome zachowanie?), wystarcza, że odpowiednie stany intencjonalne pozwolą przewidywać dane zachowania podmiotu⁵³. Teorie są instrumentami pozwalającymi przewidywać rezultaty doświadczeń. Model świadomości rozumianej jako wiele szkiców wielokrotnie przetwarzanych jest dla Dennetta rodzajem „dźwigni intuicji”, a jego teoria świadomości obejmuje nie tylko osiągnięcia neurologii i nauki o sztucznej inteligencji, ale także nasz potoczny obraz rze-

⁵² Według Lycana: „According to the inner sense theory, conscious awareness is the successful operation of an internal scanner or monitor that outputs second-order representations of first-order psychological states” (Lycan 1992: 216). Ów *internal scanner* nie oznacza pojedynczego (kartyżjańskiego) skanera ani centralnej wewnętrznej sceny, na której treści świadomości są przedstawiane, ale jest mechanizmem uwagi, który prawdopodobnie może być skierowany na przedstawiane podsystemy.

⁵³ Dennett pisze: „[...] all there is to being a true believer is being a system whose behavior is reliably predictable via the intentional strategy [...]” (Dennett 1987: 29).

czywistości. Dennett nie odpowiada na pytanie, jak zjawiska mentalne są konstituowane przez zjawiska neuronalne.

Dodatkowym atutem proponowanego przez Dennetta modelu jest uwzględnienie wpływu pamięci na świadomość oraz analizy stanów zwrócenia uwagi. Wracając do podanego w (4.) *thumbnail sketch* świadomości widzimy, że Dennett rozumie świadomość jako analogiczną do kompleksowej, obliczeniowej aktywności mózgu. Aktywność ta została zaprogramowana przez genetykę, kulturę oraz przeżycia indywidualne danego osobnika. Ewolucja wszczepiła swoistego rodzaju wirtualną maszynę do rozwijającego się równoległe do niej mózgu.

Mówiąc we wnioskach końcowych o statusie swojej teorii, Dennett podkreśla, że nie zastąpił Kartezjańskiego Teatru teorią ściśle naukową.

„Wszystko, czego dokonałem, to zastąpienie jednej rodziny metafor i obrazów poprzez drugą: teatr, świadka, Centralnego Dawcę Znaczenia i fantazję zamienić na software, na Wielość Szkiców, chaos homunkulusów. Jest to zatem jedynie wojna metafor, powiecie – ale metafory nie są «tylko» metaforami, metafory są narzędziami myślenia” (Dennett 1991: 455).

Odwołanie się do Modelu Wielu Szkiców sugeruje, że to, co jest obiektywnie determinowalne (jako jednoznaczne „przed” i „po”), nie musi koniecznie i jednoznacznie odpowiadać naszemu subiektywnemu przeżyciu. Czasowo-przestrzenne przedstawienie niektórych bodźców jest konstruktem powstałym z elementów neuronalnej aktywności i nie musi ono odpowiadać porządkowi realnie przeżywanemu przez podmiot. Coś podobnego zachodzi, gdy odruchowo cofamy rękę przy dotknięciu czegoś gorącego, jeszcze zanim poczuliśmy ból. Nie istnieje jakaś neuroanatomiczna „scena”, dzięki której stajemy się czegoś świadomi. Zamiast niej jest jedynie strumień poroździelanych równoległe zachodzących przedstawień albo „wielości szkiców”, współdziałających między sobą, samoorganizujących się nieliniarnie w odpowiednie zespoły. Szkice te są związane razem dzięki przestrzennym i czasowym wzorcom, sterowanym prawdopodobnie przez daleko sięgające zsynchronizowane neuronalne działania.

Model Wielu Szkiców przypomina metaforyczne rozumienie świadomości krytykowane przez H. Bergsona. Dla Bergsona świadomość jest żywym, konkretnym trwaniem. Trwanie jest wielością czysto jakościową, jej poszczególne stany zlewają się i przenikają⁵⁴. W ten sposób powstaje nieustan-

⁵⁴ Bergson opisuje jakościową „wielość stanów świadomości” (*multiplicity of conscious states*) i odróżnia ją od wielości zlokalizowanych w przestrzeni przedmiotów materialnych tworzących liczbę (por. Bergson 1913: 60–65, 88–81).

nie się zmieniająca, dynamiczna jedność. Następstwo wzajemnie się przenikających stanów zasadniczo różni się od przestrzennego uszeregowania obok siebie przedmiotów. Bergson neguje wynikającą z „prawa zachowania siły” tezę o bezpośrednim wytwarzaniu świadomości przez mózg, jak i tezę o braku wzajemnego zdeterminowania procesów psychicznych i neurologicznych⁵⁵.

Ewolucjonista Dennett neguje kartezjański model świadomości z kilku powodów.

(1) Jeżeli świadomość powoduje jakąś różnicę w zachowaniach, to czymś nieekonomicznym byłoby zbieranie wszystkich informacji w jednym jedynym miejscu mózgu, zanim zaczną one wpływać na zachowania. Aby informacja mogła przejść z jednego miejsca mózgu do innego potrzeba czasu. Gdyby wszystkie informacje musiały się zejść w jednym miejscu, to reakcje podmiotu byłyby odpowiednio wolniejsze.

(2) Kartezjański Teatr jest związany z istnieniem określonego punktu czasowego, w którym zdarzenia w mózgu stają się zdarzeniami świadomymi. Ponieważ informacja potrzebuje czasu na przejście z jednej części mózgu do drugiej, nie możemy określić ile tego czasu potrzeba, zanim zidentyfikujemy miejsce, w którym informacja ta stała się świadoma. Określenie czasu, w którym powstało świadome zdarzenie, jest więc związane z „kartezjańskim” miejscem jego powstania. Dennetta argumenty przeciwko jednoznacznie przestrzennie określonym miejscom w mózgu są tym samym argumentami przeciwko jednoznacznym momentom czasowym.

Założenia, z których wychodzi Dennett, nie wydają się jednak konieczne. (a) Jak widzieliśmy, może istnieć jakieś funkcjonalnie zdefiniowane miejsce świadomości, która jest fizycznie poroździelana i operuje na różnych poziomach oraz w różnych momentach czasu. (b) Argument ten można rozbudować stwierdzając, że świadomość może się pojawiać w różnych miejscach mózgu w różnych momentach czasowych (kwestią do ustalenia pozostanie: co powoduje, że jedne stany stają się świadome, a drugie nie?). Jeżeli jednak stany w różnych miejscach mogą być świadome w różnych momentach czasu, to dlaczego nie mogą być one świadome w tym samym czasie? Taki przypadek oznaczałby jednak utratę „von Neumanesque” (seryjnego) aspektu świadomości, na którym Dennettowi zależy.

⁵⁵ Por. Bergson 1913: 100–101. „Żaden obraz jednak nie jest w stanie dowieść i nie dowiedzie nigdy, że fakt psychologiczny jest uwarunkowany koniecznie ruchem molekularnym” (tamże: 104). Według Kulczyckiego: „Bergsonowska koncepcja pamięci nieświadomej i możliwości przechowywania się wspomnień «w sobie», niezależnie od mózgu, budziła zastrzeżenia W. Jamesa” (Kulczycki 1996: 160).

(3) Powiedzmy, że zgadzamy się z Dennettem, że w mózgu nie ma Kartezjańskiego Teatru i że przez jakiś bardzo krótki okres czasu nie można określić, kiedy stany stają się świadome. Pozostawia to otwartą możliwość, że istnieje większy teatr, w którym zachodzi świadomość – jest nim cały mózg. Dennettowi nie chodzi jednak tylko o odrzucenie Kartezjańskiego Teatru, czy o odrzucenie rozróżnienia: stalinowski–orwellowski. Jego ostatecznym celem jest przekonanie nas, że istnieje tylko mózg i heterofenomenologiczne opisy zachowań. Nie ma żadnych faktów dotyczących świadomych przeżyć, nie ma czegoś takiego jak „fenomenologia tego, co mentalne”. Każda oczywistość, jaką dysponujemy, gdy chodzi o uprawianą w ramach potocznego obrazu filozofię, jest – jego zdaniem – jedynie oczywistością dotyczącą heterofenomenologicznych interpretacji zachowań podmiotu (w tym przede wszystkim jego wypowiedzi słownych). Dlatego nie będziemy nigdy mogli odkryć jakichś zjawiskowych faktów, a jedynie fakty heterofenomenologiczne.

Rozwinięta przez Dennetta strategia przemawia wyraźnie na niekorzyść potocznego obrazu świata. Obraz ten należy wyeliminować, gdyż jest on tylko konstruktem ułatwiającym nadanie sensu niektórym kompleksowym schematom zachowań. Patrząc od strony ontologicznej możemy powiedzieć, że nie ma w nas nic, co by odpowiadało obiektom potocznego świata, dlatego też nie ma żadnej „luki w wyjaśnianiu”. Dennetta należy jednak zapytać: dlaczego mielibyśmy być w aż takim błędzie co do nas samych (że nie jesteśmy homunkulusami)? Jak nasze potoczne pojęcie osoby (z jego nieeliminowanymi społecznymi i etycznymi implikacjami) może być tak błędne?

Dennett powie: dlaczego mówić o subiektywności, skoro jest ona neuronalną aktywnością? Parafrazując powiedzenie Nagla można powiedzieć, że dla Dennetta subiektywność jest neuronalną aktywnością zachodzącą w mózgu podmiotu, jest ona owym *jak to jest* dla neuronów być w tym szczególnym stanie. Dennettowska pseudojaźń przechodzi w wielkości narratywne tworząc iluzję, że istnieje jakaś oryginalna jaźń, chociaż ta ostatnia jest jedynie stale podtrzymywaną konstrukcją językową. Dennett nie musi przyjmować idei, że zdawanie sobie sprawy (*awareness*) wymaga jakiegoś „świadomego modułu” albo „wewnętrznego reflektora” albo „Kartezjańskiego Teatru”. Stwierdza on jedynie istnienie czegoś opisywalnego jako zdawanie sobie sprawy, co jest częścią naszego mentalnego życia i co wydaje się być odmiennym od różnych mentalnych zjawisk, które sobie uprzytomiamy. Niemniej jednak ta szczególna cecha świadomości powinna być wyjaśniona przez jakąś teorię, czego brakuje w prowadzonych przez Dennetta analizach.

ROZDZIAŁ PIĄTY

THOMAS NAGEL – SUBIEKTYWNE WŁASNOŚCI ŚWIADOMOŚCI

5. Wprowadzenie: jakiej filozofii świadomości poszukujemy?

„Some would say that the philosophy of mind without the first-person perspective, or the first person point of view, is like *Hamlet* without the Prince of Denmark” (Shoemaker 1996a: 157).

T. Nagel wychodzi z założenia, że źródła filozofii są przedwerbalne i przedkulturowe, a jednym z najtrudniejszych jej zadań jest wyrażenie za pomocą języka owych niesformułowanych, lecz intuicyjnie (potocznie) odczuwanych problemów, w taki sposób, aby nic na tym nie utraciły¹. W ramach tak pojętej filozofii należy widzieć zarówno jego opisy języka, jaźni wraz z jej świadomością, jak i jego poszukiwania poszerzonej teorii świadomości. Zachowując proponowane przez Nagla rozumienie roli filozofii w formowaniu intuicyjnie dostępnych problemów świadomości, przyjmujemy twierdzenie:

¹ Por. Nagel 1997a: 16. „Filozofia jest dzieckiem intelektu, a kultura, która chce ją odrzucić, nigdy nie dorośnie” (Nagel, s. 17). W niniejszym rozdziale zanalizujemy fragmenty książki Nagla *Widok znikąd*, którą Ch. Peacocke określa jako „striking and highly original book” (por. Peacocke 1989: 65). J. Levin nazywa argument Nagla o subiektywności „extremely influential” (por. Levin 1986: 246; Shoemaker 1996: 99). Nagla filozofię umysłu omawia W. Dziarnowska (2001: 188–196).

V.1 Poszukiwana naukowa teoria świadomości musi znaleźć miejsce dla wyjaśnienia psychologiczno-filozoficznego rozumienia subiektywności oraz dla wyjaśnienia jakościowych treści naszych świadomych przeżyć.

V.1 oddaje wyjściowe pytanie Nagla: Jak połączyć perspektywę zanurzonej w świecie osoby z obiektywnym obrazem świata, którego osoba ta jest częścią? (por. Nagel 1997c: 7). Nagel wskazuje tym samym na istnienie nieprzezwyčajalnej granicy, obecnej w tych materialistycznych filozofiach świadomości, które przyjmują, że osoba razem ze wszystkimi jej psychologicznymi cechami nie jest niczym więcej aniżeli ciałem z jego fizycznymi atrybutami (por. Nagel 1965: 339).

W obecnym rozdziale zostaną przedstawione prowadzone przez Nagla analizy subiektywności świadomości. Jego ujęcie jaźni wymaga osobnego opracowania. Jeżeli odwołujemy się do tego ujęcia, to jedynie po to, aby ukazać proponowane przez Nagla rozumienie świadomości. Nagel jest przekonany, że to, co nazywamy „świadomym punktem widzenia” podmiotu, jest logicznie nieredukowalne do jego cech opisywanych fizycznie, lecz mimo to powiązane *koniecznie* z tymi cechami (por. Nagel 1998a: 337–338).

Kto posiada ostateczny autorytet, gdy chodzi o poznanie subiektywności świadomości zjawiskowej – przeżywający podmiot czy badający ją z zewnątrz filozof lub naukowiec? Ta epistemologiczna asymetria znajduje swój teoretyczno-historyczny wyraz w dwóch całkowicie odmiennych analizach świadomości.

Pierwsza z nich, jak wspomnieliśmy, polega na konsekwentnym używaniu perspektywy pierwszoosobowej. Sztandarowe przykłady tej orientacji to kartezjański dualizm oraz husserlowska fenomenologia – wielka próba stworzenia podwalin systematycznej, opartej na introspekcji nauki o świadomości. Teoriom wychodzącym od pierwszej osoby zarzuca się: (a) że oczywistość spostrzeżenia wewnętrznego okazała się – na podstawie badań naukowych – założeniem nie do utrzymania oraz (b) że nie dostarczają one odpowiedniej metody eliminowania błędnych obserwacji stanów wewnętrznych w przypadkach, w których mamy do czynienia ze sprzecznymi wypo wiedziami danego podmiotu.

Drugim sposobem analizowania świadomości jest przedstawiony w III i IV rozdziale naturalistyczny obiektywizm. Zbieranie danych ogranicza się do dostępnych z zewnątrz informacji, dotyczących zachowania się podmiotu.

Wspomniane sposoby badań świadomości doprowadziły do metodologicznych ekstremizmów, osłabionych częściowo dzięki powstaniu psychologii kognitywnej oraz komputerowych modeli ludzkiego umysłu. Aktualna

debata nad świadomością – w której Nagel odgrywa istotną rolę – jest dlatego tak owocna, ponieważ zacierają się w niej *aprioryczne* podziały pomiędzy metodologią stosowaną w klasycznej fenomenologii, a metodologią stosowaną w filozofii analitycznej. Świadczy o tym chociażby rosnące zainteresowanie husserlowską fenomenologią w kręgach filozofów analitycznych (por. dennettowska heterofenomenologia albo rozwijana przez F. Varełę neurofenomenologia, 6.3.2).

5.1. Fizykalna koncepcja obiektywności

„Rzeczywistość nie jest po prostu obiektywną rzeczywistością” (Nagel 1997c: 36).

Zdaniem Nagla, mówienie o osobowej perspektywie czy o odpowiadającym jej „pojęciom mentalnym” odnosi się zasadniczo do „pojęć psychologicznych” (pojęcie lęku, gniewu, bólu, mniemania, myślenia) [por. Nagel 1997c: 31–33]. Nagel przypisuje istnienie stanów wewnętrznych istotom posiadającym rozbudowane systemy pojęć psychologicznych, a także małym dzieciom i zwierzętom. Również o nas samych myślimy potocznie, iż mielibyśmy przeżycia także wtedy, gdybyśmy nie posiadali języka, w którym o nich mówimy. Nagel zajmuje się stanami mentalnymi świadomymi, na przykład uświadomionym lękiem, uświadomionym gniewem, czyli stanami, o których podmiot ma pojęcie². Nie zajmuje się on natomiast stanami, które psychoanaliza opisuje jako nieświadome. Nagel, jak zobaczymy, zgadza się z L. Wittgensteinem, że pojęcia mówiące o stanach wewnętrznych są publiczne, a nie prywatne. Pojęcia te są jednak w inny sposób publiczne, aniżeli publiczne są pojęcia służące do opisu świata fizycznego. Dlatego właśnie problemy stosunku umysł–ciało, wolnej woli i tożsamości osobowej nie są problemami dotyczącymi jedynie każdego z nas z osobna (por. Nagel 1997b: 249).

Swoje analizy Nagel rozpoczyna od przedstawienia obrazu świata, który określa zwrotem: „fizykalna koncepcja obiektywności” (Nagel 1997c: 19; por. Nagel 1991: 13–15). „Nasz opis jest tym bardziej obiektywny, im mniej zależy od swoiście ludzkiego punktu widzenia” (Nagel 1997: 213).

Koncepcja ta nie jest identyczna z naszym potocznym ujęciem tego, jak fizyczna rzeczywistość jest faktycznie zbudowana. Niemniej jednak dzięki

² Por. Nagel 1997: 204. „[...] jeśli ktoś posiada pojęcie typu świadomego stanu mentalnego i od czasu do czasu sam znajduje się w tym stanie, to będzie umiał stosować to pojęcie zarówno od wewnątrz jak i od zewnątrz, zgodnie z opisem Wittgensteina” (Nagel 1997c: 31).

tej koncepcji dochodzimy do rozumienia świata fizycznego. Podążanie w kierunku fizykalnej obiektywności oddała nas od form spostrzegania i działania charakterystycznych dla naszego gatunku. Nie bierze się wówczas pod uwagę ani cech odróżniających dany podmiot poznający od innych podmiotów, ani jego położenia przestrzennego lub czasowego. Celem analiz jest pojęcie świata, który nie jest widziany z jakiegoś konkretnego punktu widzenia. Przedmioty tego świata są pozbawione cech spostrzeniowych i doznaniowych oraz są analizowane w terminach dotyczących ich pierwszorzędnych cech (kształt, ruch). To, jaki świat jest rzeczywiście sam w sobie, poznamy dopiero wtedy, gdy zobaczymy go bez używania naszych zmysłów i pojęć. Zdaniem Nagla, uciekamy od tego, co subiektywne pod presją założenia,

„[...] że wszystko musi być czymś nie z jakiegoś punktu widzenia, lecz samo w sobie. Uchwycić to, odrywając się coraz bardziej od naszego punktu widzenia – to nieosiągalny ideał, ku któremu zmierza dążenie do obiektywności” (Nagel 1997b: 250).

U podstaw takiej koncepcji obiektywności leży ideał poznawczy urzeczywistniony w naukach ścisłych, pretendujący do ekstrapolacji na każdy rodzaj poznania rzeczywistości. Rozwój wielu dyscyplin naukowych jest wynikiem tego rodzaju obiektywnego rozumienia, które okazało się bardzo owocne epistemologicznie i heurystycznie. Naukowe poznanie świata rozwinęło się głównie dzięki teoriom, w których odwołujemy się do pojęć nie związanych bezpośrednio ze szczególnie ludzkim punktem widzenia. Poznanie takie jest następnie ujmowane w formuły matematyczne i potwierdzone przez odpowiednie badania naukowe. Tego rodzaju „obiektywne” i „teoretyczne” poznanie zawodzi jednak, gdy próbujemy z jego pomocą ująć subiektywne przeżycia jakościowe.

Wprawdzie fizykalno-obiektywne spojrzenie na świat nie implikuje jeszcze redukcji stanów psychicznych do fizycznych, ale dualistyczne ujęcie tych stanów wydaje się w ramach takiego spojrzenia zupełnie niezrozumiałe. Dualizm zakłada bowiem, że stany mentalne są zasadniczo innego rodzaju aniżeli stany fizyczne i to zarówno od strony teoriopoznawczej, jak i ontologicznej.

Omawiając fizykalno-obiektywne spojrzenie na świat Nagel podkreśla zarówno, że perspektywa subiektywna nie daje się wyeliminować, jak i to, że nie można zrezygnować z perspektywy obiektywnej. Tak rozumianego przeciwstawienia spojrzeń nie da się na dziś uniknąć. Każdy jednak zwrot ku większej obiektywności oddała nas, zdaniem Nagla, od poznania całości świata. Według Nagla termin „fizykalny” jest związany z pojęciem „obiek-

tywności”³ i odnosi się on nie tylko do tego, co można opisać za pomocą pojęć współczesnej fizyki, gdyż oczekujemy dalszych postępów naukowych. Być może zjawiska mentalne zostaną kiedyś uznane za fizyczne, a to, co fizyczne, musi być obiektywne⁴.

Nagel uwzględnia Popperowski wymóg „poznania bez poznającego podmiotu”. Jeżeli jednak nasze pojęcie tego, co fizyczne, ma, według Nagla, objąć kiedyś zjawiska mentalne, to będzie ono musiało przypisać im charakter obiektywny. Bardziej prawdopodobne wydaje się jednak, że relacje między stanami mentalnymi a fizycznymi zostaną wyrażone ostatecznie w teorii, której podstawowych terminów nie da się w sposób jasny przypisać do żadnej z tych dwóch kategorii. Nagel uważa, że poszukiwana teoria będzie miała, od strony ontologicznej, panpsychiczny charakter (por. 5.4.2).

Zakładając prowadzone przez Nagla analizy obiektywności można przyjąć ogólne twierdzenie:

V.2 W badaniach świadomości nie można się wyrzec obiektywności, ale nie do utrzymania jest obiektywne spojrzenie na świat proponowane przez nauki ścisłe.

Fizykalnie obiektywny punkt widzenia, który odcina się od subiektywnego punktu widzenia, aspiruje do spojrzenia na rzeczywistość *sub specie aeternitatis*⁵. Równoczesne istnienie tych dwóch punktów widzenia generuje podstawowe problemy w filozofii świadomości. W recenzji książki Nagla *Widok znikąd* C. McGinn zauważa: „Jesteśmy rozrywani pomiędzy tymi dwoma sposobami widzenia świata i nie możemy zadowolająco zintegrować ich w koherentne ujęcie danej rzeczy” (Nagel 1997b: 250).

Rozdźwięk pomiędzy wspomnianymi punktami widzenia powoduje, że filozofowie podkreślają znaczenie albo jednego, albo drugiego z nich, prowa-

³ Por. Nagel 1997b: 247–248. Zdaniem Nagla: „Ciąg coraz doskonalszych obiektywizacji może nas wyprowadzać coraz dalej poza osobistą, ludzką perspektywę. Jeżeli jednak pragniemy zrozumieć cały świat, nie powinniśmy zapomnieć na zawsze o naszych subiektywnych początkach; należymy przecież do świata wraz z naszymi osobistymi punktami widzenia” (Nagel 1997c: 10).

⁴ Por. Nagel 1997: 219. Możemy się tutaj odwołać do wprowadzonego przez W. Sellarsa rozróżnienia pomiędzy predykatami „fizykalny₂” (które są adekwatne do kompletnego opisu stanów świata, chociaż niekoniecznie do świata przed pojawieniem się życia) i „fizykalny₁” (które dotyczą obiektów fizycznych; por. Mehl i Sellars 1956: 252).

⁵ Od strony epistemologicznej chodzi o uzasadnianie przekonań za pomocą apriorycznych pewników (matematycznych, logicznych), tak że nie potrafimy sobie wyobrazić ich fałszywości (por. Nagel 1997c: 159, 108).

dząc tym samym do powstania zniekształconego obrazu rzeczywistości. W dalszych analizach (rozdziały VI i VII) proponujemy:

(a) utrzymanie perspektywy pierwszo- i trzecioosobowej (w ich dzisiejszym rozumieniu),

(b) uwzględnienie ludzkiej specyfiki rodzajowej przy pominięciu specyfiki ściśle indywidualnej (dotyczącej danego osobnika),

(c) pojęciowe określenie neuronalnych korelatów świadomości, podpadających pod teorię świadomości, zakładającą emergencyjny rodzaj praw łączących stany mentalne z fizycznymi.

Nagel nie odwołuje się do tak rozumianych korelatów świadomości. Jego poszukiwania „poszerzonej” teorii świadomości wychodzą od *apriorycznego* założenia o istnieniu koniecznych powiązań pomiędzy stanami mentalnymi i fizycznymi, które to powiązania muszą następnie zostać poparte znalezieniem *a posteriorycznych* praw ich dotyczących (por. 5.5.1, 5.5.2).

5.1.1. Obiektywna fenomenologia

Nagel zaznacza, że pojęcie świadomości, chociaż związane z subiektywnością, nie pozostaje ograniczone do tego, co możemy zrozumieć wyłącznie za pomocą pojęć odnoszących się do naszych własnych przeżyć. Odczucia, uczucia, a także stany intencjonalne (mimo ich całej „obiektywności”) muszą się subiektywnie manifestować, aby mogły dotrzeć do świadomości. W przeciwieństwie do ich materialnych korelatów, subiektywne procesy psychiczne nie pozwalają się ująć za pomocą pojęć odpowiednich dla fizycznego świata, który leży u ich podstaw. Poszukiwany realistyczny opis świata musi dodatkowo ująć wewnętrzne stany innych rodzajów organizmów, bez ontologicznego redukowania ich do stanów funkcjonalnych lub behawioralnych. Redukcje tego rodzaju dążą bowiem do wprowadzenia wspomnianego wcześniej fizykalnego typu obiektywności świadomości, którego istotnym kryterium teoriopoznawczym jest ontologiczna teza: tylko to istnieje, co może być zrozumiałe w całkowicie fizykalnie określony sposób (por. Nagel 1997c: 21–22).

W tym sensie Nagel uważa, że próba zrozumienia ludzkiego umysłu na podstawie zrozumienia działania komputera prowadzi donikąd (nawet, gdyby ten komputer potrafił wykonać wszystkie czynności, jakie wykonuje człowiek). Struktury i funkcje umysłu można odkryć tylko przez badanie samego umysłu. Musimy szukać takiego rozumienia „mentalnej obiektywności”, które z jednej strony uwzględnia subiektywność świadomości, z drugiej zaś strony nie traktuje naszego świadomego umysłu jako czegoś uprzywilejowanego (por. Nagel 1997c: 21–22). Dlatego należy wypracować

pojęcie punktu widzenia, które nie wychodzi z naszego punktu widzenia. Pojęcie to musi być dostępne dla podmiotów, których punkt widzenia jest różny od naszego. Nagel zaznacza, że faktycznie posiadamy takie ogólne pojęcie – pojęcie świadomości w ogóle – ale z konieczności nie uwzględnia ono specyfiki odmiennych rodzajów subiektywności. Skutkiem tego nie wszystkie aspekty rzeczywistości są możliwe do obiektywnego przedstawienia.

C. McGinn stwierdza, że Nagel nie używa spójnie terminów „subiektywny” i „obiektywny”, lecz ich swoistej mieszaniny, z którą spotykamy się w całej książce *Widok znikąd* (por. McGinn 1997: 89). Nagel wyraża się następująco:

(O) „Obiektywność jest metodą rozumienia. To przekonania i postawy są obiektywne w zasadniczym sensie tego słowa” (Nagel 1997c: 8).

Celem uzyskania obiektywnego rozumienia nas samych oraz otaczającego świata, dystansujemy się od naszego początkowego spojrzenia na życie oraz świat i tworzymy nowe ich ujęcie⁶. Termin „obiektywny” jest predykatem dotyczącym ujęć, a nie faktów lub cech. W tym sensie różnica pomiędzy tym, co subiektywne, a tym, co obiektywne, jest faktycznie sprawą „stopnia”. Zdaniem Nagla, możemy myśleć o rzeczywistości jako o szeregu koncentrycznych, progresywnie odkrywanych sfer, prowadzących nas do uwolnienia się od przygodności własnej jaźni. Fizykalna obiektywność, z jednej strony, i subiektywność – z drugiej, stanowią dwa bieguny owego sferycznego kontinuum. Pomimo ustanowień definicyjnych (por. [O]), Nagel przechodzi do przypisywania obiektywności częściom rzeczywistości.

Możemy stwierdzić, że potrafimy naukowo badać struktury naszego poznania, ale nie nasze odczucia z ich cechą *jak to jest* być świadomym organizmem⁷. Z tego względu Nagel poszukuje „[...] obiektywnej fenomenologii, niezależnej od empatii i wyobraźni” (Nagel 1997: 218). Jej celem ma być taki opis subiektywnego charakteru doznań, który będzie zrozumiały dla istot niezdolnych do przeżywania tych doznań. Możemy przykładowo próbować rozwijać pojęcia służące do wyjaśnienia osobie od urodzenia niewidomej, na czym polega widzenie. Pojęcia tego rodzaju umożliwią nam nowe dojsście do

⁶ Nagel oraz McGinn używają terminu „conception”, który tłumaczymy jako „ujęcie”. „Conceptions” oznaczają tym samym sposoby rozumienia, teoretyzowania względnie wyjaśniania tego, co mentalne. M. Gebauer tłumaczy „conception” jako „Auffassung” (por. Nagel, *Der Blick von nirgendwo*: 12, 13, 55; por. McGinn 1997: 92).

⁷ Chalmers pisze: „As Nagel (1974) has put it, there is something it is like to be a conscious organism. This subjective aspect is experience” (Chalmers 1995: 2).

naszych własnych doznań. Oparta na tych pojęciach fenomenologia pozwoli, zdaniem Nagla, na sensowniejsze sformułowanie pytań o fizyczną podstawę naszych doznań. Pewne cechy subiektywnych przeżyć, dopuszczające tego rodzaju nowe opisy, mogą się bowiem lepiej nadawać do wyjaśnień obiektywnych.

Odnosnie do poszukiwanej przez Nagla obiektywnej fenomenologii nasuwają się dwie uwagi:

(1) Stworzenie tego typu fenomenologii jest na razie bardzo odległe, zwłaszcza gdy chciałoby się np. za jej pomocą wyjaśnić osobie niewidomej, na czym polega widzenie.

(2) Jednym z empirycznych modeli ukazujących obiektywizację w ramach nowej fenomenologii mogą być badania nad zażywaniem tabletek odwykowych, działających na system nerwowy. Dzięki tabletkom głód nikotynowy podmiotu zostaje tak zredukowany, że nie potrafi on zauważyć różnicy pomiędzy samopoczuciem spowodowanym wewnętrznym działaniem tabletek, a samopoczuciem spowodowanym zewnętrznym działaniem papierosów. Tym samym może on uogólniać jakości przeżyciowe pojawiające się w wyniku działania tabletek na znane mu przeżycia wynikające z palenia. Badania te dostarczają zarazem informacji o neurologicznych podstawach określonych stanów subiektywnych, które to informacje można następnie uogólniać.

5.1.2. Obiektywność zjawiska psychicznego

C. McGinn precyzuje podane przez Nagla definicje subiektywności. Fakt lub cecha są subiektywne, jeżeli zawierają w sobie specyficzny punkt widzenia; jeżeli nie, to są one obiektywne (por. McGinn 1997: 90). Dane ujęcie jest subiektywne, jeżeli przedstawia jakiś fakt ze specyficznego punktu widzenia, wykorzystując ten punkt jako medium przedstawiania (*as a medium of representation*); w przeciwnym wypadku ujęcie jest obiektywne. Kombinując ze sobą owe rozróżnienia McGinn otrzymuje cztery możliwości:

(1) subiektywne ujęcie subiektywnego faktu (na przykład imaginatywne, zakładające poznanie przez znajomość pojęcie naszych własnych stanów świadomości).

(2) subiektywne ujęcie obiektywnego faktu (na przykład nasze perspektywiczne, spostrzeżeniowe przedstawienie jakiejś pierwszorzędnej cechy, powiedzmy rodzaju powierzchni),

(3) obiektywne ujęcie subiektywnego faktu (na przykład nasze schematyczne pojęcie świadomości w ogóle, które może także zawierać formy świadomości niedostępne naszej wyobraźni),

(4) obiektywne ujęcie obiektywnego faktu (na przykład teorie w fizyce).

Proponowany w (1)–(4) sposób użycia terminów „subiektywność” lub „obiektywność” jakiegoś faktu nie podaje rozstrzygnięcia, czy ten fakt jest ujmowalny (jest do pomyslenia) subiektywnie lub obiektywnie. Nagel nie twierdzi, że (4) jest fałszywe ani że tylko (1) jest prawdziwe.

Podkreślmy to, czego w interpretacji podanej przez McGinna brakuje, że Nagel nie odrzuca filozoficznej obiektywności, twierdzi jedynie, że wspomniana fizykalna obiektywność nie jest jedyną interpretacją obiektywności (por. Nagel 1997c: 23).

Obiektywność zjawiska psychicznego, jak ją rozumie Nagel, można traktować jako całość wszystkich subiektywnych perspektyw. Przy obiektywnym ujęciu stanów mentalnych chodzi o ich bezstronny opis, ujmujący zarazem podmiotowy (indywidualny lub gatunkowy) punkt widzenia. Opis taki otrzymamy, gdy subiektywne cechy naszych przeżyć potraktujemy jako jeden z przejawów bytu mentalnego w świecie.

Pierwszym krokiem ku obiektywizacji ujęć stanów mentalnych jest uchwycenie wszystkich ludzkich perspektyw, łącznie z naszą. Równocześnie nie odmawiamy tym perspektywom własności decydujących o ich szczególnym charakterze. Aby ten krok zrealizować, musimy nasze własne przeżycia ująć jako zdarzenia w świecie (spojrzeć na nie z heterofenomenologicznego „zewnątrz”). Przeżycia te staną się wtedy dostępne innym podmiotom, a my będziemy mieli z kolei dostęp do ich przeżyć. W rezultacie otrzymamy ogólne ujęcie subiektywnego punktu widzenia. Nagel pisze:

„Kiedy próbujemy pojąć umysły innych ludzi, nie możemy pominąć istotnego czynnika, jakim jest punkt widzenia. Musimy raczej dokonać uogólnienia i uznać samych siebie za jeden z wielu punktów widzenia”⁸.

Nie jest jasne, do czego odnosi się tutaj „punkt widzenia”: czy do pojedynczego osobnika, czy też do rodzaju osobników. Nagel nie sugeruje, iż punkt widzenia jest czymś zupełnie prywatnym, niekomunikowalnym. Prywatny język nie jest językiem, i nie ma sensu mówić, że ktoś z nas ma poprzez prywatny język dostęp do swojego własnego punktu widzenia. Pojęcia dotyczące tego, co mentalne, nie odnoszą się do prywatnych obiektów świadomości.

Zdaniem McGinna, jedynie mówienie o obiektywności ujęcia dopuszcza w spójny sposób stopnie, o których mówi Nagel: jakieś ujęcie może być mniej lub bardziej oddzielone od początkowej subiektywnej perspektywy. Nie ma jednak sensu (w tym rozumieniu) mówić, że różnica pomiędzy su-

⁸ Nagel 1997c: 27 (oryg. „[...] we cannot abandon the essential factor of a point of view”; *The View from Nowhere*: 20).

biektywnymi i obiektywnymi faktami jest sprawą stopnia. Fakty albo zawierają odniesienie do świadomości, albo nie. Możemy się zastanawiać mniej lub bardziej obiektywnie nad wrażeniem czerwonego albo nad kwadratowością jakiegoś obiektu, ale nie ma sensu mówić, że sama kwadratowość jest bardziej obiektywna aniżeli wrażenie czerwonego. Często w takich przypadkach pojawiający się błąd przypomina pomylenie natury faktu z naturą pojęcia, jakie o tym fakcie mamy. Błąd tego rodzaju można łatwo popełnić rozważając punkty widzenia o punktach widzenia. McGinn nie twierdzi, że Nagel popełnia taki błąd, tylko że za mało nas przed nim ostrzega.

Pytanie Nagla w II rozdziale *Widoku znikąd* brzmi: czy subiektywność świadomości pozwala nam oddalić się od niej i utworzyć pojęcie świadomości, które nie będzie ograniczone przez jej subiektywność? Odpowiadając Nagel podaje kilka krytycznych spostrzeżeń na temat funkcjonowania ogólnego pojęcia świadomości. Faktycznie zdajemy się operować pojęciem świadomości, a ono nie może być wyjaśnione w terminach subiektywnego przedstawienia.

Nagel mówi niewiele o treści tego rodzaju ogólnego pojęcia, które jest pojęciem niemożliwej do zrozumienia perspektywy. Dokładniejszą jej charakterystyką, chociaż metaforyczną, byłoby: perspektywa, która może być w pełni ujęta tylko przez jakąś wszechwiedzącą istotę⁹. Taka charakterystyka jest jednak również nieprzekonywująca oraz mało obiektywna.

Nadal pozostaje niejasne, czy rzeczywiście możemy posiadać ogólną ideę świadomości, posiadającą więcej treści aniżeli całkowicie niewyspecyfikowana idea jakiegoś aspektu rzeczywistości, którego ze względu na naszą ludzką konstytucję nie jesteśmy zdolni ująć? Tym samym Nagłowska idea obiektywnego ujęcia subiektywności i związana z nią idea „obiektywnej fenomenologii”, to znaczy „[...] niezależnej od empatii i wyobraźni”, wydaje się swoistym *contradictio in adiecto*. Pozostaje bowiem pytanie, jak taka obiektywna fenomenologia jest możliwa.

Odpowiadając Nagel odwołuje się wprawdzie do wyobraźni, ale mówiąc o niej mamy do czynienia z granicami określonymi przez każdorazowe struktury i perspektywy, tak iż każda „mentalna obiektywność” musi pozostać niepełna. Im bardziej nasze przeżycia są podobne do przeżyć innego organizmu, tym lepiej potrafimy ująć subiektywność jego przeżyć (por. Nagel

⁹ Nagel pisze: „Niewykluczone, że ludzka obiektywność może uchwycić jedynie część świata, jeżeli jednak odnosi sukces, to powinna umożliwić nam zrozumienie tych aspektów rzeczywistości, których istnienie jest zupełnie niezależne od naszej zdolności pojmowania – równie niezależne, jak istnienie niewyobraźalnych dla nas przedmiotów” (Nagel 1997c: 113).

1997: 209; 1991: 25). Ale na tym sprawa się kończy. Proponowana przez Nagla obiektywna fenomenologia może na razie jedynie zapewnić to, że subiektywności nie traktuje się jako czysto prywatnej cechy.

5.2. *Neuralne podłoże jedności świadomości*

Nagel zaznacza, że świadome doznania spotykamy u ludzi i zwierząt, nie jesteśmy natomiast pewni ich obecności w prostszych organizmach. To, że jakiś organizm w ogóle posiada świadome doznania, znaczy, „że istnieje coś takiego jak bycie tym organizmem”. Organizm jest świadomy – w sensie istotnym dla całej argumentacji Nagla – jeżeli istnieje jakieś *jak to jest* być tym typem organizmu¹⁰.

Bycie danym organizmem zakłada istnienie subiektywnych doznań. To właśnie w przypadku doznań związek z punktem widzenia jest ściślejszy aniżeli w przypadku pojęć. Według Nagla: „Trudno zrozumieć, na czym miałyby polegać *obiektywny* charakter doznań poza szczególnym punktem widzenia, z którego ujmuje je ich podmiot” (Nagel 1997: 212).

Powyższy cytat wymaga dwóch komentarzy.

(1) Jesteśmy przekonani, że jest jakieś *jak to jest być* dla człowieka, psa, nietoperza. Chcąc uformować ujęcie owego *jak to jest być* nietoperzem musielibyśmy przyjąć punkt widzenia nietoperza¹¹. Przedstawiając argumentację Nagla H. Kornblith zauważa, że nie należy mieszać faktu, iż jest coś takiego jak bycie nietoperzem, z twierdzeniem o pojęciowym poznaniu nietoperza przez samego siebie. Nietoperzowi brakuje pojęć, które my posiadamy, dlatego nie wiemy jak rozumie on siebie i swoje własne przeżycia (por. Korn-

¹⁰ Nagel 1997: 210. „Ów organizm ma jednak z istoty świadome stany mentalne wtedy i tylko wtedy, gdy jest coś takiego, jak bycie tym organizmem – coś co jest jakby dla tego organizmu. Możemy to nazwać subiektywnym charakterem doznań” (s. 204).

¹¹ Według Nagla: „Chcę raczej stwierdzić, że nawet po to, by wytworzyć sobie jakieś pojęcie tego, jak to jest być nietoperzem (a tym bardziej wiedzieć jak to jest być nietoperzem), trzeba przyjąć punkt widzenia nietoperza” (Nagel 1997: 211). Nagel zaznacza, że wybrał nietoperze, a nie osy czy flądry, bo „[...] gdy zbyt nisko schodzi się ku dołowi drzewa filogenetycznego ludzie w ogóle tracą wiarę, że w ogóle można tam mówić o doznaniach” (s. 206). Pewien gatunek nietoperzy, *Rousettus aegyptiacus*, wykazuje orientację przestrzenną obydwu rodzajów: wzrokową [przed zmrokiem] i echolokacyjną [w jasności]. Według Szafera, badania nie wykazały u nietoperzy zmysłu odróżniania barw. U nietoperzy *Microchiroptera* w siatkówce oka występują tylko pręciki, brakuje natomiast czopków. Pręciki mają umożliwić dostrzeganie różnicy w natężeniu światła, tj. różnic w odcieniach szarości, czopkom zaś przypisuje się rozpoznawanie barw pstrych (por. Szafer 1969: 274).

blith 1998: 52). W swojej interpretacji zwrotu *jak to jest być* Kornblith odwołuje się do jedności świadomości (a nie na przykład do samoświadomości), co odpowiada jej holistycznemu rozumieniu.

(2) N. Malcolm zaznacza, że Nagel przyjmuje argumentację L. Wittgensteina: pojęcia dotyczące zdarzeń mentalnych nie odnoszą się do „logicznie prywatnych obiektów zdawania sobie sprawy”, lecz do „subiektywnych punktów widzenia”, typowych dla danego rodzaju organizmów¹². Na podstawie tego, iż sami należymy do rodzaju ludzkiego, możemy ująć punkt widzenia innego przedstawiciela tego rodzaju, nie możemy jednak ująć punktu widzenia charakterystycznego dla przedstawiciela całkiem innego rodzaju organizmów.

Przyjmując (1) i (2) możemy powiedzieć, że warunkiem wystarczającym istnienia przeżyć *jak to jest być* w stanie *x*, jest ich związanie z perspektywą pierwszoosobową. Z jednej strony cechy przeżyć ukazują się jako monadycznie-atomistyczne i nieanalizowalne. Z drugiej strony wykazują one pewną przestrzenno-czasową, złożoną organizację. Dlatego można je analizować w ich zjawiskowych i psychologicznych aspektach. Wspomniana złożoność występuje jednak w ramach subiektywnej jedności. Nagel pisze:

„Na poziomie czysto mentalnym możemy zrozumieć, czym są «składniki» procesu umysłowego. Jesteśmy w stanie zanalizować dźwięk, zapach lub uczucie, tak aby uwyżnić jego aspekty fenomenologiczne albo psychologiczne. Możemy zanalizować pole percepcji wzrokowej albo nasze doznania cielesne za pomocą terminów, które mówią o relacjach przestrzennych, ale są przy tym *fenomenologiczne*. Wszystkie te złożoności występują jednak w ramach subiektywnej jedności” (Nagel 1997c: 64).

Według Nagla, nie wolno nam się odwoływać do tego rodzaju fenomenologicznych modeli, gdy chcemy zrozumieć obiektywną złożoność przestrzenną. Nagel nie odpowiada na pytanie, czy na tak rozumianej subiektywnej jedności możemy budować rozumienia obiektywnej złożoności.

Typy holistycznych świadomości są indywiduowane na podstawie zjawiskowych własności. W tym jednoznacznym sensie, nie możemy całkowicie zrozumieć, *jak to jest być* nietoperzem, ponieważ nie posiadamy chociażby echolokacyjnych przeżyć, jakie mają nietoperze. Jeśli założymy, że holi-

¹² Por. Malcolm 1995: 120. Załóżmy na przykład, że mam lęk wysokości. Lęk ten nie jest tylko mój, inni także mogą go mieć. Lęk, który posiadam, nie jest „logicznie prywatnym obiektem”. Wyłącznie moim w tym lęku jest mój punkt widzenia.

styczne rozumienie świadomości musi znaleźć wyjaśnienie dzięki fizykalnej teorii umysłu, to musimy uznać, że żadna z dostępnych obecnie teorii umysłu nie pokazuje nam, jak można by to zrobić.

Ten antyfizykalistyczny wniosek dotyczący świadomości możemy wyprowadzić porównując Nagla twierdzenia o subiektywności ujęć faktów mentalnych z jego twierdzeniami o generalnej obiektywności, gdy chodzi o ujęcia faktów fizycznych. Fakty fizyczne należą do sfery faktów obiektywnych i mogą być ujmowane z wielu punktów widzenia i w dodatku przez istoty wyposażone w odmienne systemy percepcyjne. Fakty o holistycznym przeżywaniu „*jak to jest* być szczególnym rodzajem świadomego organizmu” nie mogą zostać ontologicznie zredukowane do jakiegoś rodzaju faktów fizycznych, bez utracenia swojej subiektywności.

Nagel zakłada, że przeżycia typu *jak to jest być* zależą w sposób konieczny od stanów mózgu, lecz przy dzisiejszym stanie wiedzy nie mogą one zostać z tymi stanami utożsamione. Poprawnie funkcjonujący mózg jest – według Nagla – faktycznym medium dla znanych nam przeżyć psychicznych ciągłości. Nawet gdybym utracił całe moje ciało z wyjątkiem nienaruszonego mózgu, to pozostałbym nadal sobą. Można zadać pytanie: jeżeli za pomocą jakiegoś spektakularnego postępowania genetycznego dałoby się zbudować mózg, który nigdy nie był częścią istoty żywej, to czy byłby on indywidualnym podmiotem?

Nagel nie analizuje pojęcia „jaźni”, lecz stawia teoretyczno-empiryczną hipotezę o naturze tego pojęcia i twierdzi, że jest ono otwarte na uzupełnienie. *Ja* jest trwałą pojedynczą rzeczą w obiektywnym porządku natury, która leży u podstaw subiektywnych doświadczeń ciągłości wewnętrznego życia psychicznego, które nazywamy naszym własnym życiem. Tak rozumiana obiektywna tożsamość umożliwi odpowiedź na pytania o tożsamość *ja* tylko wtedy, gdy poszukiwane pojedyncze *ja* jest zarówno *nosicielem* stanów wewnętrznych, jak i *przyczyną* tego, że przeżywamy je jako ciągłe.

Zamiast zakładanej przez Nagla koniecznej zależności pomiędzy stanami mózgu a jaźnią przyjmujemy wstępnie tezę o korelacji pomiędzy mózgiem i jaźnią, rozumianą w sensie synchroniczno-emergentystycznym (por. rozdział VII). Spełniający warunki korelacji mózg jest nieodłącznym elementem jaźni, o której zakładamy, że cechuje się holistyczną jednością świadomości. Jedynie w tym sensie stwierdzamy zgodnie z Naglem:

V.3 Funkcjonujący mózg, według Nagla, jest konieczny dla istnienia holistycznej świadomości. Reszta ciała jest w sposób empiryczny powiązana w jedną całość.

Według Nagla, nie jestem tylko mózgiem: ważę więcej aniżeli 1,5 kg, jestem wyższy aniżeli 15 cm, posiadam szkielet, mięśnie itd. A jednak nieuszkodzony mózg jest jedyną częścią mnie samego, której zniszczenia bym prawdopodobnie nie przeżył. Nagel dodaje, że tylko mój mózg, a nie pozostałe części organizmu, jest *konieczny* i istotny dla pojawienia się mojej jaźni.

Ze świadomą przesadą Nagel stawia hipotezę, że ja *jestem* moim mózgiem (por. Nagel 1997c: 51), pozostawiając bez odpowiedzi pytania o to, co w przypadku wątpliwości mogłoby być uważane za ten sam mózg, oraz pytania o zależność mózgu od reszty organizmu. Okazuje się, że nieuszkodzony i funkcjonujący mózg jest odpowiedzialny zarówno za pamięć i za wszystkie pozostałe psychiczne ciągłości, jak i za jedność subiektywnej świadomości. Jeżeli ponadto stany psychiczne są stanami mózgu – a mózg tym samym nie jest żadnym czysto materialnym systemem – to także wtedy jest on poważnym kandydatem do bycia jaźnią. Według Nagla, jest tak nawet wówczas, gdy mózg nie spełnia wszystkich intuicyjnych warunków obecnych w naszym dzisiejszym wyobrażeniu jaźni.

Wychodząc z pierwszoosobowego punktu widzenia Nagel tylko sporadycznie odwołuje się do analiz świadomości prowadzonych w ramach badań neurofizjologicznych. Swoisty wyjątek stanowi kwestia jedności świadomości, którą omawia w artykule *Rozszczepienie mózgu i jedność świadomości*¹³. Przedstawiając przypadki rozdwojonych mózgów Nagel nie zaczyna od analizy jedności świadomego umysłu, lecz posługując się jego potocznym wyobrażeniem stosuje to wyobrażenie bezpośrednio do interpretacji tychże przypadków i pyta: czy badani pacjenci mają jeden umysł czy dwa?¹⁴ Nagla interesuje *explicite* to, co w przypadkach rozszczepionych mózgów dzieje się z naszym pojęciem indywidualnego umysłu. *Implicitie* jego rozważania dotyczą tematyki związanej z tezą o holistycznej jedności świadomości.

Nagel stwierdza, że gdy badane podmioty zostają poddane odpowiednim eksperymentom, to pojawiają się w nich równocześnie stany, które nie mogą zachodzić w jednym umyśle (na przykład ich uwaga kieruje się równocześnie na dwa zadania). Jego zdaniem:

¹³ Por. także Nagel 1998: 36–38. Por. też Hurley 1998: 164–200; Churchland 1993: 174–200.

¹⁴ Argumentację za „jednolitością świadomości pacjentów po komisurotomii” przedstawia M. Grabowska (2000: 424–426). Grabowska omawia również metody badania jedności świadomości, którym poddaje się tychże pacjentów (por. s. 402–408).

„W tej sytuacji trudno jest pojąć, *jak to jest być* jedną z tych badanych osób. [...] Brak interakcji w sferze doświadczenia wzrokowego i świadomego zamiaru zagraża założeniom o jedności świadomości, podstawowym dla naszego rozumienia innej jednostki jako osoby” (Nagel 1997a: 197).

Nagel zaznacza, że nasze potoczne założenia o jedności świadomości sprawdzają się u wspomnianych pacjentów w obrębie jednej półkuli, chociaż nie da się przeprowadzić odpowiednich porównań, przekraczających rozdział między półkulami. Z tego względu jesteśmy skłonni do przyjęcia hipotezy, że pacjenci ci mają po dwa umysły. Hipoteza o dwóch umysłach ma tę przewagę, że pozwala zrozumieć *jak to jest być* pacjentem z przeciętym spoidłem wielkim dopóty, dopóki nie staramy sobie wyobrazić, *jak to jest być* każdym z tych pacjentów przy jednoczesnym uwzględnieniu obydwu półkul. Zdaniem Nagla, z hipotezą tą trudno się jednak zgodzić. Na przeszkodzie stoi bowiem obserwowalna w życiu codziennym jedność w zachowaniach wspomnianych pacjentów. Dopiero podczas odpowiednich eksperymentów zachowują się oni nietypowo.

Podając próbę zrozumienia tych pacjentów w kategoriach mentalistycznych traktujemy samych siebie jako wzór jedności psychicznej. Czy wolno nam jednak rzutować dane z naszego życia mentalnego na ich życie mentalne? Używając siebie samych jako kryterium rozstrzygającego, czy o innym podmiocie można orzec, że posiada on cechującą się jednością świadomość, zapominamy – zdaniem Nagla – o ewentualności, że nasza własna jedność świadomości, „[...] może nie być niczym absolutnym a tylko innym przypadkiem integracji, mniej czy bardziej skutecznej, w systemie kontrolnym złożonego organizmu” (Nagel 1997a: 201).

Zwracamy się do siebie w pierwszej osobie, co pozwala nam myśleć o naszym organizmie jako o (w pewnym sensie) numerycznie absolutnie jednym, a nie o czymś będącym skutkiem integracji tworzących nas treści. Gdy naturalne pojęcie cechującej się jednością świadomości zostaje odniesione do nas samych, popada ono w konflikt z niektórymi faktami neurologicznymi. Przyrodki rozszczepionych mózgów dostarczają nam tego rodzaju faktów.

Reasumując możemy powiedzieć. To, czym ja sam (JB) jestem, jest w gruncie rzeczy miejscem wrażeń i doświadczeń osoby JB, moich zdolności do identyfikowania i reidentyfikowania mnie samego w pamięci i w przeżywanym. Czynnikiem podstawowym dla rozumienia siebie jako osoby jest przeżywana jedność świadomości.

Chcąc siebie identyfikować z samym sobą, nie muszę się odwoływać do danych obserwacyjnych, do których musi się odwoływać ktoś inny, aby móc ująć mnie jako osobę. „Ja jestem osobą” domaga się tego, że posiadam tę zdolność do identyfikacji jako daną. Nagłowskie „ja jestem osobą” nie doma-

ga się natomiast wiedzy, że to mój mózg umożliwia zdolność do samoidentyfikacji. W rzeczywistości nie posiadam szczegółowej wiedzy o tym, co jest za tę zdolność odpowiedzialne, a inni nie potrzebują tego znać, aby móc wiedzieć, że (ja, JB) jestem daną osobą.

W normalnych warunkach mózg spełnia zarówno wzajemnie jednoznaczny warunek dla istnienia świadomości, jak i zasadę „wszystko albo nic” rozumianą w sensie, że poważnego uszkodzenia naszych mózgów prawdopodobnie byśmy nie przeżyli. Mając do dyspozycji nawet wszystkie neurofizjologiczne dane o mózgu danej osoby, nie znalibyśmy jej treści przeżyciowych. Dopiero przyjęcie jej punktu widzenia pozwoliłoby doświadczyć *jak to jest* być nią. Czysto zewnętrzna analiza cech fizycznych mózgu nie będzie pełna oraz wystarczająca i należy ją uzupełnić o wewnętrzną analizę w ramach określonej perspektywy. Nagel oddziela opisy tego, co mentalne (związane z określoną perspektywą), od trzecioosobowych opisów mózgu.

5.3. Argumenty przeciwko fizykalizmowi

W pracach Nagla znajdujemy dwa rodzaje rozumowań w obronie tez: (a) subiektywność jest nieredukowalną cechą rzeczywistości, (b) fizykalnie obiektywne rozumienie rzeczywistości jest nieadekwatne ujęciu subiektywności i stawia samo siebie pod znakiem zapytania.

Rozumowanie pierwszego rodzaju, które możemy nazwać „pośrednim”, posiada następującą strukturę:

(a) Aby fizykalnie obiektywny punkt widzenia w ogóle mógł zostać sformułowany, potrzeba subiektywnej aktywności.

(b) Jeżeli tego rodzaju aktywność występuje, to fizykalnie obiektywny punkt widzenia nigdy nie może być całkowicie obiektywny.

(c) Ponadto sama decyzja, aby na daną rzecz spojrzeć fizykalnie obiektywnie, jest subiektywna.

(d) Tym samym spojrzenie obiektywne dopuszcza subiektywny element, chociaż dopuszczanie takie neguje¹⁵.

Argument powyższy nie podważa redukcjonizmu, chociaż uwidoczniła niespójność w ramach fizykalnie obiektywnego punktu widzenia. Nagel chce w nim jedynie pokazać pragmatyczną sprzeczność występującą w ramach fizykalnej obiektywności i materialistycznej ontologii.

¹⁵ Nagel pisze: „Subiektywność świadomości jest nieredukowalną cechą rzeczywistości, bez której nie moglibyśmy uprawiać fizyki ani niczego innego, i musi zajmować równie podstawowe miejsce w każdym wiarygodnym obrazie świata, jak materia, energia, przestrzeń, czas i liczby” (Nagel 1997c: 12).

Antyfizykalistyczne rozumowanie drugiego rodzaju można nazwać „bepośrednim” i wyodrębnić kilka jego powiązanych ze sobą form (por. 5.3.1, 5.3.2).

5.3.1. Nieredukowalność przeżyć

Zdaniem Nagla, redukcja (wyjaśniająca) zjawisk mentalnych do fizycznych musi przebiegać inaczej niż znane redukcje w fizyce czy chemii. Z redukcji zjawisk mentalnych nie można bowiem w taki sposób wykluczyć subiektywnego doznania, w jaki wyklucza się zjawiskowe cechy w redukcji fizycznej czy chemicznej – wyjaśniając je jako efekty oddziaływania przedmiotów na nasze zmysły. Według Nagla:

„Jeżeli ma się bronić fizykalizmu, trzeba podać fizykalne przedstawienie owych cech fenomenologicznych. Gdy jednak uwzględnimy w nich subiektywny charakter, rezultat taki wydaje nam się niemożliwy” (Nagel 1997: 205–206).

Trudność z redukcją psychofizyczną wynika stąd, że w innych dziedzinach nauki redukcja jest postępowaniem ku bardziej obiektywnemu ujęciu faktycznej natury przedmiotów (por. Nagel 1997: 213).

Szkic podanego przez Nagla argumentu przedstawia się następująco:

(a) Organizm posiada świadome stany mentalne wtedy i tylko wtedy, gdy istnieje owo subiektywne „jak to jest być tym organizmem” w sensie *jak to jest* być dla tego organizmu (por. Nagel 1997: 207). *Jak to jest być x* może, ale nie musi być dostępne dla innych świadomych organizmów.

(b) Redukcyjne traktowanie stanów mentalnych jest logicznie do pogodzenia z nieobecnością (dla doznającego podmiotu) subiektywnych przeżyć. Tym samym jest logicznie możliwe, iż jakiś stan może odgrywać przyczynową rolę, która dla podmiotu jest definitywnie bólem, a zarazem nie jest doznaniem czegoś subiektywnego i dlatego nie jest stanem świadomego zdawania sobie sprawy.

(c) Jednak przedstawienie przeżyć, które nie uwzględnia ich subiektywnego charakteru opuszcza coś istotnego (i jest dlatego nieadekwatną podstawą dla redukcji zjawisk mentalnych do fizycznych)¹⁶.

Drugi argument dotyczy wewnętrznego rodzaju doznań i ma postać¹⁷:

¹⁶ Według Ch. J. Maloneya: „Still, one might argue, with Nagel and Jackson, that it must be incorrect to identify qualia and physical types since an exhaustive understanding of the physical would still leave us in the dark about the phenomenal” (Maloney 1985: 35).

¹⁷ „W przypadku doznań związek ze szczególnym punktem widzenia wydaje się dużo ścisłejszy. Trudno zrozumieć, na czym miały polegać obiektywny charakter doznań, poza szczególnym punktem widzenia, z którego ujmuje je ich podmiot” (Nagel 1997: 212).

- (a) Mogę bezpośrednio poznać jakości moich doznań.
- (b) Nie mogę w ten sposób poznać cech moich stanów mózgu.
- (c) Stąd: jakości moich doznań \neq cechy moich stanów mózgu.

„Wewnętrzność” (*insideness*) jakości jest dla Nagla emergencyjną cechą fizycznej organizacji mózgu. Nagel nie przyjmuje istnienia kartezjańskiej, metafizycznej, mentalnej substancji, bowiem przyjęcie takiej substancji nadal pozostawia nas z pytaniem o istnienie specyficznej perspektywy (a tym samym subiektywności i jakości doznaniowych). Proponując swoją teorię dwóch aspektów (por. 5.4; 5.4.2) [por. Nagel 1965: 355] Nagel odrzuca monistyczną teorię identyczności oraz postulującą dwie substancje dualizm.

Trzeci argument Nagla ma następującą formę: chociaż nie wiadomo jak dużo byśmy wiedzieli o neurofizjologii nietoperza i jego interakcji z fizycznym otoczeniem, to nie będziemy wiedzieć ani nie potrafimy sobie przedstawić, *jak to jest* być nietoperzem. Argument ten jest znany w literaturze jako argument „o brakujących qualiach” (*absent qualia argument*)¹⁸. Należy stwierdzić, że istotą przekonania, iż nietoperze mają takie czy inne doznania, jest myśl o istnieniu holistycznej świadomości, do której odnosi się zwrot *jak to jest* być nietoperzem. Postrzeganie świata i siebie przez nietoperza jest całkowicie różne od tego, które my sami posiadamy lub które potrafimy sobie wyobrazić.

5.3.2. Argument z poznania

Podobnie do Nagla argumentuje F. Jackson w swoim eksperymencie myślowym z Panią Mary. Mary przebywa od urodzenia w pomieszczeniach z czarnymi, szarymi lub białymi przedmiotami (włącznie z kolorem jej skóry)¹⁹. O świecie zewnętrznym Mary dowiaduje się poprzez czarno-białe monitory. *Ex hypothesi* wie ona wszystko o neurofizycznej strukturze ludzkiego systemu wzrokowego oraz o jego funkcjonowaniu. Mary nie posiada jednak przeżyć ludzi żyjących poza jej czarno-białym otoczeniem oraz nie wie nic o czekających ją przeżyciach, gdy opuści swoje otoczenie. Mary nie wie nic o tym *jak to jest* widzieć czerwony pomidor, *jak to jest* mieć doznanie czegoś czerwonego itd. Innymi słowy: jej dotychczasowe rozumienie kolorów w terminach odnoszących się do długości fal świetlnych, do stymula-

¹⁸ Nagel pisze dosłownie: „[...] the subjective character of experience is not captured by any of the familiar, recently devised reductive analyses of the mental, for all of them are logically compatible with its absence” (Nagel, *What it is like to be a bat?*: 166).

¹⁹ Jackson 1982: 127–136. Od strony empirycznej Mary odpowiada np. osobie z wrodzoną wadą systemu wizualnego, którą to wadę udało się usunąć.

cji siatkówki i neurotransmiterów nie wystarczy, ażeby ująć istotne jakościowe cechy przeżyć kolorów. Argument Jacksona ma postać:

- (a) Mary wie wszystko, co można wiedzieć o stanach mózgu i ich cechach.
- (b) Mary nie wie wszystkiego, co można wiedzieć o odczuciach i ich cechach.
- (c) stąd na mocy zasady Leibniza: odczucia i ich cechy, stany mózgu i ich cechy.

Jackson pyta: czy Mary po opuszczeniu swojego pomieszczenia dowie się czegoś nowego o ludzkim spostrzeganiu oraz o otaczającym ją świecie? Wydaje się, że tak. Mary zdobędzie bowiem informacje typu: „*jak to jest widzieć dojrzałego pomidora*”. Świadczy to jednak o tym, że jej poprzednie poznanie nie było kompletne, pomimo że dysponowała pełną informacją na temat ludzkiego mózgu, w tym także na temat zachodzących w nim procesów skorelowanych z powstawaniem kolorów. Stąd wniosek: fizykalizm jest teorią niezupełną, gdyż nie potrafi podać adekwatnego, redukcjonistycznego przedstawienia niektórych zjawisk mentalnych. Fizykalistyczna teoria mózgu, która nie uwzględnia przeżyć jakościowych, nie ujmuje tym samym wszystkiego, co jest możliwe do poznania, gdy mówimy o umyśle. Argument Jacksona znany jest jako „argument z poznania” (*knowledge argument*).

Podobny argument można rozwinąć, gdy chodzi o zmysły smaku, dotyku, czy też inne stany mentalne zawierające qualia. Konkluzja jest ta sama: qualia pozostają poza teorią fizykalną. Stąd Jackson wnioskuje, że są one epi-fenomenami: (EP) „Wszystko co chcę wykazać to to, że jest możliwe utrzymywanie, że niektóre własności niektórych stanów mentalnych, mianowicie te, które nazywam qualiami, są takie, że ich posiadanie lub nieposiadanie nie czyni różnicy w świecie fizycznym” (Jackson 1982: 133).

Jackson konstruuje qualia jako cechy pewnych stanów mentalnych. Qualia nie są przyczyną niczego fizycznego, chociaż same są zdeterminowane przez coś fizycznego.

Dualiści wyjaśniając powyższą kwestię przyjmują, że będąc świadomi naszych przeżyć żyjemy się z ich нефizycznymi cechami. W ten sposób zdobyta wiedza uzasadnia nasze pierwszoosobowe sposoby pojmowania owych świadomych przeżyć. Dochodzimy do tego „*jak to jest widzieć czerwone*” i uczymy się wyobrazić sobie ten kolor.

Jak teorie fizykalistyczne wyjaśniają fakt, że jakości przeżycia można sobie wyobrazić tylko wtedy, gdy samemu się je miało? Załóżmy, że wyobrażenie jakiegoś przeżycia jest dosłownym odtworzeniem w tym sensie, iż w wyobrażeniu zostaną zaktywowane te same części mózgu, jak w przeżyciu pierwotnym (choć z aktywowanie nie jest w widoczny sposób identyczne

z przeżyciem pierwotnym, a szczególnie brak mu normalnych połączeń z zachowaniami). Coś można sobie wyobrazić tylko wtedy, gdy się to, albo jego elementy, najpierw przeżyło (mianowicie: aby móc stworzyć kopie odpowiedniego stanu, mózg potrzebuje oryginału tego stanu). Proponowane przez fizykalistów wyjaśnienie odrzuca tym samym epifenomenalistyczno-dualistyczne założenie Jacksona, że w przypadku przeżywania danego stanu mentalnego jest się obeznanym ze szczególnymi, нефizycznymi cechami stanów świadomych.

Założenie, że wyobrażone przeżycie „odczuwa” się tak samo jak odczuwa się przeżycie pierwotne, fizykalista wyjaśnia poprzez odwołanie się do bezpośredniego udziału fizycznej kopii pierwotnego przeżycia w procesie wyobrażania. Spornym punktem pozostanie to, czy alternatywne dualistyczne wyjaśnienie, odwołujące się do нефizykalnych własności przeżycia, jest konkurencyjnym wyjaśnieniem tego, że jakieś wyobrazić-sobie odczuwa się w ten sam sposób, jak przeżycie pierwotne. Mówiąc ogólnie, owo coś sobie wyobrazić (jako wydarzenie z określoną cechą) może mieć jakiś dowolny charakter przeżyciowy albo żaden.

Nadal pozostaje jednak pytanie: dlaczego fizyczne podobieństwo dwóch stanów mózgu powinno powodować podobieństwo stanów zjawiskowych? (do kwestii tej wrócimy w 7.1.2). Fizykaliści zakładają najczęściej, że wyobrażanie sobie jest podobne do przeżycia pierwotnego, obojętnie, czy jest to wyjaśnialne przez fizyczne podobieństwo obydwu, czy też nie.

Argument z *Mary* przypomina argument C. D. Broada z „matematycznym archaniołem” (por. 3.6.2). *Mary* ma do dyspozycji propozycjonalną wiedzę, brakuje jej jedynie „wiedzy” przeżyciowej o tym, *jak to jest* być w jakimś stanie przeżyciowym. Archanioł Broada nie posiada wiedzy propozycjonalnej o tym, że określone neuronalne zdarzenia łączą się z określonymi przeżyciami jakościowymi. Według Broada, tego rodzaju wiedza nie może być – z zasady – wyprowadzona z faktów neurologicznych. Gdyby ktoś takiemu archaniołowi zdał sprawozdanie o kolorach, to wtedy posiadałby on nieprzeżyciową wiedzę o tym, jak wygląda pomidor. Gdyby mógł go sam (w „ludzki sposób”) zobaczyć, to wówczas posiadałby także przeżyciową wiedzę. Obydwa rodzaje wiedzy nie są mu jednak dostępne, dopóki zna tylko zdarzenia fizjologiczne.

5.3.3. Fizykalistyczno-redukcyjnistyczne kontrargumenty

Argumenty Nagla i Jacksona były i nadal są szeroko komentowane i krytykowane²⁰. Mówiąc ogólnie, ontologiczno-redukcyjnistyczni materia-

²⁰ Por. Chalmers 1996: 81–84; Dennett 1991: 398–401; Horgan 1984a: 150–153; Levin 1986: 246–245; Nemirow 1990: 492–499.

liści zakładają, że omawiane przez Nagla cechy zjawiskowe to na przykład: obiektywna czerwień róży, odczuwane ciepło. Przyjmują ponadto, że cechy te nie zostają wykluczone z naszych redukcji. Czerwień, jako obiektywna cecha róży jest identyczna z pewną długością odbitej fali elektromagnetycznej, ciepło (jako obiektywna, zjawiskowa własność makroobiektów), jest identyczne ze średnią energią kinetyczną cząstek tworzących dany obiekt. Elektromagnetyczne i mikromechaniczne cechy w świecie obiektywnym są faktycznie cechami zjawiskowymi i pomimo że przy spostrzeganiu ignorujemy ich dynamiczne i mikroskopijne detale, są to te obiektywne fizyczne cechy, z którymi są powiązane nasze mechanizmy spostrzegawcze. Same redukcje są rozumiane w tym sensie, że ktoś może całkowicie zastąpić części potocznego słownictwa dotyczącego obserwowalnych cech i uczyć się ujmować swoje percepcyjne sądy bezpośrednio w terminach redukującej teorii.

W ramach teorii redukjonistycznych przyjmuje się, że spostrzegane cechy nie są jakościami „drugorzędnymi” (w tym sensie, że nie istnieją one poza umysłem ludzkiego obserwatora). Przeciwnie, cechy te są obiektywne i posiadają różnorodne obiektywne własności przyczynowe. Zdaniem redukjonistów, błędem byłoby próbować umieścić cechy zjawiskowe wewnątrz umysłu, gdyż przesunęłoby to tylko problem znalezienia dla ich miejsca w naturze.

Przykładem redukjonistyczno-materialistycznej krytyki subiektywności w rozumieniu Nagla są tezy podane przez P. M. Churchlanda, które przedstawimy w dwóch punktach.

(A) „Obiektywne” qualia (czerwień, ciepło) nie powinny zostać „sprawdzone” do umysłu obserwatora, lecz powinny być zredukowane tam, gdzie one występują, czyli na zewnątrz umysłu²¹ (pozostaje oczywiście pytanie, jaki sens ma tutaj przestrzenna lokalizacja stanów świadomości oraz ich cech)²². Podobnie mogą być traktowane „subiektywne” qualia (to znaczy mogą być zredukowane tam, gdzie występują – wewnątrz obserwatora). W tym sensie przypadki „zewnątrznej” i „wewnętrznej” redukcji nie są od siebie różne, lecz przebiegają równolegle.

²¹ Churchland mówi o „objective phenomenal properties” (jak widziane kolory), które powinny być „[...] confronted where they stand, and they should be reduced where they stand: outsider the human observer” (Churchland 1985: 19). Widzimy tutaj próbę naturalizacji „obiektywnych qualiów”, odwołując się do naszych realistycznych, potocznych przekonań o tym, co widzimy w otaczającej nas przestrzeni.

²² Wychodząc od tradycji F. Brentano, można dodać, że Churchland nie rozróżnia pomiędzy aktami świadomości z przedmiotami tych aktów, a także pomiędzy cechami aktów świadomości a cechami przedmiotu aktu świadomości.

P. M. Churchland, w odróżnieniu od D. Dennetta, nie neguje istnienia jakościów, odmawia im jednak samodzielności (niezależności). Churchland broni teorii identyczności twierdząc, że postęp naukowy doprowadzi do wyjaśnienia subiektywności świadomości. Świadomość wydaje się wprawdzie być tajemnicza i niedostępna dla współczesnej nauki, ale stan ten jest wynikiem braku odpowiedniej wiedzy o mózgu i zachodzących w nim procesach.

Największe braki, według Churchlanda, w naukowym podejściu do świadomości znajdujemy w opisach pierwszoosobowego przedstawienia przeżyć. Tego rodzaju opisy są używane do zdawania sprawy z subiektywnego przeżycia i do dostarczania pojęciowych ram w obiektywnych sprawozdaniach o świadomości, które to sprawozdania mogą być następnie porównywane w terminach obiektywnych. Każdy podmiot może subiektywnie obserwować swoje stany mentalne i posiadać swoiste przyczynowe połączenie z tymi stanami. Chociaż dany podmiot zna w jakiejś mierze neurologiczną kondycję swojego mózgu, która jest niezależna od każdej obiektywnej obserwacji, nie oznacza to, że dostęp do subiektywnego, pierwszoosobowego zjawiska musi być нефизыkalnej natury.

Zdaniem Churchlanda, eksperymenty myślowe mające potwierdzić nieredukowalność świadomości do stanów mózgu potwierdzają jedynie przednaukowy sposób poznawania stanów mentalnych obecnych w podmiocie. Eksperymenty te nie pokazują natomiast, że stany te są нефизыkalne albo niemożliwe do wyjaśnienia przez odwołanie się do stanów fizycznych. Ponieważ owe eksperymenty zdają się uwzględniać różnicę pomiędzy stanami wewnętrznymi a stanami fizycznymi, są one zdaniem Churchlanda argumentami za dualizmem (por. Churchland 1997: 224–235).

Odpowiadając na redukcjonistyczne argumenty Nagel zaznacza, że każdy redukcjonista ma „[...] swoją ulubioną analogię wziętą ze świata nauki” (Nagel 1997: 204). Tego rodzaju analogie skłaniają do szukania podobieństw pomiędzy doznaniem bólu i bólem, z jednej strony, a ciepłem i doznaniem ciepła – z drugiej. Według Nagla, takie rozróżnienie w przypadku bólu nie istnieje, ponieważ ból nie jest jedną rzeczą, a doznanie bólu drugą. Ból jest rodzajem doznania posiadającym swoiście jakościowy charakter. Jego argument za nieredukowalnością bólu wprowadza na nowo kwestię „punktu widzenia”. Subiektywny charakter doznań osoby jest poznawalny tylko z punktu widzenia istotnie podobnego do tego, który zajmuje doznająca bólu osoba. Gdy zaś chodzi o podmioty inne aniżeli osoby ludzkie, to Nagel stwierdza: „Namysł nad tym, jak to jest być nietoperzem, prowadzi nas zatem do wniosku, że istnieją fakty, z których nie zdają sprawy zdania wyrażalne w ludzkim języku” (Nagel 1997: 210).

Nie wiemy, w jaki sposób fakty o konkretnym punkcie widzenia przedstawiać w terminach fizykalnych, gdyż do istoty teorii fizykalnych należy, iż przedstawiając świat rezygnują z odnoszenia się do punktu widzenia.

(B) W *Knowing Qualia: A Reply to Jackson Churchland* formułuje argument Jacksona jak poniżej, a następnie pokazuje jego niepoprawność:

(a) Przed opuszczeniem swojego pomieszczenia Mary wie – od strony neurologicznej – wszystko o ludzkim spostrzeganiu.

(b) Mary nie wie jednak wszystkiego o całym ludzkim spostrzeganiu.

(c) Dlatego: istnieją prawdy, które umykają fizykalizmowi i proponowanemu przez niego opisom oraz wyjaśnieniom.

Churchland opisuje możliwy system, dzięki któremu nasz układ wzrokowy dochodzi do rozpoznawania kolorów. W ludzkiej korze wzrokowej istnieje obszar znany jako V4. Gdy niemowlę pierwszy raz widzi kolor, to w warstwowej strukturze V4 tworzy się specjalny wzorec neuronowy, zawierający informację o widzianym kolorze (por. Churchland 1998: 145–147; 1997: 236–238). Powstały wzorec jest reaktywowany, gdy niemowlę kolejny raz widzi ten sam kolor. Podobne wzorce są tworzone podczas spostrzegania innych kolorów (czy innych długości fal świetlnych). Obszar V4, zdaniem Churchlanda, sam podzielił się na kilka regionów, z których każdy jest neuronalnym wzorcem reprezentującym odmienny kolor. Kiedykolwiek jakiś kolor jest przeżywany, synaptyczny wpływ sensorycznego impulsu reaktywuje w V4 odpowiednią etykietkę koloru. W ten sposób kora wzrokowa identyfikuje i „etykietuje” zarazem długość fali świetlnej, która jest obecna w polu wzrokowym.

Churchland dodaje, że etykiety rozpoznające kolor muszą powstać w okresie niemowlęctwa i wymagają do tego obecności kolorów. Ludzie muszą się uczyć, jak widzieć kolory. Mary posiada jedynie etykiety dla „kolorów” czarnego, białego i odcieni szarego, bo tylko te kolory na nią oddziaływały. W chwili, gdy opuści ona swój pokój, będzie już zbyt stara, aby utworzyć nowe wzorce w V4, dlatego nie będzie zdolna do zobaczenia jakiegokolwiek innego koloru. Jej mózg będzie przetwarzał docierające z zewnątrz impulsy w terminach tego, co już zna – „czarnego” i „białego”. Mary brakuje nie tylko jakiejś zdolności (sprawności), lecz czegoś o wiele więcej – odpowiedniej struktury neuronalnej.

W tym sensie, według Churchlanda, przesłanka (b) okazuje się prawdziwa. To znaczy jest coś, czego Mary nie wie i nie może wiedzieć o innych ludziach, pomimo że zna ich neurologię. Eksperyment z Mary zakłada, że posiada ona całkowitą wiedzę o procesach spostrzegania wzrokowego oraz że wie wszystko o tych procesach. Tym samym Mary powinna także wiedzieć o swoich brakach, opisanych powyżej, co oznacza, że przesłanka (a) w argu-

mencie Jacksona byłaby prawdziwa. Argumentacja Churchlanda jest logiczną możliwością ukazującą przebieg procesów neurologicznych przy widzeniu kolorów. Równocześnie należy stwierdzić, że:

V.4 P. M. Churchland negując argument z Mary odwołuje się do tzw. wzorca neuronowego (*neural pattern*). Znalazienie takiego wzorca pozostaje pytaniem empiryczno-teoretycznym.

Rozumowanie Churchlanda pokazuje możliwą sytuację, w której wniosek Jacksona nie wynika z jego przesłanek. Churchland opisuje fizycznie możliwy sposób funkcjonowania naszego systemu wzrokowego. Mary wyrosła w określonych warunkach, dlatego nie będzie zdolna do przeżywania kolorów (i to niezdolna z przyczyn neurologicznych). Tym samym Churchland jest przekonany, że powołując się na Jacksona argument z poznania, nie potrafimy zanegować fizykalistycznych tez o świadomości.

Zdaniem Nagla stany mentalne i fizyczne są jedynie kontyngentnie identyczne (por. 5.5.1). Chociaż świadome przeżycie $[\Psi_1]$ może wystąpić przy konkretnym i pojedynczym stanie neuronalnego systemu $[\Phi_1]$, łatwo można sobie wyobrazić to samo przeżycie powstające razem z innymi systemami $[\Phi_2]$, $[\Phi_3]$. Nagel twierdzi, że stany mentalne nie mogą być jednoznacznie opisywane w terminach fizykalnych, gdyż do takiego opisu potrzeba stwierdzenia koniecznej identyczności, o której będzie można mówić dopiero w jego poszerzonej teorii świadomości.

Churchlandowi nie udaje się zanegować tego argumentu. Wskazuje on jedynie na potrzebę obiektywnych opisów stanów mentalnych, nie podaje natomiast żadnej odpowiedzi na pytanie, jak powstają stany subiektywne. Broniąc teorii identyczności Churchland odrzuca Naglowe eksperymenty myślowe wychodząc z założenia, że obecny stan wiedzy w neurologii blokuje adekwatne zrozumienie świadomości. Postawioną przez Churchlanda hipotezę o nauce widzenia kolorów i związane z tą hipotezą pytanie o tworzenie neuronowych wzorców w warstwie V4 należy podporządkować pod program poszukiwania neuronalnych korelatów świadomości. Program takich poszukiwań, z uwzględnieniem związanych z nimi trudności, określają ogólne linie metodologiczne podane w rozdziale VII. Program ten nie odpowiada na pytanie Nagla o koniecznej identyczności stanów świadomości ze stanami fizycznymi.

5.4. Panpsychiczna teoria dwu aspektów

Nagel zakłada, że subiektywności nie można opisać ani wyjaśnić bez odwołania się do perspektywy pierwszoosobowej. Churchlandowi nie udaje

się zakwestionować tego założenia. Posiadanie perspektywy pierwszoosobowej uwzględnia naszą zdolność odnoszenia się do zjawiskowych cech jako takich. Na razie bez odpowiedzi pozostaje pytanie: czy jakieś inne okoliczności dostarczają nam możliwości odniesienia się do tychże cech? Nawet gdyby się okazało, że nie istnieje obiektywny opis subiektywnych faktów lub cech oraz że tego rodzaju fakty i cechy są na sposób konieczny związane z perspektywą pierwszoosobową, to Nagel nie traktuje tych twierdzeń jako antymaterialistyczne. Również nieredukcyjny materializm identyczności egzemplarzy (por. 3.4) przyjmuje, że nie wszystkie cechy są cechami fizycznymi.

Nagłowskie ujęcie cech zjawiskowych przeciwstawia się ujęciu funkcjonalistyczno-maszynowemu, według którego subiektywne cechy mentalne, a nie tylko indywidualne mentalne zdarzenia, jak w teorii identyczności egzemplarzy, są tożsame z jakimś rodzajem cech (powiedzmy cech obliczeniowych). Oznacza to większe ograniczenie funkcjonalizmu aniżeli ograniczenie, o którym mówi W. G. Lycan. Według Nagla, poznanie funkcjonalno-teleologicznego przedstawienia jakiegoś stanu subiektywnego nie pozwala samo z siebie na poznanie, *jak to jest* mieć ten stan. Problemem bowiem jest, że podanie funkcjonalnego opisu nie ułatwia odnoszenia się do cech zjawiskowych, ponieważ cechy zjawiskowe mogą nie być charakteryzowalne w terminach funkcjonalnych (podobnie jak cechy obliczeniowe nie są charakteryzowalne w terminach fizycznych).

Tym samym upada funkcjonalistyczna krytyka tezy Nagla o nieredukowalności subiektywności. Nie oznacza to jednak odrzucenia materializmu stwierdzającego identyczność zdarzeń (egzemplarzy). Zarówno Nagel, jak i Lycan dostrzegają ewentualność, że przyjęcie jakiegoś rodzaju wewnętrznie zjawiskowych (perspektywicznych) cech stoi w konflikcie z materializmem, jednak nie oznacza to przyjęcia przez nich postawy antymaterialistycznej. Naszym zdaniem wewnętrznie zjawiskowe cechy wymagają do ich ujęcia teorii synchronicznej emergencji (por. 3.6.2).

Zbudowana przez Nagla teoria świadomego umysłu należy do teorii dwóch aspektów: cechy i zdarzenia mentalne oraz cechy i zdarzenia fizyczne są nieredukowalnymi, odmiennymi aspektami jednej i tej samej rzeczywistości, która te obydwa aspekty ukazuje. Sama rzeczywistość nie jest ani mentalna, ani fizyczna. Teoria dwóch aspektów nie jest ani dualistyczna, ani idealistyczna, ani materialistyczna (ponieważ zawiera negację odrębnej substancji mentalnej i fizycznej). Z materializmem i idealizmem łączy ją ontologiczny monizm (istnieje tylko jeden rodzaj substancji).

Nagel nie podaje jakiejś spójnej teorii przyczynowych relacji pomiędzy tym, co mentalne, i tym, co fizyczne. Jego teoria dwu aspektów prowadzi do swoistej formy panpsychizmu, na co wskazują dwa punkty.

(1) Nagel uznaje, że rzeczywistość jest bogatsza od tej, którą dzisiaj ujmujemy za pomocą obiektywnych metod²³. Jest tak dlatego, że własności mentalne kompleksowego organizmu muszą wynikać z jakichś własności jego podstawowych komponentów. Te ostatnie zaś nie mogą być tylko własnościami fizycznymi, gdyż ich połączenia i kombinacje dostarczałyby jedynie innych własności fizycznych²⁴. Cechy protomentalne zdają się być wynikiem organizacji i wewnętrznych relacji materii²⁵.

(2) Tezą panpsychizmu jest, że materialne elementy rzeczywistości mają cechy protomentalne. Nagel nazywa wprawdzie tę tezę „denerwującą” (*unsettling*) i „zaskakującą” (*startling*), niemniej jednak twierdzi, że jeżeli odrzucimy psychofizyczny redukcjonizm i radykalne (synchroniczne oraz diachroniczne) formy emergencji, wtedy mózg musi mieć protomentalne cechy (por. Nagel 1997c: 62–64). Prowadzący do panpsychizmu argument Nagla przypomina panpsychistyczne rozważania Leibniza i posiada formę:

(a) Jeżeli świadomość jest nieredukowalna, to nie potrafimy jej wyjaśnić na odpowiednim poziomie fizycznej złożoności, odwołując się jedynie do terminów fizykalnych.

(b) Jeżeli ponadto świadomość nie emerguje z tego stopnia złożoności,

(c) to musi ona być obecna na stopniu fizycznym.

Panpsychizm był zawsze jakoś obecny na marginesie argumentacji naukowej i filozoficznej. Wraz z rozwojem naukowego wyjaśniania rzeczywistości tracił jednak na znaczeniu. Jedną z krytyk panpsychistycznej teorii świadomości podaje W. James (1950: 158–160). Panpsychista musi umieć wyjaśnić, w jaki sposób elementy „zatomizowanej świadomości” tworzą nową, kompleksową i zróżnicowaną świadomość. W jaki sposób poszczególne przeżyciowe jednostki mogą się złączyć w wyższą formę przeżywania? Problem z kombinacją poszczególnych jednostek świadomości, przed którym stoi panpsychizm, pro-

²³ „W obliczu tych faktów można by dojść do wniosku, że nie pozostaje nam nic innego, jak uznać, że fizykalna koncepcja obiektywności nie jest w stanie wyczerpać całej rzeczywistości” (Nagel 1997c: 21; por. 1997d: 228). Klasycznie rozumiany panpsychizm oraz wprowadzony przez Chalmersa „protopanpsychizm” przedstawia L. Gabora w swoim wywiadzie z S. Hameroffem (por. Gabora 2001: 121–122).

²⁴ Nagel pisze: „Jeżeli dowolny dwustufuntowy kawałek wszechświata zawiera materiał potrzeby do skonstruowania osoby, i jeśli odrzucimy redukcjonizm psychofizyczny oraz radykalną odmianę emergentyzmu, to każdy przedmiot, sprowadzony do swych podstawowych składników, musi mieć własności protomentalne” (Nagel 1997c: 62).

²⁵ „Wydaje się, choć jest to niezwykle, że pewne złożone, biologicznie wygenerowane fizyczne systemy (*certain complex, biologically generated physical systems*), których każdy z nas jest przykładem, mają liczne własności niefizyczne” (Nagel 1997c: 64; oryginał: 51).

wadzi do problemu generacji świadomości zjawiskowej. Problem ten jest formalnie analogiczny do problemu powstawania świadomości z materii, co oznacza, że w porównaniu z materializmem panpsychizm nie dostarcza dodatkowych argumentów wyjaśniających.

Wspomniany problem kombinacji możemy próbować ominąć przyjmując założenie, że elementarne jednostki umysłu posiadają jakąś mentalność, przy równoczesnej negacji, że są one faktycznie świadomymi przeżyciami. Tego typu założenie stawia nas na nowo wobec problemu generowania świadomości. Musimy bowiem opisać i wyjaśnić powstawanie świadomych przeżyć z kombinacji nieświadomych obiektów (choć są one w jakimś sensie obiektami mentalnymi).

Panpsychista może twierdzić, że dwa fizycznie identyczne systemy będą posiadały (dzięki tej identyczności) te same cechy mentalne. Dlatego fizycznie identyczne systemy będą się zawsze zachowywały w dokładnie podobny sposób, gdy cechy mentalne posiadają jakiś mentalny wpływ. Podobnie jak w teoriach emergencji pojawia się jednak wtedy trudne pytanie o cel postulowania jakiegoś mentalnego wpływu i o jego charakter. Dla materialisty przyjmującego, że fizyczne cechy systemu mogą wyjaśnić wszystkie jego zachowania, cechy mentalne będą (przynajmniej od strony wyjaśniającej), cechami epifenomenalnymi.

Powyższe zarzuty wobec panpsychizmu można, zdaniem J. Perzanowskiego, odeprzeć powołując się na ontologię G. Leibniza²⁶. Leibniz, według Perzanowskiego, jest autorem współczesnego pojęcia informacji, co doprowadziło go do podstawowej intuicji, że świat ma naturę mechaniczną, dynamiczną i informatyczną. Nośnikami podstawowych porcji informacji są monady, a sam świat jest zbudowany z elementów będących jednościami sił, informacji i innych domniemanych jakości. Perzanowski zaznacza, że ostatnia teza tzw. panpsychizmu metafizycznego nadal rzadko pojawia się w nauce, choć miała wielu wybitnych obrońców. Jego zdaniem, nauka stale zmierza w kierunku przyszłej syntezy materialnej (fizycznej) strony Natury i jej strony duchowej, psychicznej bądź informatycznej.

²⁶ Por. Perzanowski 1994: 284–285. G. Leibniz opowiada się za odmienną od B. Spinozy formą panpsychizmu, wprowadzając wielość substancji zwanych monadami. Są one substancjami prostymi, zawierającymi w sobie całkowitą informację o świecie. Monady to bezcielesne automaty, przemieszczające się od jednego postrzeżenia ku drugiemu, zgodnie z ustanowioną przez Boga harmonią. Jeżeli świadomość nie jest redukowalna, to nie można jej wyjaśnić na danym stopniu fizycznej złożoności (por. 1.4) i jeżeli nie pojawia się ona dopiero na tym stopniu złożoności, to musi już być obecna na niższych stopniach.

Ponieważ w zapowiedzianym poszukiwaniu neuronalnych korelatów świadomości odwołujemy się do wyników uzyskanych we współczesnej nauce, dlatego nie odwołujemy się do związanego z monadologią „panpsychizmu metafizycznego”. Ponadto przytoczone powyżej zarzuty wobec panpsychizmu pokazują, że przynajmniej na dzisiejszym etapie naukowego wyjaśniania, trudno jego tezy uwzględnić, chociażby przy dokonywaniu obrazowań mózgu.

Zamiast *résumé* Nagła teorii dwu aspektów świadomego umysłu podajemy jej krótkie porównanie z innymi teoriami świadomości.

(A) D. Davidson twierdzi, że opisy fizykalne mogą ująć ontologię tego, co mentalne. Epifenomenaliści nie dążą do proponowanej przez Nagła syntezy perspektyw. Teoria identyczności zdarzeń odmawia temu, co mentalne, zdolności do samodzielnego wyjaśniania i wyklucza je z fizykalnych opisów. Teoria ta, jak ją rozumie Nagel, nie musi postulować identyczności pomiędzy stanem mentalnym i stanem mózgu. Wystarczy identyczność pomiędzy podmiotem z określoną perspektywą i określonym stanem fizycznym (por. Nagel 1965: 341–343).

(B) Przytaczane przez S. Kripkego argumenty przeciwko materializmowi nie godzą w Nagła teorię dwu aspektów, ponieważ oczywista pojmovalność zdarzeń mentalnych dokonujących się bez zmian w mózgu może być spowodowana faktem, iż pojęcia tego, co mentalne, odnoszą się tylko do jednej cechy umysłu. Jeżeli mógłbym wiedzieć, że mogę sobie wyobrazić występowanie moich terażniejszych przeżyć wzrokowych bez występowania jakiegoś szczególnego zdarzenia w moim mózgu, wtedy mógłbym wiedzieć, że obydwa – przeżycie i zdarzenie fizjologiczne – nie są identyczne. Ale nie mogę tego wiedzieć tylko na podstawie moich pojęć dotyczących przeżyć wzrokowych czy też na podstawie posiadania tych przeżyć. Wrażenie wzrokowe jest stanem, którego cechy zjawiskowe pozwalają mi zidentyfikować ów stan bez zewnętrznych obserwacji mnie samego. Nie oznacza to jednak, że znam tym samym pełną naturę tego wrażenia (por. Nagel 1997c: 57–62).

(C) Dlaczego Nagel ani nie odrzuca fizykalizmu, ani nie przyjmuje dualizmu? Nagel uważa, że Davidson pokazał, iż intencjonalne zdarzenia mentalne oraz wrażenia zjawiskowe są także procesami fizycznymi, chociaż, jak sam pisze, nie rozumiemy, w jaki sposób ta identyfikacja zdarzeń mentalnych i procesów fizycznych zachodzi. Subiektywne przedmioty jednostkowe, zdarzenia mentalne, znajdują swoje miejsce w obiektywnym świecie rzeczy fizycznych (wchodzą w relacje przyczynowe z przedmiotami jednostkowymi). Przyjęcie czegoś niefizycznego (to znaczy nieprzestrzennego) nie ułatwi, jak rozumiemy Nagła, wyjaśnienia subiektywnych cech stanów mentalnych. Problem leży bowiem w istotnej różnicy pomiędzy subiektywnością tego, co

mentalne, a obiektywnością tego, co fizyczne. Owa różnica powoduje, że trudno zobaczyć, jak stany mózgu mogą konstituować rzeczywistą naturę stanów świadomych.

(D) Rozróżnienie na to, co subiektywne i obiektywne, jest względne. Nasz ogólnoludzki ludzki punkt widzenia jest bowiem bardziej obiektywny aniżeli punkt widzenia jednego podmiotu, ale mniej obiektywny niż obiektywne ujęcie fizykalne (por. Nagel 1997b: 249). Wspomniane rozróżnienie służy Nagłowi do ukazania dwuaspektowości umysłu i do odrzucenia karteżanizmu. Nagel, w odróżnieniu od C. McGinna, nie traktuje subiektywnego punktu widzenia jako czegoś „tajemniczego”. Nie twierdzi również, że ponieważ stany psychiczne czy punkty widzenia nie mogą być wyjaśnione poprzez odwołanie się do redukcji fizjologicznych, to tym samym muszą one mieć w sobie coś niefizykalnego²⁷.

(E) Zbliżone do proponowanego przez Nagła rozwiązanie znajdujemy w zbudowanym przez B. Spinozę „ekstremalnym monizmie” zwanym także panteizmem albo panpsychizmem: cokolwiek istnieje jest w jednej substancji tożsamej z Bogiem. To, co mentalne, jest istotną cechą wszystkich fizycznych obiektów, a ponadto cechy mentalne i fizyczne okazują się być cechami trzeciego rodzaju materiału, który obydwie te cechy ukazuje. Spinoza wprowadza dwa nieredukowalne do siebie atrybuty (myśl i rozciągłość) tej samej substancji, które ze sobą nawzajem korespondują (tzw. *paralelizm*). Korespondencja ta jest wstawiona w szeroki teoretyczny system pojęciowy. Podobnie do Spinozy argumentuje D. Chalmers, chociaż nie powołuje się na niego. Dla Chalmersa świadome doświadczenia są obiektami, o których można teoretyzować. Szuka on praw pomostowych (*bridge laws*) opisujących relacje doświadczeń do innych obiektów oraz bardzo ogólnych zasad wyjaśnialnej koherencji. Proponowane przez niego wyjaśnianie porusza się w naszym obecnym systemie pojęć. Tym samym rodzi się pytanie – mając na uwadze ujęcie proponowane przez Nagła – czy propozycja Chalmersa dostarcza rozwiązania problemu subiektywnych zjawisk mentalnych. Możemy oczywiście zaakceptować Chalmersa *bridge laws*, ale możemy także pytać, dlaczego miałyby one obowiązywać.

O ile Nagel uważa, że połączenie pomiędzy tym, co fizyczne, a tym, co psychiczne, jest konieczne, o tyle Chalmers przyjmuje je jako kontyngentne (tzn. konieczne w takim sensie, w jakim obowiązują prawa natury).

²⁷ Nagel odrzuca taką interpretację twierdząc, że jeżeli przyjmiemy, że zdarzenia mentalne nie są zdarzeniami fizycznymi, to nie wynika z tego „[...] że możemy wyjaśnić ich miejsce we wszechświecie powołując się na substancję, której jedyną rolą jest dostarczać im medium (*to provide them with a medium*)” (Nagel 1997c: 38).

5.5. Poszerzona teoria stanów mentalnych

Nagel nie podaje jednego, koherentnego przedstawienia natury świadomego umysłu, a jego próby integracji perspektywy subiektywnej i obiektywnej kończą się na metafizycznych sugestiach. Proponowana przez niego epistemologia jest określona przez „bezkompromisowo” realistyczną metafizykę (por. Nagel 1997c: 87): to my mieścimy się w świecie, a nie świat w nas. Nie posiadamy gwarancji, że możemy poznać i zrozumieć obiektywną naturę świata. Jesteśmy ograniczonymi, powstałymi na drodze ewolucji bytami, jesteśmy tylko jednym rodzajem pośród innych rodzajów, a świat nie zamierza się zredukować do naszego poznawczego wymiaru. Nasze widzenie świata jest ustalone poprzez nasz – rodzajowy i indywidualny – sposób interakcji z nim i pozostaje pytanie, czy potrafimy transcendować zjawiska (*appearances*), aby odkryć, jaki świat jest sam w sobie.

Istnieje przedział pomiędzy naszymi przekonaniem a podstawami tych przekonań. Z jednej strony może się okazać, że istnieją cechy rzeczywistości, których nie potrafimy, z zasady, ująć, gdyż rzeczywistość nie jest ograniczona naszą przypadłościową zdolnością pojęciową. Z drugiej strony, gdy założymy odpowiednio silny sens obiektywności, możemy wpaść w niebezpieczeństwo sceptycyzmu²⁸. Nastawienia obiektywne i sceptyczne są ściśle powiązane. Obydwa wychodzą z założeń, że: (a) istnieje realny świat, który nas wszystkich zawiera, oraz (b) zjawiska są rezultatem naszej interakcji z resztą tego realnego świata. Tak powstałych zjawisk nie możemy akceptować bezkrytycznie, ale musimy próbować zrozumieć, co sami do nich poznawczo wnosimy.

Centralnym problemem epistemologii okazuje się – według Nagla – problem pierwszej osoby: o czym mam być przekonany i jak mam uzasadnić moje przekonania? Nie chodzi przy tym o nieosobowy problem w rodzaju: gdyby ktoś obserwował moje mniemania razem z określonymi założeniami dotyczącymi tego, co faktycznie jest, to czy mógłby o mnie powiedzieć, „wiem, że jest tak a tak”? Odpowiedź na pytanie, czym jest wiedza, nie pomoże mi w podjęciu decyzji, o czym powinienem być przekonany. Musimy raczej pytać, na czym właściwie polega nasz stosunek do świata i jak możemy go zmienić.

²⁸ Według Nagla: „Zgodnie z teoriami sceptycznymi, treść naszych potocznych lub naukowych przekonań o świecie wykracza tak dalece poza ich uzasadnienie, że nie da się ich obronić przed różnego rodzaju wątpliwościami. Możemy się mylić w taki sposób, którego nie potrafimy wykluczyć. Kiedy zauważymy tę niedającą się wyeliminować lukę, nie możemy nadal ufać tym przekonaniom, jeśli nie chcemy być irracjonalni” (Nagel 1997c: 84).

Nagel krytykuje próby osłabiania sceptycyzmu przez ożywianie przyczynowych teorii znaczenia²⁹ i koryguje idealistyczne tendencje obecne w niektórych teoriach świadomości (na przykład u I. Kanta, L. Wittgensteina i D. Davidsona). Teorie te – zdaniem Nagla – przypominają wczesne weryfikacjonistyczne próby uniknięcia sceptycyzmu. Tego rodzaju semantyczne teorie (w których znaczenie jest zdeterminowane przez relacje przyczynowe) są raczej odrzucane przez sceptycyzm, aniżeli same go odrzucają. Nagel nie pokazuje jednak, w czym leżą braki wspomnianych semantycznych teorii, braki niezależne od znaczenia tych teorii dla sceptycyzmu (por. 1.3.5).

Nagel broni racjonalistycznej epistemologii (por. Nagel 1997c: 101–104) i przyjmuje, że istnieje jakiś nieodzowny, nietrywialny, składnik *a priori* obecny w naszym poznaniu świata. Ludzki umysł jest zdolny do tworzenia z samego siebie hipotez o tym, jakim świat mógłby być. Ponadto umysł musi umieć odrzucać hipotezy, które początkowo wydają się prawdopodobne, choć w rzeczywistości takie nie są.

Narzuca się pytanie, co Nagel rozumie poprzez poznanie *a priori*. Ogólnie możemy powiedzieć, że poznanie *a priori* jest poznanie prawd, które nie są uzasadnione przez doświadczenie. Taka definicja nie odpowiada jednak naglowskiemu „komponentowi *a priori*”, gdyż podaje ona jedynie poznanie możliwości. Poznanie tego, co jest, mamy jedynie wtedy, gdy możliwości są sprawdzone doświadczalnie.

Nagel nie twierdzi, że rozum dostarcza niepowątpiewalnych podstaw dla przekonań, chociaż zauważa, że naukowo-empiryczne poznanie albo naukowo-empirycznie uzasadnione przekonania muszą się opierać na *apriorycznej* podstawie. Jeżeli z ograniczonej empirycznej oczywistości zostają wyprowadzone dalekosiężne wnioski, to duże znaczenie przypada bezpośrednio, *apriorycznemu* formułowaniu i selekcji hipotez (jeżeli poznanie w ogóle ma być możliwe). Zdaniem Nagla, indukcja, która jest zasadą wyjściową empi-

²⁹ Przykładem „przyczynowej” teorii jest realistyczne ujęcie znaczenia rozwinięte przez Putnama (*nonverificationist theory of meaning*). Pod wyrażeniem „znaczenie znaczenia” (*the meaning of «meaning»*) Putnam rozumie wielowymiarowy wektor, który posiada część ekstensjonalną, czyli odniesienie (*reference*). Część ta zależy od rzeczywistości i przedstawia tym samym obiektywną komponentę. „Instead of seeing meanings as entities which determine reference, they now are beginning to see meanings as largely determined by reference, and reference as largely determined by causal connections” (Putnam 1975: x). Inne części są intensjonalne; syntaktyczny marker, semantyczny marker, i semantyczny stereotyp (por. Putnam 1975a: 266–270). W swoich późniejszych pismach Putnam odchodzi od realizmu i (nawiązując do Kanta) rozwija tak zwany wewnętrzny realizm (pojęcie znaczenia i prawdy może być ujęte tylko przez odwołanie się do jakiejś teorii).

ryzmu, ma sens jedynie w połączeniu z racjonalistyczną, *aprioryczną* podstawą. Obserwowalne regularności dostarczają racji dla przekonania, że będą one powtarzalne tylko do tego stopnia, iż dostarczają oczywistości dla ukrytych *koniecznych* połączeń, które zachodzą bez względu na czas. Obecnie przechodzimy do przedstawienia tych koniecznych połączeń.

Według Kartezjusza, umysł jest nierozciągliwy, a żaden rozciągliwy przedmiot nie może być substancją myślącą. Pojęcia „umysłu”, „zdarzenia mentalnego” lub „procesu mentalnego” wykluczają ewentualność, iż ich desygnaty mogą się okazać czymś fizycznym lub że mogą one zostać poszerzone na przedmioty fizyczne. Z możliwością takiego poszerzenia spotykamy się na przykład przy przednaukowym pojęciu „krwi”, którego treść jest otwarta na uzupełnienia pochodzące z dalszych badań dotyczących jej składu chemicznego i struktury.

Podobnie argumentują współcześni filozofowie analityczni: pojęcia mentalne nie dostarczają nam tego rodzaju początkowego sposobu operowania obiektami świata czasoprzestrzennego, który spotykamy przy przednaukowym pojęciu materii fizycznej³⁰. Błędem jednak byłoby, według Nagla, gdybyśmy z naszej niemożności przedstawienia sobie tego, że fenomeny mentalne mogą mieć cechy fizyczne, wyciągnęli wniosek, że taka możliwość została na zawsze wykluczona. Jeżeli jednak zdarzenia mentalne faktycznie miałyby cechy fizyczne, to należy wyjaśnić, dlaczego tak trudno im te cechy przypisać. Rodzaj nieprzedstawialności, o który chodzi przy zdarzeniach mentalnych, jest całkiem inny, aniżeli na przykład trudności z przedstawieniem sobie, że liczby mają rodziców. Zdarzenia mentalne, inaczej niż liczby, mogą być w przybliżeniu zlokalizowane w przestrzeni i czasie. Przyjmuje się, że stoją one w relacjach przyczynowych do zdarzeń mających cechy fizyczne. Występowanie relacji przyczynowych będzie przemawiało za tym, że zdarzenia mentalne mają cechy fizyczne, jeżeli tylko potrafimy nadać sens tej idei³¹. Identyfikacja zdarzeń mentalnych ze zdarzeniami fizycznymi wymaga unifikacji cech mentalnych i fizycznych w jedną, co sprawia trudności w ich zrozumieniu.

Problem leży bowiem w odmiennym charakterze pojęć pierwszo- oraz trzecioosobowych, co jest zarazem gramatyczną manifestacją subiektywno-

³⁰ Zdaniem McGinna: „In order to solve the mind-body problem we need a minimum, a new conception of space. We need a conceptual breakthrough in the way we think about the medium in which material objects exist, and hence in our conception of material objects themselves” (McGinn 1995: 226).

³¹ Nagel powołuje się na artykuł D. Davidsona *Mental Events* (por. Nagel 1998a: 341).

ści zjawisk mentalnych. Nie wszystkie istoty świadome posiadają język. Jednak nasze przypisywanie stanów świadomych stworzeniom pozbawionym języka implikuje, że stany te są tego rodzaju, iż w przypadku człowieka uchwytujemy je tylko poprzez odmienne pojęcia. Chodzi o pojęcia, które podmiot stosuje do siebie samego, bez obserwacji swojego ciała i swoich zachowań. Mówiąc o stanach mentalnych możemy przyjąć twierdzenie:

V.5 Stany mentalne są subiektywne nie w tym sensie, że odnoszące się do nich wyrażenia są czysto pierwszoosobowe, ale dlatego, że mogą być opisane tylko przez pojęcia, w których nieobserwowalne atrybuty pierwszoosobowe i obserwowalne atrybuty trzecioosobowe są logicznie nieodłączne.

Tego rodzaju stany są modyfikacjami indywidualnego podmiotowego punktu widzenia. Kwestią wtedy jest: jak coś, będące cechą lub elementem indywidualnego punktu widzenia, może także być fizjologicznie opisywalnym zdarzeniem w mózgu? Jak może ono być rodzajem przedmiotu, którego opis fizjologiczny ani nie zawiera punktu widzenia, ani nie zawiera bezpośredniego pierwszoosobowego przypisywania? Nagel pisze: „Na dzień dzisiejszy nie posiadamy pojęciowego wyposażenia pozwalającego zrozumieć, jak cechy fizyczne i cechy subiektywne mogą być aspektami pojedynczego obiektu lub procesu” (Nagel 1998a: 342).

Widoczny tutaj problem leży w braku jakiegoś przedstawialnego wewnętrznego połączenia pomiędzy modyfikacją mojego subiektywnego punktu widzenia i modyfikacją psychofizycznej aktywności mojego mózgu. Modyfikacje te mogą sobie dokładnie odpowiadać, ale identyczność *type-type*, zgodnie z przytoczonymi analizami Kripkego, wymaga czegoś więcej. Jeżeli bowiem mają to być te same stany, to niemożliwe jest, aby jeden mógł istnieć bez drugiego. Możemy mieć empiryczne podstawy dla przyjęcia, że taka niemożliwość zachodzi. Jednak zrozumienie tej niemożliwości wymaga, aby konieczność połączenia pomiędzy wspomnianymi stanami stała się dla nas pojęciowo przejrzysta. W przypadku stanów świadomych i stanów fizjologicznych nie widzimy konieczności takiego połączenia. Ponadto konieczne połączenie pomiędzy dwoma tak odmiennymi stanami jest niewyobrażalne. Stany te wydają się być od siebie logicznie oddzielone.

Nie potrafimy dostrzec, w jaki sposób szczegółowe przedstawienie tego, co zachodzi w mózgu, może wyczerpująco wyjaśnić smak kawy – nawet gdybyśmy mogli zobaczyć jak przedstawienie to wyjaśnia wszystkie fizyczne efekty takich wrażeń. Dopóki tego rodzaju luka w wyjaśnianiu nie zostanie zlikwidowana, dopóty identyfikacja obydwu stanów pozostaje problema-

tyczna. Stąd potrzeba poszerzonego ujęcia, w którym subiektywny punkt widzenia posiadałby obiektywne charakterystyki³².

W proponowanych przez dzisiejszą filozofię analityczną ujęciach świadomości nie zakłada się koniecznego powiązania pomiędzy twierdzeniami dotyczącymi subiektywności stanów mentalnych i twierdzeniami dotyczącymi fizjologii mózgu. Nie wyklucza to jednak możliwości stworzenia nowego (w sensie „następującego po”, *successor*) pojęcia umysłu, które będzie zachowywało istotne cechy pojęcia-oryginału i będzie otwarte na naukowe odkrycie poszukiwanego rodzaju połączeń. Z czymś takim mieliśmy do czynienia w historii nauki chociażby przy pojęciach „dźwięku”, „elementu”, „przestrzeni”. Nowe pojęcie tworzymy po to, aby mówić o tych samych cechach, o których mówiło się używając starego pojęcia. Ponieważ nowe pojęcie stawia owe cechy w relacjach do innych, nowych cech, to możemy odkrywać dodatkowe powiązania. Bez poszerzonego pojęcia tego, co mentalne, brakuje widoków na przezwycięzenie wspomnianej luki w wyjaśnianiu. Nagel stwierdza:

„Wydaje mi się całkiem możliwe, że bardziej adekwatne ujęcie subiektywnego przeżycia powinno traktować je jako fenomenologiczną «stronę wewnętrzną» określonych procesów fizycznych. Potrzebujemy zmniejszenia barier prowadzących do takiego ujęcia. Jeżeli stałoby się możliwym myśleć o przeżyciach jako istotnie o tym, jak to jest subiektywnie dla jakiegoś fizycznego organizmu być w jakimś stanie fizjologicznym, to otworzyłoby nam to drogę do odkrycia koniecznej prawdy *a posteriori* o tym, jaki stan fizjologiczny jest poszczególnym rodzajem przeżycia. Ale najważniejszym krokiem jest ten pierwszy, poszerzyć pojęcie przeżycia poprzez uznanie faktu, że to, co ono *explicite* zawiera – włączając jego behawioralne i funkcjonalne implikacje – daje niekompletne przedstawienie natury przeżycia” (Nagel 1998a: 344).

Nagel omawia przypadek kontraintuicyjnego myślenia o przeżyciu jako o zdarzeniu, którego pełna natura nie jest nam obecnie dostępna. Jest to antykartezjańskie myślenie o umyśle jako o czymś, co z zasady jest jedynie częściowo dostępne dla introspekcji. Jeżeli potrafimy tak myśleć o umyśle, to czynimy pierwszy krok w kierunku poszerzonej, niedualistycznej odpowiedzi na problem umysł–ciało. Chociaż myślenie takie jest dziś czystą fantazją, to chodzi w nim o teoretyczną identyfikację zdarzeń mentalnych z jakąś wewnętrzną strukturą i funkcją, które posiadają introspektywny charakter, ale się w nim nie wyczerpują.

³² Podobnie argumentuje R. Penrose, twierdząc, że materialistyczne wyjaśnianie świadomych zjawisk (na przykład niektórych anomalii w aktywności mózgu) wymagałoby fundamentalnej rewolucji w fizyce (por. Penrose 1989: 447).

W poszukiwaniu poszerzonej teorii stanów mentalnych, nasza uwaga powinna być skierowana na relacje pomiędzy umysłem a mózgiem. Jeżeli zaś się okaże, że takiej teorii nie da się zbudować, to musimy wyjaśnić, dlaczego tak jest. Nagel uważa, że luki w wyjaśnianiu (w jej dzisiejszej formie) nie można zamknąć. Jak długo bowiem posługujemy się naszymi mentalnymi i fizykalnymi pojęciami, nie zobaczymy żadnego *koniecznego* połączenia pomiędzy fizykalnie opisywalnym mózgiem a sensorycznym przeżyciem. Nie widzimy bowiem żadnego logicznego połączenia tego rodzaju, jaki znamy z wyjaśniania w innych naukach (na przykład z wyjaśniania odwołującego się do analizy składników chemicznych). Zdaniem Nagla:

(S) „Posiadamy dobre podstawy do przyjęcia założenia, że to co mentalne superweniuje na fizycznym – tzn. że nie istnieje różnica mentalna bez różnicy fizycznej. Jednak brak wyjaśnienia superweniencji nie jest rozwiązaniem, lecz znakiem czegoś fundamentalnie nieznanego. Nie możemy traktować czystej superweniencji jako końca opowiadania, ponieważ wymagałoby to od tego, co fizyczne, wywoływania w sposób konieczny mentalnego, pozostawiając nas bez odpowiedzi na pytanie, jak tego dokonuje. Mimo wszystko musi być jakieś «jak» i naszym zadaniem jest je zrozumieć. Obecność dostrzegalnego, systematycznego połączenia, które pozostaje dla nas niezrozumiałe, domaga się zbudowania odpowiedniej teorii” (Nagel 1998a: 345).

Subiektywność świadomości blokuje wszystkie redukcjonistyczne propozycje. Gdy bowiem mamy dany jakiś fizykalny albo funkcjonalistyczny opis, jakkolwiek skomplikowany, wydaje się logicznie możliwe, że może istnieć zombi spełniający te same warunki i nie posiadający zarazem subiektywnego punktu widzenia. Nic nas nie może powstrzymać od wyobrażenia sobie, że istnieje jakiś rodzaj fizycznie lub funkcjonalnie zorganizowanego systemu, a następnie od wyobrażenia sobie, że nie istnieje subiektywne przeżycie z nim związane³³. Brak jakiegokolwiek pojęciowego połączenia pomiędzy tymi ujęciami powoduje, że pierwsze nie jest zdolne do wykluczenia drugiego.

5.6 Podsumowanie

Dziś brakuje nam idei jasnego i bezpośredniego połączenia, wyjaśniającego relacje pomiędzy stanami fizjologicznymi i zjawiskowymi. Tym samym niewyobrażalne jest wyjaśniająco-redukcyjne rozwiązanie problemu świadomości. Do dyspozycji mamy tylko ekstensjonalne, empiryczne zależności, obrazujące to, co zachodzi pomiędzy wspomnianymi stanami.

³³ W ten sposób argumentuje D. Chalmers (1996: 96–98).

Niemniej jednak Nagel każe nam szukać (w ramach naukowej teorii umysłu), trzeciego, panpsychicznego ujęcia, które bezpośrednio zawierałoby to, co mentalne i fizyczne. Dzięki takiemu ujęciu może się okazać, że połączenie mentalnego i fizycznego musi być konieczne. Tego rodzaju ujęcia nie znajdziemy przypadkowo, lecz musimy je zbudować. Sukcesy w historii nauki zależały od rozwoju pojęć teoretycznych, a nie od używania pojęć potocznych. Pojęcia teoretyczne cechują się tym, że pozwalają zastąpić czyste korelacje wyjaśnieniami redukcyjnymi. To uzasadnia ich tworzenie.

Reasumując możemy powiedzieć, że z jednej strony Nagel zaznacza, iż próba zrozumienia tego, w jaki sposób może istnieć *konieczne* połączenie pomiędzy tym, co subiektywne, i tym, co fizyczne, nie może budować na analogiach odwołujących się do redukowalności w naukach fizykalnych. Na przykład poprzez odwołanie się do analogicznych stwierdzeń, że pole elektromagnetyczne jest w sposób *konieczny* połączone z zachowaniem naładowanych cząstek, a pole grawitacyjne jest w sposób *konieczny* połączone z obecnością odpowiednich mas. Z drugiej strony widać, że Nagel jest przekonany, iż w dalekiej przyszłości będziemy potrafili obserwować mózg i powiedzieć: tak z „zewnątrz” wygląda subiektywne przeżycie smakowania czy widzenia.

Nagel podkreśla znaczenie obiektywnego spojrzenia na świat, jak i jego konieczną granicę – subiektywność stanów mentalnych. Broniąc obiektywności ujęć zdarzeń psychicznych ostrzega równocześnie przed niepożądanymi skutkami, jakie niesie dążenie do obiektywności w fizykalistycznym przedstawieniu umysłu. Jego akceptacja dysharmonii perspektywy pierwszo- i trzecioosobowej prowadzi z jednej strony do ukazania złożoności rzeczywistości, którą my sami jesteśmy, z drugiej zaś – do zakwestionowania skądinąd atrakcyjnej prostoty rozwiązań redukcyjnych i eliminacyjnych.

Dualizm perspektywy (subiektywne–obiektywne) jest natury ontologicznej, a nie tylko metodologicznej czy epistemologicznej. Jednym z podstawowych twierdzeń Nagla jest, że podmiot *X* może zrozumieć jakieś szczególnie przeżycie podmiotu *Y* tylko poprzez przyjęcie subiektywnego punktu widzenia podmiotu *Y*. Aby zrozumieć, na przykład, przeżycie widzenia czerwonego, podmiot *X* musi sobie przypomnieć podobne przeżycie (czy też być zdolnym coś takiego przeżyć lub je sobie wyobrazić). Owo *jak to jest* ma tym samym ontologiczne znaczenie, gdyż dotyczy istotnej natury przeżycia: są fakty istotne dla przeżycia, które mogą być ujmowane tylko z punktu widzenia *Y*. Właśnie z powodu tych faktów jest niewyobrażalne, aby neurofizjologia mogła kiedykolwiek całkowicie wyeksplikować jakieś przeżycie. Odrzucając fizykalizm Nagel ożywia zarazem klasyczny argument za panpsychizmem.

Wobec wspomnianych problemów z panpsychizmem (por. 5.4.2) rezygnujemy w dalszych częściach pracy z przyjmowania jego tez odnośnie do poszerzonej teorii świadomości i poprzez poszukiwanie neuronalnych korelatów świadomości próbujemy określić rodzaj teoretycznego związku pomiędzy świadomością zjawiskową i jej neuronalnym podłożem. Mając do dyspozycji metodologię poszukiwania wspomnianych korelatów podajemy warunki wstępne dla stworzenia nowej teorii świadomości, którą nazywamy „prototeorią” świadomości (por. 7.6; 7.6.1).

Postawiony przez Nagla problem subiektywności jest taki „trudny”, ponieważ przyjmuje nieprawdopodobnie wysokie standardy decydujące o rozwiązaniu tego problemu. Standardami tymi są: istnienie *koniecznego* połączenia subiektywności ze stanami fizycznymi i związana z nimi teoretyczna odpowiedź na pytanie: *jak to jest* być świadomym? Współczesna medycyna i naukowe badania nad świadomością operują bardziej praktycznymi kryteriami, dotyczącymi jednak zazwyczaj świadomości dostępu, czy przytomności. Również w historii filozofii (od I. Kanta do W. Jamesa) nie odwoływano się do takiego spojrzenia na subiektywność, jakie proponuje Nagel.

ROZDZIAŁ SZÓSTY

W POSZUKIWANIU NOWEJ TEORII ŚWIADOMOŚCI

6. Wprowadzenie

„If there were no solution, why would we have gone to all the trouble of inventing the problem” (Jay F. Rosenberg).

„We simply cannot know today whether we can eventually understand a problem like subjectivity. That may become clearer as more plausible hypotheses are tested. In sum, such philosophical challenges do not invalidate a useful thinking tool” (Baars 1998: 3).

To, czy pacjent jest świadomy – w sensie przytomny – czy nie jest, odgrywa podstawową rolę w medycynie. Mówiąc o świadomości lekarze z reguły mają na myśli przytomność pacjenta. Badania zjawiskowego aspektu świadomości są pozostawiane na boku chociażby z tego powodu, że nie jest on bezpośrednio dostępny dla intersubiektywnych metod badawczych. Nic jednak nie jest każdemu z nas tak bliskie, jak subiektywna cecha świadomości określana pytaniem: *jak to jest* być człowiekiem, czy bardziej specyficznie: *jak to jest* być mną? Wyjaśnienie tej cechy domaga się zbudowania odpowiedniej teorii świadomości.

Pierwszym krokiem prowadzącym do poszukiwanej naukowej teorii jest metodologiczne doprecyzowanie tego, co rozumiemy pod pojęciem neuronalnych korelatów świadomości¹. W obecnym i następnym rozdziale przed-

¹ Niektórzy neurologowie mówią o „neuronalnych substratach procesów mentalnych” (Parsons 1997: 256).

stawimy metodologiczne warunki, których spełnienie jest konieczne dla znalezienia tychże korelatów. Omówione w rozdziałach III, IV i V teorie rezygnują z tego kroku i albo identyfikują świadomość z bliżej anatomiczno-funkcjonalnie nieokreślonymi zdarzeniami w mózgu, albo nie skupiają się na omawianiu wzajemnej zależności pomiędzy mózgiem a świadomością.

Linia przewodnią naszego filozoficznego i neurologicznego rozumienia świadomości zjawiskowej jest pytanie o wzajemne ich skorelowanie. Stąd ogólne twierdzenie:

VI.1 Słaba teza o korelacji nie zakłada, że świadomość zjawiskowa i stany mózgu są identyczne, lecz jedynie że sobie odpowiadają.

Twierdzenie VI.1 jest mniej radykalne niż stanowisko materializmu ontologiczno-redukcyjnego. Przyjąwszy, że odkrycia naukowe nie negują materializmu, widzimy, że zarówno dualiści, jak i materialistyczni moniści zakładają istnienie – mniej lub bardziej silnie rozumianych – korelacji między świadomymi i nieświadomymi stanami umysłu z jednej strony, a stanami mózgu z drugiej strony (por. Kripke 1988: 148–149; Northoff 2001: 24–26). Ponadto dualiści przyjmują, że związek „korelacji” nie jest zwrotny, monistyczni materialiści twierdzą, że poszukiwana korelacja jest jednym z przypadków zwrotności. Przeprowadzone w poprzednich rozdziałach analizy stanów mentalnych i fizycznych oraz wyniki uzyskane w badaniach naukowych pokazują, że dziś możemy mówić jedynie o „korelacji” pomiędzy nimi. Tym samym nie mówimy ani o panpsychicznie rozumianych neuronalnych częściach świadomości, ani o przekonaniach o istnieniu duchowych części mózgu. Samo stwierdzenie korelacji nie jest wyjaśnieniem świadomości np. w sensie: w jaki sposób mózg generuje świadomość zjawiskową, lecz otwiera jedynie drogę do niego. W rozdziale VII przyjmujemy silniejszą tezę od VI.1.

W historii filozofii różnie postrzegano rolę, jaką świadomość odgrywa w naukach o systemie nerwowym. Wymienione w rozdziale III teorie świadomości omawiamy krótko w 6.2, pod kątem ogólnej zależności świadomość–mózg. Po wskazaniu w 6.1.1 na konieczność uwzględniania świadomości w badaniach neurologicznych, przedstawiamy w podrozdziałach 6.3.1, 6.3.2, 6.3.3 trzy współczesne, neurokognitywne teorie świadomości. Ich autorzy dokonują różnie rozumianej syntezy badań neurologicznych i filozoficznych (bądź psychologicznych).

Proponowany przez nas punkt wyjścia oznacza tego rodzaju filozoficzne sformułowanie problematyki świadomego umysłu, którą będzie można i którą należy uwzględnić w badaniach neurologicznych. Podkreślamy przy tym odrębność pytań z poziomu filozoficznego i neurologicznego. Argumen-

tację za koniecznością wielowymiarowego badania świadomego umysłu można znaleźć chociażby w książce U. Żegleń, zdaniem której pełny obraz umysłu uzyskamy dopiero dzięki szeroko zakrojonym, interdyscyplinarnym badaniom (por. Żegleń 2003: 285, 309–310).

Poszukiwań wspomnianych korelacji nie należy utożsamiać ze spełnieniem podanych przez Nagla (5.5), ogólnych warunków tworzenia poszerzonej teorii świadomości. Nagel zakłada panpsychizm i pyta o *apriorycznie* konieczne identyczności, odkładając tym samym rozwiązanie problemu świadomości na nieokreśloną przyszłość. Uważamy, że dziś możemy się jedynie odwoływać do wyników uzyskanych w neurofizjologii i przyjąć – gdy chodzi o teorię świadomości – dwa rodzaje emergencyjnej zależności: pomiędzy zdarzeniami fizycznymi i mentalnymi, z jednej strony, oraz pomiędzy filozoficzno-psychologiczną teorią świadomości, a teorią zdarzeń zachodzących w mózgu – z drugiej.

Nie podajemy systematycznej teorii holistycznie rozumianej świadomości, która by uwzględniała jej neurologiczne podłoże. Na boku pozostawiamy również analizy świadomości spotykane w ramach nauk o sztucznej inteligencji. Prowadzone w tym rozdziale dyskusje mają równocześnie dookreślić współczesne rozumienie neuronalnych korelatów świadomości.

6.1. Neurokognitywistyka a problem świadomości

Do kwestii uwzględniania subiektywności świadomości w badaniach neurofizjologicznych podejmiemy od strony pytania: czy nauki o systemie nerwowym muszą się odwoływać do modelu holistycznie rozumianej świadomości? Współczesna neurologia nie zna zwrotu „trudny” problem świadomości i zdaje się rozwijać bez uwzględniania subiektywności świadomości, stanowiącej jądro tego problemu. Podobny rozwój obserwujemy w badaniach nad sieciami neuronowymi, naśladującymi swoimi funkcjami i budową strukturę układu nerwowego².

Dlatego należy się zastanowić, czy poszukiwać nowej kognitywnej teorii świadomości, czy też zgodzić się z F. Crickiem, że „świadomość jest problemem naukowym”, a filozofia świadomości jest jedynie nudną zabawą³.

² Por. Crick 1994: 220–245. A. Wróbel omawia elementy sieci neuronowej mózgu oraz podaje przykłady sieci neuronowych w ośrodkowym układzie nerwowym (por. Wróbel 2000: 52–65). Obszerną literaturę na temat sieci neuronowych można znaleźć w: Tadeusiewicz 1993: 8–23.

³ „No longer need one spend time [...] enduring] the tedium of philosophers perceptually disagreeing with each other. Consciousness is now largely a scientific problem” (Crick 1996: 486).

Założenie Cricka nie jest podzielane przez wszystkich neurologów. Zdaniem A. Revonsuo, jak do tej pory nie mamy nic, co przypominałoby naukowy program badań świadomości⁴. Ani bowiem nie wychodzimy z w miarę wspólnych założeń filozoficzno-metafizycznych, ani nie posiadamy jednolitej metodologii badań. Nasze dotychczasowe analizy pokazują, że świadomość nie jest jedynie problemem w neurologii, gdyż filozoficzne analizy pojęć należą do jej szeroko rozumianego, naukowego badania (por. też 6.4.1, 6.4.2)⁵.

Stąd nasze pytanie o metodę poszukiwania neuronalnych odpowiedników świadomości przeżyciowej oraz specyficznych, neurologicznych odpowiedników dla jej konkretnych stanów. Z pytaniem tym są związane dalsze trudności dotyczące chociażby możliwości wskazania neuronalnego odpowiednika dla przeżyciowej różnicy, odpowiadającej dwóm różnym stanom świadomym.

6.1.1. Pytanie o kryterium świadomości

„Filozofia stara się przede wszystkim rozpatrzyć i zrozumieć całkiem zwykłe idee, którymi wszyscy na co dzień bezrefleksyjnie się posługujemy” (Nagel 1998: 7).

Z przedstawionych w poprzednich rozdziałach linii postępowania wynika, że określenie „trudny problem świadomości” zależy od kryterium, co rozumimy pod terminem „świadomość”. Z jednej strony mamy D. Dennetta i innych materialistów, którzy rezygnują z kryterium przypisującego świadomości subiektywność, tzn. nie zajmują się pytaniem: „*jak to jest* być świadomym organizmem?” Z drugiej strony, w ostatnich latach obserwuje się w dyskusjach filozoficznych powrót tematyki dotyczącej subiektywności (por. Searle 1998: xii–xvi; McGinn 1999: 84, 134, 229).

Patrząc na kryterium „*jak to jest* być świadomym” od strony tradycji filozoficznej należy powiedzieć, że domaga się ono od obserwatora tak zwa-

⁴ „The sad truth about the current state of consciousness studies appears to be that there is nothing even remotely resembling a scientific research program of consciousness” (Revonsuo 2000: 57).

⁵ Por. chociażby wspomnianą już teorię „Higher Order Thoughts” D. M. Rosenthala, który uważa, że zjawiskowy rodzaj przeżyć dostępnych z pierwszoosobowej perspektywy może służyć jako kryterium ważności danych neurofizjologicznych. Tym samym przeżycia te mogą być traktowane jako elementy uprawomocniające wyniki badań neurofizjologicznych, dotyczących kognitywnych i mentalnych stanów wyższego stopnia.

nego „wzucia” (*observer empathy*)⁶ w stany drugiego podmiotu. W historii filozofii do kryterium wzucia odwoływała się uprawiana fenomenologicznymi metodami etyka, estetyka oraz psychologia⁷. Zagadnienie wzucia (*Einführung*) było tematem pracy doktorskiej E. Stein. Drugiego człowieka, według Stein, poznajemy przez wzucie w niego swoich własnych, przeżyciowych treści. Nie posiadamy natomiast możliwości bezpośredniego dotarcia do jego treści wewnętrznych. Wzucie jest rodzajem poznania, niesprowadzalnym do postrzegania zewnętrznego. Zdaniem neurofenomenologów odwołanie się do kryterium wzucia jest warunkiem dla stworzenia nauki o świadomości⁸.

Neurolog B. Baars zauważa, że posługiwanie się kryterium wzucia stawia naukę przed wymaganiami, które nigdy dotąd nie zostały przez nią spełnione⁹. Zgodnie z tym kryterium, zewnętrzny obserwator powinien niejako „dzielić” przeżycia obserwowanego podmiotu. Z jednej strony praktyka dnia codziennego i rozwój neurofizjologii pokazują, że kryterium takie może być celem dociekań filozoficznych (por. Nagła poszukiwanie „obiektywnej fenomenologii”). Z drugiej strony lekarze nie odwołują się do niego przy podejmowaniu konkretnych decyzji. Kryterium wzucia nie mówi bowiem, co właściwie musimy wiedzieć o świadomości, lecz obstaje przy stwierdzeniu, że tak długo nie potrafimy sensownie mówić o jej subiektywności, jak długo

⁶ Baars 1996: 211. „Empatia” obejmuje zarówno „wzucie”, jak i „odczuwanie”, jako przeciwstawne sobie kategorie poznawcze.

⁷ Można tutaj wymienić chociażby M. Schelera czy J. Volkelta (por. Scheler 2000: 51, 58–59). Omawiając tezy Volkelta R. Ingarden pisze: „[...] to nie my wczuwamy nasz stan psychiczny w drugą osobę, lecz to my pod wpływem cudzego afektu, który nawet mimo woli uchwyciliśmy, zarazamy się jakoby uczuciowo, przejmujemy się cudzym smutkiem [...]” (Ingarden 1967: 142; 1971: 411–426). Husserłowski podział na wzucie „transcendentalne” i „światowe” omawia St. Judycki (1999: 221–231).

⁸ Por. Stein 1988: 19, 25; Varela 2000: 348. Według E. Thompsona: „Empathy is the precondition (the condition of possibility) for the science of consciousness” (Thompson 2001: 2; por. s. 16–23). T. Singer [i inni] dokonali neuroobrazowania mózgow 16 zakochanych par, mierząc aktywność mózgu każdej z partnerek. Rozróżniono dwa przypadki: (a) każda z kobiet została najpierw sama poddana krótkiemu wstrząsowi elektrycznemu, a następnie (b) jej partner. Kobiety nie widziały przy tym twarzy swoich partnerów. O momencie, w którym partner został poddany wstrząsowi, wiedziały jedynie z obserwacji wskazań aparatury. W obydwu przypadkach u wspomnianych kobiet były aktywne te same obszary mózgu z jednym wyjątkiem. Obszar mózgu rejestrujący ból cielesny był aktywowany tylko w przypadku (a), obszary mózgu sterujące emocjonalnymi reakcjami na ból (podniecenie, obawa, smutek) były aktywowane w przypadkach (a) i (b) (por. Singer 2004: 1157–1160).

⁹ Por. Baars 1996: 213. B. Baars, jak wspomnieliśmy, zwraca uwagę na konieczność uwzględniania perspektywy pierwszoosobowej w badaniach świadomości.

w jakiejś nieokreślonej przyszłości nie przeżyjemy, *jak to jest* być innym podmiotem¹⁰.

W licznych przypadkach neurologzy muszą umieć odpowiedzieć na pytanie: czy dana osoba w ogóle jest przytomna?

VI.2 W neurologii termin „świadomość” odnosi się najczęściej do przytomności, czyli do spontanicznych i uważnych zachowań oraz wypowiedzi. Termin „nieświadomość” (podobnie jak „śpiączka”) odnosi się do stanów przypominających głęboki sen, w których zewnętrzne lub wewnętrzne bodźce mogą wywołać jedynie odruchowe reakcje podmiotu.

Neurologzy badają, na przykład, reakcje podmiotu na bodźce świetlne¹¹, obrażenia jego głowy, stopień uszkodzenia mózgu. Należy dodać, że lekarz mający do dyspozycji jedynie zdjęcia z tomografii komputerowej, nie potrafi powiedzieć, czy pacjent jest przytomny, czy nie. Zdjęcia takie mogą natomiast pomóc ustalić, czy np. mamy do czynienia ze śmiercią mózgu¹².

Neurolog pyta zazwyczaj albo o istnienie jakiejś wymiernej dynamiki bądź struktury neuronalnej, świadczącej o świadomości podmiotu, albo o jakiś stopień jego „ważnego” reagowania na otoczenie¹³. Spojrzenie na podmiot zarówno w ramach potocznego obrazu świata, jak i w prostych lub złożonych badaniach medycznych, współbrzmi z tradycyjnym spojrzeniem filozoficznym przyjmującym, że obecność świadomości może być stwierdzona za pomocą praktycznych kryteriów. Należy do nich chociażby zdolność skierowania przez podmiot uwagi na określony przedmiot, stan lub zdarzenie.

¹⁰ B. Baars formułuje to następująco: „Still other philosophers claim that some aspects of consciousness, such as subjectivity, may be inherently inexplicable. But that implies a misunderstanding of the scientific enterprise. The goal is a modest increment in knowledge” (Baars 1998: 3).

¹¹ Por. Edelman 1998: 155–157. Przy tego rodzaju badaniach pomocna jest tzw. *Glasgow Coma Scale* (GCS). Standardowa skala (0–15 pkt.) uwzględnia reakcje motoryczne pacjenta, jego odpowiedzi językowe oraz ruchy oczu (por. Adamek 2003: 108). Skala ma podstawowe znaczenie w ocenie stopnia pourazowego uszkodzenia mózgu.

¹² Formalnie śmierć mózgu ustalamy na podstawie spełnienia całego szeregu kliniczno-neurofizjologicznych warunków. C. E. Reis wymienia warunki ustalone przez amerykańską *President’s Commission for the Study of Ethical Problems in Medicine*, wymagane do stwierdzenia śmierci mózgu (por. Reis 1996: 1–2).

¹³ W takich i podobnych przypadkach należy unikać wszelkich medycznych „redukcjonizmów”, które – według J. Kałuży – mogą wystąpić w dwóch odmianach: jako teoretyczne uzasadnienie jakiejś metody badawczej oraz jako pewien światopogląd (por. Kałuża 1992: 76).

Interesujący nas termin: „świadomość” odnosi się do subiektywnego *jak to jest* być. Różnicę pomiędzy VI.2 a tak rozumianą świadomością widać chociażby w przypadkach pacjentów, którzy po przebudzeniu z głębokiej narkozy potrafią sobie częściowo przypomnieć pewne (najczęściej celowo stymulowane) zdarzenia z operacji. Przypadki te wskazują na istnienie subiektywnej świadomości zjawiskowej, mimo całkowitego braku zewnętrznych, przytomnych zachowań (por. Merikle i Daneman 1996: 526–537; Merikle 2000: 497–499).

Z jednej strony, odwoływanie się do subiektywności i do kryterium wczucia może powstrzymać zarówno od teoretycznego zajmowania się świadomością, jak i od poszukiwania jej nowej teorii. Omawiane kryterium stawia bowiem podobne wymagania, jakie stawali witaliści na początku XX wieku. Ich zdaniem, mówiący o żywych organizmach biologowie powinni zaakceptować istnienie *élan vital* – bliżej nieokreślonej „siły witalnej”. Gdyby tak się wówczas stało, mogłoby to zahamować na wiele lat postęp w molekularnych badaniach nad zjawiskiem życia. Z drugiej strony, kryterium wczucia nie można oceniać w ten sam sposób, w jaki ocenia się pogląd zakładający istnienie *élan vital* lub *flogistonu* (por. 3.3.1). Na istotną różnicę pomiędzy problemem świadomości a problemem witalizmu wskazuje emergentysta C. D. Broad (por. 3.6.1). Z jednej strony, pojęciowe rozróżnienie na świadomość dostępu i świadomość zjawiskową zachowuje swoją ważność, a kryterium wczucia musi być uwzględniane w teoretycznych rozważaniach nad świadomością, chociaż w dzisiejszej praktyce lekarskiej często obywamy się bez niego. Z drugiej strony pojęcia typu „świadomość” mają więcej wspólnego z pojęciem „siły witalnej” aniżeli z pojęciami typu „atom” lub „gen”.

W dalszej części pracy poszukujemy kryteriów do określenia neuronalnych korelatów dla holistycznie rozumianej świadomości zjawiskowej i dla poszczególnych stanów tejże świadomości. Odwołujemy się przy tym zarazem do kryterium subiektywnej jedności świadomości, jak i do „pragmatycznych” kryteriów służących do określenia bycia świadomym. Takim pragmatycznym kryterium jest chociażby rozumienie świadomości w neurologii (por. VI.2). Wynikające z obu kryteriów programy badawcze nie muszą się (od strony badanego obiektu) wykluczać, chociaż można je oddzielić od strony teoretycznej i eksperymentalnej. Uwzględniając obydwa rodzaje kryteriów staramy się zbudować szeroko rozumianą metodologię, do której będzie się można odwołać przy poszukiwaniu neuronalnych korelatów świadomości. Używamy przy tym roboczej definicji korelatu:

VI.3 Korelatem świadomości jest specyficzny rodzaj aktywności mózgu, wykazujący nieprzypadkową więź z pewną określoną formą przeżycia psychicznego.

W poszukiwaniu nowej wielowymiarowej teorii świadomości uwzględniamy ujęcie jej subiektywności przedstawione w rozdziale II. Konkretnie oznacza to najpierw próbę odpowiedzi na pytania typu: czy istnieją znaczące różnice w aktywności neuronalnej pomiędzy podmiotowymi stanami świadomymi i nieświadomymi?¹⁴ „Materiałem” użytym do tworzenia nowej teorii świadomości są więc fakty neurofizjologiczne.

W filozofii analizujemy logiczne warunki możliwości wystąpienia subiektywności świadomości, zaś w naukach o systemie nerwowym bada się empiryczne warunki jej możliwości. Dlatego, z jednej strony, uprawnione jest pytanie: czy w ogóle są możliwe naukowe wypowiedzi co do subiektywności świadomości? Z drugiej okazuje się, że każde systematyczne połączenie pomiędzy filozofią a naukami o mózgu zakłada porównanie ujęć świadomości. Na podstawie tego rodzaju porównania zostają doprecyzowane pojęcia filozoficzne (choćby dzięki używaniu odpowiednich pojęć z zakresu neurologii). W podobny sposób, argumenty oparte na eksperymentach myślowych typu *jak to jest* być nietoperzem, czy eksperyment z Mary, mogą zostać przeniesione do ustalania empirycznych badań świadomości¹⁵. Tym samym badania naukowe zostają bezpośrednio skonfrontowane z filozoficzno-logicznymi warunkami dotyczącymi okoliczności i przejawów występowania określonych stanów świadomości.

6.1.2. Świadome przetwarzanie informacji

W neurofizjologii mózg jest traktowany jako organ o różnych stopniach organizacji: neurony, grupy neuronów, anatomiczne części mózgu. Z jednej strony, z badań organizacji i funkcjonowania mózgu nie wynika w sposób oczywisty i konieczny, że jest on miejscem, w którym zachodzą świadome przeżycia, chociaż, z drugiej strony, poza mózgiem trudno przyjąć jakąś alternatywę dla występowania myślenia i odczuwania (w sensie świadomościowych procesów). Ponadto nie widać jednoznacznej odpowiedniości pomiędzy holistycznie rozumianą świadomością a obszarami neurologicznymi.

¹⁴ Na istnienie takich różnic wskazuje B. Baars. „When a stimulus is presented unconsciously it activates areas in cortex involved in analysing colours, sounds, faces and the like. But when the identical stimulus is shown consciously, it also recruits regions far beyond the sensory cortex” (Baars 2003: 103). Por. też Rees 2002: 261.

¹⁵ W powstałym pod redakcją J. D. Schmahmanna opracowaniu *The Cerebellum and Cognition* zwraca się uwagę na potrzebę uwzględnienia wzajemnych korelacji pomiędzy filozoficznymi i neurologicznymi teoriami świadomości (por. Schmahmann 1997: 14–16, 258–260). Por. też Żegleń 2003: 283–284.

Oznacza to, że nie mamy jasnego pojęcia tego, co naukowo oznacza termin „świadomość”.

Dane eksperymentalne dotyczące ludzkich podmiotów uzasadniają przyjęcie założenia, że bez uwzględnienia świadomości zjawiskowej trudno zrozumieć funkcjonalną organizację mózgu. Stąd założenie:

VI.4 Zachodzące w mózgu przetwarzanie informacji może być adekwatnie zrozumiane jedynie przy uwzględnieniu różnicy pomiędzy jej przeżyciowo świadomym i nieświadomym przetwarzaniem.

Założenie VI.4 znajduje potwierdzenie chociażby w wynikach badań uzyskanych przez A. Revonsuo¹⁶. Pojęciowe odróżnienie pomiędzy świadomym i nieświadomym przetwarzaniem danych jest w kognitywnej psychologii zdefiniowane poprzez bezpośrednie odniesienie do *subiektywnego przeżycia*. Nieświadomiona informacja może mieć wpływ na skierowane na zewnątrz reakcje podmiotu, informacja ta nie ma jednak zjawiskowej reprezentacji w podmiocie¹⁷. W odróżnieniu od tego, świadoma informacja jest zjawiskowo reprezentowana (znajduje wejście do treści subiektywnych przeżyć).

A. Revonsuo dodaje, że zachowania podmiotu sterowane przez leżący u ich podstaw świadomy model świata, różnią się jakościowo od wszystkich innych zachowań. Na przykład stany mózgu pacjentów, mogących tylko nieświadomie spostrzegać bodźce sensoryczne (choćby w przypadku prozopagnozji¹⁸), są wprawdzie modulowane przez bodźce zewnętrzne, ale informacja o tych bodźcach nie znajduje dojścia do świadomego modelu świata pacjenta. Cierpiący na prozopagnozę pacjent nie potrafi używać nieświadomie reprezentowanych informacji do prowadzenia wolitywnych działań.

¹⁶ Por. Revonsuo 1998: 198–203. Mówiąc o funkcjonowaniu mózgu należy zauważyć, że „[...] ninety-eight percent of what the brain does is outside of conscious awareness” (Gazzaniga 1998: 121).

¹⁷ Możemy wprowadzić pojęciowe rozróżnienie na jawne i utajone procesy przetwarzania informacji (tzn. bez zjawiskowej reprezentacji w systemie). Nie analizujemy tutaj chociażby podziału informacji na informację przedmiotu (czyli obiektywną złożoność jakościowo-ilościową przedmiotu poznania), informację symboliczną (obiektywną złożoność znaków językowych) oraz na informację w sensie orientacji podmiotu w informacji przedmiotu (niezależnie, czy używa się języka ludzkiego czy „języka kodu biologicznego” – kodu molekularnego, kodu impulsów elektrycznych).

¹⁸ Prozopagnozja (czyli niezdolność do rozpoznawania znajomych twarzy), jest spowodowana przerwaniem połączenia spostrzeganie–pamięć. Pacjenci z prozopagnozą nie mają problemów z porównywaniem równocześnie prezentowanych twarzy, lecz problemy pojawiają się, gdy jedną z nich trzeba wcześniej zapamiętać.

Nieuświadomiona informacja ma oczywiście dla podmiotu jakąś wartość, gdy chodzi o jego biologiczne lub społeczne przeżycie (np. odruchy żreniczne czy wrodzone lub wyuczone automatyzmy zawierające „nieświadomą informację”). Inaczej jest jednak, gdy chodzi o bardziej kompleksowe działania podmiotu. Obojętnie bowiem, czy jakaś nieświadoma informacja jest obecna w jego mózgu czy też nie, nie potrafi on jej uwzględnić w swojej kompleksowej interakcji z otoczeniem. Nawet takie kompleksowe zachowania podmiotu, jak spacerowanie, są bardzo ograniczone (gdy chodzi o jego możliwości poznawcze), w porównaniu z jego całościowymi zachowaniami podlegającymi świadomej kontroli. Niemożliwe byłoby jednak naukowe opisywanie i wyjaśnianie spacerów lunatyków bez odwołania się do jakiejś neuronalnej cechy mózgu, która te zachowania umożliwia, i do jakiejś cechy, która odpowiada jakościowym cechom subiektywnych przeżyć lunatyków.

Schematy zachowania, nieopierające się na świadomym modelu otoczenia, nie gwarantują skutecznego dostosowania się do świata i nie wywołują tym samym żadnych nowych reakcji w nieznanymi warunkach zewnętrznych. W przeciwieństwie do tego świadoma informacja umożliwia powstanie: (a) globalnych, (b) zmiennych i (c) zintegrowanych sposobów zachowania¹⁹.

(a) Termin „globalne” dotyczy świadomej informacji i jej zdolności do wywołania procesów poznawczych i odpowiednio dopasowanych zachowań.

(b) „Zmienne” oznacza, że reakcje podmiotu nie są statyczne, lecz wybierane z różnych możliwości działania.

(c) „Zintegrowane” odnosi się do faktu, że różne reakcje tworzą jednolitą całość i nie są ze sobą w konflikcie.

Ponadto nasze zamiary i akty woli są skierowane na taki świat, jaki jest reprezentowany w naszym świadomym przeżywaniu. Jako przykład zachowań opartych na świadomym modelu świata Revonsuo przytacza przypadki zakłóceń zachowań w fazie snu REM²⁰. Ludzie cierpiący na tę dolegliwość wykonują gwałtowne ruchy podczas snu, odpowiadające aktualnie przeżywanej sytuacji (ucieczka, walka, ...). Ich relacje o zdarzeniach przeżytych we śnie są z reguły zgodne z obserwacjami ich zewnętrznych zachowań podczas snu. Oznacza to, że świadomy dla danego organizmu model świata nie musi odpowiadać faktycznemu, przeżywanemu na jawie otoczeniu organizmu.

¹⁹ Zachowania te znajdują swoje odwzorowanie w aktywności mózgu (por. Baars 2003: 104–111).

²⁰ Podczas fazy snu REM motoryka ciała jest zablokowana, w przypadku *REM-sleep-behavior-disorder* blokada zostaje zniesiona.

Także przeżyciowy świat utworzony podczas halucynacji służy jako „świat”, do którego organizm próbuje się dostosować.

Świadomy model świata – na przykład w śnie – ma swoją podstawę w procesach zachodzących w mózgu. Świadomy model świata jest tutaj przyczynowo niezależny od (i) spostrzegawczych wejść oraz od (ii) motorycznych wyjść. W normalnych przypadkach „bycia świadomym na jawie” świadomy model świata jest ściśle związany z (i) oraz z (ii).

Prowadzone obecnie badania w neurologii pozwalają przypuszczać, że wśród elementów kognitywnego systemu istnieją składniki funkcjonujące bez uświadomienia (nieuświadomione uczenie się, nieuświadomiona pamięć, ...). Mają one ograniczone funkcjonalne zadania i nie mają jakiejś jedynej formy. Potocznie rozumiane formy wiedzy czy przypomnienia sobie należą do stopnia świadomego i nie można ich pomijać w badaniach neurologicznych.

Wracając do VI.4 można powiedzieć: jeżeli zachowanie podmiotu, uwzględniające świadomy model świata, odróżnia się jakościowo od innych zachowań, to trudno to zachowanie adekwatnie opisać i wyjaśnić bez używania w neurologii pojęć odnoszących się do świadomości zjawiskowej. Zachowania i przeżycia, zakładające świadomy model świata, nie mogą zostać jednoznacznie opisane i wyjaśnione bez uwzględnienia tego stopnia, który jest decydujący dla adaptacyjnej roli tego zachowania lub przeżycia. Taki zaś jest stopień świadomego przeżywania i odpowiadający mu neuronalny korelat. Trudno sobie wyobrazić, jak miałyby wyglądać badanie zachowania się podmiotu pozbawionego subiektywnej świadomości, podobnie jak trudno sobie wyobrazić badanie lokomocji tego podmiotu, bez uwzględnienia jego przemieszczania się.

6.1.3. Nowe techniki neuroobrazowania

Poszukiwanie neuronalnych korelatów świadomości weszło w zakres kognitywistyki i jest uprawiane nie tylko przez neurologów czy neuroradiologów, lecz także przez psychologów i filozofów oraz, w swoim analogicznym sposobie, przez twórców sieci neuronowych. Holistycznie rozumiana świadomość nie podpada w tym sensie pod naukowe metody badawcze, jak podpadają pod nie procesy obliczeniowe lub neurologiczne. Nie mamy bowiem żadnych obiektywnych metod, aby móc ją bezpośrednio badać. Tym samym jesteśmy skazani na uwzględnianie pośrednich metod badania świadomości. (analiza pierwszoosobowych wypowiedzi podmiotu, neuroobrazowanie).

Neuropsycholog A. Łuria, opisując mózg swojego pacjenta, dysponował jedynie zdjęciami rentgenowskimi jego głowy (por. 6.3.1). Obecnie w bada-

niach systemu nerwowego najbardziej znanymi i stosowanymi technikami są: tomografia komputerowa (CT), pozytronowa tomografia emisyjna (PET), rezonans magnetyczny (MRI), czynnościowy rezonans magnetyczny (fMRI), elektroencefalografia (EEG), magnetyczna elektroencefalografia (MEG). Krótkie omówienie wymienionych „nieinwazyjnych technik” oraz ich zastosowanie przy neuroobrazowaniu mózgu, a zwłaszcza bólu podają: B. Bromm, D. Lloyd, G. Rees oraz M. Steuden²¹. Techniki te nie opierają się na bezpośrednim pomiarze neuronalnej aktywności, lecz na pomiarze zmian w metabolizmie, zapoczątkowanym zwiększoną neuronalną aktywnością, wywołaną przez odpowiednie bodźce. Na przykład podawana w PET znakowana glukoza ¹⁸F–2–DG (deoksyglukoza) jest akumulowana w aktualnie aktywnych miejscach mózgu.

Wspomniane techniki pozwalają na badanie anatomicznych obszarów żyjących mózgow, a dostępna rozdzielczość uzyskanych za ich pomocą obrazów (strukturalnych lub funkcjonalnych) umożliwia przybliżone określenie neuronalnych korelatów świadomości (na przykład wrażeń wizualnych lub odczuwania bólu)²². Uzyskane wyniki prowadzą do ponownego sformułowania niektórych problemów metodologicznych, obecnych w badaniach zranień mózgu. Omawiane metody neuroobrazowania, jak każde inne metody pomiarowe, posiadają swoje nieprzewidywalne ogranicze-

²¹ Por. Bromm 2001: 244–246; Lloyd 2003: 132–147; Rees 2002: 262–264; Steuden 1998: 38–56 (Steuden podaje także duży wybór literatury dotyczącej wspomnianych technik). Od roku 1972 używano jedynie metody tomografii komputerowej TK (CAT, *Computer Axial Tomography*, znanej także jako *Computer Assisted Tomography*), będącej pierwszą, opartą na promieniach rentgenowskich, komputerową wizualizacją ciała wraz z mózgiem. Uzyskane wyniki komputer przetwarza na dwu- lub trójwymiarowe obrazy (przekroje). Na początku lat 80. zaczęto używać metody MRI, cechującej się większą rozdzielczością i nieużywającej emisji szkodliwych promieni rentgenowskich, lecz emisji fal radiowych. Nowszą odmianą MRI jest tak zwane fMRI. Za pomocą CAT i MRI uzyskujemy strukturalne obrazy (precyzują one strukturę mózgu, za pomocą MRI można np. odróżnić korę od istoty białej). W tych samych latach wprowadzono PET, podczas której pacjentowi zostają podane radioaktywne substancje o krótkim okresie połowicznego rozpadu, emitujące pozytrony. Otrzymana od nich informacja jest następnie przetwarzana przez komputer na odpowiednie obrazy. PET albo fMRI dostarczają funkcjonalnych obrazów mózgu.

²² „First, neuroimaging allows for paradigms in which subjects (healthy humans) can be trained to perform more intricately controlled, and more cognitive, tasks [...]. Second, whereas researchers in neurology and neuropsychology must wait for accidents of nature to yield «pure» cases of selective brain damage, a neuroimaging study can be designed to efficiently localize components of the neural substrate of a mental process and to be conducted in a relatively short amount of time” (Parsons 1997: 256).

nia²³. Badania PET dostarczają jedynie pośredniego, przybliżonego wglądu w neuronalne korelaty świadomości, z przestrzenną rozdzielczością ograniczoną do kilku milimetrów. Uzyskane tą drogą obrazy można oczywiście uśredniać, biorąc pod uwagę kolejne badania tego samego podmiotu lub innych podmiotów.

Analizując uzyskane wyniki należy brać pod uwagę, czy aktywność danego obszaru mózgu nie jest rezultatem jakiegoś zmniejszenia lub zwiększenia metabolizmu, którego przyczyny leżą gdzie indziej. Ponadto glukoza lub tlen, zużyte przez hamujące interneurony, mogą wytwarzać lokalne maksima aktywności w regionach, które faktycznie są mniej aktywne z powodu występującego w nich zahamowania. Także funkcjonalna aktywność jakiegoś obszaru w mózgu może się zmieniać (jak to ma na przykład miejsce podczas reorganizacji kory wzrokowej po uszkodzeniu warstwy V1²⁴).

Szeroko stosowane badania korelatów świadomości metodą fMRI porównują jeden stan zdawania sobie przez podmiot sprawy z drugim stanem, przyjmując lokalizację tych stanów i ignorując wpływ czasu. Skupiają się one na dwóch parametrach: na przestrzennej rozpiętości aktywacji i na amplitudach odpowiedzi w zaktywowanym regionie. Uzyskane wyniki są następnie przedstawiane na wykresach albo histogramach ukazujących zmianę sygnału fMRI. Otrzymujemy w ten sposób statystyczne mapy wskazujące na lokalizację i wielkość aktywizacji. Mając daną czasową i przestrzenną rozdzielczość (*resolution* – cecha ta określa rozróżnialność dwóch punktów na obrazie), okazuje się problematycznym wydzielenie czasowej informacji o „trójdzielności czasowości” (*tripartite temporality* – co to znaczy, że przeżywana terażniejszość jakiegoś przedmiotu jest pod wpływem przeszłości i oczekiwanej przyszłości?). Określenie funkcji hemodynamicznej (*hemodynamic function*, czyli podanie warunków przepływu krwi w naczyniach krwiono-

²³ Jednym z modeli interpretowania wyników uzyskanych w fMRI jest tzw. *linear transform model*, w którym przyjmuje się, że sygnał fMRI jest w przybliżeniu proporcjonalny do pomiaru lokalnej aktywności neuronalnej, uśrednionej z przestrzennej rozpiętości rzędu kilku milimetrów oraz z rozpiętości czasowej rzędu kilku sekund (por. Heeger 2002: 143–149). „First, the relationship between fMRI and neuronal responses depends on the fMRI acquisition technique. Most fMRI experiments measure a blood oxygen level dependent (BOLD) response. BOLD fMRI provides a mixed signal that depends on blood flow, blood volume and blood oxygenation” (s. 144). „Second, the relationship between fMRI and neuronal responses depends on the behavioural and stimulation protocols, and on the fMRI data-analysis. [...] Third, the relationship between fMRI and neuronal responses depends on how the neuronal activity itself is measured and quantified” (s. 145).

²⁴ Trudności związane z pomiarowym określeniem neuronalnego korelatu dla wrażeń wizualnych omawia J. Searle (2000: 571–572).

śnych) staje się poważnym wymaganiem, o ile przepływ ten w każdym momentalnym punkcie czasowym zdaje się być „zakłócony” przez wpływ pochodzący od bezpośrednio poprzedzającej ten moment odpowiedzi podmiotu²⁵. Dlatego D. J. Heeger mówi jedynie o korelacji pomiędzy zużyciem glukozy i tlenu oraz przepływem krwi, a nie o zależnościach przyczynowych (rozumianych jako relacje dynamiczne zachodzące pomiędzy zdarzeniami w czasie)²⁶.

Badanie mózgu na poziomie komórkowym i subkomórkowym pokazuje, że prawie niemożliwe jest uzyskanie szczegółowych map neuronalnych połączeń i ich aktywności, a tym samym zajmowanie się ich kompleksowością. Pojawia się więc kolejna wątpliwość dotycząca możliwości odkrycia dostatecznie precyzyjnych neuronalnych korelatów świadomości. Otwarte pozostaje pytanie: jaka płaszczyzna opisu mózgu, jaki rodzaj abstrakcji i uogólnienia jest najbardziej odpowiedni dla celów filozoficznych? Nowe techniki dostarczają wprawdzie dodatkowego wglądu w korelacje mózg–zachowania, mózg–przeżycia, ale odwołując się jedynie do nich trudno ustalić, która część mózgu, jakie jego struktury są na sposób istotny zaangażowane w realizację jego funkcji odpowiadających świadomym przeżyciom. Dlatego wyniki uzyskane za pomocą technik neuroobrazowania należy uzupełnić albo porównać z danymi uzyskanymi w badaniach uszkodzeń mózgu.

Zdaniem A. Revonsuo, wspomniane techniki są ograniczone od strony pomiarowej i dlatego nie doprowadzą do odkrycia w mózgu neuronalnej podstawy świadomości. Jest tak dlatego, gdyż zasadniczą rolę w utrzymaniu stanu czuwania i we wspomnianych już procesach spostrzegania wzrokowego odgrywa neuronalna synchronizacja pomiędzy anatomicznie poroździelanymi grupami neuronów²⁷. Dopiero technika, potrafiąca mierzyć anatomicznie poroździelaną, szybko zmieniającą się synchronizację w odpalaniu neuro-

²⁵ Por. Ishai 2002: 832–833. W teoretycznym modelowaniu wpływu przeszłości (pamięci) na aktualne przeżywanie można się odwołać do krytykowanych przez D. Dennetta modeli świadomego umysłu (stalinowskiego i orwellowskiego).

²⁶ Według Heegera: „Although the blood flow response is correlated with glucose consumption under normal circumstances, it might be caused by something else. Likewise, the blood flow response is correlated with, but not necessarily causally related to, oxygen consumption” (Heeger 2002: 144).

²⁷ Por. Revonsuo 2001: 5–21. Według Revonsuo: „[...] the level of organization at which consciousness resides might be a highly organized macro-level electrophysiological phenomenon in the brain” (s. 5). Revonsuo przedstawia metody MEG, EEG, ERP. Omówienie i krytykę tez Revonsuo można znaleźć w: Rees 2001: 35–36 oraz Lloyd 2003: 343.

nów, może doprowadzić do odkrycia świadomości w mózgu. Z powyższym twierdzeniem Revonsuo można się zgodzić lub nie zgodzić. Z jednej strony, musielibyśmy umieć przewidzieć dalszy rozwój technik pomiarowych, z drugiej – widać, jak postawione przez niego żądanie odnośnie do technik pomiarowych zależy od jego rozumienia neuronalnego odpowiednika „świadomości”. Tego rodzaju odpowiednikiem jest dla niego synchronizacja neuronalnych pobudzeń.

Ponadto zaawansowane techniki badawcze przystępują niekiedy przedmiot badania; odbiorcy – a także niektórzy badacze spoza zaawansowanej technologii – nie dostrzegają błędów myślenia w interpretacjach wyników badań z jej użyciem. Rodzą się wtedy takie – dzisiaj popularne – pomysły, jak np. neuroteologia.

W następnych punktach: *Koneksjonizm* (6.1.4) i *Neuroplastyczność* (6.1.5), zostaną omówione dwie cechy mózgu, stawiające dodatkowe wymagania zarówno wobec technik neuroobrazowania, jak i przed metodologią poszukiwania neuronalnych korelatów świadomości. Uzyskane na drodze neuroobrazowania wyniki oraz oparte na analizach pojęciowych określenia neuronalnych korelatów świadomości stanowią dla nas punkt wyjścia do zbudowania poszukiwanej wielowymiarowej teorii świadomości.

6.1.4. Koneksjonizm

W neurofizjologii mózg jest traktowany jako organ służący do utrzymania, regulacji i integracji biologicznych funkcji organizmu. Mózg jest ukształtowanym przez formę żywą narzędziem, które jest budowane, reorganizowane i reperowane. Główne czynności mózgu są niezbędne dla przeżycia organizmu i różnią się od czynności, którymi interesuje się filozofia czy psychologia, mówiące o „świadomym” mózgu. Istnieje wiele procesów w mózgu, które nie posiadają bezpośredniego świadomego odpowiednika, chociaż go warunkują (na przykład procesy regulujące produkcję hormonów). Zadaniem mózgu, podobnie jak każdego organu, jest, aby posiadający go organizm odniósł sukces polegający na przeżyciu. Dlatego bardziej niebezpieczne dla organizmu są uszkodzenia głębszych części mózgu (filogenetycznie starszych, odpowiadających za zabezpieczenie stałych fizjologicznych warunków w organizmie, pomimo zmiennego otoczenia), od uszkodzeń zewnętrznej powierzchni mózgu (tzn. obszarów filogenetycznie młodszych, odpowiadających za tzw. wyższe czynności korowe). Uszkodzenie obszarów młodszych częściowo upośledza niektóre funkcje świadomości, nie eliminując jej całkowicie. Przebieg procesów regulacyjnych i integracyjnych ma oczywiście wpływ na psychofizyczną kondycję organizmu.

Podczas gdy mózg tworzy jedynie dwa procent masy ciała, zużywa przeciętnie 20 procent całkowitej energii dostępnej organizmowi. Mózg nie ma żadnych rezerw cukru czy tlenu i „żyje z ręki do ust”. Przy braku tlenu już po paru minutach zachodzą w nim nieodwracalne zmiany²⁸. Wymiana materii w mózgu wywiera wpływ na przebieg czynności poznawczych (gdy jesteśmy głodni, gdy mamy niski poziom glukozy we krwi albo gdy brakuje nam tlenu, to trudno nam się skupić, myśleć, rozwiązywać problemy).

Definiując neuronalne korelaty świadomości należy określić, co rozumiemy pod pojęciem neuronów, połączeń neuronalnych, „miejsca” neuronów, oraz „czasowości” procesów neuronalnych. Neurony stanowią genetycznie anatomiczną i funkcjonalną jednostkę układu nerwowego²⁹. W swoim drzewie dendrytycznym neurony odbierają impulsy synaptyczne, które następnie są przesyłane do ciała komórki i zamieniane we wzgórkach aksonalnych na sekwencje następujących po sobie – z różną częstotliwością – tzw. potencjałów czynnościowych (por. Fix 1997: 86–78). Potencjał czynnościowy można (w uproszczeniu) ująć jako zdarzenia binarne (tzn. zachodzące według zasady „wszystko albo nic”) i traktować jako neuronalny nośnik informacji³⁰. Elektrochemiczne własności neuronu są rezultatem wielości wbudowanych w błony komórkowe kanałów i pomp jonowych i różnorodności ich charakterystyk (np. w przypadku kanałów jonowych jest to tzw. bramkowanie napięciem elektrycznym albo ligandem czy molekułą swoiście aktywującą kanał jonowy).

Termin „miejsce” neuronu oznacza jego pozycję wewnątrz sieci neuronów tworzących określoną strukturę mózgową. Funkcjonalne „miejsce” neuronu jest określone przez jego połączenia z innymi neuronami. Na wyższej płaszczyźnie organizacji funkcjonalne miejsce jest definiowane przez wzory wejściowych i wyjściowych struktur mózgowych oraz wyjść sensorycznych

²⁸ Wpływ zaburzeń krążenia mózgowego na procesy poznawcze omawia B. Daniluk (por. Daniluk 2000: 384–402).

²⁹ Chodzi o tak zwaną doktrynę neuronalną (por. Fix 1997: 83). J. D. Fix podaje szczegółowe podziały neuronów oraz funkcje poszczególnych rodzajów neuronów. Według J-P. Changeux „The neuron today is situated at the convergence of two lines of research: that of the chemist and molecular biologist, who consider it a system of interacting macromolecules, and that of neurobiologist and embryologist, who, on the contrary, look upon it as a basic unit from which the brain is constructed” (Changeux 1997: 39).

³⁰ Nie jest to jedyny neuronalny nośnik informacji, gdyż szeroko rozumiana „neurotransmisja” to nie tylko potencjały czynnościowe, ale także inne nośniki informacji, niekoniecznie związane bezpośrednio z potencjałem czynnościowym. Są to chociażby „niekonwencjonalne” neurotransmitery, takie jak NO, CO.

czy motorycznych³¹. Mówiąc o neuronach mamy na myśli elementy struktur mózgowych [Φ_j – Φ_n], których aktywność może zostać w jakiś mierzalny sposób zobrazowana.

W historii neurologii znane są dwie ogólne teorie funkcjonowania mózgu: lokalistyczna (*localist*) i holistyczna (*holist*)³². Według teorii lokalistycznej, mózg jest zbiorem obszarów, z których każdy jest możliwy do zlokalizowania i odpowiada za dane stany świadomości zjawiskowej oraz za funkcje i ruchy całego ciała (por. Duus 1989: 346–351). Jednak pochodzące od różnych naukowców neuronalne mapy mózgu – przedstawiające umiejscowienie tzw. ośrodków odpowiedzialnych za poszczególne czynności w korze mózgowej na podstawie budowy strukturalnej – różnią się od siebie. W miarę precyzyjne mapy mózgu jednego podmiotu odbiegają od map innego, a także mogą one ulegać zmianom u tego samego podmiotu (na przykład na skutek uszkodzeń mózgu). Stąd trudno powiedzieć, że dwa stany neuronalne będą dokładnie podobne, chociaż od strony molekularnej moglibyśmy nawet (czysto teoretycznie) znaleźć dwa identyczne mózgi. Według teorii holistycznej mózg jest wysoce zintegrowanym organem, a żadna jego pojedyncza funkcja nie może być przypisana jakiemuś jego szczególnemu układowi.

Obecnie lokalizm i holizm zostały zastąpione „koneksjonistycznym” (*connectionism*) modelem mózgu. Zgodnie z nim, podstawowe funkcje sensoryczno-motoryczne są ściśle zlokalizowane, natomiast funkcje wyższe (rozpoznawanie obiektów, pamięć, język), są wynikiem oddziaływania pomiędzy odpowiednimi obszarami mózgu³³.

Na podstawie empirycznych danych dotyczących mózgu można próbować utworzyć jego holistyczną teorię. Holizm mózgu znajduje chociażby potwierdzenie w tym, że komórek samej tylko kory mózgowej (czyli miejsca lokalizacji tak zwanych wyższych czynności mózgowych; mowy, myślenia, percepcji, złożonej motoryki) jest około 10 miliardów (por. Edelman 1998:

³¹ Por. Durstewitz i Windmann 1998: 133–134. Mowa tu o wstępujących włóknach nerwowych (przewodzących np. impulsy bólu) i o zstępujących włóknach (przebiegających od centralnego systemu nerwowego do mięśni).

³² Debata holizm–lokalizm była żywa w XIX oraz w XX wieku i brali w niej udział neurologi: F. Gall, K. Brodmann, P. Broca, K. Lashley, W. Penfield.

³³ Do koneksjonistycznego modelu funkcji mózgu odwołuje się J. Konorski, według którego na współczesnym mu etapie rozwoju fizjologii „pytania «jak» i «dlaczego» są o wiele ważniejsze niż pytanie «gdzie» [dotyczące lokalizacji określonych funkcjonalnie «ośrodków» i okolic mózgu – *J.B.*]” (Konorski 1969: 10). Konorski posługuje się blokowymi modelami funkcjonowania części mózgu, pozbawionymi dokładniejszej lokalizacji anatomicznej.

30–35). Każda komórka nerwowa, poprzez synapsy, tworzy połączenia z innymi komórkami, co daje ogromną liczbę połączeń. Licząc możliwą kombinację takich połączeń, dostaniemy hiperastronomiczną liczbę³⁴. Tym samym otrzymamy (teoretycznie) prawie nieskończenie wiele neuronalnych systemów $[\Phi_j-\Phi_n]_h$, których nie należy traktować jako proste elementy, lecz jako związane z całym mózgiem. Zachowania się danego systemu mogą być określone i zrozumiałe tylko od strony całego mózgu, a sam mózg jest czymś więcej aniżeli prostą sumą tworzących go systemów. Tak rozumiany holizm mózgu odsuwa teoretyczno-eksperymentalny program badań świadomości w nieosiągalną przyszłość.

Tego rodzaju holizm oznacza, że neurologia nie będzie mogła używać systemu $[\Phi_j-\Phi_n]_h$ do swoich celów. Neurologia potrafi dostarczyć jedynie regionalnego pomiaru i obrazu systemu $[\Phi_j-\Phi_n]_h$, tzn. jedynie jego makroopisu. Obecność systemów $[\Phi_j-\Phi_n]_h$ trudno wykazać pomiarowo. Tym samym odkrywane eksperymentalnie makrosystemy $[\Phi_j-\Phi_n]_h$ nie mogą być identyfikowane z mikrosystemami $[\Phi_j-\Phi_n]_h$. Miejskowa (stwierdzalna eksperymentalnie) lokalizacja danego systemu nie gwarantuje, cechującej się jednością, fizycznej realizacji typu $[\Phi_j-\Phi_n]_h$. Coś sklasyfikowane przez neurofizjologię jako system $[\Phi_j-\Phi_n]_h$, nie jest tym samym, co holistyczny model mózgu klasyfikuje jako system $[\Phi_j-\Phi_n]_h$.

Liczba i gęstość sieci neuronowych nie jest jedyną cechą materii mózgu. Swoisty jest sposób łączenia się komórek nerwowych w układy funkcjonalne. Pień mózgu (*truncus cerebri*) i międzymózgowie (*diencephalon*) zawierają wyodrębnione i niekiedy widoczne gołym okiem struktury istoty szarej (czyli zawierającej ciała komórek nerwowych) zwane jądrami (*nuclei*), a z kolei kora mózgowa składa się z warstw (*laminae*). Jądra składają się z ogromnej liczby neuronów i spełniają określone funkcje w złożonej sieci wzajemnych powiązań. Ze światem zewnętrznym mózg jest połączony za pomocą neuronów czuciowych. Wchodzą one w skład narządów zmysłów i przekazują sygnały do mózgu. Większa część tkanki mózgu otrzymuje i wysyła sygnały jedynie w obrębie mózgu. Zdaniem G. M. Edelmana: „Można powiedzieć, że mózg kontaktuje się znacznie częściej z samym sobą niż z czymkolwiek innym” (Edelman 1998: 33).

Prowadzone w dalszej części niniejszego opracowania analizy opieramy o twierdzenia zbliżone do koneksjonizmu.

³⁴ Ten fakt był podstawą ataku na determinizm oparty na genach DNA. DNA człowieka – jako zjawisko chemiczne – nie może reprezentować takiego poziomu informacji. Czysto kwantytatywne proporcje to wykluczają. Np. DNA człowieka ma ok. 4 miliardów par nukleotydów.

(1) część międzymózgowia, tzw. wzgórze (*thalamus*) przekazuje i rozdziela bodźce czuciowe (oprócz węchu) do kory mózgowej (płat ciemieniowy) i innych części mózgu oraz do rdzenia kręgowego³⁵. Wzgórze otrzymuje także bodźce z kory mózgowej.

(2) Mówiąc o neuronalnych korelatach świadomości oraz o neuronalnych połączeniach wewnątrz mózgu mamy na myśli połączenia pomiędzy jego anatomicznymi częściami (mózdzek, pień mózgu, wzgórze, kora mózgowa, warstwy kory mózgowej, por. 7.3.1). W tym sensie zakładamy modułarną (funkcjonalną) budowę mózgu. Uszkodzenie jednego modułu nie eliminuje z konieczności funkcji pozostałych modułów i nie zmienia radykalnie stanu całego mózgu.

(3) Według neurologa J. E. Bogaena, wyjaśnianie świadomości musi się rozpocząć od uznania jej subiektywności. Podstawą dla pojawienia się subiektywności, swoistym „rozszanym centrum”, wywołującym subiektywność świadomości, są dla Bogaena części wzgórza, tzw. jądra międzyblaszkowe (*intralaminar nuclei*)³⁶. Bogen zaznacza, że chociaż duże zranienia w wielu częściach mózgu nie powodują utraty świadomości, a jedynie jej modyfikację, to nawet małe uszkodzenia w *intralaminar nuclei* powodują nieodwracalną utratę świadomości.

6.1.5. Neuroplastyczność

Metoda definiowania i poszukiwania neuronalnych korelatów świadomości zależy od uwzględnienia bądź nieuwzględnienia w niej tak zwanej neuroplastyczności mózgu³⁷. Ta ostatnia pozwala organizmowi na zaadaptowanie się do zmienionych warunków wewnętrznych (wynikających na przykład z uszkodzenia części mózgu), lub zewnętrznych (dążenie do osiągnięcia określonego celu poprzez wykonywanie ćwiczeń). Przykładem neuropla-

³⁵ Por. Walsh 2000: 143–148. Podstawowe wiadomości o funkcjonowaniu mózgu podaje W. Duch (2002: *passim*). Por. także Lenk 2001: 1–20.

³⁶ Zdaniem J. E. Bogaena: „[...] conscious awareness (more precisely what I call C) is engendered by neuronal activity in an immediately around the intralaminar nuclei (ILN) of each thalamus” (Bogen 1995: 52). Bogen podaje diagram obrazujący projekcje z różnych regionów mózgu do ILN (s. 55). Rolę tych części mózgu w tworzeniu subiektywności przeżytych przedstawia A. Damasio (1995: 240–241).

³⁷ Zdaniem M. Kossut: „Plastyczność dorosłego mózgu i neurorehabilitacja to procesy i działania, które jeszcze nie tak dawno nie były uznawane przez neurologię. Głębokie przekonanie o stałości struktury mózgu i funkcji układu nerwowego dorosłego człowieka utrudniało akceptację nowych danych eksperymentalnych, a nawet zauważenie plastyczności spontanicznie występującej u pacjentów z urazami mózgu” (Kossut 2002: 16).

styczności jest zdolność mózgu do regeneracji oraz przejmowania przez jedne jego układy funkcji innych układów. Na razie nie wiadomo, na ile wspomniany wpływ warunków wewnętrznych i zewnętrznych może być antycypowany przy genetycznym rozwoju organizmu. Neuronalne i behawioralne zmiany przypisywane plastyczności mózgu są obserwowane w dużych i małych skalach czasowych.

Do metod leczenia patologii mózgu (zwłaszcza niektórych form lekoopornej padaczki u dzieci) należy hemisferektomia, czyli resekcja jednej półkuli mózgu. Zdarzają się pacjenci, którzy po takiej operacji prowadzą zwyczajne życie i posiadają normalny IQ. J. R. Skoyles opisuje pacjentkę, której w wieku pięciu lat usunięto dominującą, czyli lewą, półkulę. Pacjentka ukończyła szkołę podstawową. Grała na rogu w orkiestrze szkolnej. W wieku 26 lat pracowała jako kontroler jakości i ukończyła studia administracyjne. Jej IQ utrzymywał się przez cały czas na poziomie 126 (por. Skoyles 1999: 6). Ogólnie przyjmuje się, że jeżeli dziecko do szóstego roku życia dozna ciężkiego uszkodzenia mózgu w obrębie lewej, dominującej półkuli, przeważnie bez większego wysiłku udaje się po odpowiednich ćwiczeniach przekształcić prawą półkulę w dominującą (przejmuje ona wszystkie czynności w pełnym zakresie dominującej półkuli). U osób dorosłych nie jest to możliwe (por. Duus 1989: 376).

Nowsze badania kwestionują tezę, że wszystkie utworzone w pierwszych latach życia połączenia mózgowych komórek nerwowych są na zawsze ustalone oraz że plastyczny jest tylko młody mózg (tzn. że jeżeli konkretna część mózgu zostaje uszkodzona w późniejszym wieku, jego komórki nerwowe nie tworzą nowych połączeń). Według M. Kossut:

„Zmiany siły wejść czuciowych, zmiany aktywacji układów neuromodulujących przez emocje i uwagę, zmiany bodźców występujących w otoczeniu, mogą naruszyć lokalną równowagę szlaków, struktur i zespołów neuronalnych i zmienić własności odpowiedzi neuronów. Na ogół takie szybkie zmiany są przejściowe, jednak mogą prowadzić do trwałej modyfikacji siły synaps” (Kossut 2002: 17; por. Papathanasiou 2003: 346–352).

Do pewnego stopnia mózg reorganizuje i tworzy nowe połączenia przez całe życie podmiotu. Dzięki neuroplastyczności mózg może kompensować uszkodzenia i dopasować swoje czynności do nowych sytuacji albo zmian w otoczeniu (struktura reprezentacji w mózgu jest utrzymywana poprzez dynamiczne interakcje pomiędzy zespołami neuronalnymi) [por. Łuria 1976: 28–30]. Nadal jednak nie wiemy, czy reorganizacja dorosłego mózgu oznacza tworzenie nowych połączeń pomiędzy neuronami czy też wzmocnienie już istniejących.

Neurolog M. S. Gazzaniga nie kwestionuje eksperymentalnych danych dotyczących neuroplastyczności mózgu i przyznaje, że w mózgu istnieje swoista plastyczność (uczymy się przecież obcych języków i tym podobnych rzeczy). Jego pytanie brzmi: jak ową plastyczność wyjaśnić? Jedną ze skrajnych interpretacji podkreśla rolę przeżyć $[\psi_n]$ w formowaniu połączeń w mózgu³⁸, druga skrajna interpretacja podkreśla zdeterminowanie genetyczne: rozwój mózgu przebiega według ściśle genetycznego planu. Gazzaniga skłania się ku tej drugiej: chociaż prawdą jest, że normalny rozwój mózgu wymaga występowania swego rodzaju przeżyć, to nie wynika z tego, że nie jest on zdeterminowany przez czynniki genetyczne. O plastyczności mózgu możemy mówić tylko w ramach określonych genetycznych granic³⁹.

Teza Gazzanigi ma oczywiście daleko idące skutki filozoficzne, gdy chodzi o poszukiwania neuronalnych korelatów świadomości. Rodzi się bowiem pytanie, na ile można mówić o jakichkolwiek uogólnieniach uzyskanych wyników, skoro wiadomo, że nawet bliźnięta, rozwijające się z podziału tej samej zapłodnionej komórki jajowej, nie mają takiego samego mózgu (por. Changeux 1997: 206–212).

Neuroplastyczność mózgu w sensie ścisłym oznacza zdolność jego części do przejmowania funkcji spełnianych dotychczas przez uszkodzone albo usunięte części. Neuroplastyczność można jednak rozumieć szerzej. Idealnego modelu do badań szeroko rozumianej neuroplastyczności dostarczają profesjonalni muzycy⁴⁰. Przy nauce gry na skrzypcach powiększa się między innymi

³⁸ Mózg zrównuje się wtedy z empiryczną *tabula rasa*, oczekującą na „zapisanie jej” przez przeżycia. Do tego typu ujęć można zaliczyć „neural constructivism” Sejnowskiego (por. Gazzaniga i Gallagher 1998a: 706).

³⁹ M. S. Gazzaniga pisze: „The broad scaffolding of the brain is built by genetic mechanisms, which also control almost independently the specifications of what connects to what, but the details of cortical arrangements might be left to experiential effects. However, the so-called experiential effects are merely brain activity, not necessarily encoded information from the environment” (Gazzaniga 1998: 46). Proponowane przez Gazzanigę wyjaśnienie plastyczności mózgu oraz jego teorię uczenia się przez dziecko nowych treści S. Gallagher nazywa „platońskim neuralizmem” (ze względu na platoński „natywizm”).

⁴⁰ Według T. F. Münte [i innych]: „Source analysis revealed that the cortical representation of the digits of the left hand (the fingering hand) was larger in these musicians than in controls. In the case of the right hand, in which no independent movements of the fingers are required in string players, there were no difference between musicians and controls” (Münte 2002: 474). Wspomnianie badania, które przeprowadzono metodą MRI, pokazują, że niektóre obszary mózgu (*planum temporale*, *anterior corpus callosum*, *the primary motor hand area*) różnią się w swojej strukturze i rozmiarach, gdy chodzi o mózgi muzyków i mózgi podmiotów kontrolnych.

obszar tkanek mózdzku [$\Phi_j-\Phi_n$] odpowiedzialnych za kontrolę lewej ręki [$\Omega_i-\Omega_n$]⁴¹. Geneza szeroko rozumianej plastyczności mózgu jest zazwyczaj wyjaśniana poprzez tezę o doskonaleniu się dendrytów, synaps (aktywacja słabych synaps) i neuronów (zmiany na poziomie mikroskopowym) i poprzez tezę o hamowaniu bądź przyspieszaniu istniejących połączeń pomiędzy neuronami wewnątrzkorowymi.

(2) Opisany przez O. Saksę przypadek pacjenta cierpiącego na chorobę Alzheimera może być przykładem neuroplastyczności mózgu, stymulowanej przez świadome nastawienia woli pacjenta. Do momentu wystąpienia objawów choroby, pacjent był utalentowanym muzykiem i śpiewakiem operowym. W czasie trwania choroby znalazł on „sposób” na skupienie i zorganizowanie swojej mocno nadwyreżonej neurofizjologii [$\Phi_j-\Phi_n$].

„Wszystko robi śpiewając. Jeśli coś mu przeszkodzi, nie wie, co ma robić dalej, nieruchomieje, nie poznaje ubrania ani własnego ciała. Cały czas śpiewa – ma piosenki do jedzenia, ubierania się, kąpania, do wszystkiego” (Saks 1996: 36).

Można powiedzieć tak: przeżywający momenty „funkcjonalnego paraliżu” systemów zachowań [$\Omega_i-\Omega_n$], pacjent Saksy zachowywał się i działał normalnie, gdy śpiewał. Śpiewanie stanowiło istotną część jego egzystencji. Skupienie uwagi [$O_{j_j}-O_{1n}$], skupienie postrzegania [$\Omega_{2j}-\Omega_{2n}$] i działania [$\Omega_{3j}-\Omega_{3n}$] na jednym celu, czyli na śpiewaniu, pomagało mu zjednoczyć normalnie nieskoordynowane, autonomiczne systemy mózgu [$\Phi_j-\Phi_n$], a tym samym podporządkować cały mózg (albo jego części) jednej funkcjonalnej jednostki [Ω].

Z jednej strony kwestie neuroplastyczności mózgu zmuszają do zadawania dodatkowych pytań o badania łączące ze sobą dane neurologiczne z danymi psychologicznymi. Z drugiej strony pokazują one, że refleksja nad empirycznymi i filozoficznymi zagadnieniami dotyczącymi identyfikacji neuronalnych korelatów świadomości jest dopiero w początkowym stadium rozwoju.

W dalszych analizach przyjmujemy, że naukowa mowa o neuronalnych korelatach świadomości odnosi się najpierw do korelacji występujących w: (a) normalnie funkcjonującym mózgu, ale także (b) w normalnym i w zmienionym otoczeniu. (a) Lokalizacja tak rozumianego korelatu jest zależna od ar-

⁴¹ Zdaniem W. T. Thacha, ćwiczenie na jakimś instrumencie muzycznym powoduje, że mózdzek zaczyna kontrolować poszczególne zadania i „[...] recognizes the contexts in which each «piece» of consciously initiated movement occurs. Thus with time and practice, the cerebellum largely controls the process with little or no help from the cerebrum. The cerebrum and the conscious mind are free to do and think about other things” (Thach 1997: 600).

chitektury mózgu. Raczej nie należy oczekiwać, że neuronalny korelat świadomości „koreluje” (współwystępuje albo towarzyszy procesom świadomościowym) ze świadomością w przypadkach nienormalnego funkcjonowania mózgu albo przy zmienionej architekturze mózgu. Tym samym nie możemy wyciągać bezpośrednich wniosków z braku korelacji w przypadkach określonych uszkodzeń mózgu. (b) Metodami neuroobrazowania poszukujemy neuronalnego korelatu, który znamienne i charakterystycznie współwystępuje i współtowarzyszy świadomości we wszystkich normalnie niespotykanych warunkach zewnętrznych, a nie tylko w warunkach normalnie spotykanych.

6.2. Ujęcie mózgu w materialistycznych teoriach świadomości

Przedstawione w rozdziale III redukcjonistyczne i nieredukcjonistyczne teorie umysłu w różnym stopniu uwzględniają szczegółowe zależności pomiędzy strukturami mózgu a świadomością. Pod tym względem teorie te różnią się od współczesnych, neurokognitywnych teorii, pretendujących do bycia naukowymi ujęciami świadomości (por. 6.3.1, 6.3.2, 6.3.3).

Twórcy teorii identyczności typów (por. 3.3) nie zajmowali się szczegółowymi badaniami prowadzonymi w neurologii, a przyjmowana przez nich mentalno-neuronalna identyczność dotyczy raczej cech aniżeli obiektów. W omawianej teorii przyjmuje się, że:

1. Dane stany lub cechy psychiczne oraz stany lub cechy neuronalne są empirycznie identyczne. Wyrazem psychoneuronalnej identyczności jest wypowiedź: „ból to pobudzenie włókien C” albo „pobudzenie włókien C, stąd ból”. Przy czym włókna C są traktowane jako proste transmitery bólu. Tym samym pytanie o korelaty samo się eliminuje. Jest to oczywiście uproszczenie, bo „transmisja” doznań bólowych to tylko jeden istotny, ale nie wystarczający warunek, aby zaistniał ból jako doznanie świadomościowe.

2. Używana w argumentacji terminologia neurologiczna jest jedynie „pustym miejscem” dla pojęć przyszłej neurologii. Identyczności przyjmowane w omawianej teorii można nazwać identycznościami „potencjalnymi”. Teoria identyczności przyjmuje materialistyczno-monistyczną ontologię. Jeżeli jednak stany mentalne są identyczne ze stanami neurologicznymi, to można – przyjmując symetrię relacji identyczności – mózg „zredukować” do umysłu, otrzymując jakąś odmianę idealistycznego monizmu.

Argumentacja w redukcjonistycznych teoriach umysłu a tym samym i świadomości jest natury filozoficznej, chociaż jej uzasadnienie ma być z założenia natury empirycznej. Zasadniczą część programu badawczego w tych teoriach stanowi poszukiwanie ścisłych, obejmujących wiele osobników,

praw psychofizycznych (albo tak zwanych „praw łączących”, „praw pomocowych”), umożliwiających ontologiczną redukcję takich danych, jak „ból”, do danych fizjologicznych: „pobudzenie włókien C”. Wyjaśnienie świadomości (o ile się o niej mówiło) było rozumiane jako stwierdzenie wspomnianej identyczności. Teorie te cechują się znikomą eksperymentalną podstawą. W gruncie rzeczy są one metafizycznymi ujęciami, w których zakłada się, że wszystko co fizyczne i co mentalne jest zdeterminowane.

Podobny stosunek do badań neurofizjologicznych można zaobserwować w funkcjonalizmie (por. 3.5.1). Zakładając tezę o wielorakiej realizacji, jego przedstawiciele twierdzą, że stany mentalne mogą być urzeczywistnione w różnym podłożu materialnym. Tym samym nie zadają oni pytania o neuronalne korelaty świadomości, skoro te mogą być „neuronalne” lub „krzemowe”. Funkcjonalista – teleologiczny i maszynowy – interesuje się funkcjami umysłu, pozostawiając neurologom badania struktur i stanów mózgu. Samej tezie o wielorakiej realizacji brakuje dzisiaj powszechnej akceptacji.

Materializm eliminacyjny odrzuca wyjaśnienie ludzkich zachowań, podawane w ramach „psychologii potocznej”, odwołujące się do nastawień, przekonań, pragnień podmiotu (por. 3.1). Twierdzi, że psychologia potoczna zostanie zastąpiona (wyeliminowana, a nie zredukowana) przez przyszłe nauki o systemie nerwowym, rozumiane jako ostateczne i właściwe wyjaśnienie ludzkiego zachowania. Stany, do których odwołuje się psychologia potoczna, po prostu nie istnieją. Tym samym nie ma z czym korelować obiektów, do których odnosi się neurologia. Równocześnie należy podkreślić, że materializm eliminacyjny wprowadził nowe obliczeniowe metody do nauk o systemie nerwowym rozumianym jako neuronalna sieć aktywności. Według eliminatywistów, współczesne badania mózgu mają na celu wyjaśnić mentalne zjawiska jako wynik procesów obliczeniowych (por. Churchland 1996: 485–486; por. Flohr 1997: 106–112).

Na uwagę, gdy chodzi o neuronalne korelaty świadomości, zasługuje nieredukcjonistyczna teoria identyczności zdarzeń (egzemplarzy, por. 3.4; 3.4.1). Na jej podstawie powstały zarówno współczesne teorie superweniencji tego, co mentalne, nad tym, co fizyczne, a do kręgu analitycznych zainteresowań powróciły teorie emergencji.

Naukowo stwierdzane korelacje nie implikują identyczności świadomych stanów mentalnych i stanów fizycznych, choć mogą – od strony filozoficzno-teoretycznej – implikować odpowiednio rozumianą redukcję (por. 7.6.1). Wszystko zależy od definicji korelatu (relacji bycia korelatem). Równoczesne występowanie obydwu rodzajów stanów jest także, jak wspomnieliśmy, do pogodzenia z epifenomenalizmem lub z paralelizmem. Paralelizm zakłada, że świadomość jest tylko duplikatem aktywności

mózgu⁴². Wskazując jedynie na empirycznie stwierdzalne neuronalne korelaty świadomości nie potrafimy empirycznie odrzucić paralelizmu, gdyż eksperymentalnie nie da się odróżnić identyczności od paralelizmu. Nie oznacza to oczywiście, że epifenomenalizm i paralelizm z jednej strony, a teoria identyczności z drugiej strony, są tak samo możliwe, lecz jedynie, że samo stwierdzenie korelacji nie pozwoli nam przyjąć jednej z tych teorii i odrzucić drugą.

Tym samym powoływanie się na neuronalne korelaty świadomości (na stwierdzenie współwystępowania) nie wystarczy do naturalistycznego lub innego wyjaśnienia świadomości zjawiskowej. Potrzeba dodatkowych filozoficznych i neurologicznych argumentów prowadzących do nowej teorii świadomości. Naturalizm domaga się przynajmniej jakiejś formy superwencji albo emergencji. Sama jednak superwencja, jak się ją dzisiaj rozumie, nie wystarczy do zbudowania poszerzonej teorii świadomości, o której mówi naturalista T. Nagel.

6.3. Wybrane neurokognitywne teorie świadomości

Poniżej zostaną przedstawione trzy współczesne teorie, które zaliczamy do kognitywnych teorii świadomości (por. 6.3.1; 6.3.2; 6.3.3). Ich przedstawiciele łączą badania funkcjonującego mózgu z filozoficzno-psychologiczną refleksją nad wynikami tych badań. Należy wspomnieć, że nie będziemy omawiać jednego z centralnych zagadnień kognitywistyki, a mianowicie badań tak zwanych sieci neuronowych. Sieci neuronowe są rozumiane jako dynamiczne systemy, w których powstała dzięki odpowiednim wejściom jedność oddziałuje hamująco lub wzmacniająco na inne zjawiska występujące w sieci.

6.3.1. Neuropsychologia

Neuropsychologia, w odróżnieniu od psychologii ogólnej, traktuje mózg jako centralny organ umożliwiający powstanie i funkcjonowanie psychiki. Neuropsychologowie badają zależności pomiędzy funkcjami psychicznymi, a funkcjami centralnego systemu nerwowego (anatomicznymi, fizjologicznymi i biochemicznymi)⁴³. Możemy rozróżnić kilka sposobów uprawiania neuropsychologii.

⁴² Zadanie świadomości, według paralelizmu, ogranicza się do przekładu tego, co mechanicznie powstaje w substancji mózgu. Bergsona krytykę tak rozumianego paralelizmu omawia Kulczycki (1996: 134–151)

⁴³ Neuropsychologiczne analizy są także rozumiane jako uzupełnienia epistemologicznych badań nad spostrzeganiem. „One of the most exciting developments arising from recent neuropsychological studies however, is that they are beginning to offer insight into how visual awareness is achieved” (Young i Haan 1993: 59).

Według M. Kinsbourne'a, prowadzone w ramach neuropsychologii analizy świadomości należy umieszczać pomiędzy badaniami nad stymulacjami *software* w ramach sztucznej inteligencji, a neurofizjologicznymi analizami traktującymi mózg jako *hardware* (por. Kinsbourne 1997: 335–337). Ani w nauce o sztucznej inteligencji, ani w neurofizjologii nie badamy jednak bezpośrednio świadomości. Neuropsychologowie stosują pojęcia i metody z zakresu nauk kognitywnych oraz z zakresu nauk o mózgu, aby wyjaśnić odpowiednie neuronalne organizacje i funkcje mózgu. W proponowanym przez Kinsbourne'a rozumieniu neuropsychologii, świadomość nie jest „wytworem” neuronalnej aktywności, lecz aktywność ta *jest świadoma*⁴⁴. Nie potrzeba więc pytać o neuronalne korelaty świadomości, bo w grę wchodzi identyczność stanów (doprecyzowania wymagałoby jedynie, czy jest to identyczność czasowo-przestrzenna, czy też identyczność przyczynowa, oraz czy chodzi o identyczność cech lub obiektów). Takie samo twierdzenie można przypisać Dennettowi. Jego teoria świadomości jest teorią „neuropsychologiczną” w tym sensie, iż zakłada krytykę Kartezjańskiego Teatru.

Dwojakię podejście do neuropsychologii znajdujemy w pracach jednego z jej twórców – A. Łurii⁴⁵. (a) Według niego: „Kwalifikacja zaburzeń wyższych [psychicznych – *J.B.*] czynności nerwowych, umożliwiającą ocenę wartości lokalizacyjnej objawu, jest właśnie jednym z podstawowych zadań nowej dziedziny wiedzy – neuropsychologii”⁴⁶.

Łuria jest twórcą teorii dynamicznej lokalizacji procesów psychicznych (*dynamic localization theory for mental activity*) i koncepcji mózgowych układów funkcjonalnych. W tej ostatniej opisuje funkcjonalne relacje pomię-

⁴⁴ „The human brain does not generate or produce consciousness. It is conscious (or, at least some of it, in some of it possible functional states, is conscious)” (Kinsbourne 1997: 338).

⁴⁵ Podstawowe informacje o badaniach Łurii oraz o jego kontaktach z polskimi uczonymi można znaleźć w: Borkowska i Szepietowska 2000: 14–16.

⁴⁶ Łuria 1967: 86. Łuria zaznacza, że neuropsychologia, ze względu na swą treść „[...] różni się od innych działów nauk neurologicznych tym, że jej przedmiotem nie są ani ukryte mechanizmy molekularnych lub biochemicznych podstaw czynności nerwowych (...), ani też fizjologiczne mechanizmy procesów nerwowych, lecz rola, jaką poszczególne struktury układu nerwowego odgrywają w organizowaniu czynności psychicznych człowieka” (Łuria 1976a: 67). Obserwując pacjentów z brakami neurologicznymi Łuria przyjął trzy stopnie przetwarzania informacji w mózgu: ośrodki pierwszorzędowe dostają informację z receptorów, przetwarzają ją i przesyłają do ośrodków drugorzędowych, dokonujących bardziej rozbudowanych analiz (na podstawie informacji otrzymanych z innych zmysłów), a rezultat swojego działania przekazują do ośrodków trzeciorzędowych, urzeczywistniających zadania asocjacyjne i integracyjne.

dzy anatomicznie odmiennymi obszarami mózgu i interakcje owych funkcjonalnie różnych jednostek z fizycznym i społecznym otoczeniem podmiotu. Do teorii Łurii zbliżają się twórcy modularnej teorii umysłu i mózgu oraz funkcjonalności z kręgu *cognitive science*⁴⁷.

Zachodzące w nas procesy psychiczne, które potocznie wydają się być niepodzielne, są w istocie rozmieszczone w układzie współdziałających okolic półkul mózgowych. W jednych okolicach mózgu są realizowane czynności analizy i syntezy określonych bodźców zewnętrznych i wewnętrznych, w innych dokonują się analizy sygnałów, powstających w toku działania oraz porównywania oczekiwanego wyniku z rezultatem osiągniętym w działaniu⁴⁸. Łuria podkreśla, że mózg człowieka, „[...] jest zadziwiającym, jeszcze przez nas nie poznanym aparatem”⁴⁹. Duże uszkodzenia mózgu mogą spowodować utratę posiadanej wiedzy, zdolności do rozumienia konstrukcji gramatycznych, a mimo to osoba może nadal posiadać zadziwiającą swą jasnością wyobraźnię emocjonalną oraz wrażliwość w postrzeganiu różnorodnych motywów i przeżyć u innych ludzi. Nieuszkodzone części mózgu potrafią przejąć część czynności wykonywanych wcześniej przez części, które uległy uszkodzeniu lub zniszczeniu. Stopień przejmowania czynności jest między innymi zależny od wieku pacjenta.

Również formy życia społecznego powodują, że mózg potrafi się do nich dostosować i funkcjonować w nowy sposób. Dzięki wspomnianym formom tworzą się nowe systemy funkcjonalne (por. Łuria 1967: 30–31). Łuria nie mówi, że „wyższe procesy psychiczne” są wyjaśniająco redukowalne do procesów neurologicznych czy są identyczne z nimi, używa natomiast terminu „opierania się” (por. Łuria 1967: 107), który możemy rozumieć w sensie poszukiwanych przez nas korelacji i związanych z nimi teoriami synchronicznej emergencji. Łuria jest zainteresowany w podaniu precyzyjnych miejsc („ognisk”) w mózgu, odpowiadających za wyższe czynności korowe. W tym sensie Łuria byłby – w zakresie funkcjonowania mózgu – „lokalistą” (por. 6.1.4).

⁴⁷ Ci ostatni – zgodnie ze wskazówką K. Twardowskiego – powinni wyraźnie odróżniać „funkcję” w znaczeniu matematycznym od funkcji jako czynności osoby lub podmiotu (por. Twardowski 1997: 90–91).

⁴⁸ Por. Łuria 1967: 42. Do Łurii modelu mózgu i do N. Whiteheada „filozofii procesu” nawiązuje J. Brown w swojej, opartej na *process monism* neuropsychologii procesu, por. tenże 1998: 235–242. W Polsce neuropsychologią procesu zajmuje się Maria Pąchalska.

⁴⁹ Łuria 1976: 178. Nie jest tak, że uszkodzenie w jakimś konkretnym miejscu mózgu powoduje zaburzenie jednej tylko czynności psychicznej (por. Łuria 1967: 20–23).

(b) Łuria był nie tylko teoretykiem, lecz także lekarzem. Przez długie lata zajmował się pacjentem L. ZasiECKim, u którego zranienie głowy spowodowało, że stał się bezradnym intelektualnie „afatykiem”, kimś egzystującym w nowym „świecie bez pamięci”. ZasiECKi nagle „gubił” prawą połowę swojego ciała, wydawało mu się, że proporcje jego ciała się zmieniły, że głowa stała się olbrzymia, a tułów mały. Równocześnie w całości został zachowany „[...] świat jego przeżyć, jego twórczy entuzjazm, pozostawiło w pełni nienaruszoną osobowość” (Łuria 1976: 175; por. s. 52, 162–163). Łuria nie pisze, jak tę „nienaruszoną osobowość” rozumieć.

Metody stosowane w „obiektywnej” neuropsychologii opierają się na rozkładzie zdarzeń na ich części, a następnie na formułowaniu ogólnie ważnych praw pretendujących do bycia prawami obiektywnymi i umożliwiającymi – na podstawie poszukiwanych korelatów – odpowiednie prognozy naukowe. W ten sposób powstaje, nazwijmy ją, „obiektywna neuroanatomia”, której przedstawicielem może być M. Kinsbourne. Osobowy świat podmiotu zostaje zredukowany do abstrakcyjnych schematów, a jego żywa rzeczywistość zostaje podzielona na elementarne komponenty.

Przedstawiając wewnętrzny świat ZasiECKiego, Łuria stawia go w centrum uwagi. Każda czynność ZasiECKiego jest analizowana jako czynność dla niego sensowna (w danych warunkach). Traktując patologiczny (z naszego punktu widzenia) sposób życia ZasiECKiego jako subiektywnie logiczny i sensowny, Łuria przyjmuje tym samym punkt widzenia wewnętrznego obserwatora i razem ze swoim pacjentem tworzy dla niego perspektywę sensu życia (por. Łuria 1976: 167–168). Równocześnie wskazuje na neurologiczne podłoże cierpienia ZasiECKiego, wspominając „że pozornie niewielkie uszkodzenie mózgu”, w „tylnych okolicach półkul mózgowych” (pomiędzy okolicą czuciową, słuchową i wzrokową), spowodowało te wszystkie zaburzenia (por. Łuria 1976: 34–41).

Nie mamy tutaj do czynienia z systematyczną terapią, przebiegającą według ogólnych reguł, lecz ze swego rodzaju fenomenologicznym „wczuwaniem się” w niepowtarzalną sytuację pacjenta oraz z jednostkowym wyjaśnianiem jego zachowań. Wymaga to atmosfery zaufania pomiędzy pacjentem a lekarzem. Łuria łączy opisy natury obiektywnej z subiektywnymi wypowiedziami pacjenta. Z wielkim zrozumieniem przedstawia wydarzenia z życia ZasiECKiego, które są zarazem naukowym studium konkretnego podmiotu. Ten sposób postępowania wynika z proponowanej przez Łurię neurofizjologii, w której pacjent jest traktowany jako osoba i nie jest zredukowany do „szczególnego” chorobowego przypadku.

W odróżnieniu od „obiektywnej neuroanatomii” (Kinsbourne), eliminującej indywidualną subiektywność oraz w odróżnieniu od „autobiograficznej

neuropsychologii”, rezygnującej ze stawiania twierdzeń ogólnych, staramy się podać obiektywne metody (neuroanatomiczne i filozoficzne) opisywania i wyjaśniania świadomości. Pierwszym krokiem do ich znalezienia będzie określenie neuronalnych korelatów świadomości. Przy definiowaniu tychże korelatów należy: (a) odwołać się do wyników badań konkretnych przypadków pacjentów, (b) zdawać sobie sprawę z pojęciowego rozdziwku pomiędzy holistycznie rozumianą świadomością, a zachodzącymi w mózgu chemicznymi reakcjami i elektropsychologicznymi procesami.

6.3.2. Neurofenomenologia

Termin „neurofenomenologia” (*neurophenomenology*) został wypracowany przez zmarłego niedawno neurologa F. Varełę. Oznacza on monistyczno-naturalistyczne badanie świadomości, łączące Husserlowską fenomenologię z kognitywistyką⁵⁰. Neurofenomenologowie uwzględniają w swoich analizach przeżycia pierwszoosobowe, a prowadzone przez nich badania wychodzą z założenia o nieredukowalnej naturze świadomego przeżycia („świadome przeżycie” jest swoistą odmianą „mentalnej treści”, do której odwołuje się angloamerykańska filozofia umysłu).

Varela zgadza się z J. Searlem, który podkreśla, że ontologia tego, co mentalne, jest nieredukcyjną ontologią pierwszej osoby⁵¹. Broniąc nieredukowalności świadomości Searle nie dochodzi jednak do ostatecznego rozwiązania epistemologicznego problemu dotyczącego badania świadomości. Zdaniem Vareli: „Searle chce, abyśmy zaakceptowali, że «nieredukowalność świadomości jest jedynie trywialną konsekwencją pragmatyki naszych praktyk definicyjnych» (s. 168) i dlatego, chociaż nieredukowalność świadomości jest «jasnym argumentem», nie ma on «głębszych konsekwencji» [...]”⁵².

Varela powołuje się dodatkowo na twierdzenie Searle’a, że sam fakt subiektywności, który próbujemy zaobserwować, uniemożliwia przeprowadzenie tego rodzaju obserwacji. Na pytanie dlaczego tak jest, Searle odpowiada:

„Ponieważ w przypadku świadomej subiektywności nie sposób rozróżnić obserwacji i tego co obserwowane, percepcji i percypowanego przedmiotu. Model

⁵⁰ Por. Varela 2000: 341–355. Heterofenomenolog D. Dennett krytykuje fenomenologię, która „[...] has failed to find a single method that everyone could agree upon” (Dennett 1991: 44). Zwięzłe wprowadzenie do neurofenomenologii można znaleźć w: Gallagher i Varela 2001: 18–22.

⁵¹ „Słowem, nie możemy wytworzyć sobie obrazu subiektywności jako składnika naszego obrazu świata, ponieważ owa subiektywność jest niejako samym obrazowaniem (*picturing*)” (Searle 1999: 139).

⁵² Varela 2000: 342 (Varela cytuje w tym fragmencie z: Searle 1999: 168).

wzrokowy opiera się na założeniu, że istnieje różnica pomiędzy widzianym przedmiotem a widzeniem owego przedmiotu. W wypadku «introspekcji» przeprowadzenie tego rozróżnienia jest jednak niemożliwe. Każdy introspekcyjny wgląd w mój stan świadomy jest owym stanem świadomym” (Searle 1999: 138).

Stany mentalne w teorii Searle’a – jak je rozumie Varela – nie potrafią „zbadać” samych siebie. Tym samym, zdaniem Vareli, Searle pozostawia nas z zawiniętą w pragmatyczne i metodologiczne opakowanie logiczną konkluzją o niepoznawalności stanów mentalnych.

Dodajmy, że Searle chce jedynie powiedzieć (czego Varela nie dostrzeża), iż standardowy model obserwacji nie sprawdza się w wypadku obserwacji świadomej subiektywności (model ten nie stosuje się ani do świadomości innych ludzi, ani do naszej własnej świadomości)⁵³.

Varela zaznacza, że zbudowana przez niego neurofenomenologia nie może być zaliczona do jakiejś istniejącej już szkoły filozoficznej lub neurologicznej. Stanowi ona syntezę Husserlowskiej fenomenologii, widzianej w świetle współczesnej kognitywistyki, skupiającej się na teoretycznym ujęciu ludzkiego przeżywania. Reprezentowana przez E. Husserla fenomenologia jest rodzajem refleksji nad naszą zdolnością bycia świadomym. Jego powiedzenie „z powrotem do rzeczy” jest wyrazem sprzeciwu wobec trzecioosobowej obiektywizacji i oznacza powrót do świata odczuwanego w jego bezpośredniości.

Zdaniem Vareli, podstawową inspiracją dla Husserla była chęć zbudowania prawdziwej nauki zajmującej się przeżyciami. Nauki, która nie tylko zrównoważy nauki szczegółowe, ale dostarczy im koniecznej podstawy do badania każdego rodzaju poznania wynikającego z naszego przeżywania czegoś.

Analiza fenomenologiczna nie jest introspekcjonizmem. W tym sensie Varela zgadza się z D. Denneta krytyką introspekcyjnej teorii W. Wundta i W. Titchenera. Rozwijany przez introspekcjonistów program nie zdał egzaminu. Według Vareli fenomenologiczna redukcja nie oznacza „widzenia wnętrza” (*seeing inside*), lecz tolerancyjne podejście do zawieszenia konkluzji, która dopuszcza do rozwoju nowy aspekt lub wgląd w fenomen: „W konsekwencji ruch ten nie zatrzymuje podstawowej dualności podmiot–przedmiot, lecz prowadzi do obszaru zjawisk, gdzie coraz mniej oczywiste staje się rozróżnienie pomiędzy podmiotem a przedmiotem (co Husserl nazywa «fundamentalną korelacją» [*fundamental correlation*])” (Varela 2000: 346).

⁵³ Argumentacji Searle’a można postawić zarzut, że czym innym jest immanencja refleksji, czym innym zaś – niewydolność poznawcza.

Redukcja fenomenologiczna prowadzi nas ku oczywistości, że świadomość jest nierozdzielnie połączona z tym, co stoi poza nią (używając języka Husserla jest ona „transcendentalna”)⁵⁴. Świadomość nie jest jakimś prywatnym, wewnętrznym zdarzeniem, posiadającym tego samego rodzaju istnienie jak zewnętrzny, pozbawiony świadomości świat. Podobnie dociekania fenomenologiczne nie są prywatne, lecz są nakierowane na drugich poprzez ich stałe intersubiektywne potwierdzanie. Dzięki empirycznej i intuicyjnej oczywistości jesteśmy przekonani, że nasze ludzkie przeżycia (twoje i moje), następują według pewnych określonych fundamentalnych strukturalnych zasad, które podbudowują naturę tego, co jest nam dane jako treści tych przeżyć. Dzięki fenomenologicznej redukcji, świadomość ukazuje się jako podstawa rzucająca światło na to, w jaki sposób drugorzędne idee („obiektywne” i „subiektywne”) mogą się pojawić na pierwszym miejscu.

Proponowane przez Varełę analizy świadomości różnią się od analiz prowadzonych w ramach angloamerykańskiej filozofii analitycznej. Varela nie zajmuje się jakimś prywatnymi fenomenami, lecz obszarem fenomenów, w którym to, co subiektywne, i to, co obiektywne, wylania się w naturalny sposób dzięki zastosowanej metodzie i z kontekstu jej zastosowania. Redukcjonści i funkcjoniści często pomijają ten punkt. Przeżycie jest zdarzeniem osobistym, ale to nie znaczy, że jest ono prywatne w sensie jakiegoś wyizolowanego podmiotu, jako coś nagle dochodzącego do obiektywnie danego świata. Jednym z odkryć fenomenologii było stwierdzenie, że badania struktury ludzkiej świadomości niejako koniecznie powodują przesunięcie w kierunku badania różnych poziomów, na których nasza świadomość jest wczuwająco połączona ze świadomością innych ludzi oraz ze zjawiskowym światem. Tym samym błędne jest przeciwstawianie pierwszo- i trzecioosobowego punktu widzenia.

Neurofenomenologia, zdaniem Vareli, proponuje naturalne rozwiązanie problemu świadomości, pozwalające na przeniknięcie poza „trudny” problem świadomości. Zamiast poszukiwać jakiegoś dodatkowego, teoretycznego lub pojęciowego „składnika” (*extra ingredient*)⁵⁵, umożliwiającego wyjaśnienie, w jaki sposób świadomość wylania się z materii mózgu, Varela pyta o sensowne połączenia pomiędzy dwoma nieredukowalnymi fenomenalnymi

⁵⁴ Husserlowski program redukcji transcendentalnej i związany z nim problem intersubiektywności omawia St. Judycki (1999: 23–50).

⁵⁵ Takiego „ingredient” poszukuje D. Chalmers, gdy pisze: „[...] consciousness arises from a physical basis, even though it is not entailed by that basis” (Chalmers 1996: 125). „To account for conscious experience, we need an extra ingredient in the explanation. This makes for a challenge to those who are serious about the hard problem of consciousness” (Chalmers 1995: 7).

obszarami i znajduje je w tezie o słabej emergencji (por. Stephan 2002: 81). Neurofenomenologia ma tym samym dostarczyć rozwiązania „trudnego” problemu poprzez ukazanie w nowym świetle, tego, co termin „trudny” znaczy.

Varela celowo nie używa terminu „neurofilozofia”, ponieważ w kulturze amerykańskiej filozofię identyfikuje się często z angloamerykańską filozofią umysłu. B. Baars, który w przeciwieństwie do F. Cricka i D. Dennetta, uwzględnia perspektywę pierwszoosobową w badaniu świadomości, zarzuca zbudowanej przez Varełę neurofenomenologii brak związku z neurofizjologicznymi pracami o świadomości oraz niepotrzebne komplikacje w spojrzeniu na perspektywę pierwszoosobową. Odpowiadając Varela stwierdza dwie rzeczy: (1) pomimo iż współcześni filozofowie i naukowcy mówią o perspektywie pierwszoosobowej, to właściwie jej nie używają, bo nadal nie wiadomo, jak operować danymi pierwszoosobowymi, (2) Baars nie podaje żadnych argumentów na uzasadnienie swojego twierdzenia, że nie potrzebujemy poszukiwać nowych metod badania świadomości (por. Varela i Shear 2000a: 308).

Na zakończenie należy odnotować trzy kwestie:

1. Koncepcję świadomości proponowaną przez neurofenomenologię można nazwać „pierwszoosobową w liczbie mnogiej”. Przesuwa ona punkt ciężkości wyjaśniania z pojedynczego podmiotu na intersubiektywne, aktywno-eksternalistyczne relacje, zachodzące pomiędzy świadomymi podmiotami i otaczającym je światem. Indywidualne przeżycia każdego z nas są po części konstituowane przez nasze pojęcie przeżyć innych osób i przez to, jak odbierają one nasze przeżycia. W tym sensie Varela unika dwóch rzeczy: (a) mówienia o „trudnym” problemie świadomości, gdyż ten jest nieeliminowalnym konstruktem pierwszoosobowej świadomości, (b) sporu eksternalizm–internalizm, w zakresie świadomości. Pozostaje oczywiście pytanie: czy potrafimy zbudować psychologię czy fenomenologię przeżyć, które będą nam w ten sam albo w analogiczny sposób „wspólne”, w jaki są nam „wspólne” obiekty fizyki?

2. Varela przyjmuje tezę o słabej emergencji, stawiającej minimalne warunki dla bycia cechą emergencyjną. My przyjmujemy założenie o synchronicznej emergencji.

3. W dalszej części pracy przyjmujemy założenie podobne do zasadniczej tezy Vareli, mówiącej o dwóch – jak na razie, nieredukowalnych do siebie, lecz wzajemnie ze sobą skorelowanych i na siebie wpływających – obszarach poszukiwań: o filozoficzno-psychologicznym i neurofizjologicznym (por. 7.6.1). Od strony filozoficznej, podobnie jak Varela nie poszukujemy psychofizycznych praw identyczności ani nie redukujemy filozoficznej teorii świadomości do teorii neurologicznej, lecz odwołujemy się do emergentnego ro-

dzaju zależności pomiędzy dwiema teoriami świadomości. O ile Varela, w swojej neurofenomenologii, używa zasadniczo metod fenomenologii a następnie neurofizjologii, my skupiamy się na metodach i analizach rozwijanych we współczesnej filozofii analitycznej oraz w neurologii. W odróżnieniu od Vareli szukamy najpierw określenia metodologicznych warunków dla psychoneurologicznych korelacji. O ile Varela i inni neurofenomenologowie skłaniają się raczej ku eksternalizmowi (por. Thompson 2001: 3–5), uważamy, że wstępne analizy korelatów świadomości zjawiskowej można prowadzić w ramach internalizmu. Zgadza się z Varelą, że nadal brak nam jednolitej, naukowej teorii odpowiadającej filozoficznej teorii cech emergentnych, chociaż w wielu zakresach stwierdzamy istnienie takich cech (por. Varela, Thompson i Rosch 1991: 88).

6.3.3. Neurofilozofia

Filozoficzna refleksja nad świadomością musi w swojej metodologii uwzględniać wyniki badań neurologicznych, chociaż w porównaniu z innymi naukami, może tych pojęć używać z większą swobodą. Nie oznacza to uprzywilejowanej pozycji filozofii, lecz jest jej koniecznością. Neurologia, nauki o sztucznej inteligencji, psychologia mają do dyspozycji pojęcia do opisu wielu cech świadomości, które wcześniej w filozofii były używane tylko na sposób nieokreślony (przedteoretyczny). Zajmujący się analizą tych pojęć filozof nie uprawia neurofizjologii.

Tworzone w ten sposób teorie filozoficzne mogą z kolei być używane przy badaniu wewnętrznej spójności teorii neurofizjologicznych oraz przy określaniu ontologicznych i epistemologicznych następstw, wynikających z używania pojęć należących do tychże teorii⁵⁶. Nie chodzi przy tym ani o interdyscyplinarne złożenie wyników uzyskanych w neurofizjologii oraz filozofii, ani o redukcję filozofii do neurologii. Jako metodologiczny model wspomnianego połączenia filozofii i neurologii może nam służyć zakładana przez W. Sellarsa synteza „naukowego” i „potocznego” obrazu świata (por.

⁵⁶ Według A. Klawitera „Filozofa słuca się nie wtedy, kiedy prezentuje poglądy i stanowiska wyrażone w filozofii umysłu, ani nie wtedy, kiedy zadawała się rolę komentatora wyników uzyskanych w naukach kognitywnych, sztucznej inteligencji czy neurobiologii, lecz zostaje on wysłuchany i włączony na równych prawach do grona badaczy kognitywistów wtedy, kiedy potrafi postawić problem, który jest zgodny z zasadniczym korpusem wiedzy w danym obszarze oraz możliwe jest definitywne rozstrzygnięcie tego problemu przy wykorzystaniu narzędzi badawczych powszechnie uznawanych za efektywne w tej dziedzinie” (Klawiter 1999: 136).

wyżej, 1.3.3), oraz jego rozumienie teorii naukowych⁵⁷. Różnica w stosunku do Sellarsa leży w tym, że poszukuje on nowych kategorii dla dokonania ontologicznej syntezy obrazu potocznego i naukowego. My uważamy, że przynajmniej czasowo można zrezygnować z pytania o dualistyczną lub monistyczną ontologię, i skupić się najpierw na metodologicznych zależnościach, wynikających z połączenia neurologicznego i filozoficznego rozumienia korelatów świadomości, a następnie budować odpowiadającą im teorię świadomości.

Zarówno dualiści, jak i moniści pytają, czy „epistemologiczna luka w wyjaśnianiu” pomiędzy poznaniem tego, co fizyczne, i tego, co zjawiskowe, zawiera jakąś „ontologiczną lukę” pomiędzy cechami materialnymi i zjawiskowymi. Wobec wspomnianej luki wyjaśniającej, przedwczesna jest materialistyczna próba identyfikowania cech zjawiskowych z cechami fizyczno-funkcjonalnymi. O ile dualista z zasady odrzuca teorię identityczności mózg–świadomość, nie widzimy dziś sensu dyskusowania o takiej identityczności.

Można w tym miejscu wskazać na pewną analogię z zakresem fizyki. Rozwój tej ostatniej wprowadził nowe perspektywy dla zadawania filozoficznych pytań dotyczących przyrody ożywionej i nieożywionej. Podobnie oczekujemy, że rozwój nauk o mózgu oraz ich szeroka popularyzacja⁵⁸, dostarczą nowego impulsu dla filozoficznego albo, szerzej, kognitywnego spojrzenia na świadomość. Wychodząc od naukowo-doświadczalnego zajmowania się mózgiem (na przykład badania pacjentów z rozdzielonymi półkulami mózgowymi, z uszkodzeniami mózgu lub pacjentów skarżących się na bóle fantomowe), neurologzy zadają na nowo pytania, którymi od dawna interesowała się filozofia świadomości. Jednym z nich jest chociażby pytanie o jedność świadomości. W naukach kognitywnych, w odróżnieniu od fizyki, zadaje się także zasadnicze pytanie: czy sama neurologia posiada dostateczne środki, aby skutecznie badać i wyjaśnić świadomość?

Jedno z pierwszych filozoficzno-neurofizjologicznych opracowań zagadnień związanych z nauką podstawą umysłu pochodzi od filozofa nauki K. Poppera i neurologa J. Ecclesa⁵⁹. Odwołując się do wyników badań neu-

⁵⁷ Sellarsa ujęcie teorii naukowych oraz jego rozumienie wyjaśniania naukowego omawiamy w: Bremer 1997: 30–40.

⁵⁸ Przykładem takiej popularyzacji mogą być organizowane od kilku lat w marcu na uniwersytetach wielu krajów cykle popularnonaukowych wykładów pod tytułem: „Tydzień mózgu”, czy też tworzone w Internecie strony „Brainmap database”.

⁵⁹ Por. Popper i Eccles 1977: passim. Krytykę teorii J. Ecclesa od strony neurologicznej podaje R. Vaas (1999: 132–133), a od strony cybernetyki J. Trąbka (1991: 125). Por. także Bobryk 1996: 33–36.

rofizjologicznych dualista Eccles rozwiązuje Kartezjański problem umysł–ciało poprzez postulowanie niematerialnych „psychonów”⁶⁰. Podobną od strony metodologicznej, chociaż tym razem materialistyczną próbę syntezy filozofii i neurologii podejmuje P. S. Churchland w swojej książce *Neurophilosophy*⁶¹. Zarówno Eccles, jak i Churchland poszukują jednolitej teorii umysł–mózg, wprowadzając filozofię nauki do neurofizjologii, a neurofizjologię do filozofii.

Termin „neurofilozofia” zdomował się w naukach kognitywnych. Rozumie się pod nim inspirowany przez neurofizjologię nowy kierunek poszukiwań rozwiązań klasycznego problemu umysł–ciało⁶². Pomimo że Popper i Eccles wychodzą z pozycji dualistycznych, a autorka książki *Neurophilosophy* wychodzi z pozycji materialistycznych, to proponowane w obydwu pracach program badań i metoda różnią się od dwóch innych programów badań, które teraz krótko omówimy.

(1) Przedstawimy najpierw różnicę pomiędzy wspomnianymi teoriami, a teorią proponowaną przez J. Searle’a w książce *Umysł na nowo odkryty*.

(a) W Searle’a teorii umysłu, stany mentalne są fizycznymi cechami świata, biologicznymi cechami wyższego stopnia, cechami takich systemów, jak ludzki mózg. Searle nie mówi jednak prawie nic o samej strukturze i funkcjonowaniu mózgu. Mózg jest dla niego po prostu tym, co umożliwia empiryczne występowanie umysłu. W proponowanym przez nas poszukiwaniu neuronalnych korelatów świadomości uwzględniamy, w odróżnieniu od Searle’a, wybrane przypadki neurologicznych opisów funkcjonowania poszczególnych części mózgu.

⁶⁰ Por. Bremer 1997: 48–55. Psychony są elementarnymi jednostkami świadomości, skojarzonymi z aktywnością dendronów (wiązek dendrytów). Psychony wpływają na procesy synaptyczne poprzez efektywne podniesienie kwantowo-mechanicznego prawdopodobieństwa egzocytozy, prowadzące do wytworzenia w komórce nerwowej częstszych postsynaptycznych potencjałów wzbudzonych (chodzi o tzw. długotrwałe wzmocnienie synaptyczne LTP, *long-term potentiation*).

⁶¹ Wychodząc od neurologii P. S. Churchland poszukuje jednej wspólnej nauki zajmującej się umysłem i ciałem (por. Churchland 1993: ix–x). Problemy związane z redukcją subiektywnych przeżyć do procesów w mózgu Churchland omawia na stronach 323–327 i stwierdza: „Nevertheless, this current state of science does not entail that no neuroscientific theory will ever reduce psychology” (Churchland 1993: 326).

⁶² Por. Northoff 2001: 24; Walter 1996: 1515. H. Walter odróżnia neurofilozofię ogólną (rozwiązanie problemu umysł–ciało), specjalną (rozwiązania szczegółowych zagadnień związanych z problemem umysł–ciało) oraz neurofilozofię minimalną (szukanie minimalnego konsensusu pozwalającego na wspólne badanie problemów szczegółowych przez filozofię i neurologię, por. s. 1516–1517).

(b) Searle łączy ze sobą „świadomość dostępu” i „świadomość zjawiskową”, co prowadzi do paradoksalnych rezultatów. Tak na przykład w jednym z omawianych przez siebie przypadków „automatycznego” prowadzenia samochodu zakłada on jedynie częściową nieświadomość kierowcy, a w drugim przypadku całkowitą nieświadomość⁶³.

(c) Wymieniając dwanaście fundamentalnych cech świadomości Searle nie podaje, na jakiej podstawie zalicza je do takich właśnie cech. Wychodząc od holistycznego rozumienia świadomości unikamy problemów związanych z tego rodzaju zaliczaniem.

(2) Neurofizjolog G. M. Edelman stwierdza, że udało mu się zbudować teorię mózgu, obejmującą świadome doświadczenia percepcyjne, tworzenie pojęć i funkcjonowanie języka (por. Edelman 1998: 105–106, 342). Neurologicy F. Crick i Ch. Koch są przekonani, że wspomniany już przez nas, a pojawiający się w badaniach percepcji wzrokowej „problem scalania” można rozwiązać poprzez odwołanie się do różnych części mózgu emitujących fale o tej samej częstotliwości (40 Hz)⁶⁴. Problem scalania rozpatrywany w ramach modalności wzrokowej rozpada się oczywiście na kilka problemów (por. Robertson 2003: 95–96): łączenie razem obiektów, celem utworzenia zjednoczonej sceny, łączenie cech w czasie i łączenie razem części, celem uzyskania tego co widzimy jako jeden obiekt (do problemu scalania wrócimy w: 7.5.2). Również A. Damasio jest przekonany o możliwości realizacji podobnie optymistycznego programu w ramach neurofizjologii⁶⁵. Jego zdaniem umysł człowieka nie jest wprawdzie przystosowany do badania samego siebie, ale zarówno mózg, jak i umysł nie są monolitami, lecz mają złożoną strukturę (na przykład mózg podczas różnych form uczenia się posługuje się odrębnymi systemami). Najwyższe z poziomów tej struktury wytwarzają instrumenty pozwalające na obserwację pozostałych poziomów. Damasio zaznacza, że nadal jednak nie rozumiemy w pełni działania zgrupowań neuronów w obrębie określonych struktur mózgowych i dopiero zaczynamy po-

⁶³ Brak rozróżnienia przez Searle’a wspomnianych rodzajów świadomości krytykuje N. Block (por. 1998: 154–155, 158–159).

⁶⁴ Według F. Cricka: „Der Thalamus ist eng an Aufmerksamkeitsmechanismen beteiligt. Eine spezielle Bindung – wo sie für solche Operationen wie die Auszeichnung eines Objektes (besonders bei Figur-Grund-Unterscheidungen) benötigt wird – hat die Form eines koordinierten Feuerns, oft mit Rhythmen im 40-Hertz-Bereich” (Crick 1994: 310). Błędy w łączeniu spostrzeżeń wzrokowych spowodowane przez tzw. *Balint syndrome* (uszkodzenie obydwu płatów ciemieniowych, *parietal lobes*) omawia L. Robertson (2003: 94).

⁶⁵ „Po prostu wierzę, że umysł każdego z nas, choć bezcenny i wyjątkowy, w rzeczywistości jest zjawiskiem biologicznym i pewnego dnia będzie opisywany również za pomocą terminów biologicznych” (Damasio 2000: 79).

znawać interakcje pomiędzy nie sąsiadującymi ze sobą regionami mózgu. Jakiegoś regionu mózgu nie należy traktować jako siedziby świadomości. Mózg jest wzajemnie połączonym systemem posiadającym poznanie (język, pamięć) i procesy sensoryczne (słuch, dotyk, smak) skupione w różnych miejscach⁶⁶. Świadomość, według Damasio, jest podobnie porozdzielana⁶⁷.

Pomimo przytoczonych przykładów, zarówno A. Damasio, jak i G. M. Edelman stwierdzają jedynie istnienie korelacji między mózgiem a umysłem. Ich „neurologiczne” teorie wyjaśniające umysł odwołują się do przyszłego rozwoju neurologii i nie wchodzi chociażby w kwestie holistycznego rozumienia świadomości z subiektywnym poczuciem naszego własnego „ja”. Edelman twierdzi wprawdzie, że zjawiska umysłowe są opisywalne w pojęciach biologicznych, dodaje zarazem, że dane neurologiczne nie mogą wyjaśnić myślenia. „W pewnym praktycznym sensie próba redukcji psychologii do neurofizjologii musi więc zakończyć się niepowodzeniem” (Edelman 1998: 240; 244) To jednak, zdaniem G. M. Edelmana, nie stoi w sprzeczności z jego poglądem, że psychologia poznawcza może odwoływać się tylko do biologicznego wyjaśnienia świadomości. Opisanie świadomości w terminach biologicznych jest dla Edelmana, jak go rozumiemy, warunkiem kompletności nauki.

6.4. Podstawowe założenia poszukiwanej teorii świadomości

„Filozofia bez nauki jest pusta, nauka bez filozofii ślepa” (Spinner 1969: 336).

Tworząc podstawy wielowymiarowej teorii świadomości uwzględniającej jej subiektywność należy zwrócić uwagę na następujące ogólne założenia:

⁶⁶ „My main argument against the idea of an integrative brain site is that there is no single region in the human brain equipped to process, simultaneously, representations from all the sensory modalities active when we experience simultaneously, say, sound, movement, shape and color in perfect temporal and spatial registration” (Damasio 1995: 94–95; por. Damasio 2000: 78).

⁶⁷ Do podobnych wyników dochodzi K. S. Lashley w swoich badaniach nad pamięcią, co do której także trudno ustalić jakieś konkretne, jej odpowiadające miejsce w mózgu. Zdaniem Lashleya: (1) pamięć rozsiana jest po całym mózgu, a nie zlokalizowana w jednej jego części (prawo ekwipotencjalności), (2) utrata pamięci po usunięciu odpowiedniej masy mózgu jest proporcjonalna do ilości usuniętej tkanki (por. Changeux 1997: 276). Dzisiaj przyjmuje się, że pamięć posiada dwa zasadnicze aspekty: pamięć deklaratywna (dotyczy faktów i zdarzeń, a odpowiedzialnymi za nią są hipokamp, struktury międzymózgowia i płatu skroniowego) i niedeklaratywna (odpowiedzialna za zręczności i zwyczaje podmiotu, która jest zlokalizowana m.in. w korze mózgu, prążkowiu i mózdzku, por. Eichenbaum 1997: 550–570; Zola 2003: 504–506). W tym sensie teza o ekwipotencjalności okazała się błędna.

(1) Jeżeli nasze ludzkie (rodzajowe) struktury neurobiologiczne (schematy organizacji mózgu, budowa i funkcjonowanie neuronów) są wystarczająco podobne do struktur osobników innych rodzajów, wtedy tworzymy z nimi jeden typ osobników. Tym samym, jeżeli nasze struktury posiadają jakąś cechę zjawiskową, to posiadają ją także osobniki o podobnej strukturze⁶⁸.

(2) Zgodnie z tezami realizmu poznawczego i ontologicznego, spotykane przedmioty istnieją niezależnie od nas. Pozostaje sprawą otwartą, czy świadomość jest czymś materialnym, jak tego chcą materialyści. Odwołujący się do współczesnego rozwoju nauki mniemają znajdują wprawdzie więcej zwolenników, nie potrafią oni jednak opisać ani wyjaśnić, z jakiego rodzaju „materiału” składa się świadomość. W odróżnieniu od wspomnianych monistów można powiedzieć, że stany mentalne są (na przykład) powiązane z polami elektromagnetycznymi w mózgu, chociaż nie można powiedzieć, że stany mentalne są tymi stanami⁶⁹. Również nieredukcyjni materialyści mogą się zgodzić z tezą, że pola elektromagnetyczne są jedynie zewnętrzną cechą stanów mentalnych. Także serce wytwarza pole magnetyczne, chociaż nie twierdzimy, że myślimy sercem.

(3) Z jednej strony, stwierdzalne za pomocą neuroobrazowania korelacje nie implikują ani przyczynowości, ani identyczności, z drugiej – mogą one być interpretowane jako ich oznaki. Tym samym określenie neuronalnych korelatów świadomości nie jest wystarczające ani do wykazania, że nasze świadome przeżycie jest przyczynowo spowodowane przez zdarzenia neuronalne, ani, że jest z nimi identyczne. Tak rozumiane korelacje są do pogodzenia z różnymi ontologiami. Może być tak, że – w sensie „wprzód ustanowionej harmonii” G. Leibniza – istnieje jakaś wspólna przyczyna stanów mentalnych i neuronalnych, które same są od siebie niezależne (por. 1.4, 3.1). Może się również okazać, że obydwa stany są aspektami jednych i tych samych mikroprocesów, leżących u ich podstaw (por. 5.5.1), czy też może się okazać, że korelacje te są niewyjaśnialne. Sam mózg i jego funkcjonowanie niekoniecznie implikuje przyjmowanie określonych założeń ontologicznych.

⁶⁸ Założenie to jest analogiczne do „nomologiczno-behawioralnego” założenia przyjętego przez J. Konorskiego: „[...] jeśli u człowieka psychiczne doznanie percepcji jakiegось wzorca bodźcowego przejawia się w określonym akcie behawioralnym lub w określonych potencjalach bioelektrycznych mózgu i jeżeli u zwierzęcia (np. u małpy lub kota) wzorec ten wywołuje dokładnie te same reakcje, mamy wówczas prawo twierdzić, że zwierzę to doznaje percepcji przedmiotu w ten sam mniej więcej sposób, co my” (Konorski 1969: 9).

⁶⁹ Wprowadzone przez nas założenia różnią się od tego, co na temat emergencyjnego powiązania stanów mentalnych i fizycznych pisze S. Hameroff (por. Gabora 2001: 122).

Nieokreśloność tego rodzaju nie oznacza przyjęcia opcji funkcjonalistycznej z jej neutralnością co do ontologicznego statusu stanów mentalnych. Funkcjonalizm przyjmuje zasadniczą i ostateczną neutralność (rozumianą jako charakterystyka stanów mentalnych za pomocą relacji funkcjonalnych). Wspomniana powyżej ontologiczna nieokreśloność tego, co mentalne, jest przejściowa, przyjęta ze względów metodologicznych.

(4) Wracając do filozoficznej tezy o zombi należy zauważyć: nie znamy takich kognitywnych działań, dla których nie można by sobie wyobrazić neuronalnych sytemów, funkcjonujących zupełnie bez holistycznie rozumianej świadomości. Niemniej jednak świadomość posiada subiektywną rzeczywistość, która jest skorelowana z interakcjami systemu nerwowego. Celem znalezienia adekwatnej teorii świadomości zjawiskowej, a zwłaszcza jej subiektywności, należy przyjąć istnienie jakiejś formy funkcjonalnego przetwarzania zjawisk mentalnych. Przyjmujemy, że tego rodzaju funkcjonalna forma jest konieczna dla powstania świadomości i jest związana z odpowiednimi procesami neurofizjologicznymi. Obecny opis zjawisk neurobiologicznych jest niepełny. Brakuje nam licznych szczegółów dotyczących funkcjonowania pojedynczych neuronów oraz obwodów neuronalnych. Nie umiemy stwierdzić istnienia jednoznacznych związków pomiędzy poszukiwaną formą przetwarzania danych mentalnych a budową i czynnościami mózgu. Świadomy umysł i mózg nie są ani całkowicie zintegrowane, ani całkowicie rozdzielone. Trudno także powiedzieć, jakie struktury mózgu są konieczne i wystarczające, aby można mówić o subiektywności naszej świadomości (por. 7.3.1).

(5) „Trudny” problem świadomości, jak wspomnieliśmy, można wyrazić pytaniem: czy współczesna nauka (neurologia) dysponuje odpowiednimi środkami (eksperymentalnymi i pojęciowymi), aby skutecznie wyjaśnić świadomość? Dla wielu neurofizjologów lub filozofów (na przykład dla Dennetta czy Cricka) rozróżnienie na „trudne” i „łatwe” problemy świadomości nie ma sensu. Kolejne rozwiązania „łatwych” problemów sumują się, dając w rezultacie rozwiązanie „trudnego” problemu. W swoim Modelu Wielu Szkieł Dennett nie podaje wyjaśnienia holistycznie rozumianej świadomości, lecz analizuje pojedyncze zjawiska, które normalnie są uważane za świadome (na przykład skierowanie przez podmiot uwagi na jakiś przedmiot). Proponowana przez nas metoda jest w tym sensie zgodna z Dennettowską, że omawiamy kilka neurologicznie ujmowalnych problemów świadomości (w języku D. Chalmersa są to „łatwe” problemy), których rozwiązania można szukać za pomocą współcześnie dostępnych metod naukowych. Pokazujemy jednak, gdzie leży granica wyjaśnień neurologicznych. W przeciwieństwie do D. Dennetta, a zgodnie z T. Naglem, W. S. Robinsonem i T. Rockwellem, nie uważamy tym samym,

że wspomniane rozwiązania prowadzą do eliminacji „trudnego” problemu (przynajmniej na dzisiejszym etapie poszukiwań).

(6) Neurofilozofowie, neuropsychologowie i neurofenomenologowie odwołują się do pojęć pochodzących z neurologii i używają ich do opisywania i wyjaśniania tradycyjnych problemów filozoficznych. Kognitywne badanie świadomości, jak je rozumiemy w tej pracy, stoi między innymi przed dwoma zadaniami:

(a) Uwzględniając wyniki badań neurologicznych przeanalizować metodologiczne postępowanie neurologów oraz to, co rozumieją oni ewentualnie pod „wyjaśnianiem świadomości”, jak i to, jakiego typu wyjaśnienie faktycznie osiągają⁷⁰.

(b) Badanie pojęć używanych w neurologii oraz ich wpływu na zmiany potocznego i filozoficznego znaczenia pojęcia świadomości i jej subiektywności.

Należy sobie przy tym zdawać sprawę z każdorazowych granic metodologicznych teorii neurofizjologicznych i filozoficznych. W modelach świadomości proponowanych przez biologa F. Cricka, przez neurologów A. Damasio, J. Ecclesa, G. M. Edelmana oraz przez filozofów: D. Dennetta i J. Searle’a brakuje, między innymi, jasnego rozdziału pomiędzy problematyką filozoficzną a neurofizjologiczną⁷¹. Podobne „syntezy” filozofii lub psychologii z neurologią widzimy w neurofilozofii, neurofenomenologii i neuropsychologii. Zdawanie sobie sprawy z dzisiejszych granic badawczych poszczególnych nauk jest warunkiem uzyskania przejrzystości w poszukiwaniu kognitywistycznej, filozoficzno-neurologicznej teorii świadomości. Filozofia jest przy tym traktowana na równi z innymi naukami⁷².

Można oczekiwać, że nauki szczegółowe będą odgrywały główną rolę w badaniach nad subiektywnością świadomości. A co z filozofią? Holistycz-

⁷⁰ Podobne pytania można odnieść do teorii naukowych w ogóle: czy teorie naukowe faktycznie coś nam mówią o świecie, czy też, rozumiane instrumentalistycznie, dostarczają nam jedynie środków do stawiania opowiednich prognoz. Według W. Sellarsa, wtedy możemy zaakceptować jakąś teorię, gdy umożliwia nam ona myślenie o świecie i nie jest jedynie zespołem reguł wnioskowania. Teorie wyjaśniają nie tylko, dlaczego indukcyjna interpretacja wyników obserwacji zgadza się z obserwacjami, ale także dlaczego zachowania niektórych obiektów nie przebiegają zgodnie z prawami indukcji (por. Sellars 1967: 360; 1963a: 121–122).

⁷¹ G. M. Edelman poszukuje teorii świadomości, która „[...] jednocześnie będzie integralną częścią teorii działania mózgu” (Edelman 1998: 102).

⁷² Zdaniem W. V. O. Quine’a: „Filozofii zaś, rozumianej jako próba jaśniejszego spojrzenia na rzeczy, nie da się odróżnić od dobrej i złej nauki pod istotnymi względami celu i metody” (Quine 1999: 17; por. s. 267).

nie rozumiana świadomość z jej subiektywnością jest problemem filozoficznym i naukowym. Liczni filozofowie zafascynowani rozwojem nauki dążą do zastąpienia argumentacji filozoficznej (tzn. dotyczącej głównie kwestii metafizycznych) argumentacją naukową, chcąc w ten sposób przewyciężyć zastój panujący (ich zdaniem) w filozofii świadomości. Taka postawa jest oznaką scjentyzmu. Kognitywistycznie rozumiane badania świadomości ani nie hamują postępu badań nad subiektywnością, ani nie prowadzą się do argumentacji naukowej.

Z jednej strony, poszukiwanej przez nas teorii świadomości odpowiada (pod względem metodologicznym) program rozwinięty przez D. Dennetta, który, w odróżnieniu od teorii proponowanej przykładowo przez J. Searle'a, odwołuje się często do wyników badań neurofizjologicznych. Jego heterofenomenologiczny program badawczy jest jednak zdeterminowany przez intrumentalne rozumienie teorii naukowych oraz przez heterofenomenologicznie rozumianą rezygnację z uwzględnienia pierwszoosobowych doświadczeń dotyczących subiektywności świadomości.

Z drugiej strony można się zgodzić z Searlem, że cechy subiektywne nie są procesami przetwarzania informacji, i z Dennettem, że mózg jest urządzeniem przetwarzającym informacje. Wynika stąd twierdzenie podobne do nieredukcyjnej tezy proponowanej przez D. Chalmersa: świadomość musi być jakimś dodatkowym zjawiskiem w stosunku do mechanizmów przetwarzających informacje (tzn. że jakaś jej forma musi należeć do podstawowego wyposażenia świata). Dlatego, jak wspomniano, Chalmers pyta o odpowiedni „extra ingredient” potrzebny przy wyjaśnianiu świadomości. W poszukiwaniu owego „ingredientu” możemy się odwoływać do przyszłego rozwoju neurologii, do niealgorytmicznych procesów zachodzących w mózgu albo do osiągnięć mechaniki kwantowej. Naszym zdaniem, określenie metodologii poszukiwania neuronalnych korelatów świadomości jest dziś drogą do precyzacji tego, co rozumiemy pod owym „ingredientem”.

6.5. Podsumowanie

Należy odróżnić uwarunkowania logiczne od uwarunkowań empirycznych (por. 3.4.1). O ile pytania dotyczące wyjaśnienia świadomości są pytaniami natury teoretyczno-empirycznej, neurofizjologia będzie mogła kiedyś na nie odpowiedzieć. Poszukiwanie takich odpowiedzi jest jej celem. Uwarunkowania logiczne odnoszą się między innymi do możliwych światów oraz do istot żywych w tych światach i są niezależne od aktualnego świata. Do tego ostatniego odnoszą się uwarunkowania empiryczne wraz z ich biologiczno-neurologicznymi prawami. Uwarunkowania logiczne zawierają w swym

zakresie uwarunkowania empiryczne, ale nie odwrotnie. Z naukowo-empirycznej podstawy nie można bezwarunkowo wnioskować o możliwościach logicznych⁷³. Teorie filozoficzne uzasadnione czysto logicznie nie są nieste-ty weryfikowalne za pomocą badań neurologicznych. Należy zatem badać empiryczną zrozumiałość teorii filozoficznych, odwołując się do danych pochodzących z neurologii, jak również analizować neurofizjologiczne wyniki w świetle ich epistemologicznych i ontologicznych implikacji.

Poszukiwana teoria świadomości musi zostać przeanalizowana w neurofizjologicznym kontekście i to w możliwie licznych szczegółach i pojęciowych precyzacjach. Dlatego ważne jest podanie określonego (minimalnego) neuronalnego systemu $[\Phi_j - \Phi_n]$, dla którego istnieje jakaś bezpośrednia korelacja pomiędzy jego stanami a stanami świadomości zjawiskowej $[\Psi_i - \Psi_m]$, oraz jakieś odwzorowanie jego „neuronalnych treści” w treściach subiektywnej świadomości.

Należy pamiętać, że leżące u podstaw takich korelacji ontologiczne i epistemiczne implikacje są specyficzne dla tych stanów i na razie nie można mówić o ich uogólnieniu. Stąd rodzą się chociażby pytania: (a) na ile z uzyskanych w ten sposób eksperymentalno-empirycznych wyników da się wyprowadzić redukcyjne teorie filozoficzne? (b) czy konieczne jest założenie, że różne stany mentalne posiadają te same neurologiczne podstawy?⁷⁴

Patrząc z filozoficznego punktu widzenia zauważamy chociażby konieczność oddzielenia neurologicznych rozważań dotyczących cech odczuwania przez podmiot swojego ciała (propriocepcja), od cech postrzegania przedmiotów zewnętrznych. Propriocepcja w neurofizjologii odnosi się do specyficznych „modalności” czuciowych, zawierających informacje o napięciu mięśni (część czucia proprioceptywnego jest nieuświadomiona) i o ułożeniu poszczególnych części ciała względem siebie. Co nauki o mózgu mówią nam o neuronowych strukturach zaangażowanych przy spostrzeganiu własnej jaźni z jednej strony, co zaś o zaangażowaniu tychże struktur przy

⁷³ Chalmers wskazuje na tak zwaną poznawczą asymetrię, „[...] there is an epistemic asymmetry in our knowledge of consciousness that is not present in our knowledge of other phenomena. [...] The epistemic asymmetry [...] makes it clear that consciousness cannot logically supervene. If it were logically supervenient, there would be no such epistemic asymmetry; a logically supervenient property can be detected straightforwardly on the basis of external evidence, and there is no special role for the first-person case” (Chalmers 1996: 102).

⁷⁴ Nawiązując do teorii neurofilozoficznych G. Northoff wskazuje na potrzebę: (i) uwzględnienia „różnicowanej epistemologii” i „różnicowanej ontologii” dla różnych stanów mentalnych, (ii) rezygnacji z założeń ontologicznych (*ontologische Indifferenz*) (por. Northoff 2001: 25–28).

spostrzeganiu innych przedmiotów? Czy są to te same czy inne struktury? Czy od strony neuronalnej da się przeprowadzić odróżnienia pomiędzy spostrzeganiem siebie, a spostrzeganiem innych przedmiotów? Tym samym teoria świadomości, aby być zrozumiała empirycznie, musi zmodyfikować swoje założenia co do spostrzegania przez podmiot siebie samego i innych przedmiotów. Ponadto świadomość podmiotu nie jest ograniczona do jego głowy, ale poprzez ciało obejmuje świat na zewnątrz fizjologicznych granic podmiotu.

Proponowana w następnym rozdziale metodologia poszukiwania korelatów świadomości odwołuje się zarówno do eksperymentów myślowych, jak i do wyników uzyskanych w naukach szczegółowych (te ostatnie odgrywają zasadniczą rolę w poszukiwaniu wspomnianych korelatów i w budowaniu teorii świadomości). Eksperymenty myślowe uwzględniają wyniki badań eksperymentalnych rozumianych jako logicznie uzasadnione modyfikacje warunków nałożonych przez naturę badanych zjawisk. Taki punkt wyjścia pozwala nam:

(a) na wyciągnięcie wniosków dotyczących istnienia zmiennych koniecznych albo wystarczających pod względem empirycznym albo logicznym,

(b) na porównywanie teoretycznych (postulowanych) i obserwowalnych stanów rzeczy, co z kolei jest koniecznym założeniem dla dalszego postępowania, zarówno naukowo-empirycznego, jak i pojęciowo-analitycznego.

ROZDZIAŁ SIÓDMY

NEURONALNE KORELATY ŚWIADOMOŚCI

7. Szkic programu badawczego

„Seltsamer Zufall, daß alle Menschen deren Schädel man geöffnet hat, ein Gehirn hatten” (Wittgenstein 1950, 1 września 1950).

Proponowany przez Nagla program poszukiwań poszerzonej teorii świadomości (por. 5.6.1, 5.6.2) opiera się na *apriorycznym* założeniu o koniecznym rodzaju powiązań pomiędzy fenomenologią, a fizjologią przeżycia. Od strony eksperymentalno-empirycznej programu tego nie da się dziś zrealizować. Rezygnując z tego rodzaju *apriorycznego* założenia, przyjmujemy następujące etapy tworzenia nowej teorii świadomości, którą nazywamy „prototeorią” świadomości¹:

VII.1 [A] określenie metodologii poszukiwań neuronalnych korelatów świadomości.

[B] znalezienie neurobiologicznych systemów lub stanów powiązanych ze świadomością.

[C] szkic prototeorii świadomości.

¹ Nagel powiedziałby, że [A]–[B]–[C] stanowią wstępne kroki do zbudowania poszukiwanej przez niego poszerzonej teorii świadomości. Następnymi krokami, naszym zdaniem, powinny być: (a) naukowa weryfikacja, czy tak określone korelacje są (na przykład) relacjami przyczynowymi, a jeżeli tak, to w jakim sensie i (b) zbudowanie poszerzonej, naukowej teorii świadomości, opartej na koniecznych i wystarczających relacjach: mózg–świadomość.

Kroki podobne do [A]–[B]–[C] były i są obecne przy tworzeniu teorii naukowych, chociażby w fizyce. Wyjaśnianie zjawisk poprzez odwołanie się do praw fizyki tym się różni od wyjaśniania zjawisk mentalnych, że przy tych ostatnich zasadniczą kwestią staje się zdefiniowanie rodzaju praw łączących stany psychiczne z fizycznymi. Pytanie dotyczy istnienia ogólnych praw psychofizycznych. Na podstawie dzisiejszych badań neurologicznych można założyć, że jeżeli prawa te istnieją, to są one prawami empirycznymi. Trzeba zarazem pamiętać o słabych podstawach, na których jest oparte mówienie o analogiach zachodzących pomiędzy empirycznymi prawami w fizyce, a prawami psychofizycznymi, oraz pomiędzy wyjaśnieniami w fizyce, „wyjaśnieniem” świadomości za pomocą proponowanej prototeorii świadomości (por. 7.6). Kroki [A]–[B]–[C] zostaną omówione w dalszych punktach: [A] (7.1.2, 7.2), [B] (7.3.1, 7.3.2), [C] (7.4, 7.6.1).

Przy poszukiwaniu neuronalnych korelatów świadomości [B] przyjmujemy (przeciwnie do proponowanej przez D. Dennetta trzecioosobowej heterofenomenologii), że pierwszoosobowa (homofenomenologiczna) perspektywa poznawcza nie jest (przynajmniej na razie) redukowalna do perspektywy trzecioosobowej. Perspektywa pierwszo- i trzecioosobowa są przez nas traktowane jako *konieczne* i *wystarczające* kryteria do mówienia chociażby o podmiotowej tożsamości. Przykładem równouprawnionego traktowania tych perspektyw są Naglowskie analizy subiektywności świadomości.

Nie przyjmujemy Nagla panpsychicznych hipotez o psychofizycznym „trzecim rodzaju zdarzeń”, ani jego swoistej metafizyki „obietującego materializmu”. Przyjmujemy realistycznie rozumianą, ludzką płaszczyznę poznania. Analizujemy te wyniki, które uzyskujemy za pomocą współczesnych metod badawczych stosowanych w neurologii. Nie przechodzimy na płaszczyznę „wszechwiedzącej istoty”, która być może widzi wszystkie konieczne i wystarczające związki, zachodzące pomiędzy świadomością a mózgiem.

7.1. Definicje neuronalnych korelatów świadomości

Poszukiwanie neuronalnych korelatów świadomości należy dzisiaj do zagadnień nauki o świadomości (por. Metzinger 2000: *passim*; Durstewitz i Windmann 1998; Block 1998; Grush 1998: 175–190; Revonsuo 1998: 208–213; Siegal i Varley 2002: 465–470) i jest ono obarczone zarówno problemami natury empirycznej (wynikającymi chociażby ze złożoności i wzajemnego uwarunkowania pomiędzy stanami przeżyciowymi oraz stanami

nomologiczno-behawioralnymi, a stanami neurologicznymi), jak i problemami natury pojęciowej. Neuroradiologicznie stwierdzalne korelacje pomiędzy stanami mentalnymi a fizycznymi wskazują na różnorodność powiązań obecnych w mózgu. Przykłady tych powiązań widzimy na schemacie podanym przez A. K. Engla (schemat zamieszczono na końcu książki). Engel wyróżnia osiem stanów fizjologicznych $[\Phi_1]$ – $[\Phi_8]$ oddziałujących zarówno na stany behawioralne $[\Omega_1]$ – $[\Omega_8]$, jak i na stany przeżyciowe (zjawiskowe) $[\Psi_1]$ – $[\Psi_8]$ ².

Należy zauważyć, że eksperymentalnie stwierdzalna korelacja pomiędzy (przykładowo) uszkodzeniem ręki a stanami neuronów mózgu podmiotu różni się od poszukiwanej korelacji pomiędzy świadomością zjawiskową a stanami jego neuronów.

Przyjmując koneksjonistyczną naturę mózgu oraz holistyczną naturę świadomości i używając wprowadzonych wcześniej oznaczeń podajemy dwie ogólne, intuicyjne definicje neuronalnych korelatów świadomości. Pierwsza z nich {I} odnosi się do holistycznych „świadomych stanów” (bycie świadomym jako przeciwne do bycia w stanie śpiączki albo w bezmarznieowym śnie), druga {II} odnosi się do treści specyficznych przeżyć.

VII.2 {I} Neuronalny korelat świadomości to najmniejszy system neuronów, którego aktywność w koneksjonistycznie rozumianym mózgu jest wystarczająca dla jej pojawienia się.

Stany świadomości są związane z substratem materialno-strukturalnym komórek nerwowych, a także z substratem dynamiczno-strukturalnym, który jest bardzo złożony (przepływająca krew, natlenienie, poziom glukozy itp.).

² Engel 2000: 2–5. Przykładowo neurologicznie rozumiana selekcja $[\Phi 3]$ (skierowanie uwagi, w którą są zaangażowane płaty czołowe (*lobi frontales*) i płaty ciemieniowe (*lobi parietales*), jądra poduszki (*nuclei pulvinares*) i wzgórkę górne blaszki czworaczej (*colliculi superiores*) wpływa z jednej strony na zjawiskową przejrzystość $[\Psi 3]$ i zjawiskową obecność $[\Psi 4]$, a z drugiej strony na odpowiedniość behawioralnych działań, ich skierowanie na cel $[\Omega 4]$ oraz na kontrolę działań $[\Omega 5]$. Neurologicznie rozumiana sensoryczna separacja $[\Phi 2]$ (w znaczeniu różnych miejsc odpowiedzialnych za różne modalności czuciowe, w którą zaangażowane są określone obszary kory mózgu, sensoryczne i specyficzne jądra wzgórze) ma wpływ po stronie zjawiskowej na przeżywaną jakość $[\Psi 1]$, koherencję $[\Psi 2]$ i przejrzystość $[\Psi 3]$, a po stronie behawioralnej – na zdolność do reagowania $[\Omega 2]$ i na odpowiedniość działań $[\Omega 4]$. Już te dwa przykłady stanów pokazują, że oddziaływania sensorycznych i selekcyjnych stanów neurofizjologicznych konwergują w niektórych stanach zjawiskowych i behawioralnych (por. Engel i Fries 1999: 129, 142–143; por. Engel i Singer 2001: 16–17).

W tym sensie {I} nie jest ograniczona do chemiczno-materialnej budowy mózgu, lecz do struktur żywego, funkcjonującego mózgu. Neuronalny korelat świadomości w sensie {I} to określony (minimalny) system $[\Phi_j-\Phi_n]_0$ (tzn. neuronalne stany $[\Phi_j-\Phi_n]_0$ są skorelowane ze stanami świadomości zjawiskowej $[\Psi_i-\Psi_m]$). Na razie nie wchodzimy w kwestie korelacji mózgu z behawioralnymi stanami $[\Omega_k-\Omega_l]$. Stan systemu $[\Phi_j-\Phi_n]_0$ jest wystarczający, aby przy normalnie funkcjonującym mózgu, pojawił się związany z nim stan świadomości zjawiskowej. Jakiś neuronalny korelat, powiedzmy konkretnego stanu spostrzegania wzrokowego, jest wówczas określony poprzez swój udział w holistycznie rozumianym świadomym systemie i przez swój udział w koneksjonistycznie rozumianym mózgu.

Natomiast neuronalny korelat świadomości w sensie „świadomej reprezentacji x ” (posiadanie specyficznego przeżycia), można rozumieć jako:

VII.3 {II} Neuronalny korelat świadomości (w przypadku treści świadomości) jest określonym neuronalnym systemem $[\Phi_j-\Phi_n]$, takim że neurochemiczne przedstawienie treści w $[\Phi_j-\Phi_n]$ jest wystarczające (przy normalnie funkcjonującym mózgu), dla przedstawienia tej treści w świadomości zjawiskowej $[\Psi_i]$.

Z definicji {I} i {II} neuronalnych korelatów świadomości $[\Phi_j-\Phi_n]$ nie wynika, że korelat taki jest systemem skorelowanym *jedynie* ze świadomością zjawiskową ani że neuronalny korelat świadomości jest *podstawowym* systemem mózgu odpowiedzialnym za pojawienie się świadomości w sensie {I}, ani że *jedynie* badanie określonego neuronalnego korelatu świadomości dostarczy nam wyjaśnienia danego stanu świadomości zjawiskowej. Powyższe definicje nie gwarantują, że zidentyfikowanie neuronalnych korelatów świadomości wystarczy także do wyjaśnienia rodzaju procesów leżących u podstaw świadomości. W {I} oraz {II} przyjęto metafizyczne założenie: istnieje jakiś minimalny neuronalny system $[\Phi_j-\Phi_n]$, skorelowany ze stanem świadomości zjawiskowej $[\Psi_i]$ i wystarczający do jego pojawienia się.

Jeśli zamiast o korelacjach zaczniemy mówić o identyczności: stany świadomości = stany mózgu, to mniej jasny stanie się sposób poszukiwania, za pomocą dzisiejszych technik, takich stanów mózgu.

W antyredukcyjnym argumencie z nietoperzem T. Nagel twierdzi, że świadome przeżywanie posiada subiektywne elementy, i dlatego nie podlega obiektywnemu (neurofizjologicznemu) wyjaśnianiu go. Nagel poszukuje koniecznych zależności pomiędzy świadomością a mózgiem. Jak jednak mówić o takich zależnościach bez uprzedniego zbudowania zapowiadanej przez Nagla poszerzonej teorii stanów mentalnych? Częściowej odpowie-

dzi na to pytanie udziela K. Akins, opisując relacje pomiędzy współczesną teorią tego, co mentalne, a poszerzoną teorią świadomości. Akins powołuje się na wyniki badań fizjologii nietoperza i zaznacza, że wiele niejasnych kwestii dotyczących subiektywnych stanów nietoperza zależy od pytań, które na razie pozostają bez odpowiedzi ze strony neurologii³. Rozwiązanie tych kwestii Akins widzi w znalezieniu neuronalnych korelatów świadomości.

7.1.1. Uszczegółowienie podanych definicji

Definicje {I} i {II} różnią się od definicji korelatów świadomości podanych przez D. Chalmersa⁴. Chalmers wymienia najpierw obszerną listę obiektów i procesów traktowanych przez neurologów jako korelaty świadomości⁵. Następnie powołuje się na badania neurologiczne i przeprowadza analizy pojęciowe dotyczące tychże korelatów. Skupiając się na korelatach treści zjawiskowej (*phenomenal content*) pyta o połączenie treści przedstawianych w systemie nerwowym z treściami przedstawianymi w systemie wizualnym. Szerzej omawia tzw. *minimalny system neuronalny*, który traktuje jako wystarczający dla wywołania odpowiedniego stanu świadomości. Na

³ Nie wchodząc w etiologię nietoperza K. Akins przeprowadza szczegółowe analizy jego „fenomenologii” twierdząc: „Certainly we cannot confidently declare that science must fail to unearth «that something», for we have no clear idea to what this amounts; nor can one say what the scientific approach will necessarily leave out, if it must leave out anything at all. This gives us, I think, good reason to continue on with our empirical investigations of mental representation – to look towards the disciplines of neurophysiology, psychology and artificial intelligence – without undue pessimism about the relevance of their experimental results” (Akins 1993: 270).

⁴ „An NCC is a minimal neural system N such that there is a mapping of states of N to states of consciousness, where a given state of N is sufficient, under conditions C, for the corresponding state of consciousness”. „An NCC (for content) is a minimal neural representational system N such that representation of a content in N is sufficient, under conditions C, for representation of that content in consciousness” (Chalmers 2000: 31; por. Chalmers 1996: 238–242). W pierwszym z wymienionych opracowań Chalmers skupia się zasadniczo na pytaniu: „What does it mean to be a neural correlate of consciousness?”. Zwięźle przedstawił Chalmersa rozumienia korelatów świadomości podają A. Atkinson i M. Thomas (2000: 372–374).

⁵ „A few of these include: 40-hertz oscillations in the cerebral cortex (Crick and Koch 1990), intralaminar nuclei in the thalamus (Bogen 1995), reentrant loops in thalamocortical systems (Edelman 1989), 40-hertz rhythmic activity in thalamocortical systems (Llinas et al. 1994), extended reticular-thalamic activation system (Newman and Baars 1993), [...]” (Chalmers 2000: 17). Niektóre z podanych właśnie przykładów omawia V. G. Hardcastle (2000: 259–260).

koniec Chalmers podaje szereg metodologicznych zasad, istotnych dla filozoficznego spojrzenia na świadomość.

1. Definicje {I} i {II}, w odróżnieniu od definicji podanych przez Chalmersa, są ograniczone przez deinterioryzacyjno-internalistyczne rozumienie stanów świadomości i odnoszą się do subiektywnej świadomości. Są one definicjami „dynamicznymi” w tym sensie, że podczas empiryczno-eksperymentalnego poszukiwania odpowiednich systemów neuronów, da się w nich uwzględnić chociażby strukturalny wpływ ukrwienia i dotlenienia mózgu na stany świadomości. Definicje te nie uwzględniają natomiast ani procesu powstania mózgu, ani procesu jego odtwarzania, i w tym sensie odnoszą się do „struktury” mózgu. Definicje {I} i {II} można rozszerzyć na „model” świadomości zjawiskowej, jakim jest „śpiący” mózg w fazie snu REM. Model ten dopuszcza częściową separację stanów przeżyciowych od stanów behawioralnych oraz pozwala na uchwycenie różnic i podobieństw pomiędzy stanami świadomości odpowiadającym strukturalnie „śpiącego” i „nieśpiącego” mózgu. Chalmers nie odwołuje się do takiego modelu świadomości.

2. Podstawowe rozbieżności pomiędzy {I} i {II}, a definicjami podanymi przez Chalmersa widać w {II}. Rozbieżności te dotyczą rozumienia „treści stanów świadomości”. Według Chalmersa: dla każdego przeżycia *P* istnieje jakiś neuronalny, reprezentacyjny system *N*, taki że *N* jest (pod pewnymi warunkami) minimalnym neuronalnym systemem reprezentacyjnym, którego treść jest bezpośrednio związana z treścią świadomości *P*⁶.

W podanej przez Chalmersa definicji brakuje doprecyzowania, jak należy rozumieć stałe (ciągle) połączenie (bo o takie chyba Chalmersowi chodzi) treści systemu *N* z treściami *P*. Chalmers przyjmuje, że istnieje jakieś jedno do jednego (bezpośrednie) odwzorowanie cech świadomego przeżycia *P* w cechy minimalnego neuronalnego systemu *N*. Oznacza to, że jeżeli obecność *N* wystarcza dla wystąpienia *P* (powiedzmy, dla wystąpienia przeżycia z dwoma stopniami głośności), to *N* musi się składać przynajmniej ze schematów aktywności wyjaśniających spostrzeganą różnicę w głośności. Ze względu na kwestie wyjaśniania, Chalmers domaga się swoistego izomorfizmu treściowego. Tym samym należy jednak podać jakiś sposób ustalenia relacji tożsamości struktur (braku różnorodności w strukturach), zachodzącej pomiędzy elementami *N* oraz *P*. Nigdy nie mielibyśmy bowiem wystarczają-

⁶ „A content NCC [neural correlate of consciousness – J.B.] is a neural representational system *N* such that the content of *N* directly correlates with the content of consciousness” (Chalmers 2000: 20). „The obvious remedy is to require that an NCC has to be a minimal sufficient system: that is, a minimal system whose state is sufficient for the corresponding conscious state” (s. 25).

cej racji do przyjęcia (w stosunku do treści N), że jest ona minimalnie wystarczającą treścią przeżycia P , jeżeli nie uda się nam ustalić takiego odwzorowania cech przeżycia P w cechy N .

Nie wystarczy tu stwierdzić, że jeżeli mamy przeżycie ciągłego dźwięku, to neuronalna aktywność też musi być jakoś ciągła, albo że jeżeli słyszymy dźwięk ($C^{\#}$) pochodzący z prawej strony źródła dźwięku (C^b), to neuronalna aktywność odpowiadająca ($C^{\#}$) musi być z prawej strony aktywności odpowiadającej (C^b). Pozostaje pytanie: czy istnieje reprezentacyjny, neuronalny system N , którego treści w sposób stały odpowiadają treściom świadomości? „Trudny” problem leży między innymi w tym, że treści świadomości na poziomie podmiotu różnią się zasadniczo od treści (niepodmiotowych), neuronalnych przedstawień. Dodatkowo teza o psychofizycznym izomorfizmie treści musi sprostać krytyce ze strony funkcjonalizmu z tego tezą o wielorakiej realizacji, według której osobniki neuronalnie niezomorficzne mogą pomimo tego być psychologicznie izomorficzne.

3. Nadal zbyt mało wiemy o tym, w jaki sposób różne indywidualne cechy „odpalających” neuronów są ze sobą scalone i prowadzą do uzyskania wysoce kompleksowych sprawności całego mózgu lub jego części. To oznacza, że nadal zbyt mało wiemy o różnicach pomiędzy neuronami zlokalizowanymi w poszczególnych częściach mózgu (choćby o neurotransmitalnych cechach tych neuronów), aby na tej podstawie móc określać jakościowe treści w częściach koneksjonistycznie rozumianego mózgu.

Systemy neuronalne mogą przyczynowo umożliwiać przeżycia spostrzegania niezależnie od tego, że treść podmiotowego przeżycia spostrzegania jest po prostu treścią systemu neuronalnego w jakimś momencie jego aktywności. System $[\Phi_j - \Phi_n]$ może warunkować stan $[\Psi_i]$, chociaż nie musi to oznaczać, że zjawiskowe treści stanu $[\Psi_i]$ są jedynie przyczynowymi skutkami $[\Phi_j - \Phi_n]$ ⁷. Ponadto należy odróżnić neuronalną aktywność, która jest bezpośrednio skorelowana z pojedynczym świadomym przeżyciem, od aktywności odzwierciedlającej nieświadome spostrzeganie (czy działanie), skojarzone z tym przeżyciem (por. Rees 2002: 261–262). Pomocny może się tutaj okazać wprowadzony w 2.3 model świadomości zjawiskowej. Ponadto, chcąc zidentyfikować neuronalne korelaty świadomości, musimy umieć odróżnić jaki wkład w powstanie sensorycznych przeżyć mają procesy neuronalne odpowiadające procesom świadomym, a jakie nieświadomym.

⁷ Podobnie argumentuje W. Sellars w swoim „teoretycznym mentalizmie” odniesionym do wrażeń zmysłowych (por. Sellars 1963b: 191, 188–189). Teorię Sellarsa omawiamy w: Bremer 1997: 74–80, 188–121.

4. Proponowany w niniejszej pracy filozoficzno-neurologiczny program poszukiwania neuronalnych korelatów świadomości zakłada internalistyczne ujęcie treści przeżyć świadomości. Od strony filozoficznej teorii przeżycia są traktowane jako emergujące na procesach neuronalnych zachodzących w danym czasie t (por. 1.3.4, 3.6.2). Jednym z podstawowych zarzutów wobec internalizmu jest to, że spostrzeganie przez podmiot otoczenia zależy nie tylko od momentalnej neuronalnej aktywności jego mózgu, ale też od związku mózgu z całym ciałem (choćby kwestia dostarczania krwi i tlenu do mózgu) oraz od związku pomiędzy podmiotem i jego otoczeniem (odruchowe lub przemyślane ruchy gałek ocznych, ruchy głowy oraz odpowiadające tym ruchom stany mózgu). Posiadanie przeżycia dźwięku ($C^\#$) z docierającymi dźwiękami z otoczenia wymaga zintegrowanego skierowania na niego uwagi, odpowiedniego „wsluchania się”, przypominania sobie, zwrócenia głowy i tym podobnych zachowań⁸. Aktywny udział całego ciała podmiotu w określeniu spostrzeganej treści musi więc znaleźć swoje miejsce w relacji pomiędzy kompleksowym przeżyciem postrzegania, a mózgiem podmiotu. Jeżeli jednak treść spostrzegania zależy od aktywności całego podmiotu (od jego pamięci, od ruchów jego otwartych oczu, uszu, głowy i ciała) oraz od wpływów otoczenia, to rodzi się pytanie, czy zdefiniowany przez Chalmersa minimalny neuronalny system N jest wystarczający do wywołania przeżycia P .

Stąd bierze się nasze ograniczenie do internalistycznej argumentacji, odwołujące się do modelu „śniącego” mózgu. Ponadto neuronalny system [Φ_j – Φ_n] jest odniesiony do koneksjonistycznie rozumianego mózgu. Chalmers oczekuje, że poszczególne stany lub typy świadomości są wytworami danych regionów w mózgu. W tym sensie używany przez niego model mózgu jest „lokalistyczny” (por. 6.1.4).

Problem z ustaleniem obiektywnych kryteriów dla analiz świadomości – bo jako takie kryteria rozumiemy poszukiwanie neuronalnych korelatów świadomości – leży w tym, że przyjmują one istnienie czegoś, co wymaga niejako pokazania. Nie można jednak pokazać, że jakieś neurofizjologiczne zjawisko jest skorelowane ze świadomością, bez wcześniejszego posiadania bezpośredniego sposobu identyfikacji obecności lub nieobecności świadomości. Dwie zmienne N i P mogą zostać eksperymentalnie skorelowane tylko wtedy, jeżeli obydwie mogą zostać zmierzone (nawet gdy mierzalność jest jedynie ograniczona do jakościowego „tak” lub „nie”). Istnieje wiele sytu-

⁸ Odpowiadające tym zachowaniom integracyjne mechanizmy działania mózgu omawia A. Wróbel (2000a: 461–465).

cji, w których nie jesteśmy pewni obecności świadomości. Dlatego twierdzenia o korelacji neurofizjologicznych i przeżyciowych zmiennych cechują się jedynie prawdopodobnym stopniem pewności.

Określenie metody poszukiwania neuronalnych korelatów świadomości jest przez nas rozumiane jako podstawa zbudowania – w przyszłości – wielowymiarowej teorii świadomości. Dziś takiej teorii nie mamy i możemy jedynie mówić o teorii emergencyjnej. Chalmers skupia się na samej metodologii poszukiwania korelatów, a proponowana przez niego metoda poszukiwania korelatów dotyczy korelacji pomiędzy poszczególnymi stanami (mentalnymi i fizycznymi). W niniejszej pracy przyjmujemy holistyczny model świadomości oraz koneksjonistyczny model skorelowanych z nią procesów neurologicznych. Procesy te oraz odpowiadające im stany świadomości zostają następnie ujęte w ramach tzw. prototeorii cechującej się subiektywnością świadomości.

7.1.2. Dwa rodzaje kwestii metodologicznych

Przy poszukiwaniu neuronalnych korelatów świadomości natrafiamy na trudności pochodzące z: (A) ograniczeń metod badawczych i pomiarowych, oraz (B) z konieczności uściślenia pojęć.

(A) Jednym ze źródeł danych dotyczących neuronalnych korelatów świadomości są badania uszkodzeń mózgu (powypadkowych lub pooperacyjnych, por. 6.1.1). Wyniki tych badań są jednak w tym sensie trudne do oceny, że poszukiwany korelat świadomości jest zawsze odniesiony do specyficznej architektury mózgu, do jego organicznej dynamiki i do jego normalnego funkcjonowania. Badania nad neuroplastycznością mózgu potwierdzają, że uszkodzenia zmieniają architekturę mózgu. Może się więc okazać, że system $[\Phi_j - \Phi_n]$, który wcześniej był neuronalnym korelatem świadomości, już nim po uszkodzeniu nie jest.

Z występującego w przypadku uszkodzeń braku korelacji pomiędzy danym systemem a świadomością nie możemy więc bezpośrednio wnioskować, że dany system jest neuronalnym korelatem świadomości. Czasami można wnioskować pośrednio, odwołując się do braku korelacji jako oczywistości mówiącej o braku korelacji w normalnych przypadkach. Odwołując się do badań uszkodzeń mózgu często nie potrafimy odróżnić efektów destrukcji mózgu i przerwania określonych połączeń. Nie potrafimy również wyspecyfikować neuronalnych grup w heterogenicznych regionach, jak na przykład w pniu mózgu. O technicznych ograniczeniach neuroobrazowania mózgu wspomniano już wcześniej.

Należy się zgodzić z twierdzeniem, że trudno na przykład szukać korelatów stanów świadomych w pniu mózgu, ale bez pnia mózgu nie można w ogóle mówić ani o holistycznym stanie świadomości, ani o związanych z nią pojedynczych stanach. Neurony pnia mózgu są konieczne do pojawienia się świadomości, chociaż ich aktywność nie jest wystarczająca do jej normalnego funkcjonowania i do jej wyjaśnienia.

(B) Należy zauważyć, że neuronalnych korelatów świadomości może być wiele. Korelaty zdefiniowano (por. 7.1) jako układy neuronów $[\Phi_j - \Phi_n]$ albo stany tych układów, których aktywność jest wystarczająca do pojawienia się określonych stanów świadomości. Z tego rodzaju definicją jest związanych kilka zagadnień.

1. Potrzebujemy jasnych kryteriów, według których danemu podmiotowi przypisujemy świadomość. Szukamy neuronalnego systemu $[\Phi_j - \Phi_n]$, który koreluje z określonym stanem świadomości $[\Psi]$ musimy jednak wiedzieć, kiedy u podmiotu mamy do czynienia ze świadomością zjawiskową. Jak wspomniano, nie jest to łatwą sprawą ze względu na brak bezpośredniego, trzecioosobowego dojścia do świadomości zjawiskowej.

Przy przypisywaniu jakiemuś podmiotowi świadomości zjawiskowej jesteśmy skazani na pośrednie kryteria⁹, których poszukiwanie musi odpowiadać wprowadzonemu w 2.3 modelowi świadomości. Tego rodzaju kryteria nie są sprawdzalne eksperymentalnie, co oznacza, że należy być szczególnie ostrożnym wobec wszelkich założeń dotyczących subiektywnych stanów świadomości.

Wyjątek zdają się stanowić pierwszoosobowe sprawozdania osoby przy równoczesnym obserwowaniu przez nią aparatury podłączonej do jej mózgu¹⁰. Także tego rodzaju badanie ma jednak swoje techniczne ograniczenia. Ponadto należy dodać, że:

⁹ Podczas badania świadomości dostępu u innych podmiotów, kryterium takim są zazwyczaj ich zachowania, chociaż należy zaznaczyć, że interpretując zachowania zwierząt z naszej ludzkiej perspektywy możemy ulegać iluzjom. Według D. Durstewitza i S. Windmanna: „Wenn ein Affe vor einem Spiegel Grimassen schneidet und verschiedene Bewegungen ausführt, bei denen er «sich selbst» beobachtet, so muß das nicht notwendigerweise auf das Vorhanden eines «Ich-Bewußtseins» zurückzuführen sein. Vielmehr könnte der Affe einfach dem artspezifischen und mit Freude einhergehenden Trieb folgen, [...]” (Durstewitz i Windmann 1998: 128).

¹⁰ Do eksperymentu myślowego z autocerebroskopem odwołuje się H. Feigl (1958: 456). Eksperyment ten przypomina G. Leibniza eksperyment z młynem. D. Dennett i M. Kinsbourne w dyskusji z N. Blockiem podkreślają: „Does Block then hold, perhaps, that it is only possible to determine correlations of phenomenal consciousness with brain events by autocerebroscopy? Even here we will run into problems, according to Block’s own account, for even the most careful and persistent autocerebroscopist is stuck

VII.4₁ Poprzez „autocerebroskop” mógłbym jedynie zobaczyć swój własny mózg i jego stany („odpalanie” neuronów), nadal jednak nie mógłbym zobaczyć bólu (który mogą jedynie odczuwać).

Eksperyment z autocerebroskopem zakłada obecność przedempirycznych (filozoficznych) założeń, wymaganych do interpretacji uzyskanych danych. Końcowe wnioski dotyczące neuronalnych korelatów świadomości nie będą pewniejsze od pewności, jaką przypiszemy owym założeniom.

2. Różnym rodzajom świadomych stanów mogą odpowiadać różne neuronalne korelaty świadomości (przykładowo dla świadomości wzrokowej, słuchowej, jak i dla różnych aspektów każdej z nich) [por. Durstewitz i Windmann 1998: 134–135]. Dlatego przyjmujemy, że:

VII.4₂ Raczej mało prawdopodobne jest założenie, że każdej możliwej własności i każdemu obiektowi otaczającego nas świata (w dodatku z każdej możliwej perspektywy postrzegania), odpowiada w mózgu jakaś liczba wyspecjalizowanych neuronów.

Mimo tego nie mamy problemów z ponownym rozpoznaniem już raz widzianych przedmiotów i słyszanych dźwięków czy z klasyfikacją nowych przedmiotów. Należy jednak zauważyć, że w hipokampie zarówno szczurów, jak i ludzi stwierdzono obecność tzw. *place cells*, odpalających tylko w określonym miejscu, w którym dany podmiot się znajduje¹¹. Można sobie wyobrazić, że w systemie słuchowym lub wzrokowym istnieje jakaś liczba neuronów, reagujących na elementarne pobudzenia, z których – jak z alfabetu – dają się tworzyć przedstawienia przedmiotów.

3. Jak wspomniano, dla jakiegoś konkretnego rodzaju świadomego stanu (jak na przykład bólu *x*), nie możemy wykluczyć możliwości istnienia dwóch różnych systemów, których działanie jest wystarczające dla wywołania tego stanu. Może się okazać, iż neuronalnych korelatów świadomości jest mało albo że jest tylko jeden taki korelat, rozumiany jako jakiś centralny system przedstawiający treści wzrokowej świadomości, treści słuchowej świadomości, przeżyć emocjonalnych, strumienia świadomych myśli, stanów tła świa-

relying on his own judgments of what seems to be happening, but «to assume that the subjective is exhausted by judgments is to beg the question against phenomenal consciousness [...]» (Dennett i Kinsbourne 1992a: 240).

¹¹ Por. O’Keefe 1976: 79–108 (wskazanie na artykuł J. O’Keefe zawdzięczamy D. Adamkowi). „[...] these cells were not coding for simple sensory stimuli but were instead computing the more abstract concept of place or location” (O’Keefe 1998: 1334; por. s. 1338).

domości. Przykładem tego typu systemu może być wspomniany już przez nas, a wprowadzony przez B. Baarsa rodzaj „świadomego modułu”, w jego teorii *Global Workspace*¹². Baars postuluje, że ludzkie poznanie jest realizowane poprzez wielość relatywnie prostych procesów, które są specjalnego rodzaju i przebiegają nieświadomie.

Poszukiwanym systemem może się także okazać krytykowany przez D. Dennetta Kartezjański Teatr, chociaż od strony empirycznej wydaje się mało prawdopodobne że taki jeden jedyny neuronalny system istnieje. Nie podejmując w tej sprawie arbitralnych decyzji należy uwzględnić, że definicja neuronalnego korelatu powinna dać się pogodzić z istnieniem wielorakich lub wielu neuronalnych korelatów świadomości.

4. Przy poszukiwaniu neuronalnych korelatów świadomości można zaproponować swoiście „redukcjonistyczną” metodę przyjmując, że korelat powinien być rozumiany jako określony neuronalny system $[\Phi_j - \Phi_n]_o$. Przyjęcie takiego założenia zawęża poszukiwanie do znalezienia tak „wąskiego” korelatu, jak to tylko możliwe. Mając dany szeroki, zdający się korelować ze świadomością system, izoluje się wówczas te jego części, w których owa korelacja ma swoją podstawę.

Zakładając dwie hipotezy, że świadomość koreluje z szerokim systemem albo z węższym systemem w nim zawartym, można postępować według dwóch programów badawczych¹³. (a) Możemy najpierw zbadać węższą hipotezę. Jeżeli się okaże, że tak rozumiany system koreluje ze świadomością, to wtedy szerszy system nie będzie neuronalnym korelatem świadomości. Wprawdzie nie mamy gwarancji, że takie rozumowanie jest poprawne, dostarcza nam ono jednak jednej z hipotez pomocnych przy poszukiwaniu neuronalnych korelatów świadomości. Tak chociażby postępuje stosujący kano-ny J. Milla fizjolog, który w swoich badaniach eliminuje kolejno pewne obszary ciała. (b) Oczywiście można się skupiać na badaniu szerokiego systemu, a następnie próbować go zawęzić.

5. Mogą istnieć odmienne neuronalne korelaty świadomości dla różnych rodzajów stanów świadomości. Stąd konieczność założenia, że:

VII.4. Potrzebne jest rozróżnienie pomiędzy neuronalnymi korelatami stanów tła świadomości (czuwanie, śnienie, ...) i neuronalnymi korelatami specyficznych treści świadomości.

¹² „Global Workspace theory is based on the belief that, like the cells of the human body, the detailed workings of the brain are widely distributed” (Baars 1997: ix).

¹³ Podobne analizy prowadzi J. Searle (2000: 558–559).

Może się okazać, że w grę wchodzi całkiem inne systemy. Obecny stan badań wskazuje, że neuronalny korelat dla stanów tła obejmuje procesy we wzgórzu albo wzgórzowo-korowe interakcje, podczas gdy neuronalny korelat dla specyficznych treści świadomości obejmuje procesy w korze mózgowej (por. Engel 2000: 6). Zdaniem neurologa M. Raichle'a: „Wiele opracowań wskazuje na systemy wstępujące od pnia mózgu poprzez wzgórze do kory mózgowej jako odpowiedzialne za przebudzony lub uważny stan czuwania” (Raichle 1998: 1889).

Różne rodzaje neuronalnych korelatów świadomości domagają się odmiennych metod ich badania. Należy uwzględnić chociażby różnicę pomiędzy: (a) zidentyfikowaniem systemu $[\Phi_j - \Phi_n]$ dla spostrzegania jakiejś nowej cechy lub obiektu, a (b) zidentyfikowaniem systemu przy spostrzeganiu znanej już cechy lub obiektu (por. wyżej, 2).

6. Badania nad neuronalnymi korelatami świadomości muszą dopuszczać możliwość określania neuronalnej treści. Stąd badanie neuronalnych determinantów specyficznych treści świadomości, takich jak chociażby treści słuchowej świadomości. Neuronalny korelat świadomości, w tym przypadku $[\Phi_j - \Phi_n]$, byłby neuronalnym reprezentacyjnym systemem, którego treści są skorelowane z treściami stanów świadomości $[\Psi_i - \Psi_m]$.

VII.4₄ Aby określić, czy system $[\Phi_j - \Phi_n]$ jest faktycznie neuronalnym korelatem świadomości zjawiskowej, potrzebujemy metod monitorujących przedstawieniowe treści tego systemu.

Nasza wiedza, jakie korowe systemy są odpowiedzialne za przedstawieniowe treści naszej świadomości, jest nadal niepełna. Jedną z trudności w identyfikacji takich systemów leży, według M. Raichle'a, w odróżnianiu ich od systemów odpowiedzialnych za nieświadome procesy poznawcze i emocjonalne (por. Raichle 1998: 1889–1890, 1899).

7. Pomimo obszernej literatury dotyczącej badań mózgu i mimo postępujących badań, mamy do dyspozycji jedynie pewną ilość opisanych przypadków. Są to opisy badań podmiotów z uszkodzeniami mózgu albo opisy podmiotów poddawanych różnego rodzaju stymulacjom. Wyniki tych badań nie podlegają jednoznacznej interpretacji naukowo-empirycznej. Chcąc stwierdzić, że jakiś system $[\Phi_j - \Phi_n]$ jest faktycznie neuronalnym korelatem świadomości, należy uwzględnić opisy i badania kolejnych przypadków.

VII.4₅ Współczesne metody badań pozwalają raczej na uzyskanie negatywnej wiedzy, czyli wiedzy o systemach, które nie są neuronalnymi ko-

relatami danych stanów świadomości. Uzyskanie wiedzy pozytywnej jest trudniejsze.

Celem określenia, który system neuronalny koreluje z danymi stanami świadomymi, który zaś nie, prowadzone eksperymenty muszą uwzględniać różne stany przeżyciowe. Znany jest chociażby intuicyjny związek pomiędzy zmysłowym spostrzeganiem jakiegoś przedmiotu a wyobrażaniem go sobie. Neuropsychologiczne badania pokazują, że w obydwu procesach biorą często udział te same powierzchnie korowe, łącznie z tymi samymi sensorycznymi powierzchniami (por. Grush 1998: 175–183). Tego typu udział, jak wskazują inne badania neurologiczne, nie jest jednak obecny w całej jego rozciągłości¹⁴.

W precyzyjniejszym doborze innego typu przeżyć możemy się odwołać do odpowiednich „eksperymentów myślowych”, choć szczególne znaczenie mają naukowo-empiryczne eksperymenty z normalnie funkcjonującymi mózgami. Ponadto (od strony filozoficznej) należy rozważyć różnorodność hipotez, aby zawęzić szereg minimalnych neuronalnych systemów, korelujących ze świadomością w ramach kolejnych programów badawczych. Należy być przygotowanym, że:

- (a) zostaną odkryte różne neuronalne korelaty świadomości dla różnych sensorycznych modalności,
- (b) zostaną odkryte różne neuronalne korelaty świadomości dla stanów tła oraz dla świadomych treści, chociaż nie można *a priori* wykluczyć, że
- (c) takich korelatów będzie jedynie niewiele.

Filozoficznym i neurologicznym sukcesem byłoby wyizolowanie takich systemów, które silniej aniżeli inne, korelują ze świadomością w normalnie funkcjonujących mózгах. Chociażby ze względu na neuroplastyczność mózgu najprawdopodobniej nie znajdziemy (na sposób konieczny) świadomości wzrokowej w neurofizjologicznym systemie wzrokowym $[\Phi_j - \Phi_n]_w$, a świadomości słuchowej w neurofizjologicznym systemie słuchowym $[\Phi_j - \Phi_n]_s$. Być może znajdziemy korelat holistycznej świadomości zawierającej modalności wzrokowe, słuchowe itd.

¹⁴ Neuropsycholog C. H. Vanderwolf pisze: „In support of the distinction between imagination and perception of real objects, recent studies of brain imaging indicate that imagining visual scenes or verbal jingles does not increase activity in primary sensory cortex even though a peripheral sensory input does produce increases in activity in these regions” (Vanderwolf 1998: 129).

7.2. Neuronalne korelaty marzeń sennych

W naukowym doprecyzowaniu podanego w II.5 modelu świadomości pomocne są dwa sposoby postępowania.

1. J. Searle wskazuje na różnice, które są widoczne pomiędzy „obrazowaniem” mózgow przytomnych (nieśpiących) podmiotów i podmiotów w fazie snu REM, z jednej strony, a mózgow podmiotów w fazie snu „slow-wave” z drugiej strony (por. Searle 2000: 569–570). Trudno jednak wykazać, jak dalece różnice te są związane ze świadomością. Liczne procesy zachodzą bowiem zarówno w „świadomym”, jak i w „nieświadomym” mózgu (będącym np. w fazie *slow-wave* snu). Trudno powiedzieć, na ile procesy te mają coś wspólnego z generowaniem świadomości. Obrazując mózg nieśpiącego podmiotu, którego poprosiliśmy o zwrócenie na coś uwagi, poszerzamy tym samym nasze rozumienie funkcjonowania niektórych części jego mózgu. Mając jednak do dyspozycji tylko obrazy mózgu, w wielu wypadkach nie zobaczymy różnicy pomiędzy byciem ogólnie świadomym, a byciem całkowicie nieświadomym (faza *slow-wave* snu). Tak więc już przy radiologicznym ustalaniu tego, czy podmiot jest świadomy czy nieświadomy, okazuje się, w jak wczesnych stadiach rozwoju technik neuroobrazowania mózgu jesteśmy.

2. Neuroobrazowanie mózgu w czasie fazy snu REM ukazuje obszary, w których następuje zwiększenie albo zmniejszenie metabolizmu¹⁵. Tego rodzaju badania umożliwiają neurologiczne opisy znanych przeżyciowych cech snów, co jest możliwe dzięki dostrzeżeniu relacji pomiędzy: (a) związkami specyficznych przeżyciowych stanów świadomych „obudzonego” podmiotu z określoną aktywnością w poszczególnych częściach jego mózgu, a (b) związkami specyficznych przeżyciowych stanów świadomych „śpiącego” podmiotu z określoną aktywnością w jego mózgu. Niektóre przeżyciowe cechy świadomości podczas fazy snu REM znamy z perspektywy pierwszoosobowej, albo z opowiadań innych osób. Na przykład (w gruncie rzeczy „nierzeczywiste”, bo tylko „wyśnione”) ruchy ciała podczas jakiegoś marzenia sennego mają swój neurologiczny korelat w mózdzku oraz w zwojach podstawy. W tych samych częściach mózgu mają swoje korelaty akcje motoryczne „obudzonego” podmiotu. Halucynacje wzrokowe mają swój korelat w spadku metabolizmu w pierwszorzędowej korze wzrokowej (V1) i we

¹⁵ Neurolog J. A. Hobson omawia i przedstawia graficznie aktywności poszczególnych organów „śpiącego” mózgu oraz odpowiadające tym aktywnościom świadome przeżycia „nieśpiącego” mózgu (por. Hobson 2000: 810–813; 2001: 774; Hobson i Pace-Schott 2002: 686–691).

wzroście metabolizmu w warstwach kory wzrokowej V2, V3. Silne emocje (obawa, lęk) przeżywane podczas snu mają swoje korelaty w paralimbicznym układzie, w jądrach migdałowatych, w hipokampie i w płacie czołowym. W tych samych obszarach mają swoje korelaty przeżywane na jawie stany emocjonalne.

Wychodząc z (a) oraz (b) i zakładając, że zachodzi podobieństwo pomiędzy neuronalnymi korelatami stanów „na jawie” i stanów „we śnie” można próbować przewidzieć neuronalne korelaty przeżyciowych stanów w fazie snu REM. Jeżeli przewidywania te sprawdzą się, można przyjąć, że wspomniane stany mózgu są związane ze świadomymi cechami, spotykanymi w obrazach sennych.

Tym samym proponowany model świadomości zjawiskowej przemawiałby za przyjęciem empirycznych zależności pomiędzy danymi stanami świadomości zjawiskowej, a stanami poszczególnych części mózgu.

Na podstawie tak określonych korelatów można przyjąć hipotezę, że korelaty przeżyciowych treści snów odpowiadają funkcjonalnie wyizolowanemu połączeniu pomiędzy warstwami kory V2 i V4, a systemem paralimbicznym, przy rozłączonym połączeniu z wejściem sensorycznym i wyjściem motorycznym, powstałym dzięki hamującemu działaniu prążkowiego (*striatum*) i płatu czołowego (*lobus frontalis*). Dodatkowych informacji dostarczają badania śniących podmiotów, u których wspomniane części mózgu są uszkodzone. Otrzymujemy w ten sposób informacje o funkcjonalnej strukturze mózgu oraz o tym, w jaki sposób różnorodne zjawiskowe cechy subiektywnej świadomości są skorelowane ze wzrostem lub spadkiem aktywności w różnych obszarach mózgu. Na przykład niektóre osoby, cierpiące na ślepotę barw (mózgowa achromatopsia) spowodowaną zniszczeniem neuronów w warstwie V4 (w zakręcie potyliczno-skroniowym), nie tylko nie potrafią postrzegać kolorów, ale także nie potrafią ich sobie wyobrazić ani nie przeżywają tych kolorów w swoich snach¹⁶. Neurologiczne badania „śpiących” podmiotów ukazują funkcje i cechy świadomości, które nie mogą być właściwie ujęte ani przez introspektywną fenomenologię, ani przez eksperymenty psychologiczne.

7.3. Wpływ wad mózgu na stany świadomości

Holistycznie rozumiana świadomość posiada subiektywną rzeczywistość, która jest zjawiskiem treściowym, skorelowanym z systemem nerwowym. Czym innym jest zobaczyć, jakie neuronalne uwarunkowania leżą ewentual-

¹⁶ Jeden z takich przypadków opisuje Konorski (1969: 176).

nie u podstaw tej rzeczywistości, czym innym natomiast, jak pojęciowo te uwarunkowania określić.

Pytanie o neuronalne korelaty świadomości zawiera tym samym kognitywistyczne pytanie o połączenie argumentacji różnego typu: (a) w jaki sposób subiektywne przeżycie „bycia świadomym” jest skorelowane ze zdarzeniami w mózgu? Jakie warunki określają, czy system neuronów (i jego stany) może być jakoś odwzorowany w treści świadomych przeżyć? Jedną z odpowiedzi na pytanie (a) podamy w punkcie 7.3.2, do którego wprowadza 7.3.1, zawierający opisy konkretnych neurologicznych przypadków, posiadających konsekwencje dla filozoficzno-psychologicznie rozumianego zagadnienia subiektywności świadomości. Chodzi przy tym o tego rodzaju filozoficzne postawienie problemu, które może być brane pod uwagę w badaniach neurologicznych.

Podobna rola do tej, jaką przy określaniu neuronalnych korelatorów świadomości odgrywają badania marzeń sennych, przypada badaniom podmiotów z uszkodzonym lub niedorozwiniętym mózgiem. Wyniki niektórych z tych badań omawiamy w następnym punkcie.

7.3.1. Przykłady niedorozwoju mózgu

Nawiązując w poszukiwaniu neuronalnych korelatów świadomości do zagadnień poruszonych przez T. Nagła (por. 5.5.3, 5.6.1) należy zapytać: jak duże części mózgu i w ilu fragmentach są konieczne, aby móc mówić o subiektywnej świadomości ludzkiego podmiotu? Ponadto należy jednak zapytać: obecność jakich struktur koneksjonistycznie rozumianego mózgu jest konieczna, aby móc mówić o subiektywności świadomości? Dziś nie jest możliwe udzielenie jednoznacznej odpowiedzi na te pytania. Widać to chociażby na podstawie znanych z literatury neurologicznej opisów różnych anomalii w rozwoju mózgu, a także na podstawie opisów skutków, jakie na stan świadomości wywierają uszkodzenia lub niedorozwój mózgu. Poniżej przedstawiamy kilka tego rodzaju przypadków. Pierwsze z nich (I) dotyczą filogenetycznie młodszych struktur mózgu, drugie (II) starszych jego struktur. (I) Z jednej strony, w podręcznikach neuroanatomii spotykamy „mapy” przedstawiające pofałdowaną, zewnętrzną powierzchnię normalnie rozwiniętych półkul mózgowych, wraz z podaniem obszarów korowych, odpowiadających za poszczególne czynności organizmu. Kora mózgowa jest wielowarstwową strukturą mózgu związaną z nadawaniem świadomego sensu wszystkim informacjom do niej dochodzącym. Z drugiej strony, istnieją udokumentowane opisy ludzi z małymi głowami (tzw. małogłowie, *microcephaly*) i posiadającymi normalny stopień inteligencji. Objętość mózgu jest u nich zredukowana do 880–960 cm³, przy zachowanej podstawowej, chociaż zmniejszonej ar-

chitekturze mózgu (przeciętna objętość „normalnego” mózgu wynosi 1371 cm³ u mężczyzn i 1216 cm³ u kobiet). Według J. R. Skoylesa: „Pomimo że małągłowie jest zwykle kojarzone z zahamowaniem i niską inteligencją, ni-
kły lecz znaczący procent [pacjentów – J.B.] wykazuje się przeciętną inteli-
gencją [...]” (Skoyles 1999: 4; por. Travis 2002: 1–3)¹⁷.

Skoyles zaznacza, że wielkość mózgu jest związana ze zwiększeniem
sprawności podmiotu (*expertise capacity*), a nie z jego IQ. Na przykład umie-
jętność profesjonalnego grania na skrzypcach zależy od finezyjnego skoordy-
nowania palców lewej ręki i dokładnej koordynacji pomiędzy obydwo-
ma rękami. Zawodowi skrzypkowie mają dwa do trzech razy większą powierzch-
nię kory odpowiadającą za ich lewe palce aniżeli niegrający na skrzypcach.

Neurolog J. Lorber opisuje przypadek studenta matematyki (IQ 126),
który nie skarżył się na jakieś poważniejsze dolegliwości, a jednak po zbada-
niu go metodą CAT stwierdzono, że „prawie wcale” nie ma on mózgu. Na
skutek wodogłowia (*hydrocephalus*) zamiast dwóch półkul wypełniających
jamę czaszki miał zaledwie cieniłą warstwę tkanki mózgowej (*cerebral
tissue*). Reszta czaszki była wypełniona płynem mózgowo-rdzeniowym (*liquor
cerebrospinalis*) [por. Lewin 1980: 1232]. Względnie duży brak „substancji
szarej” nie ograniczał jednak jego inteligencji ani nie wpływał ujemnie na
wykonywane przez niego czynności.

Lorber przebadał ponad 600 pacjentów chorych na wodogłowie wyodrę-
biając cztery grupy: (i) z prawie normalnym mózgiem, (ii) 50–70 procent ja-
my czaszki wypełnionych płynem mózgowo-rdzeniowym, (iii) 70–90 pro-
cent jamy czaszki wypełnionych płynem, (iv) 95 procent jamy czaszki wy-
pełnionych płynem. Ostatnia grupa obejmowała nieco mniej niż 10%
badanych. IQ połowy członków tej grupy przekraczał 100.

Lorber zgadza się, że odczytywanie wyników uzyskanych w CAT może
być obarczone błędami, ale dodaje: „Nie mogę powiedzieć, czy student ma-
tematyki ma mózg o ciężarze 50 czy 100 gramów, ale jest prawie pewne, że
jego ciężar nie jest zbliżony do normalnego, czyli do 1,5 kg”.

Opisany przez Lorbera przypadek rodzi wiele pytań. Brak dodatkowych
danych powoduje, że trudno go sklasyfikować. Nie wiemy chociażby, na ja-
ki rodzaj wodogłowia jego pacjent cierpiał: na wodogłowie wewnętrzne (wy-
wołane zamknięciem ważnych dróg przestrzeni podpajęczynówkowej) lub
zewewnętrzne (rozwija się w wypadku zaniku kory mózgowej)¹⁸, na wodogło-
wie wrodzone lub spowodowane krwotokami wewnątrzczaszkowymi.

¹⁷ Ewolucyjną zależność pomiędzy wielkością mózgu, a zdolnością organizmu do prze-
życia opisuje G. Roth (1999: 214–218).

¹⁸ Por. Duus 1989: 322–325. Raczej musiało to być wodogłowie wewnętrzne i wrodzone.

(II) Zajmującemu 10 procent objętości mózgu mózdkowi (*cerebellum*) neurologdy długo przypisywali podstawową funkcję przy koordynacji skurczów mięśni, odpowiedzialnych za wykonanie szybkich, precyzyjnych ruchów¹⁹. Poważne uszkodzenia mózdku objawiają się w zaburzeniach równowagi, w uczuciach niepewności podczas stania i chodzenia.

Badania przeprowadzone przez L. M. Parsons wskazyują na dodatkowe funkcje mózdku. Używając metody PET obserwował on aktywację mózdku podczas sensorycznych i poznawczych procesów związanych z odczuwaniem przez podmiot pragnienia. W czasie obserwacji Parsons odwoływał się do hipotetycznego modelu, w którym funkcjonowanie mózdku jest podobne do funkcjonowania serwera rozsyłającego informacje do rozgałęzionych komputerów. Gdy jesteśmy spragnieni albo właśnie ugasiłiśmy pragnienie, mózg rozpoczyna wielki „proces oceny” tego stanu, ustalając, czy wszystko jest w porządku. Mózdek czuwa nad tym, aby cały proces eliminacji pragnienia przebiegł efektywnie²⁰. Mózdek można więc rozumieć jako organ zaangażowany w procesy poznawcze i behawioralne, a nie tylko jako organ koordynujący aktywności ruchowe. Ponadto mózdek koordynuje wejścia i wyjścia do struktury korowej i podkorowej mózgu²¹.

Uzyskane przez L. M. Parsons i I. Dauma wyniki pokazują, że mózdek koordynuje i synchronizuje pracę reszty mózgu oraz moduluje zachowania motoryczne, planowanie tych zachowań i uczenie się ich. Wpływ ten musi być istotnym czynnikiem w strukturze świadomych procesów. Nie widać jakiejś oczywistej zależności pomiędzy uszkodzeniem mózdku a przeżywaniem przez osobę tego czy innego stanu (na przykład pragnienia). Omawiane badania pokazują jednak, że uszkodzenie mózdku powoduje subtelne pogorszenie szybkości przebiegu mentalnych procesów, modulowanych przez mózdek²².

¹⁹ Historyczny rozwój badań nad wpływem mózdku na funkcje motoryczne i poznawcze przedstawiają I. Daum i H. Ackermann (1997: 456–457).

²⁰ Por. Parsons 2001: 167; Fox 2000: 1–2. Zależność pomiędzy uszkodzeniami różnych części mózdku a procesami poznawczymi omawia J. D. Schmahmann (1997: 437–440). Przedstawiając wpływ mózdku na proces nauki języka, J. Fiez i M. E. Raichle stwierdzają: „[...] there is a growing consensus that the cerebellum contributes to some aspects of language processing, it is less clear exactly how the cerebellum does so” (Fiez i Raichle 1997: 249).

²¹ Skutki uszkodzenia mózdku są często niedostrzegalne na pierwszy rzut oka. „Likewise, if one is born without cerebellum, few impairments are reported. These observations imply that the function of the cerebellum is subtle and not necessary, but supportive” (Parsons 2001: 174).

²² Por. Parsons 2001: 174. Analogicznie można powiedzieć, że wiele organów ciała ma duży „margines” jakby „zapasowej” tkanki. Na przykład usunięcie jednej nerki zmniejsza jedynie margines bezpieczeństwa życia.

Znane są także przypadki osób, które urodziły się z bardzo małym mózdzkiem (*hypoplasia cerebelli*), ponieważ jednak często brak ten występuje razem z innymi brakami w mózgu, trudno stwierdzić jego jednoznaczny wpływ na zachowania i przeżycia tychże osób²³. Kontrolując rozsiane synchronizacje pomiędzy partiami mózgu, mózdzek odgrywa istotną rolę chociażby przy dostrzeganiu przez podmiot różnicy pomiędzy snem a jawą, a tym samym jego funkcjonowanie jest istotne dla mówienia o subiektywnej świadomości (por. 7.2).

Przypadki osób z małą objętością kory mózgowej, osób urodzonych bez albo z bardzo małym mózdzkiem należy odróżnić od wrodzonej wady zwanej bezmózgowiem (*anencephalus*). Rodzące się z nią niemowlęta są pozbawione części czaszki oraz kory mózgu, chociaż posiadają rudymenatny i funkcjonujący pień mózgu i międzymózgowie. Bezmózgowie jest wadą letalną (prowadzi niechybnie do śmierci niemowlęcia).

7.3.2. Określanie korelatów świadomości w zdeformowanym mózgu

„All parts of the brain may well be involved in normal conscious processes but the indispensable substratum of consciousness lies outside of the cerebral cortex, probably in the diencephalon” (W. Penfield)²⁴.

W dalszych analizach odróżniamy rolę anatomicznych czynników „warunkujących” procesy biologiczne i psychiczne, od roli czynników nadrzędnych, „kontrolujących”. Przykładowo praca serca lub nerek warunkuje i umożliwia, ale nie kontroluje procesów psychicznych. Wspomniany w 7.3.1 (I) przykład studenta ze skrajnym wodogłowie (czyli zanikiem kory mózgowej) pokazuje, że niezbywalny i konieczny neuronalny substrat holistycznej świadomości, czyli „bycia świadomym”, znajduje się poza korą mózgową. Prawdopodobnie jest on zlokalizowany w międzymózgowiu (*diencephalon*), składającym się ze wzgórza (*thalamus*), nadwzgórza (*epithalamus*), niskowzgórza (*subthalamus*) oraz z podwzgórza (*hypothalamus*). Docierające do organizmu bodźce, aby mogły zostać uświadomione, muszą

²³ Zdaniem I. Dauma i H. Ackermanna: „[...] degenerative diseases and selective lesions of the cerebellum do not necessarily give rise to overt general intellectual impairment” (Daum i Ackermann 1997: 459). D. Adamek [i inni] opisują przypadek 38-letniej, mentalnie nieznacznie upośledzonej kobiety, u której w czasie badania pośmiertnego stwierdzono aplazję (głęboki niedorozwój) mózdzku. Przypadek ten autorzy rozumieją jako „[...] an example of adaptive abilities of the CNS to the absence of a structure which plays so important role in motor and sensory function” (Adamek 1986: 98).

²⁴ Cytat pochodzi z: Bogen 1995: 54.

przejsć przez wzgórze (dlatego wzgórze nazwano „drogą do świadomości”²⁵).

Przykłady (I)

Chociaż pacjenci opisani w 7.3.1 (I) mają znacznie zredukowaną objętość kory mózgowej, w miarę normalnie radzą sobie w życiu. Niektórzy z nich mają wysoki stopień IQ. Neuronalnym korelatem ich stanu „bycia świadomym” jest, nazwijmy go, neuronalny system $[\Phi_j-\Phi_n]_n$ bezpośrednio połączony z holistycznie rozumianą, choć może jakoś ograniczoną, świadomością zjawiskową. Przypadek korelacji zachodzącej w międzymózgowiu może, przynajmniej częściowo, zostać rozwinięty na $[\Phi_j-\Phi_n]_n$. Pytanie brzmi: czy porównując wyniki badań uzyskanych na grupach pacjentów *i-iv* (7.3.1, [I]) wystarczy mówić o różnicach w lokalizacji tego systemu?

Mowa o tego rodzaju korelacji zakłada oczywiście możliwość definiowania neuronalnego systemu, niezależnie od jego związku ze świadomością (niezależnie od tego, czy podmiot relacjonuje, że jest świadomy, czy nie jest). System $[\Phi_j-\Phi_n]_n$ jest związany ze stopniami neurochemicznej aktywacji, tak iż będzie trzeba umieć empirycznie odróżnić – jeżeli nie dzisiaj to w przyszłości – czy odpowiada mu stan: (a) braku świadomości lub (b) posiadania holistycznie rozumianej świadomości.

Na uwagę zasługują stany subiektywnych przeżyć, występujące u wspomnianych w (I) podmiotów w konkretnym odcinku czasu. Chodzi o ich stany przeżywania kolorów, dźwięków, o przeżywanie stanów poznawczych, itp. Od strony zjawiskowej tego rodzaju stany mogą być wyodrębnione i określone poprzez ich przeżyciowe treści (czyli poprzez ich przeżywane treści przedstawiające świat jako właśnie taki, a nie inny). Neuronalnym korelatem takich stanów są wtedy jakieś stany neuronalnych systemów $[\Phi_j-\Phi_n]_1$, $[\Phi_j-\Phi_n]_2$, ..., określające specyficzne treści danego przeżycia. Znowu zakłada to oczywiście, że znany jest sposób definiowania reprezentujących, neurochemicznych „treści” systemów $[\Phi_j-\Phi_n]_1$, $[\Phi_j-\Phi_n]_2$ (por. 7.1). Dopiero wtedy moglibyśmy powiedzieć, że znaleźliśmy takie systemy, gdy wpływając na ich aktywność (na przykład przez elektryczne pobudzenie) możemy

²⁵ Dvus 1989: 246. Wzgórze jest podkorowym miejscem przełącznikowym i integracyjnym, do którego docierają eksteroceptywne i proprioceptywne impulsy ze świata zewnętrznego i wewnętrznego. Ponadto jest ono miejscem przełącznikowym dla wszystkich impulsów (z wyjątkiem impulsów węchowych), wychodzących z receptorów skórnych i z narządów wewnętrznych oraz tych, które są doprowadzane do wzgórza przez drogi wzrokową i słuchową. Do wzgórza dopływają impulsy z podwzgórza, mózdzku i pnia mózgu (z *formatio reticularis*), aby następnie dotrzeć do mózgu. Wychodzące ze wzgórza drogi częściowo biegną do prążkowiec (*striatum*), ale przede wszystkim do kory mózgu (s. 240–247).

określać subiektywne treści danych przeżyć albo przynajmniej jakieś granice tych treści. W grupach *i-iv* pojawiłyby się odpowiednio: $[\Phi_j-\Phi_n]_{nI}$, $[\Phi_j-\Phi_n]_{nII}$, uwzględniające wpływ połączeń międzymózgowia ze zmienną grubością warstwy kory mózgowej, odpowiadającą przypadkom opisanym w 7.3.1 (*i-iv*).

Możemy szukać neuronalnych korelatów dowolnego zjawiskowego stanu świadomości, rozumiejąc, że chodzi tutaj o stan obejmujący jako korelaty stany systemów $[\Phi_j-\Phi_n]_1$, $[\Phi_j-\Phi_n]_2$. Otrzymamy wtedy neuronalny korelat cechy zjawiskowej $[c\Psi]$.

VII.5 Jakiś stan systemu $[\Phi_j-\Phi_n]$ jest neuronalnym korelatem zjawiskowej cechy $[c\Psi]$, gdy bycie systemu w tym stanie jest bezpośrednio skorelowane z występowaniem cechy $[c\Psi]$.

Nie chodzi tutaj o sam system jako korelat cechy $[c\Psi]$, lecz o pewien stan systemu $[\Phi_j-\Phi_n]$. Na przykład pobudzenia w określonych neuronach mogą być neuronalnymi korelatami widzenia koloru czerwonego. Ponadto nie chodzi o neuronalne korelaty pojedynczych cech zjawiskowych, lecz o rodziny cech zjawiskowych (powiedzmy jedna cecha dla każdego rozróżnialnego koloru).

VII.5₁ Neuronalnym korelatem rodziny zjawisk $[Rc\Psi]$ jest neuronalny system $[\Phi_j-\Phi_n]_R$, taki że stan $[\Phi_j-\Phi_n]_R$ bezpośrednio koreluje ze zjawiskami $[Rc\Psi]$.

Neuronalny system $[\Phi_j-\Phi_n]_R$ będzie neuronalnym korelatem świadomości dla $[Rc\Psi]$, gdy istnieje odpowiednia liczba stanów $[\Phi_j-\Phi_n]_R$, jeden dla każdej cechy zjawiskowej typu $[c\Psi]$.

VII.5₂ Dla neuronalnego korelatu holistycznie rozumianej świadomości można podać prostą zjawiskową rodzinę z dwiema „cechami”: „byciem świadomym” $[\Psi_s]$ i „niebyciem świadomym” $[\Psi_{ns}]$.

Neuronalny korelat holistycznej świadomości będzie wtedy systemem z dwoma stanami, skorelowanymi z tymi dwiema „cechami”. Należy zatem zbadać ogólny warunek bycia „świadomym” mózgiem w przeciwieństwie do warunku bycia „nieświadomym” mózgiem.

Subiektywności nie da się prawdopodobnie wyjaśnić poprzez poszukiwanie jakiegoś jedyne i minimalnego neuronalnego korelatu zjawiskowej świadomości. Kwestią jest, jakie cechy systemu złożonego ze wzajemnie po-

łączonych jąder i warstw kory potrafią korelować z subiektywną świadomością. Można postawić roboczą hipotezę, że jakaś część świadomości zjawiskowej jest zlokalizowana w systemie wzgórze–kora mózgowia oraz że różne inne systemy doprowadzają informację do tego systemu, co chociażby powoduje powstanie modyfikacji odpowiadających różnym modalnościom sensorycznym.

Przykłady (II).

Neuropsychologiczne badania pacjentów z uszkodzonym mózdzkiem nie dostarczają jednoznacznych wyników, gdy chodzi o skutki tych uszkodzeń. Różnice mogą pochodzić z doboru pacjentów, z których niektórzy cierpieli nie tylko na uszkodzenia lub schorzenia mózdzku, lecz także na inne choroby mózgu. Różnice te mogą być także powodowane brakami metodologicznymi, nieadekwatną kontrolą lub nieadekwatnym porównywaniem odpowiednich grup pacjentów. Ponadto kwestią istotną jest zasada dobierania (wykształcenie, wiek, IQ) podmiotów kontrolnych przy badaniu wpływu mózdzku na procesy poznawcze.

Pamiętając o wspomnianych ograniczeniach metod badawczych (por. 6.1.3) oraz o wieloznacznościach w interpretacji uzyskiwanych wyników, można przedstawić wyniki badań mózdzku dotyczące odczuwania pragnienia. Prawie całkowity brak albo brak części mózdzku nie wpływa na samo jakościowe i subiektywne odczuwanie lub nieodczuwanie przez pacjentów pragnienia, w tym sensie, że nie odczuwaliby oni pragnienia jako pragnienia. W tym rozumieniu możemy mówić, że stan systemu $[\Phi_j - \Phi_n]_R$ jest neuronalnym korelatem świadomości zjawiskowej w sensie *jak to jest* odczuwać pragnienie. Z drugiej strony widać, że: dobrze funkcjonujący mózdzek może być traktowany jako dodatkowy neuronalny korelat: jako system $[\Phi_i - \Phi_j]^{II}$ umożliwiający sprawne funkcjonowanie systemu $[\Phi_j - \Phi_n]_R$.

Odpowiednie funkcjonowanie korelacyjnego systemu $[\Phi_i - \Phi_j]^{II}$ jest wymagane dla utrzymania aktywności w $[\Phi_j - \Phi_n]_R$, wywołującej efekty łączone zazwyczaj ze świadomością i umożliwiające, że $[\Phi_j - \Phi_n]_R$ jest skorelowany z holistyczną świadomością. W intuicyjnie rozumianym sensie $[\Phi_i - \Phi_j]^{II}$ nie musi być korelatem świadomości, jego rola jest podobna do roli płuc, których funkcjonowanie umożliwia pracę mózgu.

System $[\Phi_i - \Phi_j]^{II}$ można nazwać neuronalno-neuronalnym korelatem świadomości²⁶. Świadomy podmiot nie jest oczywiście świadomy ani

²⁶ Lateralna kora mózdzku (*cortex cerebelli*) otrzymuje wejścia od kory mózgowej (płatów przednich mózgu, odpowiedzialnych między innymi za funkcje językowe) i projektuje (poprzez jądro zębate [*nucleus dentatus*] i wzgórze) z powrotem do kory mózgu (*cortex cerebri*) tworząc tak zwany system korowo-mózdzkowy). Tego rodzaju połączeń

$[\Phi_j - \Phi_n]_R$ ani $[\Phi_i - \Phi_j]^{II}$. Obecność $[\Phi_i - \Phi_j]^{II}$ wpływa jednak pośrednio na skorelowany z $[\Phi_j - \Phi_n]_R$ subiektywny stan odczuwania pragnienia. Przy odczuwaniu pragnienia, oczekiwanie zaspokojenia go jest na każdym kroku sterowane przez dążenie mózgu do osiągnięcia stanu sprzed wystąpienia pragnienia. Nawet gdy podmiot zapomniiał, że odczuwa pragnienie, poprzez stany $[\Phi_i - \Phi_j]^{II}$ mózdzek wpływa na stany dążące do zaspokojenia pragnienia.

7.4. Stopień zależności między systemami neuronalnymi a świadomością

Wychodząc poza przykłady (I) i (II) (por. 7.3.1) należy zadać bardziej ogólne pytanie: jak silna modalnie musi być omawiana do tej pory relacja: mózg–świadomość, aby system $[\Phi_j - \Phi_n]$ mógł być traktowany jako neuronalny korelat świadomości? Przykładowo dla każdej materialistycznej teorii świadomości idealny byłby dowód, na który wskazuje J. Searle, a w którym poszukiwany neuronalny element nie jest jedynie skorelowany ze świadomością, ale jest *przyczynowo konieczny i wystarczający* (przy zachowaniu innych parametrów) dla obecności świadomości (por. Searle 2000: 569). Searle zauważa jednak, że na obecnym etapie badań bardzo trudno jest znaleźć tak rozumiany korelat świadomości. Modalnie podobnego skorelowania poszukuje T. Nagel.

7.4.1. Warunek konieczny i wystarczający wystąpienia stanu świadomości

Załóżmy, że dany stan systemu $[\Phi_j - \Phi_n]$ jest *konieczny* (w sensie: neuronalna aktywność w innym miejscu mózgu nie jest konieczna) i *wystarczający* (rozumiany jako kwestia praw natury), dla wystąpienia odpowiadającego mu stanu świadomości zjawiskowej $[\Psi_i]$ (por. Żegleń 2003: 288).

(A) Warunek ten wydaje się być zbyt silny, gdyż – jak się przekonamy w następnym punkcie – może istnieć więcej niż jeden neuronalny korelat danego stanu świadomego. Mogą, na przykład, istnieć dwa systemy, $[\Phi_j - \Phi_n]$ i $[\Phi_{j+l} - \Phi_{n+k}]$ takie, że pewien stan systemu $[\Phi_{j+l} - \Phi_{n+k}]$ jest wystarczający

nie spotykamy u innych gatunków naczelnnych. W ten sposób kora mózdzku otrzymuje informacje z kory mózgu. Według H. C. Leinera: „Yet, precisely these additional connections in the human brain, particularly those linking the cerebellum to the language areas of the cerebral cortex, are of paramount importance to theorists who seek to explain the neural basis of human cognitive and linguistic capabilities” (Leiner 1997: 541).

dla wystąpienia bólu $[\Psi_i]$, a jakiś stan $[\Phi_j - \Phi_n]$ jest także wystarczający dla wystąpienia bólu $[\Psi_i]$. Świadczą o tym chociażby badania nad neuronalnymi odwzorowaniami bólu przeprowadzone osobno przez B. Bromma i G. Reesa²⁷.

Dodatkowo może być tak, że systemy $[\Phi_j - \Phi_n]$ i $[\Phi_{j+l} - \Phi_{n+k}]$ nie są ze sobą skorelowane. Tym samym powiemy, że zarówno system $[\Phi_j - \Phi_n]$, jak i $[\Phi_{j+l} - \Phi_{n+k}]$ są neuronalnymi korelatami bólu. Nie jest jednak tak, że aktywność w $[\Phi_j - \Phi_n]$ jest konieczna i wystarczająca dla wystąpienia bólu, ani nie jest tak, że aktywność w $[\Phi_{j+l} - \Phi_{n+k}]$ jest konieczna i wystarczająca dla wystąpienia bólu.

VII.6 Jeżeli dwa systemy, $[\Phi_j - \Phi_n]$ i $[\Phi_{j+l} - \Phi_{n+k}]$ mogą być traktowane jako neuronalne korelaty stanu świadomości $[\Psi_i]$, to nie możemy wymagać, aby jeden z nich (np. $[\Phi_j - \Phi_n]$) był warunkiem *koniecznym i wystarczającym* dla wystąpienia tego stanu.

Oznacza to odrzucenie tezy S. Kripkego o konieczności (por. 3.4.1), i przyjęcie tezy o przygodności wspomnianych korelacji.

B. Bromm zauważa ponadto, że ze względu na duże różnice w anatomii mózgu obserwowane u poszczególnych osobników, nie da się przeprowadzić jakichś daleko idących generalizacji uzyskanych wyników. W celu ustalenia, z jaką dokładnością określono wyniki badań wyższych funkcji mózgu, należy powtarzać badania z tym samym podmiotem zmieniając chociażby intensywność bodźców.

Niektórzy ludzie są bardziej, inni mniej wrażliwi na ból. Porównując wyniki neuroobrazowania mózgow dwóch grup ludzi R. C. Coghill stwierdza, że u tych, którzy mówili, iż silnie odczuwali ból, części mózgu zaangażowane w przeżywanie bólu były bardziej aktywne²⁸. Z jednej strony, Coghill za-

²⁷ „Pain is processed in multiple brain areas with large overlap in space and time. All imaging studies stress the involvement of the secondary somatosensory cortex, the cingulate gyrus, and the frontal cortex in the experience of pain. Some PET, fMRI, and other BOLD studies found, in addition, subcortical structures in the medial midbrain, thalamus, lentiform nucleus, and cerebellum, which are active in subacute and chronic pain” (Bromm 2001: 246). G. Rees zaznacza, że identyczne zmiany w przeżyciu mogą mieć dwa różne neuronalne korelaty. „Thus, two different experiments show that identical changes in phenomenal experience have two different neuronal correlates in two different cortical areas of two different species” (Rees 2001: 36).

²⁸ „Pain intensity processing is distributed across multiple cerebral cortical and subcortical regions and is not confined to areas classically thought to be engaged in sensory-discriminative processing” (Coghill 1999: 1939). Por. też Craig 2002: 662–664.

uważył wzrost aktywności w korze sensorycznej (zajmującej się lokalizacją bólu i jego intensywnością) oraz w korze zakrętu obręczy (*anterior cingulate cortex*), która przetwarza nieprzyjemne odczucia spowodowane bólem. Z drugiej strony odkrył u badanych podmiotów jedynie małe różnice w aktywności wzgórza, które jest zaangażowane w transmitowanie sygnałów z rdzenia kręgowego do regionów mózgu. Może to świadczyć o tym, że niezależnie od wrażliwości podmiotu, jego rdzeń kręgowy przekazuje w podobny sposób sygnały bólowe, a dopiero gdy dotrą one do mózgu, są przetwarzane w różny sposób (gdy chodzi o ich intensywność)²⁹. Dostępne Coghillowi neuroobrazowanie dostarcza nie tylko obiektywnego potwierdzenia dla indywidualnych i subiektywnych sprawozdań dotyczących odczuwania bólu, lecz także informacji o rozłożeniu neuronalnych korelatów bólu.

Odpowiedź na pytanie o konieczność połączenia między zjawiskami mentalnymi a fizycznymi nie jest łatwiejsza, gdy zamiast o stany zacznie się pytać o korelacje zdarzeń mentalnych i fizycznych, rozumianych jak w teorii identyfikacji egzemplarzy (por. 3.4.). Jako ontologiczny warunek konieczny do tego, by mentalne zdarzenie [Ψ_j] mogło być identyczne ze zdarzeniem neurofizjologicznym [Φ_j] trzeba przyjąć, że obydwa zdarzenia zachodzą w tym samym podmiocie. Mając daną lokalizację obydwu zdarzeń (związaną z tym samym podmiotem) otrzymamy konieczny warunek ich identyfikacji w przestrzeni. Ale potrzebujemy jeszcze – zgodnie z tezą o identyfikacji zdarzeń – identyfikacji w czasie. Tutaj jednak rodzi się problem z określeniem tożsamości zdarzenia neurofizjologicznego, gdyż *prze-strzenność* zdarzenia fizycznego jest łatwiejsza do określenia aniżeli jego *czasowość*. Jeżeli bowiem udało nam się określić „przestrzennie” najmniej system [$\Phi_j - \Phi_n$], w którym zachodzą odpowiednie zmiany), to na jakie odcinki czasowe dzielić zachodzące w nim procesy? Neurofizjologiczne zdarzenie może być (prawdopodobnie) wyróżnione jako coś wydarzającego się w mózgu w czasie minut, sekund lub milisekund. Które z tych „plasterków” czasowych należy przyjąć jako interesujące nas zdarzenie neurofizjologiczne? Ponieważ indywiduacja zdarzeń fizycznych ma swoją podstawę w ich identyfikacji, to tym samym sama indywiduacja utraciła swój status wyjaśniający.

(B) Stąd nasze założenie, że system [$\Phi_j - \Phi_n$] jest *wystarczający* dla wystąpienia odpowiedniego stanu świadomości. Przyjęcie takiego warunku prowadzi jednak w dalszej konsekwencji do twierdzenia, że cały, koneksjoni-

²⁹ Według Coghilla: „Pain intensity information [...] is both a critical precursor and an integral component of other individual processes [affect, motor control, attention, feature extraction – J.B.] of the pain experience” (Coghill 1999: 1941).

stycznie rozumiany mózg (albo duże jego części) są neuronalnymi korelatami jakiegoś stanu świadomego. Odwołując się do badań przeprowadzonych przez B. Bromma należy zauważyć, że większe partie mózgu albo cały mózg będą wtedy traktowane jako korelaty bólu $[\Psi_i]$, ponieważ bycie w jakimś stanie całego mózgu wystarczy dla wystąpienia bólu.

Tego rodzaju „wystarczający korelat” nie jest jednak tym, jakiego się poszukuje chcąc zbudować nową wielowymiarową teorię świadomości. Jakiś system neuronów $[\Phi_j-\Phi_n]$ i leżący obok niego system $[\Phi_{j+l}-\Phi_{n+k}]$ będą bowiem wtedy korelatami świadomości.

7.4.2. Minimalny system neuronalny

Jakiś bliżej *określony* system jest przez nas traktowany jako wystarczający korelat świadomego stanu $[\Psi_i]$. Na przykład, wszystkie funkcjonalne badania odwzorowujące (PET, fMRI), podczas których obrazuje się neuronalny metabolizm związany z permanentnym bólem, wskazują na aktywność występującą w korze przedniej części zakrętu obręczy (*anterior cingulate cortex*, [por. Bromm 2001: 244, 247; Craig 2002: 661–663]).

VII.7 Stan *określonego* systemu $[\Phi_j-\Phi_n]$ jest *wystarczający* dla wystąpienia skorelowanego świadomego stanu. Stan jakiejś części tego systemu $[\Phi_j-\Phi_m, m < n]$ nie jest wystarczający dla wystąpienia wspomnianego stanu świadomości.

Takich określonych i wystarczających systemów (dla wystąpienia świadomego stanu $[\Psi_i]$) może być więcej niż jeden i mogą one być określonymi neuronalnymi korelatami tego stanu. Nie możemy *a priori* stwierdzić, że istnieje tylko jeden *określony* system będący korelatem stanu świadomego $[\Psi_i]$. Odpowiedź na pytanie, czy jest jeden taki system, czy też jest ich kilka, może być jedynie natury *aposteryorycznej* (eksperymentalno-empirycznej) i wynikać z kolejnych badań.

Można odróżnić dwa rodzaje neuronalnego korelatu świadomości dla stanu $[\Psi_i]$.

(i) Jakiś neuronalny korelat świadomości będzie *określonym* (tzn. najmniejszym) neuronalnym systemem $[\Phi_j-\Phi_n]_o$, tego rodzaju, że dany stan $[\Phi_j-\Phi_n]_o$ jest wystarczający dla bycia skorelowanym z odpowiednim świadomym stanem $[\Psi_i]$. Jeżeli system $[\Phi_j-\Phi_n]_o$ jest, powiedzmy, w stanie $[\Phi_j-\Phi_n]_o^1$, to powinien mu odpowiadać świadomy stan $[\Psi_i]^1$.

(ii) Ścisłe pojedyncze korelacje typu stan–stan zdają się jednak nie istnieć (a jeżeli istnieją, to trudno wyodrębnić i określić ich części neurofizjologicz-

ne). Korelacje muszą być raczej zdefiniowane w stosunku do szeregu przypadków, do jakiejś całości. Stąd nasze twierdzenie:

VII.7₁ Jeżeli system $[\Phi_j - \Phi_n]$ ma być wystarczający dla wystąpienia skorelowanego z nim świadomego stanu $[\Psi_i]$, to znaczy, że w szeregu przypadków neuronalnemu stanowi systemu $[\Phi_j - \Phi_n]$ będzie zawsze towarzyszył świadomy stan $[\Psi_i]$.

Oznacza to, że jest możliwe być w odpowiednim stanie $[\Phi_j - \Phi_n]_o$, a nie być w skorelowanym z nim stanie świadomości.

Weźmy pod uwagę należące do wzgórza jądra międzyblaszkowe, które zdaniem J. E. Boga są źródłem subiektywności (por. Bogen 1995: 57) [por. 6.1.4]. Jądra te należą do tak zwanych jąder nieswoistych (*non-specific nuclei*), będących końcowym miejscem zstępującego tworu siatkowatego (*formatio reticularis*)³⁰. Trudno przypuszczać, że gdy jakoś pobudzimy wejście do jąder międzyblaszkowych i wywołamy w ten sposób odpowiedni stan systemu jąder, będzie temu zawsze towarzyszył skorelowany stan świadomości. Wydaje się więc, że w mózgu podmiotu może występować stan neuronalnego korelatu świadomości bez odpowiedniego świadomego stanu podmiotu.

Neuronalny korelat świadomości można bowiem rozumieć o wiele szerzej, aniżeli się to przyjmuje we współczesnych hipotezach. Może być nim jakiś rodzaj rozbudowanego stanu mózgu, zawierającego (albo będącego zaangażowanym w) procesy z różnych obszarów mózgu. Nie takiego stanu oczekuje się jednak mówiąc o neuronalnym korelacie świadomości. Omawiany tutaj stan jest bowiem jakimś *całościowym* neuronalnym korelatem świadomości, który, jako stan, automatycznie jest wystarczający dla odpowiadających mu stanów świadomych. Pojawia się jeszcze inna trudność. Od wyników pomiaru globalnej aktywacji (powiedzmy, od elektroencefalograficznego pomiaru średniego napięcia, czy od pomiaru korowego przepływu krwi) nie można oczekiwać, że ukążą one wszystkie szczegóły mechaniczne lub funkcjonalne poszczególnych części mózgu. Wspomniane metody po-

³⁰ W przeciwieństwie do swoistych jąder wzgórza, jądra nieswoiste należą do tych, których drogi dośrodkowe biegną do tworu siatkowatego, prowadząc impulsy z większej liczby różnych narządów zmysłowych (por. Duus 1989: 241–242). Twór siatkowaty ze swoimi jądrami rozciąga się od rdzenia kręgowego, przez rdzeń przedłużony i most aż do przedniej części śródmózgowia. Jądra tworu siatkowatego w obrębie śródmózgowia mają projekcję do wyższych dogłowowo leżących ośrodków, przekazując impulsy do rozległych obszarów kory mózgowej, gdzie działają aktywująco. Wstępujący siatkowaty układ pobudzający zapewnia utrzymanie stanu świadomości, stanu uważnego czuwania, a także wpływa na rytm snu i czuwania (por. s. 182–186, 194–196).

miarowe nie potrafią zidentyfikować małych, ale znaczących populacji neuronów, jak na przykład obszaru w pniu mózgu obejmującego tzw. miejsce sinawe (*locus coeruleus*), albo grupy jąder rozrzuconych w centralnej szarej masie pnia mózgu (jądra szwu – *raphe nuclei*)³¹.

Nie chodzi nam jednak o jakiś *całościowy* neuronalny korelat świadomości, lecz o korelat minimalny (najmniejszy system). Możemy stwierdzić:

VII.7₂ *Minimalny* neuronalny korelat świadomości zawiera jedynie *określone* procesy skorelowane ze świadomością. Reszta *całościowego* neuronalnego korelatu świadomości zostaje przesunięta do tła, wymaganego do poprawnego funkcjonowania wspomnianego *określonego* procesu.

Przy twierdzeniach dotyczących tła, potrzebnego do poprawnego funkcjonowania określonego procesu, należy odróżnić:

- (a) podtrzymującą dynamikę biologiczną i neuronalną,
- (b) właściwą dynamikę neuronalną (dla danego stanu),
- (c) dynamikę stanu świadomości.

W jaki jednak sposób odróżnić ów określony neuronalny proces od tła, na którym on zachodzi? Wiele procesów neuronalnego przetwarzania zachodzących w tle jest nieświadomych. Chociażby wtedy, gdy bezowocnie myślimy nad jakimś problemem, nie znajdujemy rozwiązania, przestajemy się nad nim zastanawiać i nagle (przy całkiem innej okazji), rozwiązanie nam niejako „wpada” do głowy. Kolejnym przykładem może być np. efekt zwany *primingiem*, który jest traktowany jako zjawisko nieświadome. Identyfikacje pewnych słów zachodzą szybciej, jeżeli zostały one częściowo zaktywizowane przez *prime* słowa.

Być może rozróżnienie: zachodzący proces–tło, da się uzyskać dzięki założeniu, że dla neuronalnego korelatu świadomości, nie jest wymagane, żeby

³¹ Elementy te posiadają wejścia od systemu sensorycznego i są połączone (pośrednio lub bezpośrednio) z wszystkimi częściami centralnego układu nerwowego. Chociaż jądra miejsca sinawego stanowią ważne części pnia mózgu, są one małymi strukturami (o przekroju poprzecznym mniejszym niż 2 mm) i dlatego trudno je odróżnić za pomocą kamery PET. Być może w przyszłości tego typu odróżnień (np. *locus coeruleus*) będzie można dokonać metodą PET z jakimś identyfikującym jądrem znacznikiem. Jądra podkorowe (ciało migdałowe, *corpus amygdaloideum*) stanowią ważne części płatu skroniowego (*lobus temporalis*), niemniej jednak: „Separating the amygdala from adjoining brain is not a trivial task” (Gorman 2000: 498). Ciało migdałowe ma wprawdzie średnicę od 6 do 8 mm, ale z powodu bezpośredniej bliskości kory płata czołowego trudno powiedzieć, gdzie się ono zaczyna i gdzie kończy.

był on skorelowany ze świadomością w jakichś możliwych warunkach, lecz wystarczy, że koreluje on w jakimś szeregu przypadków, w których niektóre aspekty normalnie funkcjonującego mózgu są utrzymywane jako coś stałego.

Pozostaną jednak dwa pytania. (i) Co to znaczy, że owe aspekty są utrzymywane jako stałe? (ii) Jaki szereg przypadków jest wymagany, aby powiedzieć, że dany neuronalny korelat świadomości współoddziaływuje ze świadomością?

Nie należy z tego wyciągać wniosku, jak to proponuje D. Chalmers, że neuronalny korelat świadomości jest korelatem jedynie w przypadkach normalnie funkcjonujących mózgów. Domaganie się korelacji tylko w przypadkach zawierających normalnie funkcjonujący mózg w normalnych warunkach (otrzymujący bodźce podobne do tych, jakie otrzymuje w „codziennie spotykanych warunkach”) byłoby ograniczoną analizą, która miałaby sens, jeśli dysponowalibyśmy jasnym kryterium normalności. Ponadto to, co jest normalne dla stanu świadomości, może być patologią (wodogłowie, małe rozmiary głowy) neurologiczną.

Tego rodzaju „normalne” kryterium wydaje się zbyt słabe, aby za jego pomocą wyodrębnić neuronalny korelat świadomości. Może się bowiem okazać, że w ten sposób znajdziemy neuronalne korelaty świadomości na wszystkich stopniach poszczególnych układów (słuchowego, węchowego, wzrokowego, ...). Traktując system słuchowy jako część mózgu możemy powiedzieć tak: w normalnych warunkach neuronalne treści obecne w słuchowych układach od ciał kołankowatych przyśrodkowych poprzez kolejne warstwy kory słuchowej (A1–A6) będą wszystkie korelowały z treściami świadomości słuchowej (tzn. z akustycznymi charakterystykami fali dźwiękowej)³². Treści obecne w uchu wewnętrznym³³, będą do pewnego zakresu korelowały z treściami świadomości słuchowej. Można sobie bowiem wyobrazić eksperyment myślowy, że zastępując właśnie wspomniane części właściwymi „proteżami”, spowodujemy pojawienie się treści świadomości słuchowej.

³² Nie wchodzę tutaj w złożony problem zbudowania właściwej teorii słyszenia, która zdaniem A. Klawitera, powinna ustalić „[...] jak to się dzieje, że system słuchowy potrafi odekodować z materiału dźwiękowego informację o cechach przedmiotów [a nie tylko o akustycznej charakterystyce fali dźwiękowej – J.B.] znajdujących się w otoczeniu” (Klawiter 1999a: 334).

³³ Przemiana fal mechanicznych w elektryczny potencjał czynnościowy zachodzi w wyspecjalizowanych komórkach zmysłowych zwanych narządem spiralnym (*organum spirale cochleae*). W jego wnętrzu znajduje się zwój spiralny, którego ośrodkowe wypustki tworzą nerw ślimakowy wchodzący do pnia mózgu. W pierwotnym i wtórnym polu kory słuchowej impulsy słuchowe podlegają analizie, identyfikacji oraz porównaniu z wcześniejszymi doświadczeniami słuchowymi (por. Duus 1989: 158–161, 335).

Należy zauważyć, że w normalnych przypadkach cały system słuchowy (kora słuchowa, ucho wewnętrzne, ...) tworzy łańcuch przyczynowy, a rozważane części tego systemu nie są rozdzielone. Błędne jednak byłoby twierdzenie, że tylko na podstawie przyczynowej jedności systemu słuchowego wszystkie jego podsystemy – także ucho zewnętrzne i środkowe – mają być traktowane jako neuronalny korelat świadomości. W tym kontekście należy chociażby wspomnieć, że choroby ucha środkowego są domeną laryngologa, objawy w następstwie uszkodzeń nerwu ślimakowego i jego ośrodkowych połączeń zaś – neurologa.

Przedstawiając wyniki badań neuronalnych korelatów świadomości wzrokowej G. Rees stawia hipotezę, że neurony w pierwszej warstwie kory wzrokowej (V1), „[...] chociaż są konieczne do powstania normalnego wizualnego przeżycia, nie dostarczają podstawy, z siebie samych, do świadomego wizualnego spostrzegania – innymi słowy, aktywność w komórkach V1 nie wystarczy dla wyjaśnienia świadomego spostrzegania” (Rees 2002: 263)³⁴.

Potrzebujemy precyzyjniejszego kryterium dla wyodrębniania podsystemów i oddzielania określonego neuronalnego korelatu świadomości od procesów, które są z nim jedynie przyczynowo połączone. W celu otrzymania szukanego podziału należy rozpatrywać korelacje zarówno w przypadkach normalnie funkcjonujących mózgów, jak i w wielu przypadkach nienormalnych (na przykład z użyciem odpowiednich „protez słuchowych” albo przypadki uszkodzeń poszczególnych części układu słuchowego).

W ustaleniu programu badań może pomóc odwoływanie się do tworzonych do tego celu eksperymentów myślowych. Można sobie przykładowo wyobrazić, że uda się: (i) skonstruować krzemowe „neurony”, które będą – jak tego chce funkcjonalizm – funkcjonalnie identyczne z biologicznymi neuronami, (ii) dokładnie określić struktury połączeń w indywidualnym mózgu. Można sobie wyobrazić skonstruowanie krzemowego mózgu, *funkcjonalnie* identycznego z biologicznym³⁵. Nadal jednak pozostanie pytanie,

³⁴ Współczesne naukowe poszukiwania neuronalnych korelatów świadomości skupiają się w dużej mierze na badaniach systemu wzrokowego, por. Rees 2002: 262–269.

³⁵ Należy tutaj oczywiście pamiętać o wspomnianym już rozróżnieniu wprowadzonym przez K. Twardowskiego: (i) „funkcja” w znaczeniu matematycznym i (ii) „funkcja, jako „czynność, którą wykonuje osoba lub rzecz”. „Otóż czynność umysłowa jest niezamodną funkcją mózgu w pierwszym tego znaczeniu, albowiem pewne zmiany, zachodzące w mózgu, pociągają za sobą zmiany w czynności umysłowej. Nie można atoli nazwać czynności umysłowej funkcją mózgu w drugim z przytoczonych znaczeń. Nie ma bowiem wcale na to dowodów, że czynność umysłową wykonuje w zupełności i wyłącznie mózg; wiele zaś okoliczności, jak np. jedność i ciągłość życia umysłowego, przemawia raczej przeciwko takiemu przypuszczeniu” (Twardowski 1997: 91).

czy taki mózg pozwoliłby znaleźć wspomniane wyżej podsystemy i określone korelaty świadomości zjawiskowej.

7.5. Przykłady rodzajów korelatów świadomości

Czy jest jakaś pewność, że zdefiniowane jak w 7.1 {I}, {II} neuronalne korelaty świadomości w ogóle istnieją? Odpowiadając na to pytanie przyjmujemy: (A) że stany świadomości w jakiś sposób zależą od całościowego stanu mózgu. Przyjęcie (A) oznacza jedynie, że istnieje przynajmniej jakaś korelacja neuronalnych stanów ze świadomością.

7.5.1. Neuronalne korelaty świadomych przeżyć

Przy określaniu neuronalnych korelatów treści świadomości neuronalny korelat musi nie tylko odpowiadać stanowi świadomości, ale także oddawać jego treść. Mamy wtedy do czynienia z wykluczeniem przypadku, że mózg jest nieminimalnym neuronalnym korelatem treści świadomości. Jest tak dlatego, gdyż przedstawianie jakiejś treści w mózgu nie wystarczy, aby przedstawiać tę treść w świadomości. Jak wspomniano, wiele treści przedstawianych w mózgu to treści nieświadome (por. 7.4.2).

Może się okazać, że istnieją bardziej restryktywne neuronalne systemy, których treść systematycznie oddaje treści niektórych własności świadomości. Nie jest jednak oczywiste, że takie systemy muszą istnieć. Można na przykład utrzymywać, że treści świadomości są emergentnym wytworem powstałym z treści obecnych w różnych neuronalnych systemach, które razem wzięte wystarczą dla pojawienia się świadomej treści, o którą nam chodzi, ale żaden z nich z osobna nie odwzorowuje owej świadomej treści.

U podmiotów używających języka, świadome treści korespondują z treściami bezpośrednio dostępnymi dla słownych sprawozdań. Używając wprowadzonego przez N. Blocka rozróżnienia na „świadomość dostępu” i „świadomość zjawiskową” (por. 4.3.3) stwierdzamy, że może istnieć (w jakiejś ograniczonej formie) świadomość dostępu bez świadomości zjawiskowej. Można też przyjąć, według Blocka, że świadomość zjawiskowa jest częścią świadomości dostępu. Block zaznacza jednak:

„Być może stopień świadomości zjawiskowej potrafi być zawsze określony na podstawie stopnia świadomości dostępu. Chociaż świadomość zjawiskowa i dostępu odróżniają się pojęciowo, (tak samo jak pojęcia «wody» i «H₂O»), do tej pory nie wiemy, czy w mózgu zbiegają się one w jedno, czy też nie” (Block 1988: 158)³⁶.

³⁶ Odpowiadając na pytanie, czy neuronalny korelat świadomości zjawiskowej jest

Bardziej ogólnie można powiedzieć, że u świadomych podmiotów treści świadomości korespondują z treściami bezpośrednio dostępnymi dla globalnej wolitywnej kontroli zachowań. Przyjmując, że treści te dochodzą do głosu w werbalnych sprawozdaniach i ukazują się w kontroli zachowań, mamy tym samym podstawę do przekonania, że są one przedstawiane w jakimś miejscu wewnątrz kognitywnego systemu.

Zależy to oczywiście od rozumienia terminu „przedstawienie”.

(i) Jeżeli założymy, że jedynym przedstawieniem jest przedstawienie symboliczne, wtedy trudno powiedzieć, że treść ukazana w zachowaniu musi być przedstawiana.

(ii) Jeżeli systemem jest przypisywana przedstawiana treść zgodnie z ich funkcjonalną rolą (tzn. gdy stawiamy mniejsze wymagania co do „przedstawiania”), to czymś naturalnym będzie oczekiwanie, że treść ukazująca się w funkcjonalnej roli będzie przedstawiana w systemie, który odgrywa tę funkcjonalną rolę (por. wyżej, 3.5.1; 3.5.2). Trudno jednak o pewność, że znajdziemy jakiś pojedynczy neuronalny system, którego treść zawsze odpowiada treści świadomości. Może się okazać, że funkcjonalna rola jest odgrywana przez różnorakie systemy, i że dany system czasami może grać tę rolę, czasami zaś nie (por. funkcjonalistyczną tezę o wielorakiej realizacji).

Jeżeli jednak nie ma przestrzennie lokalizowalnego neuronalnego systemu, kwalifikującego się jako korelat świadomej treści, to możemy poszukiwać bardziej globalnego, funkcjonalnego systemu twierdząc, że treści świadomości korespondują z treściami dostępnymi w tym systemie. Jako przykład można podać prowadzoną przez J. Aronsona i A. Damasio krytykę Dennettowskiego rozumienia Kartezjańskiego Teatru (4.2.1). Z krytyki tej wynika, że neuronalny korelat świadomości nie musi być specyficznym, anatomicznie rozumianym obszarem w mózgu.

W przedstawionym właśnie sensie neuronalne korelaty świadomości są identyfikowane raczej poprzez cechy funkcjonalne albo fizjologiczne aniżeli poprzez cechy przestrzenno-anatomiczne. Tym samym jakiś „neuronalny system przedstawiający” też może być identyfikowany bardziej wieloznacznie: odpowiednimi neuronalnymi korelatami przedstawiającymi treści mogą być treści przedstawiane albo przez czasowo wydłużone odpalenie neuronów, al-

dokładnie tym samym korelatem, jak dla świadomości dostępu, N. Block stwierdza w końcu (krytykując wyniki uzyskane przez F. Cricka i Ch. Kocha nad śpiącymi i czuwającymi kotami): „Zweifellos gibt es zwischen den schlafenden und den wachen Katzen viele Unterschiede in den Bereichen außerhalb des visuellen Cortex. Aber man müßte schon sehr gute Gründe nennen, um irgendeine dieser Unterschiede zum neuronalen Korrelat der visuellen Phänomenologie zu zählen, und nicht zum Korrelat der, sagen wir, nicht-phänomenalen kognitiven Verarbeitung visueller Information” (s. 164).

bo przez specjalne rodzaje komórek. Tak rozumiana wieloznaczność i związany z nią brak dokładnej przestrzennej lokalizacji neuronalnego systemu nie stanowi przeszkody w zakwalifikowaniu tego systemu jako neuronalnego korelatu świadomości. Istnieje zatem cały niejednorodny szereg możliwych korelatów świadomych stanów: od specyficznie anatomicznych obszarów w mózgu po abstrakcyjnie wydzielone neuronalne systemy. Można tym samym powiedzieć, że:

VII.8₁ Specyfikacja neuronalnego korelatu świadomości okazuje się sprawą empiryczną. Można oczekiwać, że będzie to jakaś określona systemowa specyfikacja.

Twierdzenie **VII.8₁** dotyczy całej współczesnej neurofizjopsychologii, w której panuje prymat struktury nad dynamiką, co wynika z obecnie dostępnych metod badań. Przez omawianych w tej pracy neurologów nie jest brany pod uwagę fakt, że forma żywa buduje swoją anatomię, że z niej korzysta, że ją modyfikuje i nieustannie odmładza.

Odpowiednio do **VII.8₁** można powiedzieć, że u danego rodzaju organizmów istnieje ściśle połączenie między charakterystycznymi funkcjami i charakterystycznymi anatomiczno-fizjologicznymi systemami. Można tym samym oczekiwać specyficznych neuronalnych korelatów właśnie dla świadomych treści. Jeżeli nie, to musimy się zadowolić bardziej ogólnymi korelatami, możliwymi do przynajmniej częściowego indywidualizowania na poziomie kognitywnym. Ponadto możemy przyjąć, że połączenie pomiędzy treścią neuronalną a treścią świadomą nie domaga się czasoprzestrzennej odpowiedniości. Neuronalne przedstawienie kształtu nie musi oczywiście samo mieć „kształtu” (por. Sellars 1967: 340–347). Tworzone przez podmioty wizualne i akustyczne „sceny” nie są nigdy jednoznaczne, lecz posiadają całą gamę możliwości interpretacyjnych, począwszy od rozpoznania kształtów, barw, rozpoznawania wzajemnego położenia obiektów. Tej wielości danych sensorycznych rzadko kiedy podmiot jest świadomy, a nawet gdy w spostrzeganiu jest jej jakoś „abstrakcyjnie” świadomy, to jej nie spostrzega.

Można przypuszczać, że w mózgu nie ma obszaru specyficznie poświęconego holistycznej świadomości w sensie *jak to jest* być świadomym, jako przeciwstawionej temu, co nazywamy świadomym widzeniem, słyszeniem, pamięcią, świadomymi zachowaniami itd. Z założenia, że tego rodzaju anatomiczny obszar nie istnieje, nie wynika, że tym samym nie istnieje neuronalny korelat świadomości. Można wprowadzić następujące dookreślenie:

VII.8₂ Mówiąc o neuronalnym korelatcie świadomości zjawiskowej stwierdzamy jedynie, że jakiś system jest skorelowany ze świadomością, nie zaś, że system ten jest całkowicie przypisany świadomości.

Tak rozumiane ujęcie neuronalnego korelatu świadomości jest szersze aniżeli ujęcie używane w wielu pracach empirycznych dotyczących tego tematu. Często jest w nich przyjmowane, że neuronalny korelat świadomości może być ściśle związany z wizualnymi procesami, z pamięcią czy z innymi procesami. Dzieje się tak szczególnie wtedy, gdy zrezygnujemy z założenia, że istnieje pojedynczy neuronalny korelat świadomości, i zaakceptujemy tezę, że takich korelatów może być wiele i to we wszystkich modalnościach (słuchowej, wzrokowej, ...). Założenie takie prowadzi pośrednio do rezygnacji z holistycznego traktowania świadomości na rzecz jej „atomistycznego” rozumienia (por. 1.3.3).

7.5.2. Neuronalne korelaty jedności świadomości

„Praca mózgu to ustawicznie zmieniające się układy chwilowo pobudzonych grup neuronów, powodujące równie szybkie, nieliniowe zmiany efektywności odpowiednich synaps” (Wróbel 2000a: 477).

Wyniki uzyskiwane w licznych badaniach neurologicznych stawiają pod znakiem zapytania tradycyjne rozumienie jedności świadomości. Znane opisy przypadków rozdwojonych mózgów (*split brain*), rozszczepienia osobowości (*multiple personality disorder*, MPD) zdają się negować ideę świadomości cechującej się jednością. Podobnie negujące nastawienie obserwujemy u niektórych psychologów pytających o integrację różnych modalności sensorycznych (tak zwany problem scalania, por. 6.3.3). Skupiając się na systemie wzrokowym i na spostrzeganiu jakiegoś przedmiotu można problem scalania wyrazić pytaniem: jak to się dzieje, że takie informacje, jak kształt, kolor, dźwięk, ruch, powodują powstanie pojedynczego, świadomego urzeczywistnienia przedmiotu? Albo inaczej: w jaki sposób własności, pojawiające się w różnych, anatomicznie oddzielonych obszarach mózgu, łączą się razem w jedno koherentne przeżycie? Jak mózg je łączy razem?

Zakłada się przy tym, że neuronów nie można traktować jako małych, panpsychicznie rozumianych „generatorów przeżyć”. Nawet gdyby każdy neuron był traktowany jako wytwarzający jakąś część świadomego przeżycia, dalej problemem pozostałoby łączenie tych cząstkowych przeżyć.

Odpowiadając na ten problem możemy się odwołać do wspomnianej już hipotezy F. Cricka i Ch. Kocha, mówiącej o odpowiednio skorelowanej czę-

stotliwości oscylacji grup neuronów w korze mózgowej. Cechy spostrzeganego wzrokowo przedmiotu (powiedzmy jego kolor i kształt) są przetwarzane w różnych częściach mózgu, a następnie łączone razem w jedną spostrzeganą całość dzięki zsynchronizowanemu „odpalaniu” neuronów³⁷.

Przy wzrokowym spostrzeganiu obiektów aktywność grup komórek nerwowych rozsianych po różnych częściach mózgu jest synchronizowana co do milisekund. Gdy ukazujące się podmiotowi zarysy stanowią części tego samego przedmiotu, to komórki nerwowe reagujące na różne jego kontury „odpalają” z równą częstotliwością. Odpowiednie miejsca w mózgu muszą otrzymać sygnały, które z wielu zaktywowanych komórek nerwowych kodują razem jeden przedmiot. Synchronizacja neuronalnych sygnałów spełnia rolę łączącą pomiędzy owymi zaktywowanymi komórkami³⁸. Synchroniczne „odpalanie” neuronów umożliwia powstanie jednolitego spostrzegania przedmiotu, które to spostrzeganie pochodzi z aktywności różnych podsystemów mózgu. Zsynchronizowane pod względem częstotliwości „odpalanie” grup neuronów ma tym samym umożliwić powstanie jedności pomiędzy różnymi cechami spostrzegania wzrokowego. Tego rodzaju synchronizację możemy traktować jako potencjalny neuronalny korelat wzrokowej świadomości. Na podstawie przeprowadzonych przez Cricka i Kocha badań można przyjąć:

VII.9 Potencjalnym neuronalnym korelatem świadomości wzrokowej może być grupa neuronów „odpalająca” z określoną częstotliwością pomiędzy wzgórzem a warstwami kory mózgowej IV i VI.

Mówimy „potencjalnym”, gdyż hipoteza 40 Hz jest stosunkowo niepewna i do tej pory nie została jednoznacznie empirycznie potwierdzona³⁹. Neu-

³⁷ Weźmy pod uwagę dwa postrzegane przedmioty. „Die Schlüsselidee herbei ist das korrelierte Feuern. Zugrunde liegt der Gedanke, daß es nicht einfach auf die Durchschnittsfeuertate eines Neurons ankommt, sondern auf die exakten Zeitpunkte, zu denen jedes Neuron feuert. Die Neuronen, die mit den Eigenschaften des ersten Objekts zusammenhängen, werden allesamt in irgendeinem Muster zum selben Zeitpunkt feuern. Die Neuronen, die mit dem zweiten Objekt zusammenhängen, werden ebenfalls alle zusammen feuern, aber zu anderen Zeitpunkten als die erste Neuronengruppe” (Crick i Koch 1994: 260; por. s. 299; oraz Koch 2004: 55).

³⁸ Widać tutaj wyraźną różnicę względem koncepcji prezentowanej przez J. Ecclesa, w której należąca do świata, jaźń wybiera i sortuje potrzebne informacje, dostarczane jej przez korę mózgową (por. Eccles 1985: 7–9).

³⁹ Zdaniem A. Wróbla, „[...] jedni badacze uważają oscylacje za uboczny produkt pobudzenia mózgu, podczas gdy inni przypisują im rolę nośnika w procesie synchronizacji pobudzenia, zachodzącego między komórkami położonymi w dużej odległości od siebie” (Wróbel 2000a: 472).

rony mogą bowiem również „odpalać” wtedy, gdy nie wnoszą nic do stanu świadomości. Pytaniem pozostaje także, czy „nieodpalające” neurony wnoszą coś do powstawania świadomości.

Hipoteza Cricka i Kocha jest przykładem, w jaki sposób neurologia albo neuroradiologia może pomóc w znalezieniu neuronalnych korelatów świadomości. Uwzględniając koneksjonizm mózgu i holistyczne rozumienie świadomości stwierdzamy, że synchronizacja jest istotna zarówno z neurologicznego, jak i z filozoficznego punktu widzenia, a samą synchronizację możemy rozumieć szerzej jako współoddziaływanie.

Crick i Koch nie odwołują się do świadomych stanów mentalnych. Wspomniany proces łączenia tego, co spostrzegane, wydarza się niezależnie od naszej świadomości. Nieświadome spostrzeżenia wpływają na zachowania podmiotu dzięki swoim cechom, które zostają powiązane, podobnie jak w świadomym spostrzeganiu. Aby przedstawiony w świadomości obraz powodował zamierzony efekt – na przykład w słownej wypowiedzi – podmiot musi postrzegać cechy tego obrazu jako cechującą się jednością.

Spostrzeżenia, jak rozumiemy Cricka i Kocha, są łączone niezależnie od tego, czy są świadome lub nie. Proces łączenia spostrzeżeń wydarza się niezależnie od świadomości, a tym samym teoria świadomości nie musi go wyjaśniać. Sam przebieg neuronalnego przetwarzania informacji nie jest dla nas świadomy, lecz świadome są tylko określone „stany końcowe” systemu przetwarzającego informacje.

Załóżmy, że hipoteza Cricka i Kocha jest prawdziwa i że odkryto owe korelaty, powiedzmy, poprzez stymulację połączeń nerwowych. Na razie pod znakiem zapytania stoi, czy umożliwi ona wyjaśnienie świadomości. Potwierdzając tę hipotezę można powiedzieć, że stan neuronalnego systemu $[\Phi_j - \Phi_n]$ jest stanem przeżyciowo świadomym $[\Psi_i]$ albo świadoma jest jakaś kombinacja $[\Phi_j - \Phi_n][\Psi_i]$ tych stanów. Można przestudiować fizyczne cechy stanu $[\Phi_j - \Phi_n][\Psi_i]$, badając każdy pojedynczy akson, dendryt i synapsę. Hipoteza 40 Hz mało jednak pomoże przy próbie wyjaśnienia albo opisanie zjawiskowych cech $[\Phi_j - \Phi_n][\Psi_i]$. Nadal bowiem pojawia się „trudny” problem połączenia obiektywnych, fizycznych cech świadomości z jej subiektywnymi, zjawiskowymi cechami.

Kolejnymi pytaniami pozostaną: czy wspomniane zsynchronizowanie umożliwia także wybór tych informacji, które staną się świadome? czy neurony reagujące na świadomie spostrzegany obraz odpalają – powiedzmy – silniej i są bardziej zsynchronizowane w stosunku do innych odpalających neuronów? czy neurony z danej grupy tylko wtedy wykazują wzmożoną synchroniczną aktywność, gdy reagują na ten sam przedmiot?

Interesujące oczywiście jest znalezienie wyjaśnienia świadomości, zapewniające subiektywne ukazywanie się jej jedności, uwzględniające ewentualny (faktyczny albo potencjalny) brak jedności procesów zachodzących w mózgu, a skorelowanych z subiektywnością świadomości. Stawiając zagadnienie w ten sposób należy zapytać: które cechy jedności świadomości mogą być wyjaśnione przez odwołanie się do nieświadomych, neurofizykalnych albo psychologicznych procesów? Jedność świadomości może obejmować różne zjawiska, nas interesuje przede wszystkim jedność przeżyciowa, leżącą u podstaw subiektywnej świadomości.

7.6. Neurokognitywna prototeoria świadomości

Odkrycie stanów neuronalnych skorelowanych ze stanami świadomymi nie prowadzi bezpośrednio do wyjaśnienia wszystkich relacji pomiędzy procesami w mózgu a stanami subiektywnej świadomości ani nie odpowiada na pytanie: dlaczego i w jaki sposób mózg generuje świadomość? Może się okazać, że korelaty są generatorami lub/i realizatorami świadomości. Słaba korelacja – jak wspomnieliśmy wcześniej – może oznaczać jedynie paralelizm (współwystępowanie zjawisk fizycznych i psychicznych) lub synchroniczną współzmiennność. Przy takiej interpretacji korelacji poszukiwane korelaty mogą istnieć niezależnie od siebie. Kolejna interpretacja korelacji polega na tym, że bez neuronalnego korelatu nie ma świadomości: jakieś $[\Phi_j]$ jest korelatem $[\Psi_j]$, $[\Phi_j]$ jest warunkiem koniecznym $[\Psi_j]$.

Silniejszą relacją jest realizacja zawierająca aspekt modalny. „Realizator” (*realizer*) jest przez nas traktowany jako: (i) termin zawierający w sobie odniesienie do warunku wystarczającego, występującego w podanych wyżej definicjach korelatu (por. 7.1), a równocześnie (ii) jako termin odnoszący się do teorii superwencji mereologicznej i do zawartej w niej koncepcji wyjaśniania.

Urzeczywistniona własność (*realized property*) superwenuje na realizatorze, a obecność realizatora odpowiednio wyjaśnia występowanie urzeczywistnionej własności. Posiadanie realizatora jest jednym ze sposobów, w jaki dany przedmiot może posiadać urzeczywistnioną własność⁴⁰. W tym sensie rozumiany realizator jest warunkiem wystarczającym do pojawienia się superwenientnej własności.

⁴⁰ Por. LePore, Loewer 1987: 633–640. Odwołując się do analogicznie rozumianego przykładu można powiedzieć, że bycie barbituratem jest realizatorem bycia środkiem nasennym.

Przeżyciowe stany świadomości mogą posiadać kilka neuronalnych korelatów, chociaż większość z nich może nie być dla tych stanów w ogóle relewantna. Problemem jest, w jaki sposób zidentyfikować przyczynowe i treściowe (jakościowe) czynniki, istotne dla wyjaśnienia świadomości zjawiskowej. Ocenę relewantności albo nierelewantności uzyskanych danych empiryczno-eksperymentalnych można uzyskać jedynie w ramach poszukiwanej teorii świadomości. Samo znalezienie – nawet dużych ilości – jedynie paralelnych korelacji nie wystarcza do takiej oceny.

Istnieje potrzeba zbudowania teorii relacji pomiędzy stanami mózgu i stanami świadomości zjawiskowej, która umożliwiłaby uporządkowanie empirycznie stwierdzalnych zależności, a tym samym określenie, które z nich byłyby istotne dla badań świadomości zjawiskowej, które zaś nie.

Pod terminem „teoria naukowa” rozumiemy tutaj – ogólnie – zespół praw naukowych, definicji, twierdzeń oraz hipotez, dotyczących stanów mózgu i stanów świadomości. Zespół ten jest ze sobą powiązany i tworzy logicznie uporządkowaną i spójną całość. Teoria tym różni się od zbioru zdań (na przykład od opisu), że pomiędzy jej zdaniem zachodzą relacje wynikania logicznego (inferencji). Zdania teorii są uporządkowane przez wynikanie logiczne. Takie rozumienie teorii jest oczywiście logiczną idealizacją teorii empirycznych, swoistym ideałem, do którego trzeba dążyć. W tzw. naukach szczególnych teorii rzadko mają taką postać.

Zacznijmy od świadomości dostępu i wyróżnijmy dwa typy teorii związanych z dwoma rodzajami wyjaśniania. Można powiedzieć, że podmiot *X* odczuwa psychiczny dyskomfort, ponieważ dowiedział się, że ma raka. Wiedza *X-a* o raku jest *przyczynowo* relewantnym czynnikiem wyjaśniającym stan psychiczny *X-a*, do którego to stanu ma on pierwszoosobowy dostęp. Przejawy tego stanu są zazwyczaj opisywane pierwszo- i trzecioosobowo. Nazwijmy teorię, w której uzyskujemy tak rozumiane wyjaśnienie makroteorią (**MkT**). Jest ona zbudowana w ramach psychologii potocznej (**PP**) i używa jej języka. Można także powiedzieć, że *X* odczuwa dyskomfort, ponieważ zmniejszył się poziom noradrenaliny w jego mózgu. Brak wspomnianego neurotransmitera jest *przyczynowo* relewantnym czynnikiem przy wyjaśnianiu dyskomfortu *X-a*. Tego rodzaju wyjaśnienie jest podawane w ramach mikroteorii, którą nazwiemy (**MiT**). Jest ona zbudowana w ramach szeroko rozumianej neuronauki i używa jej trzecioosobowego języka⁴¹. Przedstawiane w ramach (**PP**) wyjaśnienia obejmują stany i procesy, których

⁴¹ Neuronauka (*neuroscience*), jest nową, rozwijającą się dziedziną nauki, która, w odróżnieniu od klasycznej neurobiologii, większy nacisk kładzie na poznanie tych funkcji mózgu, które leżą u podstaw naszego poznania, chorób psychicznych itp.

dotychczas używane teorie (MiT) nie muszą (w zasadzie) uwzględniać, gdyż dotyczą one innych „obiektów” aniżeli te, których dotyczy (PP)⁴². Proponowane w ramach współczesnych teorii neurokognitywnych (MiT) wyjaśnienia – jeżeli takie są podawane – dotyczą głównie sposobów zachowania się podmiotu, a nie stanów jego świadomości zjawiskowej. Bodźce, na które mózg reaguje i które przesyła, opisuje się w (MiT) w terminach neuroelektrycznych.

Należy zauważyć, że teorie MkT i MiT są rozumiane jako teorie empiryczne, tzn. ich twierdzenia uzasadnia się na podstawie doświadczenia, czyli tzw. zdań obserwacyjnych. Z jednej strony zdania obserwacyjne są zapisem bezpośredniego doświadczenia, z drugiej strony zdania obserwacyjne wynikają logicznie z adekwatnych zdań teoretycznych (zdania teoretyczne wyjaśniają fakty opisywane w zdaniach obserwacyjnych). Dany zbiór zdań obserwacyjnych jest wyjaśniany przez jedno lub kilka zdań teoretycznych. Teorie MkT i MiT dostarczają zdań teoretycznych, z których wynikają zdania obserwacyjne dotyczące subiektywnej świadomości. W tym sensie: świadomość jest przedmiotem teorii, czyli czymś, czego nie postrzegamy zmysłami, ale równocześnie czymś empirycznym, czego „doświadczamy” pierwszoosobowo.

Powstaje pytanie: Jaka zależność zachodzi pomiędzy teoriami MkT i MiT? Która z nich rzeczywiście albo bardziej adekwatnie wyjaśnia przyczynę dyskomfortu *X-a*? Czy jest bardziej prawdopodobne, że *X* będzie odczuwał dyskomfort, gdy się dowie, że ma raka (nazwijmy to „wyjaśnieniem odgórnym”), czy też, gdy się dowie, że ma raka i że jego poziom noradrenalin się obniżył? Czy też jest prawdopodobne, że *X* będzie apatyczny, jeżeli jego poziom noradrenalin się obniżył, chociaż on sam jako podmiot nic o tym nie wie (postępowanie takie możemy nazwać „wyjaśnieniem oddolnym”)? Z większym prawdopodobieństwem można powiedzieć, że *X* będzie odczuwał dyskomfort, jeżeli jego poziom noradrenalin się obniży, aniżeli można powiedzieć, że będzie on odczuwał dyskomfort, gdy się dowie o poważnym stanie swojego zdrowia. Dlatego chcąc adekwatnie wyjaśnić zmianę nastroju *X-a* spoglądamy w kierunku neuronauk.

Zarówno wyjaśnianie w teorii MiT może być i jest często uzupełniane o dane pochodzące z teorii MkT, jak i wyjaśnianie w teorii MkT może zostać uzupełnione o dane pochodzące z MiT. Tego typu „sprzężenie zwrotne” obydwu wyjaśnień obserwujemy chociażby w praktyce lekarskiej.

⁴² Dziedziny teorii MkT i MiT (czyli to, do czego dane teorie się odnoszą) różnią się od siebie i są zarazem ze sobą związane (stanowią one różne poziomy organizacji tego samego indywiduum). Stąd właśnie pojawiają się problemy dotyczące rodzaju relacji zachodzących pomiędzy tymi poziomami.

Przykład z psychicznym stanem *X-a* odnosi się do zależności rozumianych jako przyczynowe i przypomina rozumowanie w naukach ścisłych. Gdy natomiast w grę wchodzi nie tylko czynniki przyczynowe, lecz także przeżyciowe (jakościowe), to rozumowanie będzie jedynie analogiczne do powyższego. Neuronalne korelaty świadomości zjawiskowej różnią się bowiem od procesów, które są z nią jedynie związane przyczynowo. Neuronalne korelaty świadomości (w przypadku treści świadomości) są określonymi neuronalnymi systemami, których stany są stanami przedstawiającymi (reprezentującymi)⁴³. Tym samym mówimy o reprezentacji zjawiskowej (fenomenalnej) i o reprezentacji neuronalnej. Poszukiwana nowa teoria świadomości **MkiT** powstająca na podstawie danych dotyczących znalezionych neuronalnych korelatów świadomości musi:

(a) wyjaśnić, w jaki sposób liczne komponenty procesów neurologicznych, o których mówi dzisiejsza **MiT**, mogą być zintegrowane oraz w jaki sposób – z tak porozdzielanych systemów aktywności neuronalnej – może powstać wielkoskalowa jedność, o której mówi **MkT**,

(b) wyspecyfikować mechanizmy odpowiedzialne za dynamiczną selekcję neuronalnych podsystemów, ponieważ tylko część docierającej do organizmu (systemu poznawczego) informacji uzyskuje dostęp do jego świadomości.

Ponieważ teorie **MkT** i **MiT** mogą – przynajmniej na razie – zachować swoje metody badawcze, więc nie mamy tutaj do czynienia z redukcją **MkT** do **MiT**.

W następnym punkcie nie podajemy wypracowanej, naukowej teorii **MkiT** odpowiadającej na pytania (a) i (b), lecz jedynie wstępny etap jej tworzenia, polegający na ustaleniu rodzaju zależności pomiędzy teoriami **MkT** i **MiT**, zachodzącej w ramach poszukiwanej **MkiT**. W tym sensie mówimy o „prototeorii” świadomości. Odwołując się do przedstawionych w rozdziale III wyników badań nad teoriami świadomego umysłu przyjmujemy, że na dzień dzisiejszy najbardziej odpowiednim rodzajem poszukiwanej zależności jest jakiś rodzaj emergencji synchronicznej.

⁴³ Jak już wspomnieliśmy wyżej, komórki w nerkach i żołądku pełnią określoną rolę i nic ponadto. Natomiast neurony – zdaniem Damasio – na każdym poziomie układu nerwowego odzwierciedlają zjawiska występujące gdziekolwiek w organizmie, a pośrednio wszystko, z czym ten ma do czynienia. „Komórki mózgowe są urodzonymi kartografami samego organizmu i wydarzeń w nim zachodzących. [Mechanizmy w mózgu] nie tylko regulują całe nasze życie, ale także odwzorowują ciągle zmieniające się stany organizmu. Mózg ma zatem naturalne sposoby odzwierciedlania budowy i stanu całego ciała” (Damasio 2003: 8). Por. też Revonsuo 2000: 67–69.

O emergencji mówimy w dwóch, powiązanych ze sobą – gdy chodzi o wyjaśnianie – sensach: ontologicznym i metodologicznym⁴⁴. W niniejszym punkcie interesuje nas zasadniczo sens metodologiczny. Teorie emergencji są jedynie surogatami teorii *stricte* naukowych. Oczekujemy, że poszukiwana teoria **MkIT** będzie teorią naukową⁴⁵.

7.6.1. Redukcja bez identyfikacji: teoria emergencji jako surogat teorii naukowej

Na podstawie przeprowadzonych analiz dotyczących badań neuronalnych korelatów świadomości można zbudować prototeorię świadomości uwzględniającą jej subiektywność oraz odpowiadające tej subiektywności procesy neurologiczne (por. 7, [C])⁴⁶.

Powiedzmy, że **MkT** jest filozoficzno-psychologiczną teorią, zbudowaną w ramach potocznej psychologii **PP** i uwzględniającą holistyczne rozumienie świadomości zjawiskowej oraz jej stanów $[\Psi_i - \Psi_m]$. Modelem tak rozumianej świadomości zjawiskowej w teorii **MkT** są przeżycia senne w fazie snu REM (por. 2.3). Zarówno przeżycia senne, jak i przeżycia na jawie mają odzwierciedlenie w mózgu, w którym zachodzą neurologiczne procesy związane ze stanami neuronów należących do systemów typu $[\Phi_j - \Phi_n]$. Stany te są skorelowane ze świadomością zjawiskową i z poszczególnymi jej aspektami.

Zakładamy, że *O* jest filozoficzno-psychologicznym opisem świadomości, *W* jest konkretną wypowiedzią (na przykład: „widzę czerwoną płamę”, „mam wrażenie czegoś czerwonego”). Przyjmujemy, że *W* może być wyprowadzone z całości warunków początkowych, do których odnosi się *O* oraz z obowiązujących w **MkT** potocznie rozumianych, prawopodobnych (*law like*) generalizacji.

Logiczną rekonstrukcję zależności pomiędzy teoriami **MkT** i **MiT** oddaje poniższa próba jej quasi-formalnego zapisu:

⁴⁴ Zadawane w ramach teorii emergencji pytania o naturę stanów zjawiskowych, o ich fizyczną realizację (implementację), są pytaniami ontologicznymi.

⁴⁵ Zarys neurokognitywnej teorii świadomości, opartej na tzw. pamięci roboczej (*working memory*, WM) podaje W. Duch (2003: 141–151). Bardziej rozbudowaną, naukową wersję teorii świadomości opartą na WM znajdujemy u B. Baarsa (1997: passim). Podobnie do Ducha przyjmujemy relacyjną (nieredukcyjną) zależność stanów świadomości od stanów mózgu.

⁴⁶ Niniejszy punkt zawdzięcza swoją obecną formę po części precyzacjom i uzupełnieniom pochodzącym od R. Poczobuta. Autor niniejszej książki jest odpowiedzialny za wszelkiego rodzaju niejasności i niedociągnięcia, które, mogą wystąpić w dalszych częściach tego i następnego rozdziału.

(a) (MkT, O) ⊢ W,

gdzie symbol „⊢” oznacza relację wynikania logicznego (rozumianą jako swoista idealizacja)⁴⁷. Wynikanie logiczne, albo „inferencja”, jest tutaj terminem używanym określenie dowolnych, zawierających logikę klasyczną, binarnych relacji zdań języka teorii MkT.

Chodzi nam o inferencje oparte na założeniu „normalnych okoliczności”. Przyjmujemy, że inferencja jest poprawna, a wniosek prawdziwy w sytuacjach, które są najbardziej typowe spośród wszystkich okoliczności spełniających przesłanki. Wracając do naszego przykładu z parasolem (por. 3.2; 3.5) możemy jedną z relacji inferencji – dotyczącą świadomości dostępu – przedstawić następująco: A – „Pada deszcz”, B – „Zabieram parasol”, C – „Nie zmoknę”. Załóżmy, że zwykle, gdy pada deszcz, mam zwyczaj zabierania ze sobą parasola {A → B, A} ⊢ B. Oczywiście jest, że gdy pada deszcz i mam otwarty parasol, to nie zmoknę {A ∧ B → C, A ∧ B} ⊢ C. Wynikanie wyraża tutaj relację między zdaniami. Sam wniosek zaś odnosi się do mojego zwyczaju zabierania parasola (mogę go jednak czasami zapomnieć).

Coś podobnego zachodzi przy świadomości zjawiskowej: A' – „Panuje dzieńne oświetlenie, jestem zdrowy i mój system wzrokowy także”, B': „Przedemną leży dojrzały pomidor”, C': „Będę miał wrażenie czerwoności”. Czymś oczywistym jest, że gdy panuje normalne oświetlenie i jestem zdrowy, a przedemną leży dojrzały pomidor, to będę miał wrażenie czerwoności {A' ∧ B' → C', A' ∧ B'} ⊢ C'. Przedstawiona relacja wynikania odnosi się do zdań (argumentów zdaniowych) i odzwierciedla normalne warunki zewnętrzne, normalny stan mojego organizmu i umysłu⁴⁸. Może się jednak zda-

⁴⁷ Według W. Strawińskiego: „Jeśli interpretujemy symbol «⊢» jako symbol wynikania logicznego, czyli operator konsekwencji, to schemat ten [wprowadzony przez niego schemat redukcji teorii T₁ do T₂ – J.B.] stosuje się w pełni tylko do języków sformalizowanych. Z takimi językami i teoriami właściwie nie mamy do czynienia w rzeczywistej nauce empirycznej. Będziemy go tutaj jednak stosować jako pewną idealizację, gdyż pomimo rozmaitych zastrzeżeń uważamy, że metoda logicznej rekonstrukcji teorii naukowych i zależności derywacyjnych lub inferencyjnych między nimi – przy wszystkich jej ograniczeniach – nadal zachowuje swoją wartość, chociażby tylko dla wskazania, co w danym rozumowaniu uważa się za założenia, a co za wniosek” (Strawiński 1997: 132, wskazanie na książkę Strawińskiego zawdzięczam J. Perzanowskiemu). R. Poczobut w rozmowie zauważył, że na podstawie powyższego cytatu nie jest jasne, czy Strawiński synonimicznie używa terminów „wynikanie logiczne” (które jest relacją semantyczną) oraz „konsekwencja” (która jest relacją syntaktyczną).

⁴⁸ Przy tak sformalizowanych zdaniach dotyczących wynikania mamy do czynienia z apodyktyczną relacją. Zmiana warunków towarzyszących może jednak sprawić, że stan rzeczy opisywany we wniosku nie zajdzie. Jest to warunek z klauzulą *ceteris paribus*

rzyć, że mam halucynacje czegoś czerwonego, dlatego w praktyce sam wniosek jest mniej pewny aniżeli przesłanki A' oraz B'.

Na podstawie intuicyjnie rozumianej zależności pomiędzy psychologią i filozofią z jednej strony, a neurologią – z drugiej można przyjąć ogólne twierdzenie:

VII.10 Jeżeli jakaś „makroteoria” (**MkT**) pozwala na przewidzenie *ex post* zjawiskowego stanu podmiotu na podstawie jego innego stanu, to związana z nią naukowa „mikroteoria” (**MiT**) winna umożliwić to samo.

W pierwszej części VII.10 chodzi o przewidywalność *ex post*, tzn. coś musiało najpierw zostać przeżyte (zobaczone), aby można to było przewidzieć. Nie jest to przewidywalność jak się ją rozumie w ścisłych naukach empirycznych⁴⁹. Stany zjawiskowe (fenomenalne) są przewidywane *ex post*, a nie według nomologiczno-dedukcyjnego modelu, spotykanego w naukach ścisłych⁵⁰.

Niech $[\Psi_j]$ będzie którymś z układu zjawiskowych stanów mentalnych $[\Psi_i - \Psi_m]$, gdzie j jest wybrane i ustalone z przedziału $[i, \dots, m]$, ($i \leq j \leq m$) zaś $[\Psi_k]$ niech będzie jakimś stanem z tego samego układu, gdzie k jest wybrane i ustalone z przedziału $[i, \dots, m]$ ($i \leq k \leq m$). Niechaj układowi stanów mentalnych odpowiada pewien układ stanów neurofizjologicznego systemu $[\Phi_j - \Phi_n]$ ⁵¹. Terminy x i y reprezentują zdania odpowiadające zjawiskowym stanom $[\Psi_j]$ oraz $[\Psi_k]$, a terminy x' i y' reprezentują zdania odpowiadające stanom neuronalnego systemu $[\Phi_j - \Phi_n]$, wtedy VII.10 zapisujemy formalnie jak niżej:

(b) (**MkT**, $x \vdash y$) \leftrightarrow (**MiT**, $x' \vdash y'$).

(o ile wszystko inne będzie bez zmian, warunek ten będzie spełniony). Klauzula *ceteris paribus* obowiązuje dla wszystkich praw wyższego rzędu.

⁴⁹ Według E. Nagla: „Zwolennicy teorii emergencji skłonni są czasem szczególnie wykorzystać fakt, iż w oparciu o teorie fizyczne nie można przewidzieć [«wyprowadzić ściśle w drodze dedukcji» s. 322 (J.B.)] tak zwanych własności wtórnych” (Nagel 1961: 324).

⁵⁰ Przewidywalność *ex post* polega na tym, że jeśli doświadczyliśmy pewnego stanu zjawiskowego, spowodowanego wystąpieniem innego stanu, to możemy w przyszłości oczekiwać, że zawsze ilekroć wystąpi ten ostatni stan, wystąpi również ów stan zjawiskowy. Nie dysponując owym doświadczeniem nie wiedzielibyśmy, co przewidywać.

⁵¹ W przyszłości może się okazać, że istnieją jednostkowe $[\Phi_k]$ oraz $[\Phi_j]$ wybrane z układu $[\Phi_j - \Phi_n]$ i skorelowane odpowiednio z $[\Psi_j]$ oraz $[\Psi_k]$.

Symbol „ \leftrightarrow ”⁵² odpowiada metarównoważności (dotyczy metajęzyka) symbol „ \vdash ” odpowiada relacji wynikania logicznego.

Na poziomie przedmiotowym można mówić o tzw. pociąganiu z koniecznością (*ontological entailment*), rozumianym jako rodzaj związku pomiędzy dwoma szeregami faktów, który istnieje wtedy, gdy istnieją prawa pomostowe, pozwalające na przejście od teorii dotyczącej jednego szeregu faktów do teorii dotyczącej drugiego szeregu faktów. Mówiąc o *entailment* chodzi nam o konieczność nomologiczną, a nie logiczną.

Uwzględniając w (b) język psychologii potocznej i neurologii otrzymujemy:

(c) $[(MkT, O) \vdash W] \leftrightarrow [(MiT, O') \vdash W']$.

MiT jest wystarczająco ogólną, neurologiczną teorią mózgu, *O'* jest opisem (w języku teorii **MiT**) warunków początkowych, *W'* jest opisem warunków końcowych (w języku teorii **MiT**).

Musimy umieć wyjaśnić, w jakim sensie inferencja w teorii **MiT** jest związana z inferencją w teorii **MkT**. Poniżej podajemy kilka możliwości rozumienia relacji „ \leftrightarrow ” oraz „ \vdash ”.

(1) Formuła (b) nie zakłada redukcyjnej zależności pomiędzy opisem oraz wyjaśnianiem świadomości w psychologii potocznej **PP**, a jej opisem i wyjaśnianiem w ramach neurologii współczesnej **MiT**. Tego rodzaju założenie świadczyłoby o możliwości analitycznej redukcji **MkT** do **MiT** (na przykład logiczno-behawioralnego „tłumaczenia” zdań jednej teorii na język drugiej, por. 3.2). Redukcja byłaby wtedy rozumiana jako redukcja jednej teorii do drugiej, nie zaś jako redukcja terminów odpowiadających subiektywnym przeżyciom czy zjawiskom – do terminów odpowiadających obiektom lub ich stanom neurobiologicznym.

Trzeba w tym miejscu zaznaczyć, że od teorii empirycznych oczekuje się nie tylko wyjaśnienia faktów, ale także ich przewidywania. W naukach empirycznych przewidywanie jest możliwe ze względu na znajomość makro- i mikropraw, które odnoszą się do konkretnych struktur i do sytuacji powta-

⁵² Przedstawiciele materializmu redukcyjnego wprowadzają w miejsce symbolu „ \leftrightarrow ” symbol jednostronnej redukcji teorii **MkT** do teorii **MiT**, obojętne, czy redukcja taka byłaby rozumiana w sensie J. G. Kemeny’ego i P. Oppenheima (czyli przez „eliminację”) czy też w sensie E. Nagla. Wspomniane rodzaje redukcji omawia W. Strawiński (1997: 126–141). Zagadnienie to wymaga dalszych analiz.

rzalnych. Zależność „ \leftrightarrow ” jest metarównoważnością, gdyż w zdaniu (b) nie twierdzimy ani: (i) że istnieją ściśle „makroprawa”, umożliwiające przewidywanie stanu $[\Psi_j]$ ze stanu $[\Psi_i]$, ani (ii) że istnieją empirycznie stwierdzalne prawa pomostowe (por. E. Nagla „zasady korespondencji”, 3.6.2), którym podlegają pary stanów, powiedzmy $[\Psi_i]$ oraz $[\Phi_i]$ (w prawach pomostowych występują dwustronne implikacje). Na wspomniane właśnie zależności wskazuje metarównoważność „ \leftrightarrow ”. Symbol „ \leftrightarrow ” rozumiemy jako emergencyjną zależność pomiędzy dwiema teoriami, a tym samym (b) rozumiemy jako emergencyjne prawo pomostowe⁵³.

(2) Zapytajmy, czy „ \leftrightarrow ” można rozumieć jako relację nomologiczną (por. 3.3)? Czy istnieje jakieś psychofizyczne prawo pomostowe, łączące inferencję na poziomie makroteorii z inferencją na poziomie mikroteorii? Powiedzieliśmy, że na poziomie psychologii potocznej trudno mówić o faktycznych „makroprawach”. Dlatego pytamy, czy „ \leftrightarrow ” w (b) lub (c) można rozumieć w sensie konieczności wynikającej z mikropraw. Będzie to jedynie oznaczało, że (c) wynika z praw mikroteorii **MiT**, w sensie: $[(\mathbf{MiT}, x' \vdash y') \Rightarrow (\mathbf{MkT}, x \vdash y)]$.

Logicznie jest jednak możliwe, że choć **MiT** jest prawdziwa, to jednak – gdy chodzi o twierdzenie (b) – w niektórych przypadkach zawodzi. Jest bowiem możliwe, że:

[i] teoria **MiT** jest prawdziwa, ale niepełna – gdy chodzi o przepowiadanie odpowiednich stanów lub zdarzeń, które może być dokonane za pośrednictwem teorii **MkT**.

Stosownym przykładem może być przypadek opisany w eksperymencie myślowym z Mary. Mary dysponuje rozbudowaną teorią **MiT**, czyli zupełną wiedzą neurologiczną⁵⁴, jednak nie potrafi – bez odwołania się do danych z teorii **MkT** – „przewidzieć”, *jak to jest* widzieć czerwone.

Od analiz tego rodzaju wychodzą teorie emergencji. Na przykład C. D. Broad twierdzi, że cech stanów mentalnych (które wyłoniły się na odpowie-

⁵³ Por. Strawiński 1997: 186–189. Należy zaznaczyć, że prawa pomostowe możemy rozumieć jako prawa emergentne lub jako prawa umożliwiające redukcję w sensie E. Nagla. Powiedzmy za Naglem, że teoria T_2 ma być zredukowana do teorii T_1 (por. wyżej, 3.6.2). Teoria T_2 posiada słownik S_2 , za pomocą którego opisujemy własność F, T_1 posiada słownik S_1 , do za pomocą którego opisujemy własność G. Wtedy prawo pomostowe ma postać: $\forall x(Fx \leftrightarrow Gx)$ – „dla każdego obiektu x, x posiada własność F jeżeli (i możliwe tylko jeżeli), x posiada G”. Redukcja poprzez równoważnościowe prawa pomostowe jest do pogodzenia z różnymi stanowiskami na temat relacji psychofizycznej.

⁵⁴ Emergentne założenie „zupełności wiedzy” i założenie „nieprzewidywalności” (niededukowalności) analizuje W. Strawiński (1997: 182–184).

dnim stopniu zorganizowania materii żywej), nie da się wydedukować na podstawie znajomości stanów fizycznych (które należą do niższych poziomów organizacji).

Wprowadzona przez nas relacja (b) może być konieczna nomologicznie, lecz kontyngentna logicznie. Konieczność nomologiczna naturalna (oparta na prawach natury) nie implikuje konieczności logicznej. Jedyne częściowo podobnie argumentuje Chalmers w swoim *The Conscious Mind*⁵⁵. Chalmers operuje mocniejszym pojęciem redukcji niż Nagel. Według Chalmersa redukcja wymaga logicznej superweniencji (czyli konieczności logicznej). Według Nagla redukcja wymaga tylko praw pomostowych, które nie mają cechy logicznej konieczności (jako prawa są one konieczne tylko nomologicznie).

(3) A więc formuła (b) oznacza emergencyjną zależność pomiędzy teoriami **MkT** i **MiT**, która to zależność przypomina teorie superweniencji nomologicznej. Tym samym w tezie o superweniencji konieczność rozumiemy jako zależność nomologiczną.

Metarównoważność „ \leftrightarrow ” w (b) oraz (c) mówi jedynie, że gdy jakoś świadomie widzimy czerwoną płamę, albo gdy odczuwamy, *jak to jest być* nami, to w języku psychologii potocznej **PP**, a także w języku neuronauk, musi istnieć przynajmniej jedno zdanie, które odnosi się do tego konkretnego zdarzenia. Nie chodzi przy tym o przetłumaczalność jednego opisu na drugi, lecz raczej o założenie dotyczące ogólnych możliwości wyrażania w obydwu językach. Tego rodzaju metarównoważność musi być uwzględniana w badaniach świadomości i jej subiektywności, a swoje teoretyczne podłoże znajduje ona w podanych przez nas w poprzednich rozdziałach metodach określania neuronalnych korelatów świadomości.

(4) Zależność pomiędzy inferencją „ \vdash ” w teorii **MkT** a inferencją „ \dashv ” w teorii **MiT** rozumiemy wszak jako zależność typową dla teorii emergencyjnych. Inferencje w **MkT** i **MiT** odwołują się do odmiennych narzędzi poznawczych i do odmiennych warunków badawczych. Pierwsza z nich posługuje się introspekcją (która w ramach **PP** ma duży status wiarygodności) oraz odwołuje się do potocznych obserwacji opisywanych w trzeciej osobie. Dru-

⁵⁵ W rozdziale IV Chalmers twierdzi, że własności zjawiskowe nie są na sposób konieczny realizowane przez własności fizyczne (materializm jest błędny: „[...] the failure of logical supervenience directly implies that materialism is false: there are features of the world over and above the physical features”, (s. 123), oraz zaznacza „[...] there is no a posteriori necessary connection between physical facts and phenomenal facts”, tamże, s. 94). W rozdziale III natomiast twierdzi, że własności zjawiskowe nie są logicznie superweniencyjne na własnościach fizycznych (nie istnieje redukcyjne wyjaśnienie świadomości: „[...] there is no a priori entailment from physical facts to phenomenal facts” [s. 94]).

ga z nich odwołuje się głównie do języka neuronauk – na przykład – do opisów skanowania mózgu, ale także do trzecioosobowych opisów obserwowalnych zachowań.

7.6.2. Wyjaśniająca wartość proponowanej prototeorii

Redukcjonistyczne teorie identyczności typów (*type-type identity theories*), przedstawione w rozdziale III, nie odnoszą się do szczegółowych danych empirycznych, lecz odwołują się najczęściej do psychofizycznych „praw pomostowych” typu: „Jeżeli X widzi czerwoną plamę, to w jego mózgu zachodzi proces P ”, bądź też odwrotnie: „Jeżeli w mózgu X 'a zachodzi proces P , to widzi on czerwoną plamę”. Typy również mogą być emergencyjnie związane, mamy wtedy do czynienia z prawami pomostowymi bez identyczności. Jak wspomnieliśmy, tak rozumiane prawa psychofizyczne mają postać równoważności.

Podane w rozdziale VII twierdzenia dotyczące neuronalnych korelatów świadomości pokazują, że dziś nie potrafimy istnienia tego rodzaju „praw” empirycznie stwierdzić. Jeżeli z jednej strony uwzględnimy filozoficzno-psychologicznie rozumianą subiektywność świadomości, a z drugiej strony wyniki współczesnych badań neurologicznych, to rodzą się wątpliwości co do istnienia takich ściśle redukcyjnych praw – odnoszących się zarówno do pojedynczych stanów świadomych i stanów neurologicznych, jak i do typów tych dwóch rodzajów stanów. Tym samym rodzą się wątpliwości co do możliwości redukcji przez identyfikację filozoficzno-psychologicznie rozumianej świadomości do jej neuronalnego podłoża.

Wspomniane prawa mogą natomiast być prawami emergencji, a redukcja rozumiana jako słaba (czyli bez stwierdzania identyczności). Dlatego zamiast o ściśle redukcyjnych prawach, mówimy o prawach emergencyjnych, do wyeksplikowania i uściślenia których prowadzi znalezienie neuronalnych korelatów stanów świadomych. Praw redukcyjne w sensie Nagla mogą być interpretowane jako prawa emergencji.

Zaproponowana w poprzednim punkcie prototeoria opiera się – ściśle biorąc – na filozoficznie rozumianej, emergencyjnej zależności pomiędzy stanami lub zdarzeniami neurologicznymi i psychologicznymi, z jednej strony, oraz na emergencyjnej zależności pomiędzy dwiema teoriami tychże zdarzeń – z drugiej. Tym samym jej zdolność do wyjaśniania i przewidywania jest ograniczona. Dopiero poszukiwana teoria **MkiT** będzie miała adekwatną moc wyjaśniającą.

Odwołując się do zaproponowanej prototeorii możemy jednak badać zależności pomiędzy filozoficzno-psychologiczną teorią świadomości, a teorią

neurologiczną, poszukując zarazem sposobu ich późniejszego połączenia w ramach jednej teorii. Będzie wtedy można wyjść z jednej teorii – powiedzmy z poszerzonej teorii **MkiT** – i szukać wyjaśnienia oraz przewidywania *ex post* faktów i zjawisk wyjaśnianych i przewidywanych dzisiaj w ramach obydwu teorii. Mamy bowiem dwa rodzaje zjawisk $[\Psi]$ i $[\Phi]$, których na razie nie potrafimy połączyć *a posteriori* (nie potrafimy ani zredukować poszczególnych zjawisk, ani rodzajów zjawisk $[\Psi]$ do rodzajów zjawisk $[\Phi]$, ani odwrotnie), dlatego przyjmujemy, że zależność pomiędzy teoriami **MkiT** i **MiT** nosi charakter zależności emergencyjnej. Teoria **MkiT** jako wielowymiarowa teoria unifikująca będzie mogła zachować schemat emergentystyczny. Emergentyzm będzie ontologicznym założeniem takiej teorii, chociaż obecnie nie wiemy, jaki to będzie model emergencji⁵⁶.

Jeśli potrafimy podać warunek wystarczający wystąpienia zjawiskowego stanu $[\Psi_j]$, to umiemy go wyjaśnić i przewidzieć (w rozumieniu teorii emergencji). Jeżeli stan neuronalnego systemu $[\Phi_j-\Phi_n]$ jest warunkiem wystarczającym zjawiskowego stanu $[\Psi_j]$, oraz wiemy, że $[\Phi_j-\Phi_n]$ zaszedł, to możemy przewidzieć wystąpienie $[\Psi_j]$. Nie potrafimy jednak na razie stwierdzić, jak to przebiega empirycznie, jaki empiryczny mechanizm za tym stoi, ale jesteśmy coraz bliżej takiego stwierdzenia.

Odwołujemy się wtedy do następującego modelu emergencji $[\Psi_j]$ względem stanów systemu $[\Phi_j-\Phi_n]$.

1. $[\Phi_j-\Phi_n]$ jest warunkiem wystarczającym dla $[\Psi_j]$. Zawsze, gdy występuje $[\Phi_j-\Phi_n]$, występuje $[\Psi_j]$. Oznaczając występowanie $[\Phi_j-\Phi_n]$ przez Z , a $[\Psi_j]$ przez X , otrzymujemy schemat prawa: (P) $\forall x[Z(x) \Psi X(x)]$, które wyraża słabą konieczność nomologiczną.
2. Stany $[\Phi_j-\Phi_n]$ nie są identyczne z $[\Psi_j]$ ($[\Phi_j-\Phi_n] \neq [\Psi_j]$).
3. Zdania (teorie) dotyczące $[\Phi_j-\Phi_n]$ nie są analitycznie przekładalne na zdania (teorie) dotyczące $[\Psi_j]$.

Identyczność, którą odrzucamy w (2), jest identycznością intensjonalną (tożsamościową). Wprowadzony przez nas model emergencji dopuszcza jednak identyczność słabą (ekstensjonalną, zakresową). Mielibyśmy z nią do czynienia, gdyby się okazało, że „predykaty zjawiskowe” są koekstensywne z „predykatami neuronalnymi”. Takiej możliwości nasz model emergencji nie odrzuca (tzn. nasz model emergencji uwzględnia słabą identyczność – identyczność ekstensjonalności).

⁵⁶ Np. W. Duch tworzy naukową teorię stanów świadomości opartą na emergentystycznej ontologii (por. Duch 2003: 136–153). Według Ducha świadomość ma do odegrania pewną rolę funkcjonalną, ale nie można jej do tej roli sprowadzić (Duch jedynie nieguje epifenomenalizm w odniesieniu do świadomości zjawiskowej).

Przechodząc na płaszczyznę ontologiczną możemy powiedzieć, że nasz model nie prowadzi do eliminacji ani do identyfikacji (własności i praw opisywanych w **MkT** z własnościami i prawami opisywanymi w **MiT**), mówi jednak o determinacji. Odwołując się do twierdzeń R. Poczubota możemy powiedzieć, że być może determinacja polega tutaj, na „pociąganiu z koniecznością nomologiczną” (por. Poczubot, *Uwagi*: 2–5).

(i) Kompletnie wyznaczenie składu i układu mikrostruktury $[\Phi_j - \Phi_n]$ pociąga za sobą w sposób konieczny wystąpienie określonych makrowłasności $[\Psi_i - \Psi_m]$.

(ii) Własności pociągane z koniecznością są mikrodeterminowane, tzn.: wystąpienie mikrostruktury $[\Phi_j - \Phi_n]$ i pojawienie się jej odpowiedniego stanu jest warunkiem wystarczającym wystąpienia makrowłasności rodzaju $[\Psi_i]$. Innymi słowy: takie same neuronalne mikrostruktury, takie same własności zjawiskowe (por. wyżej, 3.6, [5(b)]).

Otrzymujemy w ten sposób jedną z wersji relacji superwencji mereologicznej, którą nazywa się także „twierdzeniem o mikrodeterminacji”. Własność mikrodeterminowana (superwencyjna) jest pociągana z koniecznością przez mikrostrukturę obiektu $[\Phi_j - \Phi_n]$, któremu przysługuje. Po przekroczeniu pewnego określonego progu złożoności wewnętrznej (w niniejszej pracy jest to próg „minimalnego systemu”), obiekt przejawia własność emergentną.

Wartość obejmującej te zjawiska teorii emergencyjnej jest nie tylko heurystyczna (każdorazową relację pomiędzy stanami $[\Phi_j - \Phi_n]$ i stanami $[\Psi_i - \Psi_m]$ oraz uśrednioną relację pomiędzy typami tych stanów trudno zweryfikować empirycznie), lecz stawia sobie także cele wyjaśniające. Przedstawiona prototeoria należy do współczesnych metod szukania rozwiązań, dla których nie ma pewności znalezienia rozwiązania optymalnego, a nawet często jedynie prawidłowego (por. 7.1.2, 7.3.2).

Emergencyjna teoria świadomości jest rozwiązaniem przybliżonym i służy interpretacji odkrywanych nowych faktów i związków między nimi oraz formułowaniu dalszych hipotez. Tym samym otrzymane wyjaśnienia różnią się zasadniczo od empirycznych rozwiązań proponowanych przez teorie neurologiczne oraz od *apriorycznych* rozwiązań proponowanych przez filozoficzne teorie redukcjonistyczne. Wnioski wyciągane z *apriorycznej* zależności teorii **MkT** i **MiT** będą akceptowalne tylko na tyle, na ile spójnie i owocnie poszerzą one nasze dzisiejsze rozumienie świadomości.

Ścisłe materialistycznie rozumiana redukcja teorii **MkT** do teorii **MiT** wymaga, w najsilniejszym tego słowa znaczeniu, aby wszystkie zdania teorii **MkT** były przetłumaczalne na zdania teorii **MiT**, zaś obiekty teorii **MiT** były ontologicznie bardziej podstawowe aniżeli obiekty teorii **MkT**. W proponowanej przez nas emergencyjnej prototeorii niektóre aspekty teorii świado-

mości mogą być wyjaśnione przez odwołanie się do teorii neurologicznej, inne zaś nie (por. kwestie związane z różnymi stopniami jedności świadomości, 7.5.2). Świadomość nadal jednak pozostaje „zagadkowa” chociażby w swoich aspektach ontologicznych.

Nawet wtedy, gdy jakaś współczesna teoria stanów świadomości MIT dostarcza wyczerpujących opisów subiektywności, nadal brakuje wyjaśnienia jej natury, chociażby w sensie odpowiedzi na pytanie: w jaki sposób i dzięki jakim różnorodnym procesom neurologicznym nasza świadomość widzenia niebieskiej plamy dochodzi do przedstawienia jednorodnej, niebieskiej powierzchni przede mną? Mimo iż wiemy i widzimy, że stany świadome $[\Psi_i]$ posiadają jakąś jednorodną treść, nie wiemy jak ją osiągają. Częściową odpowiedzią może być twierdzenie, że stany neurologiczne, obecne na funkcjonalnym poziomie spostrzegania, konstytuują wystarczającą podstawę dla powstania świadomości na poziomie subiektywnym.

Przedstawiona w 7.6.1 prototeoria jest otwarta na fakt, że jeżeli miałyby się okazać, iż w naszej świadomości jest coś, co w sposób zasadniczy nie jest dostępne dla nauki (szczególnie dla neuronauki) – w takim sensie, że jest niewyjaśnialne (w sensie wyjaśniania redukcyjnego) i nieprzewidywalne, to mimo tego osiągnęliśmy to, co filozofowie nazywają *postępem poznawczym*. Może on polegać na tym, że w bardziej precyzyjny i pełniejszy sposób potrafimy opisać i wyjaśnić, dlaczego na pewne pytania zasadniczo nie ma naukowo zadowalającej odpowiedzi.

W przedstawionej prototeorii da się uwzględnić dalsze neurologiczno-eksperymentalne i kognitywistyczne poszukiwania korelacji pomiędzy świadomością zjawiskową a mózgiem. Wystarczy tu chociażby przypomnieć, że selektywna zmiana przeżywania kolorów musi być skorelowana z jakiegoś rodzaju szybkimi i dokładnymi zmianami schematów neuronalnych (wymaga odpowiedzi od szeregu bodźców, które wcześniej były antycypowane jako sygnał z określoną wartością) na inny szereg bodźców, które teraz są antycypowane jako posiadające wartość sygnału.

Ewentualna czasowa nieprzewidywalność (w sensie logicznej dedukowalności) i niewyjaśnialność (w sensie wyjaśniania redukcyjnego) – w odniesieniu do aktualnego stanu naszej wiedzy – zostaje tym samym wprowadzona w szerszy horyzont metodologiczny. Swoista perspektywa wyjaśnienia „trudnego” problemu subiektywnej świadomości zdaje się leżeć w znaczącym poszerzeniu badań empirycznych. Z metodologicznego punktu widzenia warto się do tej perspektywy przygotować poprzez analizy prowadzących do niej badań w logice, psychologii, neurobiologii i informatyce.

7.7. Podsumowanie

Poszukiwanie neuronalnych korelatów świadomości jest dziedziną kognitywistyki i obejmuje również badanie świadomości na poziomie empirycznym. W tym rozdziale poszukiwanie to ograniczyliśmy do wybranych przykładów z neurologii i do filozofii umysłu. Zadając w ramach kognitywistyki pytanie o subiektywność świadomości filozof umieszcza problemy związane ze świadomością w szeregu stanowisk teoretyczno-empirycznych. Jednym z celów nauk empirycznych jest badanie korelacji, chociaż najczęściej są to słabszego rodzaju korelacje aniżeli te przez nas omawiane.

Przedstawiony program poszukiwania korelatów świadomości jest uproszczony. Bardziej naukowo rozbudowany program dotyczący badań podmiotu żyjącego w swoim otoczeniu będzie musiał uwzględnić dodatkowe czynniki związane z pozostałymi zachowaniami tegoż podmiotu. Można przypuszczać, że:

(i) znalezione lub bliżej określone neuronalne korelaty świadomości, okażą się poszukiwanymi systemami, umożliwiającymi pojawianie się holistycznie rozumianej świadomości i poszczególnych stanów świadomości,

(ii) w dalszych etapach badań umożliwią one wyjaśnienie świadomości w sensie: w jaki sposób i na ile mózg ją tworzy.

Zgodnie z szkicem programu badawczego (por. 7), znalezienie neuronalnych korelatów holistycznej świadomości oraz jej poszczególnych stanów pozwoli zrealizować takie cele, jak:

(1) Wskazanie systemów umożliwiających monitorowanie i chociażby przybliżone przewidywanie treści świadomości w przypadkach, w których nie możemy liczyć na sprawozdania werbalne podmiotu. Chodzi tutaj na przykład o pytanie dotyczące obecności holistycznej świadomości oraz o obecności świadomych przeżyć u pacjentów chorych na „zespół zamknięcia” (*locked-in syndrome*), polineuropatię, leżących w śpiączce („koma”), czy pacjentów cierpiących na *deaf-hearing* czy *optic ataxia*⁵⁷. Powyższe zespoły chorobowe cechują się podobnymi objawami zewnętrznymi (zwła-

⁵⁷ *Locked-in syndrome* – rzadko spotykana choroba neurologiczna, cechująca się całkowitym paraliżem mięśni ruchowych we wszystkich częściach ciała, z wyjątkiem tych, które kontrolują ruchy oczu. Pacjenci są świadomi, chociaż nie potrafią mówić ani się poruszać. *Polineuropatia* – progresywne uszkodzenie nerwów obwodowych (może mieć tło zapalne, toksyczne, genetyczne), prowadzące do utraty zdolności do poruszania się i do odczuwania. Pacjenci nie reagują na wypowiedane do nich słowa, na muzykę. *Optic ataxia* – chorzy nie potrafią skierować ręki ku jakiemuś przedmiotowi, pomimo że go widzą.

szcza trzy pierwsze) i są wyróżniane w zależności od rejonu uszkodzenia mózgu⁵⁸.

Dzięki znalezionym korelatom świadomości będzie można coś powiedzieć o stanach mentalnych pacjentów pod narkozą czy podczas głębokiego snu. Trzeba jednak pamiętać, że w przypadkach, w których struktura mózgu zasadniczo różni się od struktury mózgu człowieka zdrowego (choćaby w niektórych przypadkach śpiączki), uzyskana oczywistość wyników i wyciągniętych z nich wniosków będzie pozostawiała wiele do życzenia.

(2) Wyodrębnione układy [Φ_j – Φ_n] stanowią wstępny etap w rozwoju kognitywnych nauk o świadomości. Wiedząc bowiem, które układy są neuronalnymi korelatami stanów świadomości, można „obiektywnie” badać mechanizmy ich pracy. Można na przykład próbować określić „obiektywne” kryteria bólu niezależnie od tego, co dany podmiot twierdzi o intensywności jego przeżywania.

(3) Gromadzenie danych o neuronalnych korelatach świadomości pozwala na usystematyzowanie rodzajów relacji zachodzących pomiędzy stanami neuronalnymi a stanami świadomymi oraz na uchwycenie ogólnych praw rządzących tymi relacjami (na razie widzimy, że są to relacje emergencji). Tak określone korelaty pomogą w zbudowaniu poszerzonej teorii świadomości, o której pisze Nagel.

W punktach [A]–[B]–[C] przeprowadzonych rozważań skupiliśmy się na metodologicznych pytaniach typu:

– Według jakiego kryterium należy określać, czy dany podmiot jest świadomy w sensie całościowego stanu świadomości?

– Według jakiego modelu mamy opisywać przeżycia, aby mogły być bezpośrednio skojarzone z procesami zachodzącymi w mózgu?

– Jaki stopień opisu neuronalnej aktywności jest na dzień dzisiejszy odpowiedni, aby pokazać korelacje pomiędzy tą aktywnością, a jakościowymi treściami świadomości?

Na razie brakuje jednoznacznie zadowalającej, naukowej odpowiedzi na pytania typu:

– Czy możemy mówić o koniecznych i wystarczających neuronalnych korelatach świadomości, czy też korelacje są przypadkowe?

– Czy system, który osiągnął funkcjonalny stopień równy ludzkiemu mózgowi, może pozostać bez świadomości?

⁵⁸ Na pierwszy rzut oka objawy śpiączki i *locked-in syndrome* wyglądają podobnie. W odróżnieniu od *locked-in syndrome* (gdzie świadomość jest zachowana), śpiączce odpowiadają uszkodzenia na poziomie grzbietowych części pnia mózgu. Gdy uszkodzenie występuje parę milimetrów dalej w kierunku brzuszonym, to wywołuje ono pojawienie się *locked-in syndrome*.

– Jeżeli dominujące w dzisiejszej filozofii analitycznej materialistyczno-nieredukcjonistyczne (w sensie teorii identyczności typów) ujęcia świadomości są prawdziwe, to w jaki sposób mózg generuje świadomość?

Neuronaukowe odpowiedzi na te pytania będą zarazem odpowiedzią na pytanie, czy poprawne i prawdziwe jest przyjęte w tej pracy metodologiczne założenie o holizmie świadomości i założenie o jej emergencji. Mózg da się monitorować, świadomość zaś nie. Holistycznie rozumiana świadomość jest wyrazem przetransformowanych doznań w danym momencie i w danych okolicznościach. Widzimy jednak, że skorelowana ze świadomością złożoność struktury neuronalnej znajduje swoje emergencyjne odzwierciedlenie w holistycznej świadomości.

ZAKOŃCZENIE

„[Menon]: W jaki jednak sposób będziesz szukał Sokratesie czegoś, o czym nie wiesz zupełnie, co to jest? Co z tego, czego nie znasz, weźmiesz za punkt wyjścia? A jeżeli przypadkiem na to natrafisz, w jaki sposób poznasz, że to właśnie to, czego szukasz?” (Platon 1991: 80e).

Nakreślone ramy poszukiwania nowej teorii świadomości rzucają dodatkowe światło na kwestie związane z wyjaśnieniem jej subiektywności. Pokazują one, że nie tylko świadomość i mózg, ale także zachodzące pomiędzy nimi relacje są o wiele bardziej złożone i zagadkowe, aniżeli było to przyjmowane w wielu teoriach umysłu – zarówno neurologicznych, jak i filozoficznych. Czy zatem, jeżeli chodzi o wyjaśnienie świadomości lub o przewidywanie jej konkretnych stanów, pozostaje tylko sceptycyzm albo ucieczka w proponowany przez C. McGinna mistycyzm?

W przedstawionym w rozdziałach III, IV, V krytycznym przeglądzie problematyki umysł–ciało i związanej z nią subiektywności świadomości podzielono teorie umysłu na dualistyczne i monistyczne. Teorie dualistyczne, które jedynie wspomniano, przyjmują formę interakcjonistyczną i nieinterakcjonistyczną: umysł i mózg różnią się od siebie i albo oddziałują na siebie przyczynowo, albo nie. Teorie monistyczne przyjmują formę silnie ontologicznie redukcjonistyczną i niereducjonistyczną: albo świadomy umysł nie jest niczym więcej aniżeli mózgiem (albo *vice versa*), albo obydwa są aspektami czegoś bardziej podstawowego. Wspomniany przegląd teorii i ich krytyczne analizy ukazują nie tylko słabe i silne strony każdej z nich, ale także to, że w sporach pomiędzy filozoficznymi ujęciami jedynie rzadko chodzi o jednoznaczny dowód czy zaprzeczenie. Żadna ze wspomnianych teorii nie jest w stanie przedstawić dowodu, że sama jest teorią poprawną, a inna niepoprawną.

Problem subiektywności wrażeń zmysłowych został nazwany problem „trudnym”. Wychodzi się przy tym z założenia, że rzekomo łatwiej wyobrazić sobie to, w jaki sposób stany intencjonalne mogą powstawać w jakimś materialnym systemie, a wiele trudniej zaś – jak stany przeżyciowe mogą w nim powstawać i dlatego w ogóle w nim powstają. Także pełna znajomość wszystkich fizyczno-neurologicznych procesów zachodzących w świadomym mózgu, a także znajomość wszystkich fizjologicznych różnic pomiędzy „świadomym” i „nieświadomym” (np. będącym pod narkozą) mózgiem, nie dostarczają nam dziś ściśle naukowo rozumianego wyjaśnienia świadomości zjawiskowej, dlatego przyjmujemy wyjaśnianie emergentystyczne.

Nawet gdybyśmy umieli opisać wszystkie fizjologiczne warunki występowania świadomości (oraz wszystkie warunki, w których brak świadomości), to nie wiedzielibyśmy, dlaczego i w jaki sposób świadomość się pojawia. Wszystkie fizjologiczne i funkcjonalne opisy systemów fizycznych – zarówno mózgu, jak i komputera – są logicznie do pogodzenia z założeniem, że stany świadomości w nich nie występują (por. eksperymenty myślowe o „nieobecnych qualia”, „o nietoperzu”). Nawet gdybyśmy potrafili wyjaśnić budowę i działanie mózgu jako systemu kognitywnego, to nie wynika z tego, że również potrafilibyśmy zrozumieć, co to znaczy, że jest on „systemem świadomym”.

Proponowana w 7.6.1 prototeoria nie rozwiązuje problemu świadomego umysłu w tym sensie, że nie dostarcza jednoznacznych naukowych wyjaśnień i nie umożliwia jednoznacznych przewidywań. Niemniej jednak ma ona wartość wyjaśniającą. Przyjmując proponowaną metodologię poszukiwania neuronalnych korelatów świadomości można za jej pomocą pokazać, w jaki sposób rozwiązuje ona problemy, z którymi pozostałe teorie gorzej sobie radzą, zwłaszcza gdy chodzi o uwzględnianie danych empirycznych. Ponadto wspomniana prototeoria – jak uważamy – nie wprowadza dodatkowych problemów, które mogłyby być lepiej rozwiązane przez znane, monistyczne teorie redukcyjne. Dodatkowo nie potwierdza ani *type-type* redukcyjnej zależności pomiędzy stanami subiektywnej świadomości a stanami mózgu, ani redukcyjnej (w sensie E. Nagla) zależności pomiędzy filozoficzną teorią świadomości a teorią neurologiczną. Stwierdzenie ewentualnej redukcyjno-identyfikującej zależności pomiędzy świadomością a mózgiem wymaga dalszych analiz językowych i prowadzenia dalszych badań eksperymentalnych. Na dzisiaj możemy jedynie mówić o nieredukowalności eksplanacyjnej.

Neurologia ukazuje różnice pomiędzy poszczególnymi podmiotami, gdy chodzi o struktury mózgu (por. 7.3.1). Na przykład grubość warstwy kory mózgowej i jej powierzchnia mogą być ograniczone bez wyraźnego wpływu

na codzienne zachowania (opisywane przez psychologię potoczną i inne nauki) i na inteligencję podmiotu. Dlatego pojawia się problem generalizacji danych uzyskanych technikami neuroobrazowania. Ponieważ każdy mózg rozwija i „uczy” się inaczej, więc nawet bliźnięta jednojajowe nie posiadają identycznych mózgów. Jeżeli jest możliwe, że podmioty X i Y są dokładnie w tym samym stanie mentalnym w czasie t , to nie wynika z tego automatycznie, że stan ten zależy od dokładnie tego samego stanu ich całego mózgu, nawet gdybyśmy potrafili pominąć wpływy otoczenia. Z drugiej strony, identyczne stany mózgu mogą odpowiadać różnym stanom mentalnym, gdy otoczenie podmiotu było różne. Świadomość – zgodnie z eksternalizmem, którego w tej pracy nie uwzględniliśmy – nie jest bowiem tylko w głowie. Nawet gdybyśmy sami obserwowali nasze aktualne neuronalne korelaty świadomości za pomocą tej czy innej techniki autoobrazowania, to treści naszych stanów mentalnych zależałyby chociażby od naszej pamięci, a ich neuronalne, dynamiczne korelaty musiałyby być ujęte w jakimś synchronicznym czasie t .

Na podstawie przeprowadzonych badań korelatów możemy wyjść z założenia, że:

(a) Internalistycznie rozumiane neuronalne korelaty świadomości zjawiskowej występują w mózgu, chociaż żadna ze współczesnych teorii świadomości nie dostarcza wyjaśnienia tej świadomości w sensie odpowiedzi na „trudny” problem świadomości: w jaki sposób mózg generuje subiektywną świadomość czy też w jaki sposób świadomość oddziałuje na mózg i jego funkcjonowanie?

(b) Stwierdzone korelacje świadomości z systemami neuronalnymi można potraktować jako relacje wystarczające. Nie oznacza to empirycznego ograniczenia w ich poszukiwaniach, chociaż ogranicza kognitywistyczne uogólnienia co do natury świadomości. Powyższe twierdzenie wyklucza silną redukcję (w sensie identyfikacji) świadomości do stanów fizycznych. Może się jednak okazać, że istnieją pojedyncze redukcje stanów mentalnych do stanów fizycznych albo redukcja przyczynowa (zgodna z nieredukowalnością eksplanacyjną).

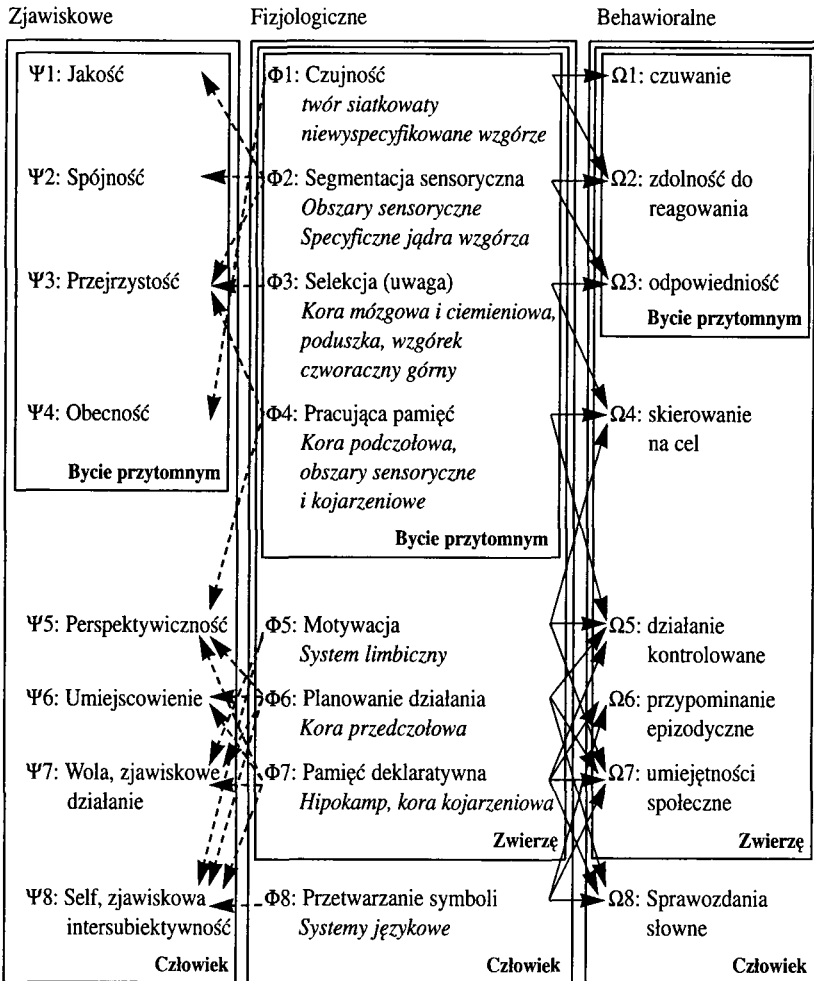
(c) Dziś można jedynie stwierdzić istnienie synchroniczno-emergencyjnej zależności pomiędzy świadomymi stanami mentalnymi a stanami w mózgu. Można także przyjąć emergencyjny rodzaj zależności pomiędzy filozoficzno-psychologiczną teorią świadomości a jej teorią neurologiczną. Powoływanie się na emergencyjną naturę cech świadomości nie jest wystarczające dla uzasadnienia *type-type*-naturalistycznej teorii świadomości, gdyż w zasadzie emergencyjność nie wyjaśnia świadomości w sensie, jakiego oczekujemy, gdy mówimy o „trudnym” problemie. Dlatego T. Nagel poszukuje poszerzonej

teorii świadomości, która byłaby teorią identyczności, traktującą o koniecznych związkach pomiędzy świadomością a mózgiem.

Trudno oczywiście wypowiadać się na temat przyszłego rozwoju nauk kognitywnych. Może kiedyś trzeba będzie stwierdzić, że nasze poszukiwania nowej wielowymiarowej teorii świadomości **MkiT** nigdy nie zakończą się jakąś jednoznaczną konkluzją. Być może praca filozofa, podobnie jak naukowca, nigdy nie doprowadzi do takiego naukowego wyjaśnienia świadomości, jakiego oczekujemy z dzisiejszego punktu widzenia i jakie znamy z nauk ścisłych. Wystarczy wtedy, że teza o subiektywności świadomości będzie posiadała filozoficzną zrozumiałość i indukcyjne poparcie.

Zjawiskowe, fizjologiczne i behawioralne aspekty świadomości¹

Pierwszo- i trzecioosobowe kryteria wyjaśniania świadomości (Engel, Singer)

¹ Engel 2000: 6; por. Engel i Singer 2001: 16.

LITERATURA

- Adamek D., S. Żulichowski, J. Kałuża. 1986, *Long survival with cerebellar aplasia and degenerative changes of CNS. A case report*, „Neuropatologia Polska”, 24:1, ss. 89–100.
- Adamek D. 2003, *Neuroprotekcja w urazach mózgu i rdzenia*, w: M. Śmiałowska (red.) *Neuroprotekcja. XX Zimowa Szkoła Instytutu Farmakologii PAN* (Mogilany), ss. 107–125.
- Ajdukiewicz K. 1960, *Postępowanie człowieka*, w: tenże, *Język i poznanie*, I. (Warszawa: PWN), ss. 317–364.
- Akins K. 1993, *A Bat without Qualities*, w: M. Davies, G. W. Humphreys (eds.), *Consciousness: Psychological and Philosophical Essays* (Oxford: Blackwell), ss. 258–270.
- Antony M. V. 1992, *The where and when of what*, *Behavioral and Brain Sciences*, w: „Behavioral and Brain Sciences”, 15:2, ss. 201–202 („Open peer commentary” do: Dennett & Kinsbourne, *Time and the Observer*).
- Armstrong D. M. 1968, *A Materialist Theory of Mind* (London: Routledge and Kegan).
- Aronson J., E. Dietrich, E. Way. 1992, *Throwing the conscious baby out with the Cartesian bath water*, „Behavioral and Brain Sciences”, 15:2, ss. 202–203 („Open peer commentary” do: Dennett & Kinsbourne, *Time and the Observer*).
- Atkinson A. P., M. Thomas A., Cleeremans. 2000, *Consciousness: mapping the theoretical landscape*, „Trends in Cognitive Sciences”, 4:10, ss. 372–382.
- Baars B. J. 1996, *Understanding subjectivity: Global workspace theory and the resurrection of the observing self*, „Journal of Consciousness Studies”, 3:3, ss. 211–216.

- Baars B. J. 1997, *In the Theater of Consciousness* (Oxford: Oxford University Press).
- Baars B. J. 1997a, *In the Theatre of Consciousness. Global Workspace Theory, A Rigorous Scientific Theory of Consciousness*, „Journal of Consciousness Studies”, 4:4, ss. 292–309.
- Baars B. J. 1998, *Metaphors of Consciousness and Attention in the Brain*, „Trends in Neurosciences”, 21:2, ss. 58–62, <<http://www.phil.vt.edu/ASSC/baars/baars1.html>>, ss. 1–7.
- Baars B. J. 2002, *The conscious access hypothesis: origins and recent evidence*, „TRENDS in Cognitive Sciences”, 6:1, ss. 47–52.
- Baars B. J. 2003, *How Brain Reveals Mind. Neural Studies Support the Fundamental Role of Conscious Experience*, „Journal of Consciousness Studies”, 10:9–10, ss. 100–114.
- Bartels A., Zeki, S. 1998, *The theory of multistage integration in the visual brain*, „Proceedings of the Royal Society” (Biological Sciences), 265, ss. 2327–2332.
- Baumeister R. F. 1998, *The Self*, w: D. T. Gilbert, S. T. Fiske, G. Lindzey, *The Handbook of Social Psychology* (Boston: The Graw–Hill Company), ss. 680–740.
- Bayne T. 2000, *The Unity of Consciousness*, „Australasian Journal of Philosophy”, 78, ss. 248–54.
- Bayne T., D. Chalmers. 2003, *What is the Unity of Consciousness?*, w: A. Cleeremans, *The Unity of Consciousness* (Oxford: Oxford Unity Press), ss. 23–58, <<http://www.u.arizona.edu/~chalmers/papers/unity.html>>, ss. 1–28.
- Beckermann A. 1999, *Philosophie des Geistes*, w: H. J. Sandkühler, *Enzyklopädie der Philosophie*, t. 2 (Hamburg: Meiner), ss. 1154–1159.
- Beckermann A. 2000, *The Perennial Problem of the Reductive Explainability of Phenomenal Consciousness: C.D. Broad on the Explanatory Gap*, w: T. Metzinger (ed.), *Neural Correlates of Consciousness: Empirical and Conceptual Questions* (Cambridge/MA: MIT Press), ss. 41–55.
- Bergson H. 1913, *O bezpośrednich danych świadomości* (Warszawa: Wende i S–ka).
- Berkeley G. 1957, *A Treatise Concerning the Principles of Human Knowledge* (N. York: Bobbs–Merill).
- Block N. 1978, *Troubles with functionalism*, w: C. W. Savage (ed.), *Perception and Cognition, Minnesota Studies in the Philosophy of Science* (Minneapolis: University of Minnesota Press), Vol. IX, ss. 261–385.
- Block N. 1988, *How Not to Find the Neural Correlate of Consciousness*, „Current Issues in Philosophy of Mind: Supplement to «Philosophy»”, 43, ss. 23–34.

- Block N. 1990, *Inverted Earth*, w: J. Tomberlin (ed.) *Action, Theory and the Philosophical Perspectives 4* (California: Ridgeview Publishers), ss. 53–79.
- Block N. 1992, *Begging the question against phenomenal consciousness*, „Behavioral and Brain Sciences”, 15:2, ss. 205–206 („Open peer commentary” do: Dennett & Kinsbourne, *Time and the Observer*).
- Block N. 1994, *What is Dennett's Theory a Theory of?*, „Philosophical Topics”, 22:1&2, ss. 23–40.
- Block N. 1995, *Functionalism*, w: J. Kim, E. Sosa (eds.) *A companion to metaphysics* (Oxford: Blackwell), ss. 188–194.
- Block N. 1995a, *On a confusion about a function of consciousness*, „Behavioral and Brain Sciences”, 18, ss. 227–287.
- Block N. 1998, *Das Neuronale Korrelat des Bewußtseins*, w: F. Esken, D. Heckmann (Hrsg.) *Bewußtsein und Repräsentation* (Paderborn [u.a.]: Schöningh, 1998), ss. 153–167.
- Block N. 1998a, *Holism, Mental and Semantic*, *The Routledge Encyclopedia of Philosophy* (N. York: Routledge), <<http://wwwserver.nyu.edu/gsas/dept/philo/faculty/block/papers/MentalSemanticHolism.html>>, ss. 1–7.
- Bobryk J. 1996, *Akty świadomości i procesy poznawcze* (Wrocław: Wydawnictwo Leopoldinum). Monografie Fundacji na Rzecz Nauki Polskiej.
- Bogen J. E. 1995, *On the Neurophysiology of Consciousness: I. An Overview*, „Consciousness and Cognition”, 4, ss. 52–62.
- Borkowska A. i E. M. Szepietowska. 2000, *Rozmowa z Profesorem Marcelem Klimkowskim*, w: A. Borkowska, E. M. Szepietowska (red.), *Diagnoza neuropsychologiczna. Metodologia i metodyka* (Lublin: Wydawnictwo Uniwersytetu Marii Curie-Skłodowskiej), ss. 11–19.
- Bremer J. 1997 *Rekategorisierung statt Reduktion. Zu Wilfrid Sellars' Philosophie des Geistes* (Göttingen: Vandenhoeck & Ruprecht).
- Bremer J. 1998, *W. Sellars' modifizierte Identitätstheorie*, „Forum Philosophicum”, 3, ss. 157–183.
- Bremer J. 2001, *Problem umysł–ciało. Wprowadzenie* (Kraków: WAM).
- Broad C. D. 1925, *The Mind and its Place in Nature* (London: Paul Kegan).
- Bromm B. 2001, *Brain Images of Pain*, „New Physiological Sciences”, 16, ss. 244–249.
- Brown J. 1998, *Fundamentals of Process Neuropsychology*, „Brain and Cognition”, 38:2, ss. 234–245.
- Burge T. 1979, *Individualism and the Mental*, „Midwest Studies in Philosophy”, 4, ss. 73–121.
- Burge T. 1996, *Zwei Arten von Bewußtsein*, w: T. Metzinger, *Bewußtsein:*

- Beiträge aus der Gegenwartsphilosophie* (Paderborn: Schöningh), ss. 583–594.
- Byrne A., D. R. Hilbert. 1997, *Colors and Reflectances*, w: A. Byrne, D. R. Hilbert (eds.), *Readings on color. The philosophy of color* (Cambridge, MA: MIT Press), ss. 263–288.
- Chalmers D. 1995, *Facing up to the problem of consciousness*, „Journal of Consciousness studies”, 2:3, ss. 200–219. <<http://www.u.arizona.edu/~chalmers/papers/facing.html>>, ss. 1–20.
- Chalmers D. 1996, *The Conscious Mind: In Search of a Fundamental Theory* (Oxford: Oxford Univ. Press).
- Chalmers D. 1996a, *Fehlende qualia, Schwindende Qualia, Tanzende Qualia*, w: T. Metzinger (Hg.), *Bewußtsein: Beiträge aus der Gegenwartsphilosophie* (Paderborn: Schöningh), ss. 367–389.
- Chalmers D. 2000, *What Is a Neural Correlate of Consciousness?*, w: T. Metzinger (ed.), *Neural Correlates of Consciousness: Empirical and Conceptual Questions* (Cambridge/MA: MIT Press), ss. 17–39.
- Chalmers D. 2003, *Zagadka świadomości*, „Świat Nauki”, 1 (styczeń), ss. 102–112.
- Changeux J-P. 1997, *Neuronal Man. The Biology of Mind* (Princeton: Princeton University Press).
- Churchland P. S. 1993, *Neurophilosophy. Toward a Unified Science of the Mind-Brain* (Cambridge/MA: The MIT Press).
- Churchland P. S. 1996, *Die Neurobiologie des Bewußtseins. Was können wir von ihr lernen*, w: T. Metzinger, *Bewußtsein: Beiträge aus der Gegenwartsphilosophie* (Paderborn: Schöningh), ss. 463–490.
- Churchland P. M. 1985, *Reduction, Qualia and direct introspection of brain states*, „The Journal of Philosophy”, 82, ss. 8–28.
- Churchland P. M. 1989, *A Neurocomputational Perspective. The Nature of Mind and the Structure of Science*, (Cambridge/MA: Bradford Book).
- Churchland P. M. 1988, *Matter and Consciousness* (Cambridge/MA: MIT Press).
- Churchland P. M. 1997, *Die Seelenmaschine. Eine philosophische Reise ins Gehirn* (Heidelberg: Spektrum Akademischer Verlag GmbH). Oryginal: tenze, *The Engine of Reason, The Seat of the Soul* (Cambridge: MIT Press, 1995).
- Churchland P. M. 1998, *Knowing Qualia: A Reply to Jackson*, w: P.M. Churchland, P. S. Churchland, *On the Contrary: Critical Essays, 1987–1997* (Cambridge, MA: MIT Press), ss. 143–157.
- Clark A., D. Chalmers. 1998, *The Extended Mind*, „Analysis”, 58:1, ss. 7–19.
- Coghill R. C., Ch. N. Sang, J. M. A. Maisong. 1999, *Pain Intensity Proces-*

- ing *Within the Human Brain: A Bilateral Distributed Mechanism*, „Journal of Neurophysiology”, 82, ss. 1934–1943.
- Craig A. D. 2002, *How do you feel? Interoception: the sense of the physiological condition of the body*, „Neuroscience”, 3:8, ss. 655–666.
- Crick F. 1994, *Was die Seele wirklich ist. Die naturwissenschaftliche Erforschung des Bewußtseins* (München: Artemis & Winkler). Oryginał: tenże, *The Astonishing Hypothesis* (N. York: Maxwell Macmillan, 1994).
- Crick F. 1996, *Visual Perception: Rivalry and Consciousness*, „Nature”, 379, ss. 485–486.
- Crick F., C. Koch. 2003, *Problem świadomości*, „Świat Nauki”, 1, ss. 10–17.
- Crick F., C. Koch. 1990, *Towards a neurobiological theory of consciousness*, „Seminars in the Neuroscience”, 2, ss. 263–75.
- Crick F., C. Koch. 1991, *Understanding awareness at the neuronal level*, „Behavioral and Brain Sciences”, 14:4, ss. 683–685.
- Czarnecki P. 2002, *Koncepcja umysłu w filozofii D. Dennetta*, <http://www.kognitywistyka.net/artykuły/pcz_umysl_wg_dennetta.pdf>, ss. 1–69.
- Czuang-tsy, 1953, *Prawdziwa Księga Południowego Kwiatu* (Warszawa: PWN).
- Damasio A. R. 1992, *The selfless consciousness*, „Behavioral and Brain Sciences”, 15:2, ss. 208–209 („Open peer commentary” do: Dennett & Kinsbourne, *Time and the Observer*).
- Damasio A. R. 1995, *Descartes' Error. Emotion, Reason and the Human Brain* (London: Macmillan Publishers). Tłum. polskie: Damasio A., *Błąd Kartezjusza* (Poznań: Rebis, 2002).
- Damasio A. R. 2000, *Jak rodzi się świadomość?*, „Świat Nauki” (styczeń), ss. 76–81.
- Damasio A. R. 2003, *Jak mózg tworzy umysł*, „Świat Nauki” (styczeń), ss. 4–9.
- Daniluk B. 2000, *Deficyty poznawcze u osoby z postępującym schorzeniem mózgu o etiologii naczyniowej. Neuropsychologiczne studium przypadku*, w: A. Borkowska, E-M. Szepietowska (red.), *Diagnoza neuropsychologiczna. Metodologia i metodyka* (Lublin: Wydawnictwo UMCS), ss. 383–405.
- Daum I., H. Ackermann 1997, *Neuropsychological Abnormalities – Fact or Fiction?*, w: J. D. Schmahmann (ed.), *The Cerebellum and Cognition* (San Diego: Academic Press), ss. 455–471.
- Davidson D. 1986, *Mental Events*, w: tenże, *Essays on Action and Events* (Oxford: Clarendon Press), ss. 207–225.
- Davies W. M. 1996, *Experience and Content. Consequences of a Continuum Theory* (Aldershot: Avebury).

- Dawkins R. 1976, *The Selfish Gene* (Oxford: Oxford Univ. Press). Tłum. polskie: Dawkins R. 1998. *Samolubny Gen* (Warszawa: Prószyński i S-ka).
- Dawkins R. 1982, *The Extended Phenotype* (S. Francisco: Freeman). Tłum. polskie: Dawkins R. 2003. *Fenotyp rozszerzony* (Warszawa: Prószyński i S-ka).
- Dennett D. 1969, *Content and Consciousness* (London, Routledge & Kegan).
- Dennett D. 1979, *On the absence of phenomenology*, w: D. F. Gustafson, B. L. Tapscott (eds.), *Body, Mind and Method* (Dordrecht: D. Reidel), ss. 93–113.
- Dennett D. 1982, *How to study human consciousness empirically?*, „Synthese”, 53, ss. 159–187.
- Dennett D. 1987, *The intentional stance* (Cambridge/MA: MIT-Press, Bradford Book).
- Dennett D. 1988, *Quining qualia*, w: A. Marcel, E. Bisiach (eds.) *Consciousness in Modern Science* (Oxford: Oxford Univ. Press), ss. 42–77. <<http://ase.tufts.edu/cogstud/papers/quinal.htm>>, ss. 1–24.
- Dennett D. 1991, *Consciousness explained* (Boston: Little, Brown and Company).
- Dennett D. 1992, *The Self as a Center of Narrative Gravity*, w: F. Kessel, P. Cole, D. Johnson (eds.), *Self and Consciousness: Multiple Perspectives* (Hillsdale, NY: Erlbaum), ss. 103–115. <<http://ase.tufts.edu/cogstud/papers/selfctr.htm>>, ss. 1–7.
- Dennett D., M. Kinsbourne 1992a, *Time and the observer: The where and when of consciousness in the brain*, „Behavioral and Brain Sciences”, 15, ss. 183–247.
- Dennett D., M. Kinsbourne, 1995, *Multiple drafts: An eternal golden braid?*, „Behavioral and Brain Sciences”, 18:4, ss. 810–811. (Autors' Response).
- Dennett D. 1998, *Brainstorms. Philosophical Essays on Mind and Psychology* (Cambridge/MA: MIT-Press).
- Dretske F. 1994, *Differences that Make No Difference*, „Philosophical Topics”, 22:1&2, ss. 41–57.
- Duch W. 1996, *From brain to mind to consciousness – without hard problems* [Symposium kognitywne '96: Świadomość a percepcja. Instytut Fizyki UAM] <www.phys.uni.torun.pl/publications/kmk/96consc.pdf>, ss. 1–10.
- Duch W. 2002, *Wstęp do kognitywistyki*, <www.phys.uni.torun.pl/~duch/Wyklady/kog-m/03.htm> [wykład].
- Duch W. 2003, *Neurokognitywna teoria świadomości*, w: A. Klawiter, W. Dziarnowska (red.), *Subiektywność a świadomość. Studia z kognitywistyki i filozofii umysłu*, 1 (Poznań: Zysk i S-ka), ss. 133–153.

- Durstewitz D., S. Windmann. 1998, *Neuronale Korelate von Bewußtsein*, w: F. Esken, D. Heckmann (Hrsg.), *Bewußtsein und Repräsentation* (Paderborn [u.a.]: Schöningh), ss. 127–151.
- Duus P. 1989, *Diagnostyka topograficzna w neurologii* (Warszawa: PZWL).
- Dziarnowska W. 2001, *Jakości doznaniowe – subiektywny składnik świadomości czy funkcjonalny aspekt przytomności?*, w: P. Orlik (red.), *Meandry podmiotowości* (Poznań: Wydawnictwo Naukowe Instytutu Filozofii UAM), ss. 173–202.
- Eccles J. 1995, *Die Psyche des Menschen* (München: Ernst Reinhardt Verlag). Oryginał: tenże, *The Human Psyche* (Berlin: Springer, 1980).
- Edelman G. M. 1998, *Przenikliwe powietrze, jasny ogień. O materii umysłu* (Warszawa: PIW). Oryginał: tenże, *Bright air, brilliant fire. On the matter of the mind* (N. York: Basic Books, 1992).
- Eichenbaum, H. 1997, *Deklarative Memory: Insights from Cognitive Neurobiology*, „Annual Review of Psychology”, 48, ss. 547– 572.
- Engel A. K., P. Fries, P. König, H. Brecht, W. Singer. 1999, *Temporal Binding, Binocular Rivalry, and Consciousness*, „Consciousness and Cognition”, 8, 128–151.
- Engel A. K. 2000, *Explananda for a consciousness theory* [niepublikowany maszynopis, cytowany za zgodą autora] <<http://www.cognitive-science.net/Download/Download-St.Johann/Engel00-2.pdf>>, ss. 1–6.
- Engel A. K., W. Singer. 2001, *Temporal binding and the neural correlates of sensory awareness*, „Trends in Cognitive Sciences”, 5:1, ss. 16–25.
- Farah M. J. 1992, *The distributed pineal gland*, „Behavioral and Brain Sciences”, 15:2, s. 209 („Open peer commentary” do: Dennett & Kinsbourne, *Time and the Observer*).
- Feigl H. 1958, *The «Mental» and the «Physical»*, H. Feigl, M. Scriven, G. Maxwell (eds.) *Minnesota Studies in the Philosophy of Science*, 2 (Minneapolis: University of Minnesota Press), ss. 370–497.
- Feigl H. 1960, *Mind–Body, Not a Pseudoproblem*, w: S. Hook (ed.), *Dimensions of Mind* (N. York: N. York University Press), ss. 67–77.
- Fiez J. A., M. E. Raichle. 1997, *Linguistic Processing*, w: J. D. Schmahmann (ed.), *The Cerebellum and Cognition* (San Diego: Academic Press), ss. 233–254.
- Fix J. D. 1997, *Neuroanatomia* (Wrocław: Wydawnictwo Medyczne). Wyd. polskie pod redakcją J. Morysia.
- Flanagan O. 1992, *Consciousness reconsidered* (Cambridge/MA: MIT Press).
- Flohr H. 1997, *Denken und Bewußtsein*, w: H. J. Sandkühler (Hrsg.), *Interaktionen zwischen Wissenschaften und Philosophien* (Bremen: Zentrum für Philosophische Grundlagen der Wissenschaften), ss. 101–118.

- Fox P. 2000, *Research Imaging Center study sheds light on thirst and the brain*, <<http://www.uthscsa.edu/opa/issues/new33-10.html>>, ss. 1–3.
- Fraassen B. C., van 1980, *The Scientific Image* (Oxford: Clarendon Press).
- Frege G. 1977, *Myśl – studium logiczne*, w: tenże, *Pisma semantyczne* (Warszawa: PWN), ss. 101–129. Oryginał: tenże, *Der Gedanke. Eine logische Untersuchung* (1918).
- Gabora L. 2001, *Mikrotubule, anestetyki i świadomość kwantowa: Wywiad ze Stuartem Hameroffem*, „Zagadnienia Filozoficzne w Nauce”, 28/29, ss. 120–144.
- Gallagher S., J. Cole. 1995, *Body Image and Body Schema in a Deafferented Subject*, „Journal of Mind and Behavior”, 16, 369–390. <<http://www2.canisius.edu/~gallagher/gall&cole95.html>>, ss. 1–16.
- Gallagher S., F. Varela. 2001, *Redrawing the Map and Resetting the Time: Phenomenology and the Cognitive Sciences*, w: S. Crowell, L. Embree, S. J. Julian (eds.), *The Reach of Reflection: Issues for Phenomenology's Second Century*, <www.electronpress.com>, ss. 17–36.
- Garrett B. 1998, *Personal Identity and Self-consciousness* (London, N. York: Routledge).
- Gazzaniga M. S. 1998, *The Mind's Past* (Berkeley: University of California Press).
- Gazzaniga M. S., S. Gallagher. 1998a, *The Neural Platonist*, „Journal of Consciousness Studies”, 5–6, ss. 707–717.
- Gazzaniga M. S. 2003, *Podzielony mózg – odłona druga*, „Świat Nauki” (styczeń), ss. 26–31.
- Gennaro R. J. 1996, *Consciousness and Self-consciousness. A defense of the higher-order thought theory of consciousness* (Amsterdam/Philadelphia: John Benjamins Publishing Company).
- Glicksohn J. 1995, «*Multiple Drafts*» of subjective experience viewed within a microgenetic framework for cognition and consciousness, „Behavioral and Brain Sciences”, 18:4, ss. 807–808. („Continuing commentary” do: Dennett & Kinsbourne, *Time and the Observer*).
- Gorman J. M. 2000, [i inni] *Neuroanatomical Hypothesis of Panic Disorder, Revised*, „American Journal of Psychiatry”, 157, ss. 493–505.
- Grabowska A. 2000, *Asymetria półkul mózgowych*, w: T. Górka, A. Grabowska, J. Zagrodzka (red.), *Mózg a zachowanie* (Warszawa: WN PWN), ss. 400–428.
- Grimes T. R. 1988, *The myth of supervenience*, „Pacific Philosophical Quarterly”, 69, ss. 152–160.
- Grush R. 1998, *Wahrnehmung, Vorstellung und die sensomotorische Schleife*, F. Esken, D. Heckmann (Hrsg.), *Bewußtsein und Repräsentation* (Paderborn [u.a.]: Schöningh), ss. 169–192.

- Guttenplan S. 1995, *Externalism/internalism*, w: S. Guttenplan (red.), *A Companion to the Philosophy of Mind* (Oxford: Blackwell), ss. 289–290.
- Guttenplan S. 1995a, *Holism*, w: S. Guttenplan (ed.), *A Companion to the Philosophy of Mind* (Oxford: Blackwell), ss. 347.
- Hagenbüchle, R. 1998, *Subjektivität: Eine historisch-systematische Hinführung*, w: R. L. Fetz, R. Hagenbüchle, P. Schulz (Hrsg.), *Geschichte und Vorgeschichte der modernen Subjektivität* (Berlin: De Gruyter), ss. 81–88.
- Hallett G. 1977, *A Companion to Wittgenstein's «Philosophical Investigations»* (Ithaca [u.a.]: Cornell Univ. Press).
- Hardcastle V. G. 2000, *How to Understand the N in NCC*, w: T. Metzinger (ed.), *The Neural Correlates of Consciousness: Empirical and Conceptual Questions* (Cambridge/MA: MIT Press), ss. 259–264.
- Hebb D. O. 1980, *Essay on Mind* (Hiltsdale/NY: Erlbaum).
- Heeger D. J., D. Ress. 2002, *What does fMRI tell us about neuronal activity?*, „Nature”, 3, ss. 142–151.
- Hershfield J. 1998, *Lycan on the subjectivity of the mental*, „Philosophical Psychology”, 11:2, ss. 229–238.
- Hobbes T. 1984, *Leviathan* (Frankfurt/M: Suhrkamp).
- Hobson J. A., E. Pace-Schott, R. Stickgold. 2000, *Dreaming and the Brain: Toward a Cognitive Neuroscience of Conscious States*, „Behavioral and Brain Sciences”, 23:6, ss. 793–842.
- Hobson J. A. 2001, *Sleep*, w: R. A. Wilson, F. C. Keil (eds.), *The MIT Encyclopedia of the Cognitive Sciences* (Cambridge/MA: The MIT Press), ss. 772–775.
- Hobson J. A., E. Pace-Schott 2002, *The cognitive neuroscience of sleep: Neural systems, consciousness and learning*, „Neuroscience”, 3:9, ss. 679–693.
- Horgan T. 1984, *Functionalism, qualia and the inverted spectrum*, „Philosophy and Phenomenological Research”, 44, ss. 453–469.
- Horgan T. 1984a, *Jackson o Physical Information and Qualia*, „Philosophical Quarterly”, 34, ss. 147–152.
- Hume D. 1963, *Traktat o naturze ludzkiej* (Warszawa: PWN). Oryginał: tenże, *A Treatise of Human Nature* (Oxford: Clarendon Press, 1978).
- Hurley S. L. 1998, *Consciousness in Action* (Cambridge/MA: Harvard University Press).
- Ingarden R. 1948, *Spór o istnienie świata* (Kraków: Wydawnictwo PAU), t. II.
- Ingarden R. 1967, *Poglądy J. Volkelta na wczucie*, „Studia estetyczne”, 4, ss. 135–149.

- Ingarden R. 1972, *Książeczka o człowieku* (Kraków: Wydawnictwo Literackie).
- Ingarden R. 1971, *O poznawaniu cudzych stanów psychicznych*, w: tenże, *U podstaw teorii poznania* (Warszawa: PWN), ss. 407–427.
- Ishai A. 2002, *Streams of Consciousness*, „Journal of Cognitive Neuroscience”, 14:6, ss. 832–833.
- Jackson F. 1982, *Epiphenomenal Qualia*, „Philosophical Quarterly”, 32, ss. 127–136.
- Jacoby L. L., L. G. Allan, J. C. Collins, L. K. Larwill. 1988, *Memory influences subjective experience: noise judgements*, „Journal of Experimental Psychology”, 14, ss. 240–247.
- James W. 1904, *Does «Consciousness» Exist?*, „Journal of Philosophy, Psychology, and Scientific Methods”, 1, ss. 477–491.
- James, W. 1950, *The Principles of Psychology* (N. York: Dover Publications).
- James, W. 2001, *Doświadczenia religijne* (Kraków: Nomos).
- Johnson M. K., J. A. Reeder. 1997, *Consciousness as Meta-Processing*, w: J. D. Cohen, J. W. Schooler, *Scientific Approaches to Consciousness* (Mahwah: Lawrence Erlbaum Associates, Publishers), ss. 261–293.
- Judycki S. 1997, *Współczesna filozofia umysłu: uwagi metakrytyczne*, w: J. Perzanowski, A. Pietruszczak (red.), *Byt, logos, matematyka* (Toruń: Wydawnictwo Uniwersytetu Mikołaja Kopernika), ss. 31–46.
- Judycki S. 1999, *Intersubiektywność i czas. Przyczynek do dyskusji nad późną fazą poglądów Edmunda Husserla* (Lublin: Towarzystwo Naukowe KUL).
- Judycki S. 2001, *Idealizm i antyrealizm*, w: U. M. Żegleń (red.), *Pragmatyzm i filozofia Hilarego Putnama* (Toruń: Wydawnictwo Uniwersytetu Mikołaja Kopernika), ss. 117–128.
- Judycki S. 2002, *Introspekcja jako problem filozoficzny*, „Roczniki Filozoficzne”, 1, 263–301.
- Kaluża J. 1992, *Humanistyczne i ahumanistyczne aspekty redukcjonizmu we współczesnej medycynie*, w: J. A. Janik (red.), *Seminarium: Nauka-Religia-Dzieje, VI* (Kraków, Uniwersytet Jagielloński, 1992).
- Kant I. 1957, *Krytyka czystego rozumu* (Warszawa: PWN). Oryginał: tenże, *Kritik der reinen Vernunft*. Tłumaczył i wstępem opatrzył R. Ingarden.
- Kim J. 1990, *Supervenience as a philosophical concept*, „Metaphilosophy”, 21, ss. 1–27.
- Kim J. 1993, *Psychophysical Supervenience*, w: tenże: *Supervenience and Mind. Selected Philosophical Essays* (Cambridge: Cambridge Univ. Press), ss. 175–193.

- Kim J. 2002, *Umysł w świecie fizycznym. Esej na temat problemu umysłu i ciała oraz przyczynowania mentalnego* (Warszawa: Wydawnictwo IFiS PAN). Tłumaczenie: R. Poczobut.
- Kinsbourne M. 1997, *What Qualifies a Representation for a Role in Consciousness*, w: J. D. Cohen, J. W. Schooler, *Scientific Approaches to Consciousness* (Mahwah: Lawrence Erlbaum), ss. 335–355.
- Klawiter A. 1999, *Wprowadzenie: O naturze problemu kognitywnego*, w: A. Klawiter, L. Nowak, P. Przybysz, *Umysł a rzeczywistość. Poznańskie Studia z Filozofii Humanistyki*, 5 (16) (Poznań: Zysk i S-ka), ss. 135–138.
- Klawiter A. 1999a, *O słyszeniu przedmiotów*, w: A. Klawiter, L. Nowak, P. Przybysz, *Umysł a rzeczywistość. Poznańskie Studia z Filozofii Humanistyki*, 5 (16) (Poznań: Zysk i S-ka), ss. 327–339.
- Kloch J. 1996, *Świadomość komputerów? Argument «Chińskiego pokoju» w krytyce mocnej sztucznej inteligencji według Johna Searle'a* (Kraków/Tarnów: OBI/Biblos).
- Koch Ch. 2004, *Francis Crick – ein Leben für die Wissenschaft. Von der Molekularbiologie zur Bewusstseinsforschung*, „Neue Zürcher Zeitung”, 239, s. 55.
- Konorski J. 1969, *Integracyjna działalność mózgu* (Warszawa: PWN).
- Korb K. B. 1993, *Stage Effects in the Cartesian Theater: A review of Daniel Dennett's Consciousness Explained*, „Psyche”, 1:4, ss. 33–43.
- Kornblith H. 1998, *What is it like to be Me?*, „Australasian Journal of Philosophy”, 76, ss. 48–60.
- Kossut M. 2002, *Wstęp do neuroplastyczności*, „Neurologia i Neurochirurgia Polska”, 36, Supl. 1., ss. 11–21.
- Kripke A. S. 1988, *Nazywanie a konieczność* (Warszawa: PAX). Oryginał: tenże, *Naming and Necessity* (Oxford: Basil Blackwell, 1980).
- Krzyżewski K. 1989, *Perspektywy i ograniczenia neobehawioryzmu. Neobehawiorystyczne ujęcie problematyki podmiotu zachowania* (Kraków: Wydawnictwo Uniwersytetu Jagiellońskiego).
- Kuhn T. 1968, *Struktura rewolucji naukowych* (Warszawa: PWN).
- Kuhn T. 1985, *Dwa bieguny* (Warszawa: PIW).
- Kulczycki J. 1996, *Problem psychofizyczny w filozofii H. Bergsona* (Kraków: Wydawnictwo Naukowe WSP).
- Kutschera F. v. 1982, *Grundfragen der Erkenntnistheorie* (Berlin: Walter de Gruyter).
- Lau J. 2002, *Externalism about Mental Content*, w: <<http://plato.stanford.edu/entries/content-externalism/>>, ss. 1–12.
- Leibniz G. W. 1995, *Zasady filozofii, czyli monadologia*, w: tenże, *Główne pisma metafizyczne* (Toruń: Comer).

- Leiner H. C. i A. I. Leiner. 1997, *How fibers subserve computing capabilities: Similarities between brains and machines*, w: J. D. Schmahmann (ed.), *The Cerebellum and Cognition* (San Diego: Academic Press), ss. 535–553.
- Lenk H. 2001, *Kleine Philosophie des Gehirns* (Darmstadt: Wissenschaftliche Buchgesellschaft).
- LePore E., B. McLaughlin (eds.) 1985, *Actions and Events. Perspectives on the Philosophy of Donald Davidson* (London: Basil Blackwell).
- LePore E., B. Loewer 1987, *Mind Matters*, „Journal of Philosophy”, 84, ss. 630–641.
- Levin J. 1986, *Could Love be like a Heatwave? Physicalism and the Subjective Character of Experience*, „Philosophical Studies”, 49, ss. 245–261.
- Levine J. 1983, *Materialism and qualia: The explanatory gap*, „Pacific Philosophical Quarterly”, 64, ss. 354–361.
- Levine J. 1997, *Recent Work on Consciousness*, „American Philosophical Quarterly”, 34, ss. 379–404.
- Lewin R. 1980, *Is Your Brain really Necessary?*, „Science”, 210:12, ss. 1232–1234.
- Lewis D. K. 1983, *An Argument for the Identity Theory*, w: tenże, *Philosophical Papers*, 1 (N. York: Oxford University Press), ss. 99–107.
- Llinas R., Ribary U. [i inni] 1998, *The neuronal basis for consciousness*, „Philosophical Transactions. R. Soc. London Ser. B”, 353, ss. 1841–1849.
- Lloyd D. 2003, *Radiant Cool. A Novel Theory of Consciousness* (Cambridge/MA: A Bradford Book).
- Loar B. 1990, *Phenomenal States*, „Philosophical Perspectives”, 4, ss. 81–108.
- Locke J. 1955, *Rozważania dotyczące rozumu ludzkiego* (Warszawa: PWN).
- Logothetis N. K., J. D. Schall. 1989, *Neural correlates of subjective visual experience*, „Science. American Association for the Advancement of Science”, 245, ss. 761–763.
- Lowe E. J. 2000, *The Introduction to the Philosophy of Mind* (Cambridge: Cambridge Univ. Press).
- Lund D. H. 1994, *Perception, Mind and Personal Identity. A Critique of Materialism* (Lanham: University Press of America).
- Lycan W. G. 1987, *Consciousness and Experience* (Cambridge/MA: MIT Press).
- Lycan W. G. 1990, *What is the «subjectivity» of the mental?*, w: J. Tomberlin (ed.), *Action, Theory and Philosophy of Mind*, „Philosophical Perspectives”, 4 (California: Ridgeview Publishing), ss. 109–130.
- Lycan W. G. 1992, *UnCartesian materialism and Lockean introspection*, „Behavioral and Brain Sciences”, 15:2, ss. 216–217 („Open peer commentary” do: Dennett & Kinsbourne, *Time and the Observer*).

- Lycan W. G. 1996, *Consciousness and Experience* (Cambridge/MA: MIT Press/Bradford Books).
- Lycan W. G. 1997, *Consciousness as internal monitoring*, w: N. Block, O. Flanagan (eds.), *The Nature of Consciousness* (Cambridge/MA: MIT Press), ss. 755–771.
- Luria A. 1967, *Zaburzenia wyższych czynności korowych wskutek ogniskowych uszkodzeń mózgu* (Warszawa: PWN).
- Luria A. 1976, *Świat utracony i odzyskany. Historia pewnego zranienia* (Warszawa: PWN).
- Luria A. 1976a, *Problemy neuropsychologii i neurolingwistyki. Wybór prac* (Warszawa: PWN).
- Madell G. 1981, *The Identity of the Self* (Edinburgh: Edinburgh Univ. Press).
- Malcolm N. 1995, *Subjectivity*, w: G. H. Wright (ed.), *Norman Malcolm – Wittgensteinian Themes. Essays 1978–1989* (Ithaca and London: Cornell Univ. Press), ss. 118–132.
- Maloney Ch. J. 1985, *About Being a Bat*, „Australasian Journal of Philosophy”, 63, ss. 26–49.
- Maloney Ch. J. 2001, *Functionalism*, w: R. A. Wilson, F. C. Keil (eds.), *The MIT Encyclopedia of the Cognitive Sciences* (Cambridge/MA: The MIT Press), ss. 332–334.
- Manzotti R. 1998, <http://www.manzotti.lira.dist.unige.it/consciousness/inverted_spectrum.html>.
- Marciszewski W. 2002, Jak się ma kod neuronowy do kodu maszynowego? <<http://www.calculumus.org/x/doIX-2002/kodneur.html>>, ss. 1–15.
- Marciszewski W. 2004, Dlaczego nie każde rozumowanie da się zmechanizować?, <<http://www.calculumus.org/forum/5/wol-marc.html>>, ss. 1–22.
- McGinn C. 1977, *Anomalous Monism and Kripke's Cartesian Intuitions*, „Analysis”, 37:2, ss. 78–80.
- McGinn C. 1983, *The Subjective View. Secondary Qualities and Indexical Thoughts* (Oxford: Clarendon Press).
- McGinn C. 1991, *The Problem of Consciousness. Essays Towards a Resolution* (Oxford: Blackwell).
- McGinn C. 1995, *Consciousness and Space*, „Journal of Consciousness Studies”, 2, ss. 220–230.
- McGinn C. 1997, *Nagel: The View from Nowhere*, tenże: *Minds and Bodies. Philosophers and their Ideas* (N. York, Oxford: Oxford University Press), ss. 88–99.
- McGinn C. 1999, *The Mysterious Flame. Conscious Mind in a Material World* (N. York: Basis Books).

- McTaggart J. 1927, *The Nature of Existence*, Vol. II (Cambridge: Cambridge University Press).
- Mehl P., W. Sellars. 1956, *The Concept of Emergence*, H. Feigl, M. Scriven, G. Maxwell (eds.), *Minnesota Studies in the Philosophy of Science*, 1 (Minneapolis: University of Minnesota Press), ss. 239–252.
- Mele A. R. 1997, *Real Self-Deception*, „Behavioral and Brain Sciences”, 20, ss. 91–136.
- Merikle P. M., M. Daneman 1996, *Memory for Unconsciously Perceived Events: Evidence from Anesthetized Patients*, „Consciousness and Cognition”, 5, ss. 525–541.
- Merikle P. M. 2000, *Subliminal Perception*, w: A. E. Kazdin (ed.), *Encyclopedia of Psychology*, t. 7 (N. York: Oxford University Press), ss. 497–499.
- Metzinger T. 1996, *Einleitung: Das Problem des Bewußtseins*, w: T. Metzinger (Hrsg.), *Bewußtsein. Beiträge aus der Gegenwartsphilosophie* (Paderborn: Schöningh), ss. 15–53.
- Metzinger T. (ed.) 2000, *The Neural Correlates of Consciousness: Empirical and Conceptual Questions* (Cambridge/MA: MIT, Press).
- Münte T. F., E. Altenmüller, L. Jäncke 2002, *The musician's brain as a model of neuroplasticity*, „Nature”, 3, ss. 473–478.
- Nagel E. 1970, *Struktura nauki* (Warszawa: PWN, 1970). Oryginał: tenże, *The Structure of Science* (New York: Harcourt, Brace World, INC., 1961).
- Nagel T. 1965, *Physicalism*, „The Philosophical Review”, 74, ss. 339–356. Przedruk: C. V. Borst (ed.), *The Mind and Brain Identity Theory* (London: Macmillan, 1970), ss. 214–230.
- Nagel T. 1991, *Bewußtsein*, w: tenże, *Die Grenzen der Objektivität* (Stuttgart: Reclam), ss. 11–38.
- Nagel T. 1997, *Jak to jest być nietoperzem?*, w: tenże: *Pytania ostateczne* (Warszawa: Aletheia), ss. 203–219. Oryginał: tenże, *What Is It Like to Be a Bat?*, reprinted in: *Mortal Questions* (Cambridge/MA: Cambridge Univ. Press, 1979), ss. 165–190.
- Nagel T. 1997a, *Rozszczepienie mózgu i jedność świadomości*, w: tenże, *Pytania ostateczne* (Warszawa: Aletheia), ss. 183–202. Oryginał: tenże, *Brain Bisection and the Unity of Consciousness*, „Synthese”, 22 (1971), ss. 396–413.
- Nagel T. 1997b, *Subiektywne i obiektywne*, w: tenże, *Pytania ostateczne* (Warszawa: Aletheia), ss. 237–257. Oryginał: *Mortal Questions* (Cambridge University Press, 1979).
- Nagel T. 1997c, *Widok znikąd* (Warszawa: Aletheia). Oryginał: *The View from Nowhere* (Oxford: Oxford Univ. Press, 1986).

- Nagel T. 1997d, *Panpsychizm*, w: tenże, *Pytania ostateczne* (Warszawa: Aletheia, 1997), ss. 221–236. Oryginał: tenże, *Mortal questions*, Cambridge, 1979.
- Nagel T. 1998, *Co to wszystko znaczy? Bardzo krótkie wprowadzenie do filozofii* (Warszawa: Aletheia). Oryginał: *What Does it All Mean?* (Oxford: Oxford University Press, 1987).
- Nagel T. 1998a, *Conceiving the Impossible and the Mind–Body Problem*, „Philosophy”, 73, ss. 337–352.
- Neisser U. 1997, *The Roots of Self-Knowledge: Perceiving Self, It and Thou*, J. G. Snoggrass, R. L. Thompson (eds.), *The Self accros Psychology* (New York: The N. York Academy of Science), ss. 19–33.
- Nemirow L. 1990, *Physicalism and the Cognitive Role of Acquaintance*, W. Lycan (ed.) *Mind and Cognition: A Reader* (Oxford: Blackwells), ss. 490–499.
- Newell R. W. 1986, *Objectivity, Empiricism and Truth* (London: Routledge & Kegan Paul).
- Northoff G. 1995, *Qualia im Knotenpunkt zwischen Leib und Seele: „Argumentatives” Dilemma in der gegenwärtigen Diskussion über die Subjektivität mentaler Zustände*, „Journal for General Philosophy of Science”, 26, ss. 269–295.
- Northoff G. 2001, *Neurophilosophie – ein neuer Ansatz in der Philosophie* „Information Philosophie”, 2001:1, ss. 24–35.
- O’Brien G., J. Opie. 1998, *The Disunity of Consciousness*, „Australasian Journal of Philosophy”, 76, 378–95.
- O’Keefe J. 1976, *Place units in the hippocampus of the freely moving rat*, „Experimental Neurology”, 51, 78–109.
- O’Keefe J., N. Burgess [i inni] 1998: *Place cells, navigational accuracy, and the human hippocampus*, „Philosophical Translations of the Royal Society in London”, 353, ss. 1333–1340.
- Papathanasiou I. 2003, *Nervous System Mechanisms of Recovery and Plasticity Following Injury*, „Acta Neuropsychologica”, 1:3, ss. 345–354.
- Parsons L. M. 2001, *Integrating cognitive psychology, neurology and neuro-imaging*, „Acta Psychologica”, 107, ss. 155–181.
- Parsons L. M., P. T. Fox. 1997, *Sensory and Cognitive Functions*, w: Schmahmann, J. D. (ed.), *The Cerebellum and Cognition* (San Diego: Academic Press), ss. 255–271.
- Peacocke Ch. 1989, *No Resting Place: A Critical Notice of The View from Nowhere, by Thomas Nagel*, „The Philosophical Review”, 98, ss. 65–82.
- Penrose R. 1989, *The Emperor’s New Mind. Concerning Computers, Minds, and the Laws of Physics* (Oxford: Oxford University Press), tenże, *Nowy umysł cesarza* (Warszawa: PWN, 1995).

- Perzanowski J. 1993, *Locative ontology*, „Logic and Logical Philosophy”, 1, ss. 9–94.
- Perzanowski J. 1994, *Teofilozofia Leibniza*, w: G. W. Leibniz, *Pisma z teologii mistycznej*. Tłumaczenie i opracowanie naukowe M. Frankiewicz (Kraków: Znak), ss. 241–351.
- Perzanowski J. 1995, *W stronę psychoontologii*, „Filozofia Nauki”, 1–2, ss. 15–24.
- Piłat R. 1993, *Czy istnieje świadomość?* (Warszawa: Wydawnictwo IFiS PAN).
- Platon, 1991 *Gorgias, Menon* (Warszawa: PWN). Tłumaczenie: P. Siwek.
- Poczobut R. 2000, *Superweniencja. Zarys problematyki*, „Filozofia Nauki”, 2, ss. 25–44.
- Poczobut R. 2002, *Odmiany emergencji w zastosowaniu do ontologii umysłu*, „Roczniki Filozoficzne”, 1, ss. 403–427.
- Poczobut R. [bdn.], *Uwagi do tekstów W. Strawieńskiego i E. Kałuszyńskiej*, <<http://www.filozofia.pl/seminarium/KomentarzeEMERG.doc>>, ss. 1–8.
- Popper K., J. Eccles. 1977, *The Self and Its Brain* (N. York: Springer).
- Putnam H. 1975, *Introduction*, w: tenże, *Mind, Language and Reality* (Cambridge: Cambridge University Press), ss. vii–xvii.
- Putnam H. 1975a, *The Meaning of «Meaning»*, w: tenże, *Mind, Language and Reality: Philosophical Papers*, Vol. II (Cambridge: Cambridge University Press), ss. 215–271. Tłumaczenie: H. Putnam, *Znaczenie wyrazu «znaczenie»*, w: tenże: *Wiele twarzy realizmu i inne eseje*, Warszawa WN PWN, 1998, ss. 93–184.
- Putnam H. 1981, *Reason, Truth and History* (Cambridge: Cambridge University Press).
- Putnam H. 1998, *Wiele twarzy realizmu i inne eseje* (Warszawa: WN PWN).
- Putnam H. 1999, *Pragmatyzm. Pytania otwarte* (Warszawa: Aletheia).
- Putnam H. 1999a, *Repräsentation und Realität* (Frankfurt: Suhrkamp). Oryginał: H. Putnam, *Representation and Reality* (Cambridge/MA: MIT Press 1998).
- Quine W. V. O. 1999, *Słowo i przedmiot* (Warszawa: Aletheia). Oryginał: *Word and Object* (MIT, 1960).
- Quine W. V. O. 1979 *Zwei Dogmen des Empirismus*, w: tenże, *Von einem logischen Standpunkt* (Frankfurt/M: Ullstein), ss. 27–50.
- Raichle M. E. 1998, *The neural correlates of consciousness*, „Philosophical Transactions. R. Soc. London Ser. B”, 353, ss. 1889–1901.
- Railton P. 1995, *Realism*, w: J. Kim, E. Sosa (eds.), *A Companion to Metaphysics* (Oxford: Blackwell), ss. 433–437.
- Ramachandran V. S. 1998, *Consciousness and body image: lessons from phantom limbs, Capgras syndrome and pain asymbolia*, Phil. Trans. R. Society London (B) 353, ss. 1851–1859.

- Ravenscroft I. 1997, *Folk Psychology as a Theory*, <<http://www.plato.stanford.edu/entries/folkpsych-theory/>>, ss. 1–7.
- Rees G. 2001, *Can Philosophy Discover Consciousness In the Brain?*, „Journal of Consciousness Studies”, 8:3, ss. 35–40.
- Rees G., G. Kreiman, C. Koch. 2002, *Neural Correlates of Consciousness in Humans*, „Neuroscience”, 3:4, ss. 261–270.
- Reis C. E. 1996, *Brain Death Criteria*, <<http://www.medstudents.com.br/neuro/neuro5.htm>>, ss. 1–2.
- Rescher N. 1991, *Thought Experimentation in Presocratic Philosophy*, T. Horowitz, G. Massey (eds.), *Thought Experiments in Science and Philosophy* (Maryland: Rowman & Littlefield), ss. 31–42.
- Revonsuo A. 1995, *Conscious and nonconscious control of action*, „Behavioral and Brain Sciences”, 18:2, ss. 265–266.
- Revonsuo A. 1998, *Wie man Bewußtsein in der kognitiven Neurowissenschaft ernst nehmen kann*, w: F. Esken, D. Heckmann (Hrsg.), *Bewußtsein und Repräsentation* (Paderborn [u.a.]: Schöningh), ss. 193–216.
- Revonsuo, A. 2000, *Prospects for a Scientific Research Program on Consciousness*, w: T. Metzinger (ed.), *The Neural Correlates of Consciousness: Empirical and Conceptual Questions* (Cambridge/MA: MIT, Press), ss. 57–75.
- Revonsuo A. 2001, *Can functional brain imaging discover consciousness in the brain?*, „Journal of Consciousness Studies”, 8:3, ss. 3–23.
- Rey G. 1994, *Dennett's Unrealistic Psychology*, „Philosophical Topics”, 22:1&2, ss. 259–289.
- Robertson L. C. 2003, *Binding, spatial attention and perceptual awareness*, „Neuroscience”, 4:2, ss. 93–102.
- Robinson W. S. 1996, *The Hardness of the hard Problem*, „Journal of Consciousness Studies”, 3, ss. 14–25.
- Rockwell T. 1996, *Awareness, mental phenomena and consciousness. A synthesis of Dennett and Rosenthal*, „Journal of Consciousness Studies”, 3:5–6, ss. 463–476.
- Rorty R. 1972, *Functionalism, Machines and Incorrigibility*, „Journal of Philosophy”, 69:8, ss. 203–220.
- Rosenthal D. M. 1992, *Time and consciousness*, „Behavioral and Brain Sciences”, 15:2, ss. 220–221 („Open peer commentary” do: Dennett & Kinsbourne, *Time and the Observer*).
- Rosenthal D. M. 1996, *Identity Theories*, w: S. Guttenplan, *A Companion to the Philosophy of Mind* (Oxford: Blackwell), ss. 348–355.
- Rosenthal D. M. 1996a, *Mehrfache Entwürfe und unumstößliche Tatsachen*, w: T. Metzinger (Hrsg.), *Bewußtsein. Beiträge aus der Gegenwartsphilosophie* (Paderborn: Schöningh), ss. 423–438.

- Rosenthal D. M. 1998, *The Kinds of Consciousness*, <<http://web.gc.cuny.edu/cogsci/dr.htm#publications>>, ss. 1–12.
- Rosenthal D. M. 2002, *Consciousness and the Mind*, „Iyyun”, 51, ss. 227–251.
- Roth G. 1999, *Kleine Gehirne – grosse Gehirne*, „Naturwissenschaftliche Rundschau”, 52:6, ss. 213–219.
- Russell B. 1959, *The Problems of Philosophy* (Oxford: Oxford University Press).
- Ryle G. 1970, *Czym jest umysł?* (Warszawa: PWN). Oryginał: tenże, *The Concept of Mind* (1949).
- Sacks O. 1996, *Mężczyzna, który pomylił swoją żonę z kapeluszem* (Warszawa: Wydawnictwo Zysk i S-ka). Oryginał: tenże, *The Man Who Mistook His Wife for a Hat* (N. York: Summit Books, 1985).
- Salter D. 1995, *Consciousness and timing*, „Behavioral and Brain Sciences”, 18:4, ss. 808–810. (Continuing commentary do: Dennett & Kinsbourne, *Time and the Observer*).
- Scheler M. 2000, *Grammatik der Gefühle. Das Emotionale als Grundlage der Ethik* (München: Deutscher Taschenbuch Verlag).
- Schmahmann J. D., J. C. Sherman. 1997, *Cerebellar Cognitive Affective Syndrome*, w: J. D. Schmahmann (ed.), *The Cerebellum and Cognition* (San Diego: Academic Press), ss. 433–440.
- Searle J. R. 1980, *Minds, brains, and programs*, „Behavioral and Brain Sciences”, 3, ss. 417–57.
- Searle J. R. 1987, *Indeterminism, Empiricism and the First Person*, „The Journal of Philosophy”, 84, ss. 123–46.
- Searle J. R. 1999, *Umysł na nowo odkryty* (Warszawa: PIW). Oryginał: tenże, *The Rediscovery of the Mind* (Cambridge/MA: MIT-Press, 1992).
- Searle J. R. 1991, *Intentionalität. Eine Abhandlung zur Philosophie des Geistes* (Frankfurt/M: Suhrkamp). Oryginał: tenże, *Intentionality. An essay in the philosophy of mind* (N. York: Cambridge University Press, 1983).
- Searle J. R. 1998, *The Mystery of Consciousness* (London: Granta Books).
- Searle J. R. 2000, *Consciousness*, „Annual Reviews of Neuroscience”, 23, ss. 557–578.
- Sellars W. 1962, *Time and the World Order*, w: H. Feigl, M. Scriven, G. Maxwell (eds.), *Minnesota Studies in the Philosophy of Science*, Vol. 3., (Minneapolis: Univ. of Minnesota Press), ss. 527–616.
- Sellars W. 1963, *Philosophy and the Scientific Image of Man*, w: tenże, *Science, Perception and Reality* (London: Routledge & Kegan Paul), ss. 1–40.
- Sellars W. 1963a, *The Language of Theories*, w: tenże, *Science, Perception and Reality* (London: Routledge & Kegan Paul), ss. 106–126.

- Sellars W. 1963b, *Empiricism and the Philosophy of Mind*, w: tenże, *Science, Perception and Reality* (London: Routledge and Kegan), ss. 127–196.
- Sellars W. 1967, *Scientific Realism or Irenic Instrumentalism. A Critique of Nagel and Feyerabend on Theoretical Explanation*, w: tenże, *Philosophical Perspectives* (Springfield, Il: Charles C. Thomas Publisher), ss. 337–369.
- Sellars W. 1975, *The Structure of Knowledge*, w: H-N. Castañeda (ed.) *Action, Knowledge and Reality: Studies in Honor of Wilfrid Sellars* (Indianapolis: Bobbs–Merrill), ss. 295–347.
- Shallice T. 1972, *Dual functions of consciousness*, „Psychological Review”, 79, ss. 383–393.
- Shoemaker S. 1982, *The inverted spectrum*, „The Journal of Philosophy”, 79, ss. 357–381.
- Shoemaker S. 1983, *On an Argument for Dualism*, w: S. Shoemaker, C. Ginet, *Knowledge and Mind. Philosophical Essays* (Oxford: Oxford University Press), ss. 233–259.
- Shoemaker S. 1984, *Functionalism and Qualia*, w: tenże, *Identity, cause and mind. Philosophical essays* (Cambridge: Cambridge University Press), ss. 184–205.
- Shoemaker S. 1984, *Some varieties of functionalism*, w: tenże, *Identity, cause and mind. Philosophical essays* (Cambridge: Cambridge University Press), ss. 261–286.
- Shoemaker S. 1991, *Qualia and Consciousness*, „Mind”, 100, ss. 507–524.
- Shoemaker S. 1993, *Lovely and Suspect Ideas*, „Philosophy and Phenomenological Research”, 53, ss. 905–910.
- Shoemaker S. 1996, *Qualities and Qualia*, w: tenże, *The first-person perspective and other essays* (Cambridge: Press Syndicate of the Univ. of Cambridge), ss. 97–120.
- Shoemaker S. 1996a, *The first-person perspective*, w: tenże, *The first-person perspective and other essays* (Cambridge: Press Syndicate of the Univ. of Cambridge), ss. 157–175.
- Siewert Ch. 1998, *The Significance of Consciousness* (Princeton: Princeton Univ. Press).
- Siegal M., R. Varley 2002, *Neural systems involved in «Theory of mind»*, „Neuroscience”, 3:6, ss. 463–471.
- Singer T. [i inni] 2004, *Empathy for pain involves the affective but not sensory components of pain*, „Science”, 303, ss. 1157–1162.
- Skoyles J. R. 1999, *Human Evolution Expanded Brains to Increase Expertise Capacity, Not IQ*, „Pscycology” <<http://psycprints.ecs.soton.ac.uk/>>, ss. 1–13.

- Smart J. J. C. 1959, *Sensations and Brain Processes*, „The Philosophical Review”, 68, ss. 141–156.
- Sousa R. 2002, *Twelve Varieties of Subjectivity*, w: J. L. Larrazabal, L. P. Miranda, *Knowledge, Language and Representation* (Dordrecht: Kluwer), ss. 135–151, <www.chass.utoronto.ca/~sousa/subjectivity/html>, ss. 1–14.
- Sperry R. W. 1969, *A Modified Concept of Consciousness*, „Psychological Review”, 76, ss. 532–536.
- Sperry R. W. 1990, *Forebrain commissurotomy and conscious awareness*, w: C. Trevarthen (ed.), *Brain circuits and functions of the mind. Essays in honor of Roger W. Sperry* (Cambridge: Cambridge Univ. Press.), ss. 369–388.
- Spinner H. F. 1969, *Wege und Irrwege der Wissenschaft*, „Soziale Welt”, 20, ss. 329–358.
- Stein E. 1988, *O zagadnieniu wczucia* (Kraków: Znak).
- Stephan A. 1999, *Emergenz. Von der Unvorhersagbarkeit zur Selbstorganisation* (Dresden: Dresden University Press).
- Stephan A. 2002, *Emergentism, Irreducibility, and Downward Causation*, „Grazer Philosophische Studien”, 65, ss. 77–91.
- Studen M. 1998, *Przegląd technik badawczych ośrodkowego układu nerwowego*, w: A. Herzyk, D. Kądziaława (red.), *Związek mózg–zachowanie w ujęciu neuropsychologii klinicznej* (Lublin: Wydawnictwo Uniwersyte-tu Marii Curie Skłodowskiej) ss. 37–60.
- Strawiński W. 1997, *Jedność nauki, redukcja, emergencja* (Warszawa: Fundacja Aletheia).
- Strawson G. 1994, *Mental Reality* (Cambridge/MA: MIT-Press).
- Strian F. 1996, *Schmerz. Ursachen, Symptome, Therapien* (München: C.H. Beck).
- Stubenberg L. 1998, *Consciousness and qualia* (Amsterdam: John Benjamins Publishing Company).
- Szafer W. 1969, *Kwiaty i zwierzęta. Zarys ekologii kwiatów* (Warszawa: PWN).
- Tadeusiewicz R. 1993, *Sieci neuronowe* (Warszawa: Akademicka Oficyna Wydawnicza).
- Tennant N. 1995, *Antirealism*, w: J. Kim, E. Sosa (eds.), *A Companion to Metaphysics* (Oxford: Blackwell), ss. 14–18.
- Thach W. T. 1997, *Context-respose linkage*, w: J. D. Schmammann (ed.), *The Cerebellum and Cognition* (San Diego: Academic Press), ss. 599–611.
- Thompson E. 2001, *Empathy and Consciousness*, „Journal of Consciousness Studies”, 8:5–7, ss. 1–32.
- Travis J. 2002, *Sizing Up the Brain*. „Science News Online”, 162:20, <<http://www.sciencenews.org/20021116/bob8.asp>>, ss. 1–5.

- Trąbka J. 1991, *Mózg i jego jaźń* (Kraków: Drukarnia Uniwersytetu Jagiellońskiego).
- Twardowski K. 1997 (1897), *Psychologia wobec fizjologii i filozofii*, w: T. Rzepa (red.) *Psychologia w szkole lwowsko-warszawskiej* (Warszawa WN PWN).
- Tye M. 1996, *Ten problems of consciousness* (Cambridge/MA: MIT Press).
- Tye M. 2003, *Qualia*, „Stanford Encyclopedia of Philosophy”, <<http://plato.stanford.edu/entries/qualia>>, ss. 1–14.
- Vaas R. 1999, *Why Neural Correlates of Consciousness are Fine, but Not Enough*, „Anthropology and Philosophy”, 3:3, ss. 121–141.
- Vanderwolf C. H. 1998, *Brain, Behavior, and Mind: What do we Know and What can we Know?*, „Neuroscience & Behavioral Reviews”, 22:2, ss. 125–142.
- Van Gulick R. 1992, *Time for more alternatives*, „Behavioral and Brain Sciences”, 15:2, ss. 228–229 („Open peer commentary” do: Dennett & Kinsbourne, *Time and the Observer*).
- Varela F. J., E. Thompson, E. Rosch. 1991, *The Embodied Mind. Cognitive Science and Human Experience* (Cambridge: MIT Press).
- Varela F. J. 2000, *Neurophenomenology: A Methodological Remedy for the Hard Problem*, w: J. Shear (ed.), *Explaining Consciousness* (Cambridge/MA: MIT Press), ss. 337–358.
- Varela F. J., Shear J. 2000a, *Editors' Rejoinder to the Debate*, w: F.J. Varela, J. Shear, *The View from Within. First-person approaches to the study of consciousness* (Thorverton: Imprint Academic), ss. 307–311.
- Velmans M. 1992, *Is consciousness integrated?*, „Behavioral and Brain Sciences”, 15:2, ss. 229–230 („Open peer commentary” do: Dennett & Kinsbourne, *Time and the Observer*).
- Voorhees B. 2000, *Dennett and the Deep Blue Sea*, „Journal of Consciousness Studies”, 7:3, ss. 53–69.
- Wallace A. B. 2000, *The Taboo of Subjectivity: Towards a New Science of Consciousness* (Oxford: Oxford University Press).
- Walsh K. 2000, *Neuropsychologia kliniczna* (Warszawa: WN PWN).
- Walter H. 1996, *Minimale Neuropilosophie*, w: Ch. Hubig, H. Poser (Hrsgs.), *Cognitio humana – Dynamik des Wissens und der Werte* (Leipzig: Allgemeine Gesellschaft für Philosophie in Deutschland), ss. 1515–1522.
- Wheelwell D. 1998, *And the meme raths outgrabe*, „Journal of Consciousness Studies”, 5:3, ss. 362–373.
- Wiener D. 2000, *Koncepcja świadomości Antonio Damasio przedstawiona w książce „The Feeling of what happens: body emotion in making of consciousness”*, <<http://kognitywistyka.prv.pl>>, ss. 1–15.

- Wilson R. A., F. C. Keil 2001 (eds.), *The MIT Encyclopedia of the Cognitive Sciences* (Cambridge/MA: The MIT Press).
- Wittgenstein L. 1950, *Taschennotizbuch*, w: *Wittgenstein's Nachlass. The Bergen Electronic Edition 2000* (Oxford University Press, University of Bergen, The Wittgenstein Trustes), 175.5v.
- Wittgenstein L. 1984, *Das Blaue Buch* (Frankfurt/M: Suhrkamp), ss. 15–116. Werkausgabe, Band 5. Tłumaczenie polskie: *Niebieski i brązowy zeszyt* (Warszawa: Spacja, 1998).
- Wittgenstein L. 1995, *Notatniki 1914–1916 (Fragmenty)*, w: W. Sady, *Uwagi o religii i etyce* (Kraków: Znak), ss. 58–74. Oryginał: tenże, *Tagebücher 1914–1916* (Frankfurt/M: Suhrkamp, 1989), ss. 89–187. Werkausgabe, Band 1.
- Wittgenstein L. 1995a, *Uwagi różne*, w: W. Sady, *Uwagi o religii i etyce* (Kraków: Znak), ss. 108–206. Oryginał: tenże, *Vermischte Bemerkungen* (Frankfurt/M: Suhrkamp, 1984), ss. 451–573. Werkausgabe, Band 8.
- Wittgenstein L. 1997, *Tractatus logico-philosophicus* (Warszawa: WN PWN). Oryginał: *Tractatus logico-philosophicus* (Frankfurt/M: Suhrkamp, 1989), ss. 9–85. Werkausgabe, Band 1.
- Wittgenstein L. 2000, *Dociekania Filozoficzne* (Warszawa: WN PWN). Oryginał: *Philosophische Untersuchungen* (Frankfurt/M: Suhrkamp, 1989). Werkausgabe, Band 1.
- Woleński J. 2002, *Co to znaczy, że doznania są prywatne?*, „Kwartalnik Filozoficzny”, 30:1, ss. 5–20.
- Wróbel A. 2000, *Neuron i sieci neuronowe*, w: T. Górská, A. Grabowska, J. Zagrodzka (red.), *Mózg a zachowanie* (Warszawa: WN PWN), ss. 40–67.
- Wróbel A. 2000a, *W poszukiwaniu integracyjnych mechanizmów działania mózgu*, w: T. Górská, A. Grabowska, J. Zagrodzka (red.), *Mózg a zachowanie* (Warszawa: WN PWN), ss. 460–485.
- Wszolek S. 1997, *Nieusuwalność metafizyki. Logiczno-lingwistyczne aspekty debaty Rudolfa Carnapa z Ludwigiem Wittgensteinem i Karlem R. Popperem* (Tarnów: Biblos).
- Wykowska A. 2002, *Czy uprawnione jest stanowisko materializmu nieredukcjonistycznego?*, „Kognitywistyka i Media w Edukacji”, 6:1–2, ss. 25–56.
- Young A. W., E. F. De Haan 1993, *Impairments of Visual Awareness*, w: M. Davies, G. W. Humphreys, *Consciousness. Psychological and Philosophical Essays* (Oxford: Blackwell), ss. 58–73.
- Zola, S., L. R. Squire 2003, *Genetic of Childhood Disorders: Multiple Memory Systems*, „Journal of the American Academy of Child & Adolescent Psychiatry”, 42:4, ss. 504–506.

- Żegleń, U. M. 2001, (red.), *Pragmatyzm i filozofia Hilarego Putnama* (Toruń: Wydawnictwo Uniwersytetu Mikołaja Kopernika).
- Żegleń, U. M. 2003, *Filozofia umysłu: dyskusje z naturalistycznymi koncepcjami umysłu* (Toruń: Wydawnictwo Adam Marszałek).
- Zeugner, G. 1965, *Barwa i człowiek* (Warszawa: Arkady). Oryginał: tenże, *Farbenlehre für Maler* (Berlin: VEB-Verlag, 1963).

INDEKS NAZWISK

- Ackermann H. 293, 294
Adamek D. 17, 236, 285, 294
Ajdukiewicz K. 12, 56
Akins K. 279
Aleksander S. 137, 138
Antony M. 182, 189
Armstrong D.M. 56, 92, 99, 105,
110, 127
Aronson J. 170, 171, 173, 189, 307
Atkinson A. 279
- Babiński J. 12
Baars B.J. 12, 73, 74, 88, 161, 173,
231, 235, 236, 238, 240, 262,
279, 286, 316
Bartels A. 40
Baumeister R.F. 40
Bayne T. 24, 40, 41, 68, 70–72
Bergson H. 16, 23, 55, 56, 190, 191,
255
Beckermann A. 136, 142
Berkeley G. 54
Block N. 22, 24, 37, 39, 80, 81, 120,
122, 130, 141, 174–177, 186,
189, 266, 276, 284, 306, 307
Bobryk J. 264
Bogen J.E. 249, 279, 294, 302
Borkowska A. 16, 256
- Bremer J. 33, 34, 73, 93, 106, 117,
119, 122, 135, 178, 264, 265, 281
Brentano F. 213
Broad C.D. 12, 136, 137, 139,
141–144, 148, 212, 237, 320
Broca P. 247
Brodmann K. 247
Bromm B. 242, 299, 301
Brown J. 257
Burge T. 13, 24, 25, 44, 46–49
Byrne A. 130
- Carruthers P. 127
Chalmers D. 13, 14, 16, 17, 24, 28,
36, 40, 41, 47–51, 68, 70–72,
77, 87, 89, 91, 124, 130, 199,
212, 218, 221, 227, 261, 269,
271, 272, 279, 280, 282, 283,
304, 321
Changeux J.-P. 246, 251, 267
Churchland P.M. 89, 99–102, 148,
213–216
Churchland P. S. 206, 254, 265
Cieśliński C. 29
Clark A. 47–49
Coghill R. C. 299, 300
Cole J. 69
Craig A.D. 299, 301

- Crick F. 16, 101, 161, 233, 234,
 262, 266, 269, 270, 307–311
 Czarnecki P. 29, 151, 153
 Czuang-tsy 9
- Damasio A.R. 16, 161, 162, 173,
 249, 266, 267, 270, 279, 307,
 315
 Daneman M. 237
 Daniluk B. 246
 Daum I. 293, 294
 Dawkins R. 178
 Davidson D. 117, 141, 220, 223, 224
 Dennett D. 43, 50, 52, 60, 74, 76,
 77, 80, 81, 89, 91, 101, 122,
 134, 148, 151–192, 214, 234,
 244, 256, 259, 260, 262,
 269–271, 276, 284–286
 Duch W. 14, 249, 316, 323
 Durstewitz D. 247, 276, 284, 285
 Duus P. 69, 247, 292, 295, 302, 304
 Dziarnowska W. 13, 77, 130, 193
- Eccles J. 89, 133, 264, 265, 270,
 310
 Edelman G.M. 236, 248, 266, 267,
 270, 279
 Einstein A. 52
 Eichenbaum H. 267
 Engel A.K. 277, 287, 333
 Ey H. 42
- Farah M.J. 159
 Favell J.E. 31
 Feigl H. 25, 56, 108, 284
 Fichte J.G. 36
 Fiez J. 293
 Fix J.D. 246
 Flanagan O. 181
 Flohr H. 254
- Fraassen B., van 35, 36
 Frege G. 46, 89, 108
 Freud Z. 12, 75
 Fries P. 277
- Gabor L. 16, 218
 Gall F. 247
 Gallagher S. 69, 251
 Garret B. 64
 Gazzaniga M.S. 16, 158, 239, 250,
 251
 Gebauer M. 199
 Gennaro R.J. 83
 Glickson J. 165, 166
 Gorman J.M. 303
 Grabowska M. 206
 Grimes T.R. 119
 Grush R. 276, 288
 Gulick R., van 175, 189
 Guttenplan S. 37, 45
- Haan E.F. 255
 Hagenbüchle R. 12
 Hallet G. 64
 Hameroff S. 16, 218, 268
 Hardcastle V.G. 279
 Hebb D. 28
 Heeger D.J. 243, 244
 Hershfield J. 21, 131
 Hilbert D.R. 130
 Hobbes T. 49, 54
 Hobson J.A. 289
 Horgan T. 80, 111, 212
 Huei-tsy 9
 Hume D. 25
 Hurley S.L. 117
 Husserl E. 16, 56, 194, 260, 261
- Ingarden R. 16, 26, 42, 56, 235

- Jacoby L.L. 31
 Jackson F. 89, 106, 133, 141, 143,
 209–212, 216
 James W. 16, 23, 55, 191, 218, 229
 Janusz R. 17
 Johnson M. K. 29
 Joyce J. 180
 Judycki S. 35, 93, 94, 145, 235, 261
- Kałuża J. 17, 236
 Kant I. 25, 34, 43, 44, 54, 55, 64,
 66, 72, 82, 113, 223, 229
 Kartezjusz 27, 44, 53–55, 61, 72,
 89, 169, 224
 Keil F. C. 12
 Kemeny J. G. 319
 Kim J. 117–119, 145
 Kinsbourne M. 89, 152, 161, 163,
 165, 166, 168, 169, 171, 172,
 175–177, 182, 256, 258, 284,
 285
 Klawiter A. 82, 263, 304
 Kloch J. 50
 Koch C. 101, 161, 266, 279, 307,
 309–311
 Konorski J. 247, 268, 290
 Korb K. B. 154
 Kornblith H. 203, 204
 Kossut M. 249, 250
 Kripke S. 102, 110–116, 122, 143,
 220, 225, 232, 299
 Krzyzewski K. 17, 40
 Kuhn T. 49, 56
 Kulczycki J. 23, 191, 255
 Kutschera F. 36
 Laplace P. S. 135
 Lashley K. 247, 267
 Lau J. 45
 Leibniz G. 53, 54, 97, 105, 111,
 114, 211, 218, 219, 268, 284
- Leiner H. C. 298
 Lenk H. 249
 LePore E. 117, 312
 Lewin R. 292
 Lewis D. K. 99, 104, 105
 Levin J. 193, 212
 Levine J. 80, 87, 116, 141, 143
 Llinas R. 74, 279
 Lloyd D. 242, 244
 Loar B. 79
 Locke J. 25, 27, 83, 130, 131, 133
 Loewer B. 312
 Logothetis N. K. 22
 Lorber J. 292
 Lowe E. J. 94, 120
 Lund D. H. 114
 Lycan W. 63, 80, 88, 130–134, 148,
 168, 189, 217
- Łuria A. 241, 256–258
- Madell G. 90
 Malcolm N. 204
 Maloney Ch. 17, 121, 209
 Manzotti R. 50
 Marciszewski W. 133
 Marr D. 180
 McGinn C. 14, 35, 85, 88, 90, 101,
 109, 113, 123, 141, 143, 197,
 199–202, 221, 224, 234, 329
 McLaughlin B. 117
 Mead C. H. 17
 Mehl P. 197
 Mele A. R. 160
 Merikle P. M. 237
 Merleau-Ponty M. 16
 Metzinger T. 21, 79, 276
 Meyer-Abich A. 40
 Mill J. 286
 Monroe M. 188

- Moore E. 12, 127
Morgan C. L. 137
Münste T. F. 251
- Nagel T. 14, 29, 65–67, 80, 81, 89,
90, 92, 99, 115, 123, 127, 133,
141–143, 147, 148, 158,
192–229, 233, 235, 255, 269,
275, 278, 291, 298, 318–321,
327, 330, 331
Nagel E. 41, 142, 145, 146, 318,
320, 330
Neisser U. 31
Neumann J., von 179
Newman 279
Niemirów L. 212
Northoff G. 77, 80, 232, 265, 272
- O'Brien G. 40
O'Keefe J. 285
Opie J. 40
Oppenheim P. 319
- Quine W.V.O. 33, 37, 38, 117, 270
- Pace-Schott E. 289
Papathanasiou I. 250
Parsons L. M. 231, 242, 293
Paçhalska M. 257
Peacocke Ch. 193
Penfield W. 247, 294
Penrose R. 16, 133, 226
Perzanowski J. 17, 27–29, 219, 317
Pilát R. 15, 70
Platon 49, 329
Poczobut R. 17, 117, 119, 135, 139,
316, 317, 324
Popper K. 264, 265
Putnam H. 35, 44–49, 93, 122, 223
- Raichle M. 287, 293
Railton P. 35
Ramachandran V. S. 68, 69
Ramsey F. P. 104, 105, 121
Ravenscroft I. 45
Reader 29
Rees G. 171, 238, 242, 244, 281,
299, 305
Reis C. E. 236
Rescher N. 49
Revonsuo A. 25, 74, 75, 234, 239,
240, 244, 245, 276, 315
Rey G. 151, 162–164
Ribary U. 74
Robertson L. 266
Robinson W. 123–128, 269
Rockwell T. 176, 183, 184, 186,
269
Rorty R. 91
Rosch E. 263
Rosenberg J. F. 231
Rosenthal D. M. 24, 27, 44, 83, 111,
174, 183, 234
Roth G. 292
Russell B. 12
Ryle G. 91, 151, 158
- Sacks O. 12, 70, 252
Schall J. D. 22
Scheler M. 235
Schmahmann J. D. 238, 293
Searle J. 14, 41, 46, 50, 56, 60, 61,
62, 64, 70, 80, 85, 87, 88, 90,
128–130, 132, 136, 164, 165,
234, 243, 259, 260, 265, 266,
270, 271, 286, 289, 298
Sejnowski T. J. 251
Sellars W. 33, 34, 56, 73, 93, 106,
107, 178, 197, 263, 264, 270,
281, 308

- Shallice T. 30
 Shear J. 262
 Sherrington Ch. 89
 Shoemaker S. 80, 90, 91, 112, 121, 193
 Siegal M. 276
 Siewert Ch. 30
 Singer T. 235, 277, 333
 Skoyles J.R. 250, 292
 Smart J.J.C. 56, 89, 92, 99, 102–107, 110
 Sokrates 329
 Sousa R. 59, 60, 67, 69
 Sperry R. 89, 137, 158
 Spinoza B. 219, 221
 Stein E. 16, 235
 Stephan A. 40, 41, 130, 135–140, 143
 Steuden M. 242
 Strawiński W. 135, 317, 319, 320
 Strawson G. 77, 91
 Strian F. 124
 Stubenberg L. 25, 77, 83
 Szafer W. 203
 Szepietowska E-M. 16, 256
- Tadeusiewicz R. 233
 Thach W. T. 252
 Thompson E. 45, 235, 263
 Titchener W. 260
 Tourette G. G. 12
 Travis J. 292
 Trąbka J. 264
 Turing A. 179
 Twardowski K. 61, 155, 257, 305
 Tye M. 45, 51, 61, 77, 79, 80, 106
- Vaas R. 264
 Vanderwolf C. H. 288
 Varela F. 90, 195, 259–263
- Varley R. 276
 Velmans M. 173, 182
 Volkelt J. 235
 Voorhees B. 154
- Wallace A. B. 12
 Walsh K. 16
 Walter H. 265
 Warhol A. 188
 Werner M. 165
 Wheelwell D. 179
 Whitehead N. 257
 Wiener D. 173
 Wilson R. A. 12
 Windmann S. 247, 276, 284, 285
 Wittgenstein L. 11, 12, 33, 34, 43, 55, 61, 64–66, 195, 204, 223, 275
 Woleński J. 56
 Wróbel A. 172, 233, 282, 309, 310
 Wszolek S. 33
 Wundt W. 260
 Wykowska A. 117
- Young A. W. 255
- Zasiiecki L. 258
 Zeki S. 40
 Zeugner G. 125
 Ziemiański S. 17
 Zola S. 267
- Żegleń U. M. 12, 27, 35, 56, 142, 153, 233, 238, 298

INDEKS RZECZOWY

- achromatopsja 290
anozognozja 173
argument z poznania (*knowledge argument*) 100, 210–212, 216
- behawioryzm 12, 15, 28, 30, 40, 87, 91, 96, 117, 119, 146
– logiczny 94, 95, 98, 319
– metodologiczny, 95, 96
– nomologiczny, 95, 96, 120, 147, 155, 268, 277, 321
- ból/doznanie bólu 11, 13, 60–61, 64, 78, 94, 114, 116, 126, 128, 214, 235, 253, 264, 299, 300
- ciało 14, 25, 31, 49, 64, 68, 82, 88, 102, 123, 156, 162, 179, 194, 205, 225, 246, 258, 272, 289
– obraz ciała/schemat ciała 69, 70, 76
- deaf-hearing* 326
- dualizm/dualiści 12, 54, 91–94, 119, 139, 214, 220, 232, 264, 329
– dualizm słaby 89, 91, 92
- eksperymenty myślowe 46–53, 74, 78, 130, 214, 238 273, 285, 288
- eksternalizm 44–49, 66, 79, 111, 122, 123, 262, 331
– aktywny 47, 48, 49, 123
- emergencja 20, 23, 41, 72, 90, 92, 116, 122, 198, 237, 257, 262, 268, 283, 306, 320, 323, 331
– słaba 138–141, 262
– diachroniczna 135, 137, 140, 218
– synchroniczna 139, 146, 148, 205, 217, 218, 257, 262, 315
– prawa emergencji 136, 142, 322
- empatia/wczucie 199, 202, 235, 258
- epifenomenalizm 56, 89, 102, 119, 138, 144, 211, 212, 255, 323
- fenomenologia 12, 17, 56, 81, 106, 155, 185, 204, 260–262, 279, 307
– obiektywna 90, 199–200, 202, 235
- fizykalizm 44, 78, 97, 111, 113, 119, 123, 127, 138, 143, 208, 212, 216, 220
- funkcjonalizm 15, 24, 30, 50, 71, 73, 80, 89, 101, 119–123, 148, 217, 254, 269, 281, 305, 307

- teleologiczny/maszynowy 122, 217, 254
- Glasgow Coma Scale 236
- hardware/software 48, 121, 157, 178, 190, 256
- heterofenomenologia 153–156, 172, 192, 201, 271, 276
- holizm 36–38, 51
 - metodologiczny 36, 38, 57, 59
 - ontologiczny 36, 38, 51
 - ścisły 40, 57, 72
- intencjonalność/stany intencjonalne 13, 14, 25, 27, 34, 45, 49, 54, 77, 78, 83, 91, 106, 109, 111, 123, 133, 151, 153, 161, 174, 185, 188, 198, 220, 330
- interakcjonizm 89, 162, 329
- internalizm 35, 44–49, 51, 66, 118, 262, 263, 280, 282, 331
- introspekcja 12, 62, 77–79, 88, 94, 100, 101, 132, 133, 155, 176, 182, 194, 226, 260, 321
- jaźń (*self*)/ja (rodzaje) 29, 55, 67, 93, 123, 155–159, 170, 180, 186, 205–207, 192, 206, 310
- kognitywistyka 12, 13, 15, 17, 48, 153, 162, 233, 241, 255, 257, 259, 260, 264, 271, 326
- locked-in syndrome* 62, 326, 327
- luka w wyjaśnianiu (*explanatory gap*) 21, 54, 87, 88, 109, 141, 192, 225–227, 264
- makroteoria/mikroteoria 313–325, 338
- materializm 12, 105, 109, 120, 135, 153, 162, 226, 319, 321,
 - eliminatywny 89, 93, 101, 134, 182, 189, 254
 - ontologiczny 92, 99–102
- memes* 177–180
- mózg (*brain*) 12–17, 19, 20, 23, 28, 32, 38, 43, 44, 51, 53, 71–76, 78, 84–93, 98–106, 108, 111, 113–116, 119–122, 124, 131, 132, 141, 142, 144, 147, 151, 152, 157, 161–175, 177–180, 183, 184, 186–192, 205–215, 218, 220, 225–228, 231–272, 275–331
 - części mózgu
 - ciało/jądro migdałowe (*corpus amygdaloideum*) 303
 - hipokamp (*hippocampus*) 158, 285, 290
 - jądra (*nuclei*) 248, 297
 - międzybłazkowe (*intralaminar nuclei*) 249, 279
 - miejsca sinawego (*locus coeruleus*) 303
 - nieswoiste (*nuclei nonspecifici*) 302
 - poduszki (*nuclei pulvinares*) 277
 - szwu (*raphe nuclei*) 303
 - zębate (*nucleus dentatus*) 297
 - kora mózgowa (*cortex cerebri*) 74, 249, 267, 291, 294, 297, 310, 331
 - kora śródwęchowa (*entorhinal cortex*) 162
 - kora wzrokowa (*area striata*) 215, 289

- kora zakrętu obręczy (*anterior cingulum cortex*) 300, 301
- kora mózdzku (*cortex cerebelli*) 297
- międzymózgowie (*diencephalon*) 248, 294, 296
- mózdzek (*cerebellum*) 249, 252, 289, 292, 293, 297, 298, 299
- narząd spiralny (*organum spirale cochleae*) 304
- nerw ślimakowy (*n. cochlearis*) 305, 306
- pień mózgu (*truncus cerebri*) 248, 249, 284, 303
- płaty
 - ciemieniowe (*lobi parietales*) 266, 277
 - czołowe (*lobi frontales*) 277, 290, 299
 - skroniowe (*lobi temporales*) 267, 303
- płyn mózgowo-rdzeniowy (*liquor cerebrospinalis*) 292
- prążkowie (*striatum*) 290
- szyszynka (*corpus pineale*) 162
- śródmózgowie (*mesencephalon*) 302
- twór siatkowaty (*formatio reticularis*) 302
- wzgórkki górne blaszki czworaczej (*colliculi superiores*) 277
- wzgórze (*thalamus*) 248, 266, 279, 294, 295, 297, 299
- zakręt potyliczno-skroniowy boczny (*gyrus occipitotemporalis lateralis*) 290
- modele mózgu,
 - holistyczny 247, 257
 - lokalistyczny 247, 282
 - koneksjonistyczny 245–249, 277, 282, 291, 311
 - neuroplastyczność mózgu 108, 249–252, 283, 288
 - rozszczepienie mózgu 158, 206, 264, 309
 - wady mózgu
 - bezmózgowie (*anencephalus*) 294
 - małogłowie (*microcephaly*) 291
 - wodogłowie (*hydrocephalus*) 292, 304
 - niedorozwój mózdzku (*hypoplasia cerebelli*) 294
- Multiple Personality Disorder 36, 309
- naturalizm 32, 90, 91, 132, 213
 - biologiczny 56, 90, 153
 - kartezjański 91
 - nieredukcyjny 135,
- neuronalne korelaty świadomości 15, 20–22, 32, 43, 57, 75, 76, 85, 98, 108, 116, 125, 143, 149, 174, 198, 216, 231, 233, 237, 251, 255, 265, 272, 285, 297, 312, 326, 336
 - korelaty rodziny zjawisk 296
 - całościowy korelat 302, 303, 306
 - minimalny system neuronalny 278–279, 301–303, 324
- neuroobrazowanie (*neuroimaging*) 76, 241–245, 268, 283, 299, 331
- optic ataxia* 326
- panpsychizm 114, 197, 216–220, 228–229, 232, 276

- paralelizm 43, 58, 221, 254, 255
- pierwszy plan/tło (*foreground/background*) 68, 186–188, 286–288, 303
- polineuropatia 326
- potoczny/naukowy obraz świata
33–34, 37, 56, 92, 93, 159, 171, 192, 236, 263, 321,
- prawa pomostowe (*bridge laws*) 142, 145, 221, 254, 276, 320, 322
- problem umysł–ciało 12–14, 29, 32, 34, 53, 99, 125, 265, 329
- problem scalania (*binding problem*) 171–173, 266, 309
- propriocepcja 69, 272, 295
- przytomność 22, 29, 231, 236–237, 289
- przyczynowość skierowana ku dołowi (*downward causation*) 89, 92, 137–140, 144, 149
- psychologia potoczna (*folk psychology*) 34, 45, 93, 105, 159, 208, 254, 316, 331
- punkt widzenia
- pierwszoosobowy 14, 45, 62, 63–67, 69, 82, 88, 90, 94, 96, 128, 142, 148, 167, 173, 194, 204, 214, 216, 217, 224, 241 259, 262, 271, 289, 313
 - trzecioosobowy 14, 56, 62–67, 80, 96, 97, 100, 114, 132, 133, 142, 147, 151–153, 159, 208, 224, 225, 260, 276, 284, 313, 322
 - *sub specie aeternitatis* 197
 - ludzki (rodzajowy) 36, 50, 62, 66, 81, 92, 93, 116, 129, 156, 178, 180, 195, 198, 201, 202, 204, 212, 215–215, 239, 260, 268, 276, 286
- qualia (jakości doznaniowe) 14, 23, 25, 77–81, 90, 91, 139, 164, 181, 211, 77–81, 91, 213
- nieobecne qualia (*absent qualia*) 129, 130, 210, 330
 - odwrócone spektrum (*inverted spectrum*) 50, 52, 129, 130
 - znikające/tańczące qualia 130
- realizm 172, 174, 222, 276
- realizm – antyrealizm 34–36, 189
 - poznawczy – ontologiczny 84, 268
 - wewnętrzny 35, 223
- redukcja/redukcjonizm 21, 32, 39–42, 49, 56, 65, 76, 78, 80, 87–90, 95, 97, 101, 103, 118, 120, 122, 134–141, 147, 196, 208, 212–216, 218, 227, 253, 257, 259, 268, 276, 319, 321–323
- fenomenologiczna 260, 261
 - (*via* prawa pomostowe) 142, 144, 145
 - ontologiczna 135, 144, 148, 181, 198, 205, 209, 212, 213, 232, 254, 329
 - teorii 145–146, 315, 324
 - wyjaśniająca 116, 119, 129, 143, 144, 148, 149, 174, 209, 227, 257, 330, 331
- reprezentacja (stany/cechy przedstawiające) 133, 200, 280, 287, 307
- świadomość (*consciousness*)
- holistyczna 38–44, 49, 57, 60, 67–70, 72, 75, 81, 83–85, 87, 89, 96, 127, 148, 161,

- 170–172, 204–206, 210, 233, 237, 238, 241, 259, 266–270, 277, 283, 289, 294–297, 308, 309, 311, 316, 326, 328
- jedność świadomości/jażni 14, 37, 40, 42–44, 48, 57, 60, 68, 70–72, 84, 123, 131, 158, 160–161, 191, 203–208, 237, 309–312
 - kognitywne teorie świadomości
 - neurofenomenologia 259–263, 270
 - neurofilozofia 263–267, 270
 - neuropsychologia 255–259, 270
 - poszerzona teoria świadomości 193, 216, 222–227, 275, 278–279
 - wielowymiarowa (nowa) teoria świadomości 15, 19, 39, 238, 245, 267, 283, 301, 323, 332
 - prototeoria świadomości 43, 57, 93, 149, 229, 275, 283, 325, 330
 - model świadomości zjawiskowej (*REM model*) 73–76, 84, 240, 281, 289, 290, 316
 - rodzaje świadomości
 - dostępu (*access*) 24–26, 56, 63, 71, 73, 84, 174, 177, 229, 237, 266, 306, 313, 317
 - zjawiskowa (*phenomenal*), *passim*
 - samoświadomość 23, 24, 42, 81–83, 172, 184, 204
 - „trudny” problem świadomości 13, 14, 21, 22, 33, 36, 42, 54, 55, 57, 65, 67, 77, 123–127, 229, 233, 234, 261, 262, 269, 270, 281, 330
- twardy (*rigid*) desygnator 46, 112, 114
- umysł (*mind*)
- modele świadomego umysłu
 - modułarny 160, 257
 - orwellowski/stalinowski 166–168, 176, 186, 192, 244
 - Kartezjańskiego Teatru 151, 161–164, 190–191, 286, 307
 - Funkcjonalnego Teatru 171–174
 - Wielu Szkiców (*Multiple Drafts*) 51, 52, 73, 152, 153, 160, 164–166, 168–170, 184, 189–190, 269
 - teorie umysłu
 - holistyczna teoria stanów mentalnych 104, 105
 - identyczności typów (*type-type*) 97–99, 139, 147, 161, 225, 253, 264, 278, 322, 328, 330, 331
 - identyczności zdarzeń (*token-token*) 97, 99, 108–111, 147, 148, 220, 254, 300
 - Global Workspace Theory 73, 74, 286
 - *higher order thoughts* (HOTs) 83, 183, 186, 234
 - teoria zmysłu wewnętrznego (*inner sense theory*) 130–134
 - teoria dwu aspektów (*dual-aspect theory*) 89, 148, 216–222
- wieloraka realizacja (*multiple realizability*) 98, 108, 145, 254

zombi 26, 50, 51, 78, 79, 89, 99,
111, 173, 227, 269
zsynchronizowana oscylacja 74,
101, 266, 279, 310, 311

SUMMARY

What it is like to be conscious? Analytical theories of mind and the problem of neural basis of consciousness

The book, which is under consideration, is a critical review of contemporary Anglo-American solutions to the body-mind problem, especially to one aspect of this problem: the problem of consciousness: “Without consciousness the mind-body problem would be much less interesting. With consciousness it seems hopeless” (Nagel *What Is It Like to Be a Bat?*: 165). Consciousness is marked by “subjectivity” and this feature – among others – forms one of the parts of the body-mind problem, the “hard problem”. The book’s point of departure is N. Block’s distinction, between “phenomenal consciousness” and “access consciousness”. According to Block “what makes a state phenomenally conscious is that there is something «it is like» (Nagel 1974) to be in that state” (Block 1995a: 228). “Access consciousness” – which refers to intentional states – does not have such a “what it is like” moment, which could be used for its identification. The distinction between phenomenal consciousness and access consciousness is methodological rather than empirical.

The subjective character of our conscious experiences – so runs one of the main theses of the book – is not captured by any of the contemporary familiar, reductive analyses (e.g. *type-type theory*) of the mental states. All of such analyses are logically compatible with the absence of subjective experiences (the *zombie* problem). Subjectivity seems not to be analyzable in terms of any explanatory theory of functional or intentional states, since such states could

be ascribed to computers or intelligent robots that behaved like people although they experienced nothing. For similar reasons, subjective character is not analyzable in terms of the causal role of experiences in relation to typical human behavior (see the theories of *logical* or *nomological behaviorism*).

The main authors critically discussed in this book belong to two groups: on the one hand, philosophers who fall within the wide area of analytical philosophy; on the other, contemporary authors within the wide domain of neural sciences. The book itself is a study in the cognitive sciences. The author not only presents the various theories but also shows the difficulties faced by each theory and suggests the directions in which to search for the resolution of those difficulties.

The first chapter departs from two points: (i) Block's distinction between phenomenal and access consciousness and (ii) W. Sellars' distinction between the "manifest image" of the world and the man in the world (described and explained in *folk psychological* terms) and the competing "scientific image" (described and explained in scientific terms). These distinctions are used as a basis for (a) a short presentation of the basic philosophical terms used in the contemporary philosophy of mind, and (b) a characterization of the psychological and the metaphysical view of phenomenal consciousness. According to D. Chalmers and T. Bayne's *Phenomenal Unity Thesis* – "necessarily, any set of phenomenal states of a subject at a time is phenomenally unified" (Bayne and Chalmers 2003: 8) the chapter offers an analysis of this "holistic" understanding of consciousness and a discussion of its ontological and methodological aspects. The holistic understanding of phenomenal consciousness, in its methodological dimension, is presupposed in the remainder of the book: There exists a single state of consciousness characterized as: *what it is like* to be a conscious subject. The chapter concludes with short supplementary descriptions of the philosophical externalism and internalism debate (H. Putnam, T. Burge, M. Tye, J. Searle), the so-called "thoughts experiments" used in the setting up the hard problem of consciousness (among other things: "the inverted spectrum argument", "zombie-argument", "Chinese room argument"), and the understanding of subjectivity in the history of philosophy. The rest of the book adopts the internalist position: the mental states of phenomenal consciousness depend on the subject's brain states.

The second chapter develops a more precise description of the philosophical-psychological conception of the subjectivity of consciousness. The subjectivity of consciousness – as the core of the "hard" problem of consciousness and as the expression of the unity of consciousness – is characterized, *inter alia*, from the first person point of view and from its spatial and temporal dimensions. The chapter introduces a "model" of phenomenal consciou-

ness: the dreams and the dreaming brain during the REM phase of sleep. In such a brain both sensory inputs and motor outputs are blocked. This model should help us in the search for the “neural correlates of consciousness”, i.e. to make clear the differences and the similarities in the neuro-imaging of the sleeping and waking brain. The theses that subjectively conscious mental states and events could cause subjects behavior, and that they may – in some cases – be given functional descriptions are not called into question. But it is argued that this kind of characterizations does not exhaust the analysis of phenomenally conscious mental states. Any reductionist program – e.g. an identification of mental states with neurological states – has to be based on an analysis of what is to be reduced. If the analysis omits something essential – in our case, subjectivity – the body-mind problem will be falsely posed.

The third chapter supplies the critical analysis of the best known materialistic – reductionist and nonreductionist – theories of mind: *logical* and *nomological behaviorism*, *type-type* and *token-token theory* (J. J. C. Smart, U. T. Place, D. M. Armstrong, P. Churchland, D. Davidson). The analysis is concentrated on the investigation of the concepts of consciousness developed in these theories. One of the problems with such theories concerns the kind of necessity attaching to claims about identity (see S. Kripke’s conception of “rigid designations”). The chapter also undertakes a critical overview of functionalism (N. Block, S. Shoemaker, W. Robinson), Inner Sense Theory (W. Lycan), and theories of emergence (C. D. Broad, S. Stephan).

Chapters four and five depict two antipodal theories of consciousness: the “functional” or “eliminative” theory created by D. Dennett and the double-aspect, panpsychical theory created by T. Nagel. One of the questions – with regard to Dennett – is how to preserve the first person point of view and the holistic conception of phenomenal consciousness in his Multiple Drafts Model (see: Dennett – Block and Dennett and Aronson’s discussion of this matter). The two main questions – with regard to Nagel, one of the well known defender of subjective point of view, – are: (a) how to empirically locate the necessary connection between the physical and the mental for which he is searching, and (b) how to empirically improve – using contemporary techniques – his panpsychical assumptions.

Chapters six and seven outline a broad philosophical-neurological framework that can serve as a methodology for searching for the neuronal correlates of consciousness. This search is compared with three contemporary neurocognitive theories: neuropsychology (A. Luria, M. Kinsbourne), neuropneomenology (F. Varela), and neurophilosophy (P. S. Churchland, A. Damasio, F. Crick, Ch. Koch). These considerations issue in two definitions: the first one of the neuronal correlates of holistic consciousness; the second one, of

the special states of this kind of consciousness, based on D. Chalmers' definitions of such correlates. In contrast to Chalmers, however holistic consciousness is treated first. The definitions are elucidated, guided by the results of the psychological and neurological investigations (P. Duus, A. Engel, M. Gazzaniga, A. Grabowska, J.A. Hobson, M.E. Raichle, J. E. Bogen, A. Wróbel), the neuroplasticity of the brain is taken into account and an attempt is made to develop the outline of a prototheory of consciousness. This theory is based, on the one hand, on an emergent dependency between a *folk psychological* theory of consciousness and neuropsychological theory, and, on the other hand, on the emergent dependency between mental states and physical states of the brain. The main point of the proposed prototheory is captured by the phrase: "Reduction without identification". This theory is a contemporary philosophical "substitute" for the intended ultimate scientific theory of consciousness. Such a future – multidimensional – theory should give us the possibility to predict, more or less exactly, the kinds of states of consciousness which are present in patients with "locked-in syndrome", "stiff-men syndrome", polyneuropathy, and coma. In such cases, access consciousness is totally or partially blocked.

The results of these investigations show – as a more universal thesis – that it is useless to base the defense of materialistic theories of mind on any analysis of mental states that fails to deal explicitly with their subjective features. We do not have any reason to suppose that a reduction (or identification) which seems plausible, when no attempt is made to account for phenomenal consciousness, can be extended to include such consciousness. Without any idea, therefore, of what the subjective character of our experience is, we cannot know what is required of a neurological theory of the mind.

J. B.

Połączone Biblioteki WFiS UW, IFiS PAN i PTF

U.84172



39084172000000



**WYDAWNICTWO
INSTYTUTU FILOZOFII I SOCJOLOGII
POLSKIEJ AKADEMII NAUK**

ISBN 83-7388-069-0