

CHALMERSA OSOBLIWA TEORIA ŚWIADOMOŚCI

Ned Block przeprowadza następujący eksperyment myślowy: wyobraźmy sobie, że chiński rząd przekonał się do funkcjonalizmu i zobowiązał obywateli do realizacji ludzkiego umysłu przez godzinę. Miliard Chińczyków miałby symulować neurony, komunikując się ze sobą radiowo. Można by ich połączyć radiem ze sztucznym ciałem. Czy taki układ byłby świadomy? Dla Blocka jest jasne, że nie. Świadomość nie jest w jego koncepcji organizacyjnie niezmiennicza. Coś umyka: jakość przeżyć świadomych.

David Chalmers jest autorem jednej z szerzej dyskutowanych współcześnie książek na temat świadomości¹. Trzeba jednak podkreślić, że dyskusja ta skupia się zazwyczaj na krytycznym aspekcie jego książki, wymierzonej przeciwko tradycyjnemu funkcjonalizmowi w filozofii umysłu. Funkcjonalizm taki legitymizuje teorie świadomości, które Chalmers uważa za niewystarczające, dlatego że nie są one w stanie poradzić sobie z tak zwanym trudnym problemem świadomości. Otóż mimo że wyjaśniają wiele zjawisk związanych z funkcjonowaniem świadomego mózgu i określają rozmaite neuronalne korelaty świadomości, nie potrafią odpowiedzieć na trudne pytanie: dlaczego przeżycia świadome mają w ogóle subiektywne jakości? Teorie funkcjonalistyczne mają ujmować mózg jedynie jako układ przetwarzający informacje, ale nie odpowiadają na pytanie, dlaczego przetwarzaniu informacji towarzyszą w ogóle jakiegokolwiek przeżycia. Przecież wiele procesów przetwarzania informacji przebiega bez przeżyć świadomych: mój kalkulator wylicza sumę dwóch dużych liczb całkowitych, ale zwykle nie przypisuję mu świadomości jakościowej.

Większość dyskusji wokół książki Chalmersa ogniskowała się na jego argumentach krytycznych wobec koncepcji dominujących w filozofii umysłu. Nic dziwnego, skoro były to argumenty wymierzone w stanowiska uznawane za stosunkowo najlepiej uzasadnione. Jego własna teoria, która miała odpowiedzieć na trudne pytanie, dlaczego subiektywne jakości towarzyszą przeżyciom świadomym, poddawana była mniej dokładnemu rozbirowi. Zapewne też i dlatego, że sam autor przyznawał, że jest ona jedynie szkicem dojrzałej teorii i że opiera się w pewnej mierze na spekulacjach. Jej główne tezy można streścić – w największym uproszczeniu – następująco: subiektywne jakości nie

¹ Zob. D.J. Chalmers, *Świadomy umysł: w poszukiwaniu teorii fundamentalnej*, tłum. M. Miłkowski, Wydawnictwo Naukowe PWN, Warszawa 2010.

występują z koniecznością logiczną wraz z przeżyciami świadomymi, lecz na mocy praw przyrody naszego świata; być może przeżycia świadome występują w całej przyrodzie (protopanpsychizm), wszędzie tam, gdzie pojawia się strukturalnie pojmowana informacja; a pewną rolę być może odgrywają zjawiska kwantowe (z tego ostatniego elementu Chalmers później zupełnie zrezygnował). Wyłaniająca się stąd teoria może być, ontologicznie rzecz biorąc, bądź teorią podwójnego aspektu (głoszącą, że cała rzeczywistość ma naturę świadomą i fizyczną zarazem), bądź pewną odmianą dualizmu².

Zamierzam pokazać, że pozytywna teoria Chalmersa niestety zupełnie nie potrafi sobie z jego własnym pytaniem poradzić. Nie widać wcale, jak jego tezy ontologiczne miałyby rzucać światło na to pytanie. Gdzie indziej już argumentowano, że stanowi ona co najwyżej bardziej zawiłą odmianę funkcjonalizmu, różniącą się od niego tylko werbalnie³. Tu zamierzam pokazać inną zbieżność teorii Chalmersa z funkcjonalizmem: teoria ta nie tylko zachowuje większość osiągnięć teorii kognitywistycznych, które często, choćby i implícite, zakładają funkcjonalistyczne ujęcie świadomości, ale również podnosi wyjątkowo kontrowersyjne założenia funkcjonalizmu do rangi koniecznych (w naszym świecie) zasad metafizycznych. Ponieważ zagadnienia związane z modalnością niektórych twierdzeń filozoficznych są interesujące same w sobie, nie twierdzą bynajmniej, że określając pewne zasady jako konieczne (logicznie, metafizycznie czy przyrodniczo), Chalmers uprawia złą filozofię. Twierdzą jedynie, że robi to zupełnie niepotrzebnie ze względu na teorię świadomości; jego koncepcja nie zyskuje bowiem w ten sposób nic, co pomagałoby w wyjaśnieniu tak zwanego trudnego problemu.

Struktura obecnego tekstu jest następująca: W pierwszej części rekonstruję funkcjonalistyczne koncepcje związane z teorią niezależności organizacji układów złożonych od ich materialnej realizacji. Teoria ta często rodzi żywiołowy sprzeciw, gdy zostanie odniesiona do układów świadomych o charakterze niebiologicznym, w szczególności zaś do przetwarzania rozproszonego. Napomknę też o niejednoznacznym rozumieniu terminu „własność funkcjonalna” w tej dyskusji. W drugiej części krótko przypominam najważniejsze teorie świadomości w funkcjonalizmie, aby nakreślić tło dla teorii Chalmersa. Niektóre z nich, jak pokażę, starają się odpowiedzieć na pytanie Chalmersa o jakości przeżyć świadomych – wbrew temu, co autor ten sugeruje.

W trzeciej części pokazuję, że Chalmers, zamiast negować funkcjonalistyczne założenia, zachowuje je jako „zasadę niezmienności organizacyjnej”.

² Dalsze szczegóły tej teorii podawać będę już dalej w tekście w związku z problemami, jakie stanowisko to rodzi.

³ Zob. M. Miękowski, *Jak udawać dualistę, wprowadzając epicykle do funkcjonalizmu*, „Przegląd Filozoficzny – Nowa Seria” 20(2011) nr 1(77), s. 27-45.

Założenia te mają obowiązywać jednak w mniejszym zakresie niż w tradycyjnym funkcjonalizmie, bo jedynie w naszym rzeczywistym świecie, a nie we wszystkich światach możliwych. Dyskusja na temat modalności wiedzie jednak na manowce, bo pozorny dualizm Chalmersa de facto opiera się na mocniejszych założeniach niż poglądy wielu fizykalistów. Co gorsza, Chalmers dodatkowo uznaje, że przetwarzanie informacji – a nawet bycie nośnikiem informacji – wystarcza do wzbudzenia przeżyć (proto)świadomych. Oto więc mój kalkulator ma świadome przeżycia, gdy liczy sumę dwóch liczb.

W czwartej części argumentuję, że Chalmers musi albo uznać, że trudny problem świadomości jest pseudoproblemem, albo odrzucić swoją teorię jako całkowicie równoważną klasycznemu funkcjonalizmowi. Różni się ona bowiem jedynie nazwą.

KLASYCZNY FUNKCJONALIZM

Funkcjonalizm w filozofii umysłu to teza głosząca, że wszystkie stany i procesy⁴ umysłowe są konstytuowane przez role takich stanów i procesów, a także przez relacje między nimi a zachowaniem oraz między nimi a bodźcami docierającymi do organizmu⁵. Mówiąc swobodnie, role funkcjonalne to relacyjnie ujęte charakterystyki pewnych stanów i procesów. Na przykład, do odkorkowywania butelek może służyć coś, co ma pewną określoną rolę: usuwanie korka z szyjki butelki – niezależnie, czy rola ta jest realizowana przez korkociąg ręczny, korkociąg pneumatyczny, wykonany ze stali czy też z innego sztywnego materiału. Podobnie krzesłem nie jest coś, co jest wykonane z drewna czy ze stali i szkła, lecz coś, co ma funkcję (rolę funkcjonalną) czegoś, na czym można usiąść, zginając nogi mniej więcej pod kątem prostym, i na czym można oprzeć plecy, lecz nie ręce. Podobnie jak wytwory ludzkie, stany i procesy umysłowe należy ujmować właśnie w kategoriach ról, jakie one odgrywają, gdyż materialny substrat jest – zdaniem funkcjonalistów – drugorzędny.

⁴ Ten fragment definicji („stany i procesy”) należy potraktować jako zmienną, pod którą można podstawiać różne kategorie ontologiczne. W tradycyjnych debatach w filozofii umysłu używa się pojęcia „własność”. Pojęcie to wydaje mi się nieadekwatne, gdyż psychologia w wyjaśnieniach postuluje też stany umysłowe (np. stan depresji, mający też składową afektywną toniczną) i procesy (np. zapamiętywanie), a nie tylko własności (np. bycie inteligentnym). Niżej będę mówił o stanach i procesach, lecz można by równie dobrze mówić także o własnościach, oddziaływaniach, działaniach (operacjach) czy składnikach.

⁵ Korzystam tu z nieco zmienionej definicji podanej przez Witolda Hensla. Zob. W. H e n s e l, *W poszukiwaniu jakości*, „Przegląd Filozoficzno-Literacki” 2003 nr 4(6), s. 15-31.

Różne odmiany funkcjonalizmu można scharakteryzować, wskazując właściwy rodzaj ról. Zwykle, ale nie zawsze, są to role przyczynowe⁶. I tak standardowy funkcjonalizm obliczeniowy (maszynowy) uwzględnia role obliczeniowe utożsamiane z określonymi rolami przyczynowymi w systemie poznawczym, a bodźce i zachowanie określa mianem „wejścia” i „wyjścia”; funkcjonalizm analityczny zaś głosi, że istotne jest funkcjonalne zredukowanie terminów umysłowych do innych terminów, oznaczających w ostatecznym rozrachunku bodźce i zachowanie zewnętrzne. Zadaniem filozofa w tym drugim wypadku jest przeprowadzenie odpowiedniej procedury redukcji terminów jednego rodzaju do terminów drugiego rodzaju. (Warto w tym miejscu zaznaczyć, że Chalmers, pisząc o funkcjonalizmie, na myśli ma zarówno funkcjonalizm obliczeniowy, jak i analityczny).

Przy takim ujęciu procesów i stanów umysłowych narzucają się dosyć oczywiste pytania. Czy jest to tylko wersja behawioryzmu, co najwyżej postulująca większą rolę zmiennych pośredniczących? Widmo behawioryzmu funkcjoniści chcą odegnąć, by dokonać legitymizacji badań prowadzonych w psychologii poznawczej i kognitywistyce⁷, ale zbieżność tych teorii nie jest czysto przypadkowa⁸. A jeśli funkcjonalizm jest behawioryzmem, to czy neguje on istnienie świadomości, jak zdarzało się co radykalniejszym behawiorystom, na przykład Watsonowi⁹?

Drugi rodzaj pytań dotyczy kwestii ogólniejszej: od czego odróżnia się stany i procesy umysłowe mające role? Zaczniemy od założenia, że chodzi o ich role przyczynowe, bo funkcjonalizm oparty na przyczynowości jest bodaj najbardziej popularny; taki funkcjonalizm uznamy za paradygmatyczny. Niemniej jednak założenie to jest prawdziwe tylko wtedy, gdy przyczynowość realnie zachodzi. Jeśli ktoś uznaje, niczym Bertrand Russell¹⁰, że termin „przyczyna” należy wyrugować z języka nauki jako przeżytek, to oczywiście przyczynowa wersja tezy funkcjonalizmu nie przypadnie mu do gustu. Dziś taki pogląd wcale nie należy do odosobnionych¹¹; w takim jednak wypadku

⁶ Zob. np. R. C u m m i n s, *The Nature of Psychological Explanation*, MIT Press, Cambridge, Massachusetts, 1983.

⁷ Zob. J.A. F o d o r, *Nauki szczegółowe (albo: niejednorodność nauki jako hipoteza robocza)*, tłum. M. Gokieli, w: *Analityczna metafizyka umysłu*, tłum. M. Gokieli i in., red. M. Miłkowski, R. Poczobut, Wydawnictwo Instytutu Filozofii i Socjologii PAN, Warszawa 2008, s. 56-75.

⁸ Zob. W.M. H e n s e l, *Dwa funkcjonalizmy Hilary'ego Putnama, czyli kawałek historii z mo-
ralem*, „Diametros” 2011, nr 3(29), s. 31-49.

⁹ Zob. I. K u s t r z e b a, *Znak i znaczenie w psychologii behawiorystycznej*, Państwowe Wydawnictwo Naukowe, Warszawa 1971.

¹⁰ Zob. B. R u s s e l l, *On the Notion of Cause*, „Proceedings of the Aristotelian Society” 13(1912), s. 1-26.

¹¹ Zob. np. D. R o s s, D. S p u r r e t t, *Notions of Cause: Russell's thesis revisited*, „The British Journal for the Philosophy of Science” 58(2007) nr 1, s. 45-76.

można nadać tezie funkcjonalistycznej inny wyraz, na przykład zastępując termin „przyczyna” terminem „zależność typu X”; musi to być zależność dająca duże możliwości opisu dynamiki układów złożonych. Taka zamiana jednak wiele nie zmienia, jeśli chodzi o status tez funkcjonalistycznych: jeśli powiemy, że teza funkcjonalistyczna głosi, iż procesy i stany umysłowe to stany X-uzależnione od innych procesów i stanów oraz X-zależne od bodźców i X-warunkujące zachowania, to nie odejdziemy od funkcjonalizmu. Dlatego też nieprzyczynową wersję funkcjonalizmu traktuję jako mniej typowy wariant tego stanowiska (w gruncie rzeczy nieistotny ze względu na dalszą argumentację w tym tekście).

Idźmy dalej. Jeśli uznać, że wszystko, co istnieje, zawsze ma jakąś rolę przyczynową¹², to znaczy mogłoby oddziaływać przyczynowo¹³, wówczas teza funkcjonalizmu okazuje się banałem. Tak, stany i procesy zachodzące w mrowiskach są konstytuowane przez role przyczynowe, podobnie też jak stany i procesy zachodzące w serach, drozdach, gwiazdach i włosach. Jeśli jednak uznamy, że istnieją byty niemające szczególnych ról przyczynowych – a więc byty epifenomenalne – to funkcjonalizm wykluczy ze świata umysłowego jedynie epifenomeny. Jak się okaże, obrona świadomości przed funkcjonalizmem często polega na wykazywaniu możliwości epifenomenalizmu w odniesieniu do świadomości; słaba to jednak obrona, skoro definicja funkcjonalizmu – przy wskazanym wyżej jej rozumieniu – właśnie sugeruje istnienie takich bytów epifenomenalnych. W istocie przyjmowanie epifenomenalizmu wynika ze zgody na tezę funkcjonalistyczną! Gdyby bytów epifenomenalnych nie było, funkcjonalista głoszący swoją tezę naruszałby maksymę jakości Grice’a¹⁴: niepotrzebne byłoby zastrzeżenie, że stany funkcjonalne konstytuowane są przez role przyczynowe, bo byłby to jedyny możliwy sposób konstytucji. Teza funkcjonalizmu byłaby niepotrzebnie wielosłowna i rozwlekła. Tak więc wydaje się, że teza ta musi być głoszona przy jednoczesnym założeniu istnienia bytów epifenomenalnych. Jeśli więc Chalmers obstaje za epifenomenalizmem, to nie zaprzecza bynajmniej funkcjonalizmowi.

Byty bezwzględnie epifenomenalne wydają się ekstrawagancją ontologiczną: dlaczego mielibyśmy postulować coś, co nigdy z niczym nie oddziałuje (w wersji nieprzyczynowej: coś, co od niczego innego nie zależy)? Użycie brzytwy Ockhama narzuca się samo. Ale może nie zawsze warto jej użyć?

Cień rzucany przeze mnie na klawiaturę jest epifenomenem: to nie on powoduje, że nie widzę klawiszy, lecz fakt, że moja głowa zasłania światło za-

¹² Bądź zależnościową, jeśli odrzucamy realność przyczynowości.

¹³ Pojęcie roli wprowadzono, by uwzględnić nie tylko faktyczne oddziaływania przyczynowe, ale także takie, które mogłyby zachodzić.

¹⁴ Zob. H. Grice, *Logika a konwersacja*, w: *Język w świetle nauki*, tłum. B. Stanosz, red. B. Stanosz, Spółdzielnia Wydawnicza Czytelnik, Warszawa 1980, s. 91-114.

rowe. Innymi słowy, cień jest tylko towarzyszącym skutkiem, a nie przyczyną; w tym sensie jest epifenomenem. Cień jednak nie jest z u p e ł n i e epifenomenalny: coś go jednak powoduje, a zatem ma fizyczną przyczynę. Pewne zjawiska mogą towarzyszyć innym, nie będąc jednak kluczowe ze względów eksplanacyjnych. Przekonującym przykładem czegoś, co nie jest istotne dla charakterystyki przyczynowych roli śledzi w kuchni polskiej, jest fakt, czy śledzie te kiedykolwiek były wyśmiewane przez foki czy pingwiny¹⁵. Mówiąc inaczej, między śledziami a ich środowiskiem z pewnością zachodzi bardzo wiele oddziaływań fizycznych, które nie mają nic wspólnego z tym, czy śledzie okażą się smaczne lub pożywne. Analogicznie moje procesy umysłowe mogą występować na przykład w określonej odległości od gwiazdy Alfa Centauri, która oddziałuje na nie grawitacyjnie, ale oddziaływanie to można pominąć w wyjaśnianiu mojego myślenia. Krótko mówiąc, kiedy chcemy zrozumieć proces lub stan umysłowy, musimy naszkicować jego strukturę przyczynową, która nie będzie obejmować wszystkich możliwych oddziaływań przyczynowych zachodzących w danym fragmencie czasoprzestrzeni, lecz jedynie te, które są relewantne dla dalszego przebiegu procesów umysłowych.

Po tej wstępnej charakterystyce funkcjonalizmu warto zauważyć, że często źle stawia się pytanie, czy coś jest własnością funkcjonalną, a więc własnością charakteryzowaną w kategoriach przyczyn i skutków (tak jak w tezie funkcjonalizmu). Własności zupełnie niefunkcjonalne musiałyby być zupełnie epifenomenalne. Istnienie tych ostatnich i tak z definicji nie ma znaczenia dla naszych teorii, bo własności te od niczego nie zależą i nic od nich nie zależy.

Pojawiające się niekiedy w literaturze przedmiotu pomysły, by były niefunkcjonalne utożsamić z tak zwanymi własnościami wewnętrznymi, nie są przekonujące. Jeśli własność wewnętrzną przedmiotu rozumieć jako taką, która może istnieć niezależnie od tego, czy cokolwiek innego niż ów przedmiot również istnieje, to nic nie wskazuje na to, że tego rodzaju własności nie mogą być członami relacji przyczynowych. Wielu autorów uważa wręcz, że paradygmatyczne przypadki oddziaływań przyczynowych zachodzą między własnościami wewnętrznymi¹⁶. To, czy dany byt jest samodzielny (samoistny), tak jak własność wewnętrzna danego przedmiotu, nie przesądza o tym, czy nie można podać jego konstytucji w kategoriach przyczynowych. Dosyć oczywiste jest to, że w fizycznym układzie złożonym może istnieć własność zależna od jego

¹⁵ Ten przykład podają za: N. Block, *Advertisement for a Semantics for Psychology*, „Midwest Studies in Philosophy” 10(1987) nr 1, s. 628.

¹⁶ Zob. J.A. Fodor, *Methodological Solipsism Considered as a Research Strategy in Cognitive Psychology*, „Behavioral and Brain Sciences” 3(1980) nr 1, s. 63-72. Warto jednak zauważyć, że – wbrew Fodorowi – relacje przyczynowe między własnościami zewnętrznymi są dosyć często postulowane w fizyce czy biologii. Zob. R.A. Wilson, *Boundaries of the Mind: The Individual in the Fragile Sciences. Cognition*, Cambridge University Press, Cambridge 2004.

części, lecz nie od innych przedmiotów w świecie – na przykład masa. Można jednak zanalizować masę pewnego przedmiotu jako sumę mas wszystkich jego części. Innymi słowy, funkcjonalista będzie mógł powiedzieć, że masa jest konstytuowana przez części przedmiotu, a więc *z a l e ż n i e* od jego części. W przypadku innych własności konstytucja może mieć też charakter przyczynowy (w tym oczywiście nie ma).

Nie chcę tu wchodzić w dosyć złożone zagadnienie, jak należy definiować własności wewnętrzne (lub wręcz kilka rodzajów takich własności)¹⁷. Kluczowe jest to, że takie pragmatyczne zainteresowanie wyjaśnieniem wyznacza to, co jest relewantne dla określenia roli przyczynowych śledzi w kuchni polskiej. Tymczasem w dyskusjach zapomina się o koniecznej relatywizacji charakterystyki ról przyczynowych do określonej struktury przyczynowej, która jest postulowana w celu opisu określonego zjawiska¹⁸. Zjawisko wyjaśniane, czyli pewne explanandum, musi zostać dokładnie opisane. Dopiero znając charakterystykę explanandum, możemy stwierdzić, czy explanans rzeczywiście je wyjaśnia, a także, co jest składnikiem explanansa. W tym przypadku explanans jest postulowany model przyczynowy explanandum. Funkcjonalizm musi postulować pewną strukturę przyczynową systemu poznawczego, która będzie wyjaśniać zjawiska umysłowe (na przykład poznawcze bądź emocjonalne). Dopiero ze względu na opis tych wyjaśnianych zjawisk można mówić o relewancji pewnych innych bytów wobec nich: niektóre byty nie będą w ogóle częściami explanansa (tak jak fakty o pingwinach śmiejących się ze śledzi), a więc w tym sensie będą nieistotne przyczynowo. To nie znaczy jednak, że pingwiny wymykają się z sieci związków przyczynowych w świecie. Niestety, debaty nad bytami bezwzględnie epifenomenalnymi czy też własnościami całkowicie samoistnymi toczą się zwykle w oderwaniu od tej relatywizacji explanansa wobec explanandum i w konsekwencji rozważa się, mówiąc nieco obrazowo, czy można by zagwarantować rozważanym pingwinom status nieprzyczynowy w sposób logicznie konieczny.

Mówiąc krótko, funkcjonalista *z a w s z e* coś pominie jako nieistotne dla procesów umysłowych (tak jak dla smaku śledzi nieistotne jest rozbawienie pingwinów) w postulowanej strukturze funkcjonalnej systemu poznawczego. Jednak jeśli będzie to jakoś przeżycia świadomego, to na zarzut przeciwnika funkcjonalista bez trudu będzie mógł tę jakość wpisać w charakterystykę wyjaśnianych zjawisk umysłowych, na przykład twierdząc chociażby, że wszelkie

¹⁷ Zob. B. Weatherston, D. Marshall, *Intrinsic vs. Extrinsic Properties*, w: *The Stanford Encyclopedia of Philosophy*, red. E. Zalta, edycja zima 2012.

¹⁸ Współczesna metodologia wyjaśnień przyczynowych jest stosunkowo złożoną i dojrzałą teorią. Zob. J. Woodward, *Making Things Happen*, Oxford University Press, Oxford 2003; J. Pearl, *Causality: Models, Reasoning, and Inference*, Cambridge University Press, Cambridge 2000.

świadome myśli są dostępne poprzez modalności zmysłowe¹⁹. Istnieje bowiem dosyć duża dowolność charakterystyki tego, co ma być konstytuowane przez role przyczynowe; co więcej, same bodźce i reakcje można rozumieć wężej i szerzej: jedni behawioryści charakteryzowali je w kategoriach fizjologiczno-mięśniowych, inni zaś bardziej abstrakcyjnie. Z tej perspektywy staje się jasne, dlaczego na przykład własności wewnętrzne stały się przedmiotem tak gorących dyskusji w kontekście przeżyć świadomych. Żeby polemizować z funkcjonalistą, jego przeciwnik stara się uniemożliwić „funkcjonalizację” niektórych bytów, twierdząc, że nie są one w ogóle funkcjonalizowalne. Jeśli potrafi taką mocną tezę obronić, funkcjonalizm upada. Dyskusja przypomina wówczas, mówiąc w ogromnym uproszczeniu, debatę między zwolennikami i przeciwnikami strukturalizmu w dowolnej dziedzinie: rzecz w tym, czy pewne byty istnieją w izolacji, czy nie.

Tymczasem klasyczny funkcjonalizm może być problematyczny z zupełnie innych względów. Charakterystyczne dla niego jest twierdzenie, że dla realizacji jego funkcji umysłowych kluczowa jest organizacja systemu poznawczego, nie zaś jego materialny nośnik²⁰. Wyraża się to w tezach o „wielorakiej realizacji” czy też „niezależności od substratu”. Początkowo sądzono, że za pomocą takich tez można bronić autonomii nauk szczegółowych, zajmujących się właśnie organizacją układów złożonych²¹.

Należy w tym miejscu zaznaczyć, że można przyjmować funkcjonalizm, nie sądząc, że organizacja układu będzie pod każdym względem niezależna od jego materialnego nośnika. Co więcej, organizacja niektórych układów bardzo zależy od ich nośników. Mówiąc krótko, tezy o wielorakiej realizacji są motywowane funkcjonalistycznie, ale nie wynikają z samej tezy funkcjonalizmu. Dopiero klasyczny funkcjonalizm jest koniunkcją tezy o wielorakiej realizacji i ogólnej tezy funkcjonalizmu.

Teza o wielorakiej realizacji głosi, że możliwe będzie zrealizowanie tej samej funkcji przez różne układy fizyczne. Muszą się one różnić pod względem organizacji prowadzącej do realizacji funkcji. Jednak tradycyjne przykłady wielorakiej realizacji – na przykład ból u człowieka i u żaby, będący tą samą własnością różnie realizowaną neurologicznie – są błędne²². Złudzenie

¹⁹ Zob. np. R.S. J a c k e n d o f f, *Consciousness and the Computational Mind*, MIT Press, Cambridge, Massachusetts, 1990.

²⁰ Nie jest to, oczywiście, zastrzeżone wyłącznie dla funkcjonalistów. Na przykład biologowie teoretyczni uprawiający tak zwaną biologię relacyjną, jak Robert Rosen (zob. R. R o s e n, *Life Itself: A Comprehensive Inquiry into the Nature, Origin, and Fabrication of Life*, Columbia University Press, New York 1991), czy też zwolennicy ogólnej teorii systemów również przyjmują taką tezę.

²¹ Zob. F o d o r, *Nauki szczegółowe (albo: niejednorodność nauki jako hipoteza robocza)*.

²² Zob. W. B e c h t e l, J. M u n d a l e, *Multiple realizability revisited: Linking cognitive and neural states*, „Philosophy of Science” 66(1999), nr 2, s. 175-207; T.W. P o l g e r, *Evaluating the*

wielorakiej realizacji bierze się tutaj stąd, że realizowany ból jest identyfikowany bardzo niedokładnie (mimo że ludzie wyraźnie różnią się zachowaniem od żab), a realizująca struktura neurologiczna szczegółowo. Znalezienie niekontrowersyjnego przykładu wielorakiej realizacji okazuje się niezwykle trudne. Co więcej, sztandarowy przykład funkcjonalistów – komputer – realizowany za pomocą różnych układów elektronicznych czy mechanicznych, można bardzo różnie analizować, w zależności od tego, jak będzie się identyfikować ich funkcję²³. Jeśli funkcję scharakteryzujemy tak, aby uwzględnić szybkość przetwarzania, to r ó ż n e procesory czy układy elektroniczne będą realizować r ó ż n e funkcje. Jeśli uwzględnimy na przykład tylko oprogramowanie komputera, pomijając szybkość jego działania, to szybkość procesora okazuje się nerelevantna dla funkcji. Jest więc tym samym, co śmiech pingwina dla smaku śledzia: po prostu nieistotnym szczegółem. W takim wypadku wieloraka realizacja nie zachodzi. Mówiąc krótko, wielorakość jest dosyć słabo uchwytna.

Funkcjonalistyczne przeświadczenie, że organizacja bywa istotniejsza od nośnika czy substratu, można wyrazić nieco porządniej, mówiąc o „niezależności od substratu”²⁴. Nie chodzi tu oczywiście o całkowitą niezależność – tego funkcjonałiści nie mają na myśli, kiedy sądzą, że ta sama funkcja może być zrealizowana przez różnego rodzaju tranzystory: germanowe, krzemowe i z arsenku galu. Oczywiście, różne cechy tranzystorów wyznaczają ich zastosowania, a te cechy nie są zupełnie niezależne od tego, z czego tranzystor zbudowano; wręcz przeciwnie, tylko dzięki temu cechy te występują. Jednak z różnych tranzystorów można zbudować całkowicie równoważne urządzenia: bramki logiczne, które będą realizować te same funkcje logiczne, na przykład koniunkcję czy negację. Chalmers określa taką niezależność od substratu dokładniej „niezmiennością organizacyjną” (ang. organizational invariance)²⁵: podkreśla, że procesy obliczeniowe, czyli procesy będące w centrum zainteresowania funkcjonalizmu obliczeniowego, a więc i komputacjonizmu w kognitywistyce, mają taki właśnie charakter.

evidence for multiple realization, „Synthese” 2008, t. 167, nr 3, s. 457-472; L.A. Shapiro, *How to Test for Multiple Realization*, „Philosophy of Science” 75(2008) nr 5, s. 514-525; t e n ż e, *The Mind Incarnate*, MIT Press, Cambridge, Massachusetts, 2004; R. E n d i c o t t, *Resolving arguments by different conceptual traditions of realization*, „Philosophical Studies” 2010, t. 159, nr 1, s. 41-59.

²³ Por. M. M i ł k o w s k i, *Explaining the Computational Mind*, MIT Press, Cambridge, Massachusetts, 2013, s. 67-71.

²⁴ Por. D.C. D e n n e t t, *Darwin's Dangerous Idea: Evolution and the Meanings of Life*, Simon & Schuster, New York 1995, s. 50.

²⁵ Zob. D.J. C h a l m e r s, *A Computational Foundation for the Study of Cognition*, „Journal of Cognitive Science” (2011), nr 12, s. 325-359.

FUNKCJONALISTYCZNE TEORIE ŚWIADOMOŚCI

Najbardziej spektakularnym kontrargumentem przeciwko funkcjonalistycznemu ujęciu świadomości związanemu z niezmiennością organizacyjną jest argument Neda Blocka²⁶. Block przeprowadza następujący eksperyment myślowy: wyobraźmy sobie, że chiński rząd przekonał się do funkcjonalizmu i zobowiązał obywateli do realizacji ludzkiego umysłu przez godzinę. Miliard Chińczyków miałby symulować neurony, komunikując się ze sobą radiowo. Można by ich połączyć radiem ze sztucznym ciałem. Czy taki układ byłby świadomy? Dla Blocka jest jasne, że nie. Świadomość nie jest w jego koncepcji organizacyjnie niezmiennicza. Coś umyka: jakość przeżyć świadomych.

Możemy teraz wyrazić wątpliwość Chalmersa wobec teorii funkcjonalistycznych: wydaje się, że przeżycia świadome mają jakości (zwane przez niego z łaciny *qualiami*), które nie towarzyszą w sposób konieczny logicznie procesom obliczeniowym²⁷. Wiemy, że mogą im towarzyszyć, bo sami jesteśmy świadomi, ale możliwe jest też istnienie zombi: istot zachowujących się tak samo jak my i reagujących na te same bodźce, przetwarzających te same informacje, a pozbawionych świadomości.

Chalmers odróżnia tu przytomność (ang. *awareness*) od świadomości: ta pierwsza polega na informacyjnej dostępności stanów poznawczych, druga zaś na posiadaniu *qualiów*. Zombi mogą być przytomne, nie będąc świadome.

Zobaczmy, jak się to ma do pięciu wybranych teorii świadomości w stylu funkcjonalistycznym²⁸. Teoria globalnej przestrzeni roboczej²⁹ rzeczywiście charakteryzuje stany świadome jako te, które są dostępne w globalnej przestrzeni dla innych procesów obliczeniowych. Świadomość to w tym ujęciu właśnie przytomność w rozumieniu Chalmersa. Baars podkreśla, że samo znajdowanie

²⁶ Por. N. B l o c k, *Troubles with Functionalism*, w: tenże, *Collected Papers: Consciousness, Function, and Representation*, t. 1, MIT Press, Cambridge, Massachusetts, 2007, s. 71-73.

²⁷ Nie chcę tutaj zajmować się samymi *qualiami*, które charakteryzowane są albo mylnie (konjunkcja przypisywanych im własności prowadzi do sprzeczności), albo na tyle beztreściowo, że z użycia terminu „*quale*” nic dodatkowego nie można wynioskować, poza tym, że chodzi o przeżycie świadome. Zob. D. C. D e n n e t t, *Quining Qualia*, w: *Consciousness in Contemporary Science*, red. A. J. Marcel, E. Bisiach, Oxford University Press, Oxford 1988, s. 42-77; K. F r a n k i s h, *Quining diet qualia*, „*Consciousness and Cognition*” 21(2012) nr 2, s. 667-676.

²⁸ Ograniczam się tutaj do bardzo krótkiego ich scharakteryzowania. Charakteryzowane teorie wybieram za artykułem: M. M i ł k o w s k i, *Obliczeniowe teorie świadomości*, „*Analiza i Egzystencja*” 11(2010), s. 133-154.

²⁹ Zob. B. J. B a a r s, *In the Theater of Consciousness: The Workspace of the Mind* Oxford University Press, New York-Oxford 1997; por. B. J. B a a r s, S. F r a n k l i n, *Consciousness Is Computational: The Lida Model of Global Workspace Theory*, „*International Journal of Machine Consciousness*” 01(2009) nr 01, s. 23. Filozoficznym odpowiednikiem tej teorii jest teoria „wielokrotnych szkiców” Dennetta (zob. D. D e n n e t t, *Słodkie sny: filozoficzne przeszkody na drodze do nauki o świadomości*, tłum. M. Miłkowski, Prószyński i S-ka, Warszawa 2007).

się w globalnej przestrzeni nie jest równoznaczne z byciem przeżyciem świadomym. Teoria ta nie ujmuje jakościowego charakteru zdarzeń świadomych, lecz ma pokazać architekturę procesów przetwarzania informacji, które prowadzą do powstania między innymi świadomych przeżyć. Wyjaśnia, w jaki sposób stany obliczeniowe stają się świadome, a więc pozwala rozróżniać stany nieświadome i świadome. Jest więc teorią stanów świadomych, a nie teorią świadomego podmiotu (i nie wyjaśnia na przykład różnicy między jawą a snem).

Teoria świadomości jako myśli wyższego rzędu (ang. Higher-Order Thought – HOT) Davida Rosenthala również ma wyjaśnić proces uświadamiania, a jednocześnie swoistą strukturę samoświadomości, która towarzyszy świadomości (przynajmniej u człowieka)³⁰. Zgodnie z tą teorią myśl jest uświadomiona, gdy jest przedmiotem innej myśli. Same myśli drugiego rzędu rzadziej bywają uświadamiane, lecz są przyczyną uświadamiania innych. Tu rzeczywiście nie ma mowy o żadnych jakościach przeżyć, a jedynie o pewnej dostępności – ujmowanej w kategoriach bycia przedmiotem innej myśli (czyli dostępności dla innej myśli).

Nieco inny charakter ma teoria informacyjnej integracji. Jej twórcą jest współpracownik Geralda Edelmana³¹, Giulio Tononi³². Jego zdaniem świadomość określają dwie istotne cechy: (1) ma ona naturę informacyjną; (2) jest silnie zintegrowana. Tononi opracował miarę złożoności integracyjnej sieci przesyłających informacje, mającą służyć do eksperymentalnego pomiaru stopnia świadomości. Świadomość ma być efektem integracji informacji w sieci nerwowej; neuronalne korelaty świadomości mają się więc cechować swoistym stopniem zintegrowania (synchronizacji). Tym razem nie mówi się w ogóle o dostępności, lecz raczej o innej cesze organizacyjnej. Innymi słowy, Chalmersowskie rozróżnienie przytomności i świadomości słabo do tej teorii przystaje; można by jednak na obronę Chalmersa powiedzieć, że o przeżyciach nie mówi się tutaj w ogóle, bo teoria ta de facto stara się określić, czy dana organizacja jest na tyle złożona, aby mogła wytwarzać świadomość. Jak okaże się w części trzeciej niniejszych rozważań, własna teoria Chalmersa stoi w wyraźnej sprzeczności z ujęciem Tononiego, gdyż Chalmers nie wymaga, aby świadomy system informacyjny miał określoną złożoność.

Czy funkcjonalistyczne teorie nie starają się jednak w ogóle wyjaśnić świadomych przeżyć? Wbrew temu, co sugeruje Chalmers, istnieją takie teorie, których celem jest ujęcie, na czym polega jakościowość przeżycia ziele-

³⁰ Zob. D. R o s e n t h a l, *Consciousness and Mind*, Oxford University Press, Oxford–New York 2005.

³¹ Znany polskim czytelnikom z popularnej pozycji: G. E d e l m a n n, *Przenikliwe powietrze, jasny ogień: o materii umysłu*, tłum. J. Rączaszek-Leonardi, Państwowy Instytut Wydawniczy, Warszawa 1998.

³² Zob. G. T o n o n i, *An information integration theory of consciousness*, „BMC Neuroscience” 5(2004), s. 42.

ni. Taką teorią jest na przykład przywołana już wcześniej w przypisie koncepcja Raya Jackendoffa³³, który uważa, że konieczne jest także zbadanie struktury nośników informacji, czyli formatu reprezentacji świadomych. Sam stopień organizacji nie wystarcza. Należy odwołać się do modalności percepcyjnych.

Zdaniem Jackendoffa, tylko reprezentacje tak zwanego pośredniego szczebla mogą być treścią świadomości. Czyste dane percepcyjne, docierające z organów zmysłowych, nie są nam dostępne w świadomości (stanowią najniższy poziom reprezentacji); podobnie nie jesteśmy w stanie operować abstrakcyjnymi myślami w oderwaniu od modalności zmysłowych. Na przykład wyrazami operujemy zawsze w reprezentacji fonologicznej (lub graficznej, jeśli są to wyrazy języka tylko pisanego); nie są nam one dostępne w postaci czysto pojęciowej. Koncepcja Jackendoffa, choć powstawała nieco wcześniej od popularnych obecnie, jest bliska ujęciu Baarsa: bycie świadomym polega w niej na znajdowaniu się w pamięci krótkoterminowej (ang. short-term memory – STM) charakterystycznej dla danej modalności zmysłowej; reprezentacje z pamięci krótkoterminowej, na które skierowano uwagę, znajdują się w samym centrum świadomości. W ten sposób pojęcie dostępności (a więc i przytomności) zostaje powiązane z empirystycznym przekonaniem, że dla świadomości kluczowa jest percepcja. Istotne jest też to, że nie wszystkie dane percepcyjne są nam dostępne w reprezentacji zmysłowej, a więc nie wszystko jest nam dostępne świadomie. Niemniej, zgodnie z tą teorią neoempirystyczną, nie istnieją żadne informacje dostępne w świadomości, które nie miałyby natury quasi-percepcyjnej. Myśl świadoma ma zawsze naturę zmysłową.

Jak widać, istnieje jednak teoria obliczeniowa i funkcjonalistyczna, która mówi coś o przeżyciach. Otóż jakości przeżyć to jakości modalności zmysłowych: świadomość dlatego ma jakości, że informacje są przetwarzane przez systemy obsługujące dane percepcyjne. Możemy obstawać, że to nie jest wyjaśnienie trudnego problemu świadomości. Pojawia się jednak pytanie: jeśli wiemy, dlaczego zombi ma określoną modalność zmysłową, i to percepcyjną, posiada pamięć krótkoterminową i reprezentacje pośredniego szczebla, to czego będzie zombi brakować? Jeśli znamy strukturę przyczynową procesów uwagowych (bo takie właśnie wchodzi w grę w teorii Jackendoffa i w jej nowszej wersji u Prinza), to dlaczego coś miałyby nam umykać?

Eksperymenty Blocka czy Chalmersa przypominają Leibnizjański młyn z *Monadologii*³⁴: choćbyśmy odtworzyli umysł w postaci wielkiego młyna, wielkiego mechanizmu, to nigdzie nie zobaczymy percepcji. Istnieją rozmaite

³³ Zob. J a c k e n d o f f, dz. cyt. Zwolennikiem i kontynuatorem tej koncepcji jest Jesse Prinz w jego najnowszej książce (zob. J.J. P r i n z, *The Conscious Brain: How Attention Engenders Experience*, Oxford University Press, Oxford–New York 2012).

³⁴ Zob. G.W. L e i b n i z, *Monadologia*, tłum. H. Elzenberg, Uniwersytet Mikołaja Kopernika, Toruń 1991.

wersje tego eksperymentu myślowego: bodaj najbardziej popularne, oprócz odmian, w których występują zombi i Chińczycy, jest wyobrażenie sobie neuroložki Marii, która wychowana w czarno-białym świecie miałaby nie wiedzieć, jak to jest widzieć czerwone pomidory³⁵. Mniej więcej z tym samym mamy do czynienia i w tym przypadku: Maria ma bowiem znać tylko młyn, nie znając wcześniej percepcji. Do Marii jeszcze wrócimy.

NIEZMIENNOŚĆ ORGANIZACYJNA I PANPSYCHIZM

Chalmers stara się nas przekonać, że zombi są tylko logicznie, lecz nie przyrodniczo możliwe w naszym świecie. Są możliwe logicznie, gdyż nie widzimy jawnej sprzeczności w stwierdzeniu, że byłyby pozbawione jakości przeżyć świadomych³⁶. A więc świadomość nie wyczerpuje się w funkcjonalnym opisie, ponieważ sama funkcji – pojmowanej w kategoriach ról przyczynowych czy ról funkcjonalnych – po prostu w ogóle nie ma. Jest tak dlatego, że świadomość jest zupelnie epifenomenalna: umyka funkcjonalizmowi, ale możemy ją postulować jako towarzyszący przytomności cień. Jeśli to jest poważne potraktowanie świadomości, to nie wiem, dlaczego zdaniem Chalmersa behawioryzm jest niepoważny. Wszak skrajni behawioryści, jak Watson czy Skinner, byli właśnie takimi epifenomenalistami.

Wydaje się, że pokusa epifenomenalizmu bierze się z pragnienia obalenia funkcjonalistycznego podejścia do świadomości przez wskazanie, że nigdy nie wystąpi ona w ramach struktury przyczynowej (czy zależnościowo-funkcjonalnej) zjawisk psychicznych związanych z przytomnością. To bardzo mocna teza, bo prowadzi wprost do kłopotów z wyjaśnieniem, dlaczego w ogóle wiedza o świadomości jakościowej jest możliwa: wiedza też musi zostać zanalizowana w kategoriach nieprzyczynowych, co Chalmers zresztą czyni³⁷. Jednak jego analiza – uprawiana w kategoriach bezpośredniości (czy też wiedzy przez znajomość) – jest zwodnicza; z faktu, że coś jest znane bezpośrednio, nie wynika, że znajomość ta nie ma podstaw przyczynowych, które zupełnie ją tłumaczą. Dopóki mamy jakąkolwiek zależność między jakimś zjawiskiem psychicznym a świadomością w rozumieniu Chalmersa (czyli z jakościowymi przeżyciami świadomymi), dopóty nie jest to epifenomenalizm, który wystarczałby do wbicia osinowego kołka w funkcjonalistycznego wampira. Wystarczający byłby dopiero absolutny epifenomenalizm, ale takie stanowisko jednocześnie podważyłoby sensowność

³⁵ Zob. F. Jackson, *Czego nie wiedziała Maria?*, tłum. T. Ciecierski, „Przegląd Filozoficzno-Literacki” 2003 nr 4(6), s. 9-14.

³⁶ Na to fizykalista może odpowiedzieć, że widzi sprzeczność, bo zombi zgodnie z jego definicją świadomości (np. w teorii Baarsa) będzie świadomy, będąc przytomny.

³⁷ Por. Chalmers, *Świadomy umysł: w poszukiwaniu teorii fundamentalnej*, s. 354.

tej teorii świadomości: jeśli nic od takiej świadomości nie zależy, to jakim cudem może się ona obronić przez zasadą niemnożenia bytów ponad konieczność? Mówiąc krótko, świadomość mogłaby być zupełnie epifenomenalna, gdyby od niej kompletnie n i c nie zależało, łącznie z wiedzą na jej temat. Jakakolwiek zależność wiedzy o świadomości od owej świadomości konstytuowałaby bowiem jej rolę, a przez to dawałaby podstawę do jej funkcjonalizacji. Chalmers stara się przyjąć epifenomenalizm, ale jest to epifenomenalizm selektywny: świadomość nie ma mocy sprawczej, a więc przyczynowej, ale ma moc ujawniania się w bezpośredniej jej znajomości. Sелеktywny epifenomenalizm to jednak za mało. Gdyby jednak przyjął epifenomenalizm absolutny, wówczas jego teoria nie mogłaby postulować bytu takiego jak świadomość, bo nawet jako wytwór innych procesów nie mogłaby ona na nic wpływać w jakikolwiek sposób.

Jaką więc teorię chce nam Chalmers zaoferować w zamian? Otóż nie neguje on bynajmniej teorii funkcjonalistycznych, które opierają się na pojęciu dostępności czy przytomności. Za pomocą serii eksperymentów myślowych Chalmers stara się dowodzić, że w naszym świecie – bo nie na mocy pojęć – przytomności t o w a r z y s z y świadomość. A to dlatego, że obowiązuje „zasada niezmienności organizacyjnej”. Nie ma mowy o zależności, a jedynie o współwystępowaniu przytomności i świadomości. Ta nowa zasada gwarantuje współwystępowanie przeżyć z rolami funkcjonalnymi myśli, co w konwencjonalnym funkcjonalizmie zapewniała po prostu przyczynowość. Głosi ona – w skrócie – że przeżycia określonego rodzaju zawsze towarzyszą układom funkcjonalnym o tej samej (abstrakcyjnie pojętej) organizacji, niezależnie od materialnej realizacji tak zorganizowanych układów. Pojawia się jednak zasadnicza wątpliwość, czy koncepcja ta nie opiera się na sztuczce werbalnej: nazywamy przyczynowość bądź zależność „współwystępowaniem” czy „towarzyszeniem”, ale de facto nic się nie zmieniło. Można odnieść wrażenie, że strukturalnie teoria ta nie odbiega od funkcjonalizmu, lecz zamiast ról przyczynowych oferuje klasyczne dualistyczne rozwiązanie problemu z przyczynowością: harmonię przedustawną. Sęk w tym, że harmonie przedustawne w siedemnastym wieku mogły uchodzić za interesującą i płodną hipotezę, lecz dziś metodologia nauk – a Chalmers cały czas twierdzi, że tworzy podstawy n a u k o w e j teorii świadomości – wymaga dokładnego określenia, czym one są.

Zasada niezmienności organizacyjnej nie jest nową tezą o świadomości czy przeżyciach, lecz „łatanie dziury” po przyczynowej roli przeżyć, którą Chalmers wcześniej odrzucił, chcąc poradzić sobie z postulowaną przez siebie niesprzecznością pojęcia zombi. Korzystając z pomysłowych eksperymentów myślowych, Chalmers pokazuje następnie, że nieintuicyjne byłoby odrzucenie zależności qualiów od organizacji funkcjonalnej systemu poznawczego³⁸. Istota większości

³⁸ Zob. tamże, rozdz. 7, s. 411-455.

z nich polega na tym, że wyobrażamy sobie izomorficzny odpowiednik funkcjonalny świadomego człowieka, który jednak zbudowany jest z elementów elektronicznych (np. półprzewodnikowych). Następnie buduje się szereg pośrednich stadiów między człowiekiem a jego odpowiednikiem funkcjonalnym i argumentuje, że wyznaczenie gdziekolwiek granicy, mówiącej, że tu właśnie zanikają przeżycia świadome i zostaje tylko przetwarzanie informacji, byłoby arbitralne.

Niemniej arbitralność ta dotyczyć ma jedynie, o dziwo, świata rzeczywistego, a nie wszelkich światów możliwych. Innymi słowy, Chalmers twierdzi, że w innych światach świadome przeżycie może w ogóle się nie pojawić mimo istnienia tej samej struktury przyczynowej. Jest to dosyć osobliwe: w innych światach, a więc w światach, gdzie nie obowiązuje zasada niezmienności organizacyjnej, taką granicę można wyznaczyć. Ale ona także jest intuicyjnie arbitralna. Mówiąc krótko, jego argumentacja zdaje się dosyć pokrętna.

Chalmers przypisuje tezie o związku świadomości z przytomnością charakter konieczności przyrodniczej, a nie logicznej. Podczas gdy fizykaliści, jego zdaniem, *m u s z ą* twierdzić, że związek ten będzie konieczny logicznie, dualista może przyjąć słabszą odmianę tej konieczności. Chalmers jednak zupełnie zapomina o możliwości stanowiska, zgodnie z którym wszystkie związki pojęciowe traktuje się jako potencjalnie podważalne, mimo że jest ono dominujące w filozofii analitycznej od czasów Quine'a. Co więcej, nasze definicje podlegają rewizjom, mimo że mówimy o tych samych przedmiotach: teorie naukowe dotyczyły tych samych przedmiotów, choćby przyjmowały różne ich definicje, jak pokazywał już Putnam³⁹. Trudno zatem dociec, dlaczego mielibyśmy uważać, że jakiś szczególnie nieobalalny charakter przysługuje definicjom pojęć takich jak „świadomość” czy „przytomność”. Poza tym, zdaniem niejednego fizykaliisty – czy klasycznego funkcjonalisty – niesprzeczne wyobrażenie zombi nie jest możliwe, bo zombi na mocy swojej organizacji funkcjonalnej jest świadomy, gdyż takie właśnie rozumienie pojęcia świadomości przyjmuje się w teoriach funkcjonalistycznych. Wyobrażenie to jest sprzeczne; brak sprzeczności w wyobrażeniu niczego w istocie nie dowodzi, a jedynie wynika z przyjęcia innej teorii świadomości, w której inaczej się to pojęcie rozumie.

Naturalistyczni fizykaliści najczęściej w ogóle nie przypisują mocnego modalnego charakteru swoim twierdzeniom: skoro można uważać, że dowiemy się zapewne jeszcze sporo o wodorze, modyfikując tym samym nasze teorie chemiczne, a może nawet zmieniając znaczenie wyrazu „wodór”, to można identycznie traktować również pojęcie „świadomość”, które wraz z badaniami naukowymi również zostanie sprecyzowane w ramach kolejnych teorii. Wszelkie ogólne tezy na temat fizycznej realizacji świadomości opierają się, zdaniem

³⁹ Zob. H. Putnam, *Wiele twarzy realizmu i inne eseje*, tłum. A. Grobler, Wydawnictwo Naukowe PWN, Warszawa 1998.

większości fizykalistów – i to fizykalistów naturalistycznych, z którymi Chalmers polemizuje – jedynie na skończonej liczbie świadectw empirycznych, które nigdy w pełni nie uzasadniają tez z kwantyfikatorem uniwersalnym. Podobnie jak Hume, tacy fizykaliści uznają te tezy za przygodne, a nie konieczne, ale nie czyni to z nich dualistów, mimo że przyjmują stanowisko istotnie słabsze od stanowiska Chalmersa. Z tego też względu dualizm Chalmersa wydaje się dosyć naciągany, a skądinąd interesujące zagadnienia modalności i analizy pojęciowej – zupełnie niezwiązane z podstawowym tematem jego książki. W dyskusji tej bowiem nie o modalność chodzi, lecz o naturę świadomości. Wypowiadanie tez mocniejszych niż przygodnie prawdziwe o empirycznie dostępnych zjawiskach wydaje się po prostu zupełnie irracjonalne, zgodnie z Kazimierza Ajdukiewicza zasadą racjonalności, iż sądy należy uznawać z taką mocą, z jaką są one uzasadnione. A prawa nauk empirycznych, z definicji kwantyfikowane uniwersalnie, nigdy nie będą uzasadnione tak, by można było poważnie się na mówienie o ich l o g i c z n e j konieczności.

Zostawmy kwestię dualizmu – realnego czy pozornego – na boku. Warto zauważyć, że Chalmersa teoria świadomości nie opiera się wyłącznie na reformułowaniu tez klasycznych funkcjonalistów, lecz zawiera istotne novum – spekulację na temat jakościowo-świadomego charakteru wszelkiej informacji (rozumianej strukturalnie; nie w sensie Shannona, jak błędnie sugeruje Chalmers⁴⁰, tylko tak jak logon u MacKaya⁴¹). Zgodnie z tą spekulacją świadomy bądź przynajmniej protoświadomy charakter ma wszystko, w czym tylko da się wskazać dowolne zróżnicowanie. Tak więc wszystko – poza izolowanymi punktami czy podobnie idealnymi bytami – będzie wyposażone w jakościową protoświadomość. Teoria Chalmersa hojnie przypisze więc praktycznie wszystkim bytom świadomość, w tym cząstkom elementarnym, choć nie odpowie, w jaki sposób taka świadomość ma się złożyć na świadomość na przykład nietoperzy czy ludzi. Taka spekulacja przede wszystkim nie ma żadnej wartości eksplanacyjnej, bo nie musimy wyjaśniać, dlaczego atomy mają świadome przeżycia, a jedynie wynika z odrzucenia przez Chalmersa pojęcia emergencji jako niejasnego. A skoro całości takie jak mózgi mają świadomość, to świadome muszą być części, sądzi Chalmers. No cóż, w podręcznikach logiki taką strukturę argumentu nazywa się błędem: nie jest prawdą, że skoro cały samochód może pojechać do Małkini, to pojedyncze koło samochodu może pojechać do Małkini. Rozumowanie takie jest po prostu niepoprawne, a żeby je odrzucić, nie trzeba uznawać żadnych niepewnych koncepcji meta-

⁴⁰ Zob. W.F. H a r m s, *The Use of Information Theory in Epistemology*, „Philosophy of Science” 65(1998) nr 3, s. 472-501.

⁴¹ Zob. D.M. M a c K a y, *Information, Mechanism and Meaning*, M.I.T. Press, Cambridge, Massachusetts, 1969.

fizycznych, lecz wystarczy przyjąć prostą tezę, że całości miewają własności inne od własności części (np. własność bycia taką właśnie całością).

Co więcej, skoro świadomość nie wymaga zintegrowanego przetwarzania informacji, to teoria Chalmersa neguje teorię Tononiego, która dostarcza również coraz lepszych miar ilościowych świadomości w neuronauce⁴². Spekulacja zastępuje u Chalmersa empirię – i tak jak Hegel dedukował filozoficznie, ile będzie planet, tak autor *Świadomego umysłu* śmiało podważa cały dział badań neuropsychologicznych związany z określaniem minimalnej złożoności układu nerwowego wymaganej dla istnienia świadomości. A zdawałoby się, że czasy gabinetowego uprawiania filozofii już minęły. Oczywiście, można by odzecz, że Tononi nie przedstawia danych empirycznych, lecz tylko teorię, więc Chalmers może ją zanegować. Chodzi tu wszak o założenie teorii Tononiego, że świadome są tylko odpowiednio wysoko złożone systemy przetwarzania informacji. Nie wiadomo jednak, czemu negacja tego założenia miałaby służyć i jak jest ona uzasadniona: jakie mianowicie zjawiska ma wyjaśnić, które bez negacji tego założenia pozostają niewyjaśnione (zauważyć warto, że Tononi może bronić swojego założenia dosyć wiarygodną, choć przygodną tezę, że świadomość łączy się z odpowiednio dużą zmiennością zachowań będących reakcjami na różne bodźce, a takie zachowania u bardzo prostych organizmów po prostu nie występują). Na mocy (proto)panpsychizmu obliczający kalkulator, zeszyt i drożdże mają świadome przeżycia. Cały świat przenikają świadome jakości przeżyć. Tyle że jakości te nic nie robią, tylko są; wszak są epifenomenalne.

Osobliwy ten świat Chalmersa. Trzeba jeszcze uznać fundamentalne prawa psychofizyczne, które łączą świat przeżyć z całym światem oddziaływać, aby teoria była kompletna; co prawda nie wiadomo, dlaczego to robimy, skoro przeżycia nic nie wywołują ani nic ich nie wywołuje. Szczerze mówiąc, trudno wyobrazić sobie gorszą teorię świadomości.

POŻEGNANIE Z MARIĄ

Wcześniej napomknąłem już o znanej z eksperymentu myślowego Jacksona neurolożce Marii⁴³. Załóżmy, że Chalmers ma rację i argument z wiedzy jest

⁴² Zob. A.K. Seth i in., *Theories and measures of consciousness: an extended framework*, „Proceedings of the National Academy of Sciences of the United States of America” 103(2006) nr 28, s. 10799-10804; A.K. Seth i in., *Measuring consciousness: relating behavioural and neurophysiological approaches*, „Trends in Cognitive Sciences” 12(2008) nr 8, s. 314-321; A. Seth, *Explanatory Correlates of Consciousness: Theoretical and Computational Challenges*, „Cognitive Computation” 1(2009) nr 1, s. 50-63.

⁴³ Zob. J a c k s o n, dz. cyt.

przekonujący. Załóżmy, że dysponująca całą wiedzą fizyczną Maria istotnie nie wie, jak to jest widzieć pomidora. Ale czy spekulacja Chalmersa – zasada niezmienności organizacyjnej, panpsychizm czy fundamentalne prawa psychofizyczne – da nam powody, aby uważać, że będzie to wiedziała Maria-dualistka? Nie wydaje się. Maria może co najwyżej wiedzieć, że pomidor będzie miał świadome przeżycia, które będą mu konieczne – lecz tylko przyrodniczo konieczne – towarzyszyć. Maria może też wiedzieć, iż powinna mieć jakieś przeżycia w kontakcie z różnymi treściami informacyjnymi. Ale jakie? Tego teoria wcale nie mówi; co gorsza, w ogóle nie wiadomo, co do niej należałoby dodać, żeby Maria mogła wiedzieć, jak to jest widzieć pomidora, zanim go zobaczy.

Być może ten przykład jest jednak zwodniczy; to nie argument przeciwko fizykalizmowi czy dualizmowi, lecz wykorzystanie faktu, że przeżycia mają subiektywny charakter, a więc wiążą się ze swoim podmiotem tak, że inne podmioty nie mają do nich bezpośredniego dostępu. Skoro Maria nigdy nie miała takich przeżyć, to zdrowy rozsądek podpowiada, że nie będzie mogła przewidzieć, jakie przeżycia będzie mieć. Nie znaczy to jednak, że dualizm lub fizykalizm jest fałszywy. Co innego bowiem mieć teorię, a co innego reprezentować coś w sposób charakterystyczny dla przeżyć wzrokowych i pamięci epizodycznej: Maria nie ma dostępu do odpowiednio zakodowanego przeżycia w pamięci roboczej – i tyle⁴⁴. Ludzie – całe szczęście, bo prowadziłyby to do częstych halucynacji – nie potrafią wzbudzać w sobie dowolnych reprezentacji wyłącznie na podstawie opisu, ale gdyby Maria była robotem, mogłaby sobie chwacko z tym radzić⁴⁵. Nie ma tu miejsca na szczegółowe omawianie tego argumentu; chcę jedynie podkreślić, że błędem jest zbyt szybkie przechodzenie od zdroworozsądkowego przeczcucia do tez metafizycznych. Dualista musi również wyjaśnić, dlaczego nawet Maria-dualistka nie byłaby w stanie sobie wyobrazić, jak to jest widzieć pomidora (jeśli uważa, że Maria-fizykalistka by tego nie zdołała uczynić).

Chalmers musiałby uznać, że Maria-dualistka pozna jednak charakter jakościowy przeżyć, ale nie wiadomo, co uzasadniałoby jego sąd: żaden aspekt jego teorii nie tłumaczy jakości przeżyć, a jedynie mówi, że one występują. Ale nie mówi jak i e! Skoro tak, to jeśli przykład z Marią wystarcza do odrzucenia funkcjonalizmu i fizykalizmu, wówczas wystarcza on też do odrzucenia teorii Chalmersa, bo (nawet po uzupełnieniu treści praw psychofizycznych)

⁴⁴ Por. P r i n z, dz. cyt., s. 324. Swoją drogą, wyobrażenia i pamięć epizodyczna są realizowane przez te same mechanizmy mózgowe, tak więc wyobrażanie sobie obrazowo możliwe jest właśnie dzięki pamięci epizodycznej. Zob. S.E. T a y l o r, *Envisioning the Future and Self-Regulation*, w: *Predictions in the Brain: Using Our Past to Generate a Future*, red. M. Bar, Oxford University Press, Oxford 2011, s. 135-143.

⁴⁵ Zob. D e n n e t t, *Słodkie sny: filozoficzne przeszkody na drodze do nauki o świadomości*.

teoria ta w żaden sposób nie umożliwia Marii przewidzenia, co odczuje. Jeśli Chalmers zaprotestuje, twierdząc, że ten przykład nie wystarcza do obalenia jego teorii, to wtedy też nie wystarcza on do odrzucenia funkcjonalizmu.

Chalmers stoi przeto przed dylematem: albo uzna, że tak naprawdę głosi jedynie wersję funkcjonalizmu ze wszelkimi nieintuicyjnymi konsekwencjami w postaci grupowej świadomości miliarda Chińczyków, i przyzna, że trudny problem świadomości jest pseudoproblemem, albo uzna, że jego własna teoria jest nie do przyjęcia, bo trudnego problemu świadomości nie tłumaczy. Ja zaś sugeruję, że jego teoria jest nie do przyjęcia także z innych powodów: po prostu nie jest nawet zaczątkiem poważnej teorii świadomości, która nie powinna głosić epifenomenalizmu, lecz rzetelnie wyjaśniać znane zjawiska, dobrze je opisując i wyróżniając.