

Jakub GOMUŁKA

## ZAGADNIENIE ŚWIADOMOŚCI WE WSPÓŁCZESNEJ ANALITYCZNEJ FILOZOFII UMYŚLU

*Początek tego, co zwiemy dzisiaj kognitywistyką, związany jest z wytworzeniem się paradygmatów badawczych obiecujących całkowitą naturalizację umysłu i psychiki, a więc sprowadzenie tych fenomenów do poziomu innych faktów dostępnych normalnym metodom matematycznego (i fizykalistycznego) przyrodoznawstwa. Jednym z tych paradygmatów był behawioryzm, wyrosły na badaniu wrodzonych i wyuczonych zachowań zwierząt, drugim zaś komputacjonizm, którego podstawę stanowiły algorytmika i teoria przetwarzania informacji, związane z rozwojem komputerów.*

Problem świadomości rozważany jest w filozofii zachodniej praktycznie od jej początków<sup>1</sup>. Bywał on jednak ujmowany bardzo różnorodnie ze względu na różnorodne konteksty, w jakich występował. W niniejszej pracy zamierzam skupić się na jednym z tych kontekstów (choć może lepiej powiedzieć: na jednej rodzinie kontekstów), a mianowicie na współczesnej analitycznej filozofii umysłu. Ścisłej rzecz biorąc, zamierzam zaprezentować toczony w niej spór o możliwość fizykalistycznej redukcji świadomości. W pierwszej części artykułu przedstawię historyczne korzenie owego sporu, a w drugiej najważniejsze argumenty za i przeciw nieredukowalności świadomości, w trzeciej zaś omówię i rozwinę pewne ciekawe stanowisko pośrednie. Moje rozważania dotyczyć będą przede wszystkim dyskusji na temat tak zwanej świadomości fenomenalnej, a także tak zwanych qualiów.

Uprawianie analitycznej filozofii umysłu wiąże się dzisiaj niemal automatycznie z realizacją Hellerowskiego postulatu metodologicznego filozofowania w kontekście nauki<sup>2</sup>. Trudno bowiem, podchodząc do tej dziedziny poważnie, całkowicie zamknąć się na teorii i wyniki wypracowywane w ramach szeroko pojętego programu badań kognitywistycznych. W zakres tego programu wchodzi prace prowadzone w obrębie informatyki, lingwistyki, teorii informacji, teorii sterowania, psychologii poznawczej, etologii zwierząt, a przede wszystkim neurobiologii. Przejawem wciąż wzrastającego znaczenia tej ostatniej jest stop-

<sup>1</sup> Por. K. G l o y, *Wprowadzenie do filozofii świadomości. Problematyka i historia zagadnienia świadomości oraz samoświadomości*, tłum. T. Kubalica, Wydawnictwo WAM, Kraków 2009, s. 89-122.

<sup>2</sup> Zob. M. H e l l e r, *Jak możliwa jest „filozofia w nauce”?*, „Studia Philosophiae Christianae” 22(1986) nr 1, s. 9-19 (przedruk w: tenże, *Filozofia i Wszechświat*, Universitas, Kraków 2006).

niowe pojawianie się coraz większej ilości dyscyplin, których nazwy zawierają przedrostek „neuro-”; obok neuropsychologii pojawiły się takie, jak: neurolingwistyka, neuroinformatyka, neurosocjologia, neuroekonomia czy neuroetyka<sup>3</sup>. Rzecz jasna, pojawiła się także neurofilozofia, którą – jak wskazał między innymi Łukasz Kurek – można uznać za przykład filozofii w nauce (w odróżnieniu od filozofii neuronauk)<sup>4</sup>. Kompromisowa propozycja rozwiązania sporu o redukcję świadomości, którą przedstawię pod koniec tekstu, choć daleka od negowania wagi osiągnięć nauk empirycznych dla filozoficznych rozważań, może być odczytywana jako próba obrony autonomii i specyfiki refleksji filozoficznej.

### HISTORIA PROBLEMATYKI ŚWIADOMOŚCI W ANALITYCZNEJ FILOZOFII UMYSŁU I KOGNITYWISTYCE

Założeniem postulatu filozofowania w kontekście nauki jest niemożliwość rozgraniczenia zagadnień naukowych (fizycznych, biologicznych i im podobnych) od zagadnień filozoficznych. Problem umysłu od początku swojej nowożytnej historii w wieku siedemnastym wydaje się jedną z najbardziej wyrazistych egzemplifikacji tej niemożliwości. Podjęcie zadania polegającego na wyznaczeniu miejsca dla sfery doświadczenia subiektywnego i treści umysłowych w obrębie nowożytnego obrazu świata doprowadziło Kartezjusza do uformowania znanej koncepcji dualizmu interakcyjnego, w którym świadomość znalazła się poza światem zjawisk fizycznych, choć była z nim powiązana przyczynowo. Jednak warunkiem zaistnienia samego problemu było powstanie owego obrazu świata, co wiązało się ściśle z powstaniem matematycznego przyrodoznawstwa. Zatem wbrew wrażeniu, jakie można dzisiaj odnieść po lekturze popularnych omówień, również dualizm psychofizyczny może być dowodem na uwikłanie filozofii w naukę.

Propozycja Kartezjusza natychmiast wywołała jednak tak zwany problem psychofizyczny, to znaczy problem związku między rzeczywistością fizyczną a rzeczywistością psychiczną. Jak się bowiem wydaje, zarówno *res extensa*

<sup>3</sup> Niektórzy badacze zwracają uwagę na fakt, że zagadnienie umysłu było w historii ujmowane za pomocą szeregu metafor, których formę narzucały aktualnie wiodące dziedziny nauki i technologii. Arystoteles przyrównywał władzę zapamiętywania do woskowej tabliczki, Leibniz używał mechanistycznych metafor zegarów i młyna, Freud stosował metaforykę hydrauliczną i elektromagnetyczną, później pojawiły się metafory związane z łącznością telefoniczną, a w końcu metafora komputerowa (por. np. J. R. S e a r l e, *Umysł, mózg i nauka*, tłum. J. Bobryk, PWN, Warszawa 1995, s. 40; S. W r ó b e l, *Umysł, gramatyka, ewolucja. Wykłady z filozofii umysłu*, PWN, Warszawa 2010, s. 113). Można powiedzieć, że dziś, w dobie dynamicznie rozwijających się nauk biologicznych, powszechnie przyjmowaną „metaforą” umysłu jest działający mózg.

<sup>4</sup> Por. Ł. K u r e k, *Neurofilozofia jako filozofia w kontekście nauki*, w: *Oblicza racjonalności*, red. B. Brożek, J. Mączka, W.P. Grygiel, M.L. Hohol, Copernicus Center, Kraków 2011, s. 63.

wpływa na *res cogitans*, wywołując wrażenia i prowadząc do powstania wielorakich reprezentacji świata fizycznego, jak też *res cogitans* oddziałuje na *res extensa*, bo przecież jesteśmy w stanie realizować nasze świadome plany i zamierzenia. Problem związku psychofizycznego nigdy nie został przez Kartezjusza rozwiązany do końca. Francuski filozof mówił jedynie o zachodzeniu związków przyczynowych zarówno „w dół”, czyli od umysłu do świata, jak i „w górę”, czyli od świata do świadomości. W swoich pismach wskazywał co prawda na miejsce zachodzenia owych relacji (słynna szyszynka), ale przyznawał, że nie rozumie sposobu, w jaki dwie wyróżnione przez niego sfery rzeczywistości kontaktują się ze sobą<sup>5</sup>.

Dualizm psychofizyczny nie jest we współczesnej filozofii stanowiskiem zbyt popularnym, co ma zapewne związek z faktem, że nie udało się jak dotąd przedstawić zadowalającego rozwiązania problemu interakcji. Ten ostatni można poglądowo zaprezentować za pomocą następującego trylematu:

1. Rzeczywistość fizyczna jest domknięta przyczynowo.
2. Między rzeczywistością fizyczną a umysłem zachodzą interakcje.
3. Umysł i rzeczywistość fizyczna to odmienne sfery bytu<sup>6</sup>.

Nie można bez popadania w sprzeczność utrzymywać, że trzy powyższe tezy są jednocześnie prawdziwe, dualista musi więc odrzucić którąś z nich. Odrzucenie pierwszej nie musi prowadzić do podważenia jednej z najsilniejszych zasad matematycznego przyrodoznawstwa – zasady zachowania masy-energii (co oczywiście nie jest uważane za racjonalną alternatywę), można bowiem próbować tworzyć koncepcje wymiany informacji w oparciu o efekty kwantowe<sup>7</sup>, droga ta jednak nie jest wolna od istotnych trudności. Odrzucenie drugiej prowadzi do przyjęcia jakiejś formy dualizmu nieinterakcyjnego (np. epifenomenalizmu), jest to jednak rodzina stanowisk dość nieintuicyjnych. Odrzucenie trzeciej definitywnie likwiduje problem, ale oznacza *de facto* odrzucenie dualizmu na rzecz jakiejś formy monizmu (nie musi to być monizm materialistyczny).

Początek tego, co zwiemy dzisiaj kognitywistyką, związany jest z wytworzeniem się paradygmatów badawczych obiecujących całkowitą naturalizację umysłu i psychiki, a więc sprowadzenie tych fenomenów do poziomu innych faktów dostępnych normalnym metodom matematycznego (i fizykalistycz-

<sup>5</sup> Por. R. Descartes, *Listy do księżniczki Elżbiety*, tłum. J. Kopania, PWN, Warszawa 1995, s. 10.

<sup>6</sup> Podobne sformułowanie trylematu psychofizycznego można znaleźć w: J. Bremer, *Wprowadzenie do filozofii umysłu*, Wydawnictwo WAM, Kraków 2010, s. 29-32.

<sup>7</sup> Czynnikiem tak John Eccles z Friedrichem Beckiem (zob. J. Eccles, F. Beck, *Quantum Processes in the Brain: A Scientific Basis of Consciousness*, w: *Neural Basis of Consciousness*, red. N. Osaka, John Benjamins Publishing, Amsterdam 1994). Na efekty kwantowe w swoim wyjaśnieniu zjawiska świadomości powołują się także Roger Penrose i Stuart Hameroff (zob. np. R. Penrose, S.R. Hameroff, *Consciousness in the Universe: Neuroscience, Quantum Space-Time Geometry & Orch OR Theory*, „Journal of Cosmology” 2011, t. 14), ich teoria jednak nie jest dualistyczna.

nego) przyrodoznawstwa. Jednym z tych paradygmatów był behawioryzm, wyrosły na badaniu wrodzonych i wyuczonych zachowań zwierząt, drugim zaś komputacjonizm, którego podstawę stanowiły algorytmika i teoria przetwarzania informacji, związane z rozwojem komputerów. Pierwsza z tych perspektyw pozwoliła spojrzeć na zjawiska umysłowe jak na zespół wrodzonych i nabytych dyspozycji do określonych zachowań. Druga zaszczerpiła w powszechnej świadomości tak zwaną metaforę komputerową, według której umysł jest niezmiernie (choć jednak skończone) skomplikowanym programem przetwarzającym informacje dostarczane z otoczenia i odpowiednio sterującym naszym ciałem, a mózg to w swej istocie biologiczny komputer (ang. *wetware*), pewna instancja maszyny Turinga, na której ów program jest realizowany.

Wspomnianą na początku ideę neuronauki, a więc bardziej współczesny sposób myślenia o umyśle, można w dużym uproszczeniu potraktować jako pewnego rodzaju rozwiniętą syntezę tych dwóch podejść. Zaawansowane i szczegółowe badania żywej tkanki nerwowej (rzecz jasna, przede wszystkim u zwierząt), a także rosnąca liczba eksperymentów dotyczących funkcji poznawczych u osób z fizycznymi uszkodzeniami mózgu doprowadziły do wytworzenia wiarygodnego modelu rzeczywistego przetwarzania przez mózg informacji płynących z otoczenia. Model ten nadal pozostaje bardzo fragmentaryczny, miejscami jednak jest już na tyle szczegółowy, że umożliwia na przykład prace nad elektronicznymi protezami oczu. Według niego wiele z funkcji przetwarzania informacji dokonywanych jest nie szeregowo – jak w tradycyjnej architekturze opartej na idei maszyny Turinga – lecz równolegle, w sposób rozproszony, w strukturach, które modeluje się za pomocą tak zwanych sieci neuronalnych czy też, mówiąc ściślej, neuropodobnych. „Uczenie” sieci neuropodobnej polega w skrócie na ustaleniu wag (ang. *weight*) połączeń między poszczególnymi węzłami sieci (czemu w rzeczywistym mózgu może odpowiadać siła połączeń między poszczególnymi neuronami), co dokonuje się na zasadzie podobnej do behawiorystycznych zasad warunkowania<sup>8</sup> i jest w gruncie rzeczy funkcjonalnie identyczne z wyrabianiem właściwych reakcji na dany bodziec.

Powstanie i postępy badań kognitywistycznych, a później neurokognitywistycznych w sposób istotny wpłynęły na to, co działo się w filozofii umysłu, zwłaszcza w jej odmianie analitycznej, znacząco bardziej otwartej na wpływy nauk przyrodniczych. Klasyczny komputacjonizm przyczynił się do popularności współczesnej funkcjonalistycznej wizji umysłu<sup>9</sup> oraz różnych

<sup>8</sup> Wypada uściślić, że w przypadku sieci neuropodobnych, w przeciwieństwie do koncepcji behawiorystycznej, wnętrze „czarnej skrzynki” jest dostępne dla badań i teoretycznie istotne. Jednakże semantycznie relewantne treści pojawiają się tylko na wejściu i wyjściu – wewnętrzny rozkład wag połączeń między węzłami sieci nie stanowi zapisu informacji w sensie reprezentacji symbolicznej.

<sup>9</sup> Jej twórcą jest Hilary Putnam (zob. np. H. P u t n a m, *Minds and Machines*, w: *Dimensions of Mind*, red. S. Hook, New York University Press, New York 1960), który sam później się z niej wycofał

teorii identyczności<sup>10</sup>. Z kolei psychologia behawiorystyczna odnalazła swoje filozoficzne przedłużenie w tak zwanym behawioryzmie logicznym<sup>11</sup>. Były to zasadniczo próby zmierzające do zredukowania umysłu do zjawisk naturalnych bądź przynajmniej dające się uzgodnić z zamysłem przeprowadzenia takiej redukcji.

Ujęcia teoretyczne redukujące sferę umysłową do obszaru badań fizyki napotykały jednak (i wciąż napotykają) opór ze strony przedstawicieli licznych stanowisk filozoficznych. Można sądzić, że podstawowym źródłem tego oporu jest ta sama intuicja, która skłoniła Kartezjusza do wydzielenia specjalnej sfery *res cogitans* – intuicja zasadniczej odmienności natury subiektywności wobec wszystkiego, co napotykamy „na zewnątrz”, a co składa się na tak zwaną obiektywną rzeczywistość. Intuicja ta nie zawsze ujawnia się z całą wyrazistością; czasem, jak na przykład w pracach Johna Searle’a, przesłaniana jest tezami dotyczącymi języka i intencjonalności<sup>12</sup>.

Jedną z pierwszych prac w dziedzinie analitycznej filozofii umysłu, w których dążenie do ochrony autonomii sfery subiektywnej znalazło swój pełny wyraz, był głośny artykuł Thomasa Nagela *Jak to jest być nietoperzem?* z roku

---

(zob. np. t e n ż e, *Funkcjonalizm – kognitywistyka czy fantastyka naukowa?*, tłum. M. Iwanicki, w: *Analityczna metafizyka umysłu. Najnowsze kontrowersje*, red. M. Miłkowski, R. Poczobut, IFiS PAN, Warszawa 2008). Źródeł funkcjonalizmu można jednak doszukiwać się już u Arystotelesa.

<sup>10</sup> Istnieje teoria identyczności typów (ang. type–type identity theory) mówiąca o tym, że typy zjawisk psychicznych odpowiadają typom zjawisk fizycznych (zob. np. U. T. P l a c e, *Is Consciousness a Brain Process?*, „British Journal of Psychology” 1956, t. 47, nr 1, s. 44-55; J. J. C. S m a r t, *Doznania a procesy mózgowe*, tłum. M. Szczubiałka, w: *Fragmenty filozofii analitycznej. Filozofia umysłu*, tłum. T. Baszniak i in., red. B. Chwedeńczuk, Aletheia, Warszawa 1995), a także teoria identyczności egzemplarzy (ang. token–token identity theory), według której każde pojedyncze zjawisko psychiczne odpowiada jakiemuś zjawisku fizycznemu, ale typom znanych nam zjawisk psychicznych nie odpowiadają żadne wyróżnialne typy zjawisk fizycznych (zob. np. D. D a v i d s o n, *Zdarzenia mentalne*, tłum. T. Baszniak, w: tenże, *Eseje o prawdzie, języku i umyśle*, tłum. J. Gryz i in., PWN, Warszawa 1992, s. 163-193).

<sup>11</sup> Paradygmatem behawioryzmu logicznego w filozofii jest koncepcja Gilberta Ryle’a (zob. G. R y l e *Czym jest umysł?*, tłum. W. Marciszewski, PWN, Warszawa 1970, por. w szczególności s. 84-101).

<sup>12</sup> Mam tu przede wszystkim na myśli tak zwany argument z chińskiego pokoju (omawiany dalej w tekście głównym) oraz rozróżnienie na intencjonalność wewnętrzną, pochodną i „jak gdyby”. W obu przypadkach amerykańskiemu filozofowi chodzi tak naprawdę o nieredukowalność subiektywności, czego chyba on sam nie jest do końca świadom. Uzasadnienia takiego odczytania filozoficznych propozycji Searle’a podaję w artykule *Problem intencjonalności na przykładzie dyskusji między Dennettem i Searlem oraz zarys propozycji jego rozwiązania* (tekst przyjęty do druku w tomie materiałów po konferencji „Świadomość w dyskursie filozoficzno-kulturowym”, 12 XII 2011, Instytut Filozofii Uniwersytetu Śląskiego, Katowice). Argument z chińskiego pokoju prezentowany był przez Searle’a na kilka sposobów (zob. np. J. R. S e a r l e, *Minds, Brains and Programs*, „Behavioral and Brain Sciences” 3(1980) nr 3, s. 417-457, tłum. polskie: t e n ż e, *Umysły, mózgi i programy*, tłum. M. Szczubiałka, w: *Fragmenty filozofii analitycznej. Filozofia umysłu*; por. t e n ż e, *Umysł, mózg i nauka*, s. 28n.); rozróżnienie rodzajów intencjonalności pojawia się na przykład w jego pracy *Umysł na nowo odkryty* (tłum. T. Baszniak, PIW, Warszawa 1999, s. 111-113).

1974<sup>13</sup>. Autor – wbrew współczesnym mu stanowiskom redukcjonistycznym – postawił w nim tezę, że problem umysł–ciało nie jest analogiczny do wcześniejszych (i uważanych za zasadniczo rozwiązane) problemów fizykalizmu, takich jak kwestie własności istot żywych czy chemicznych własności substancji. Tym, co o tym decyduje, a o czym nie wspominają teorie redukcyjne, jest istnienie świadomości w sensie zdolności odczuwania i posiadania stanów mentalnych. „Gdyby nie istniała świadomość – pisał Nagel – problem stosunku umysłu do ciała byłby mniej interesujący. Ze świadomością – wydaje się beznadziejny”<sup>14</sup>. Jest tak, ponieważ w przypadku próby poznawania tego, czym jest świadomość, nie możemy – jak twierdził – zastosować standardowej procedury obiektywizacji poprzez abstrakcję od szczególnego punktu widzenia, z którego obserwujemy zjawiska, gdyż to właśnie nasz szczególny punkt widzenia zjawisk jest tym, co chcemy wyjaśnić<sup>15</sup>.

Nagel nie uważał, że zasadnicza odmienność subiektywności polega na absolutnej prywatności doznań – to uwikłałoby jego stanowisko w problem solipsyzmu. Moje wrażenia są w pewnym sensie dostępne innym przedstawicielom gatunku *homo sapiens* – wszyscy rozumiemy ludzkie formy zjawiskowe, wszyscy mamy bowiem ten sam (bądź bardzo podobny) mechanizm poznania zmysłowego, podobne ciała i jesteśmy zdolni do wyobrazeniowego współodczuwania stanów wewnętrznych innych osób<sup>16</sup>. Dobrze wiemy, jak to jest być człowiekiem, natomiast trudno nam wyobrazić sobie, jak to jest być nietoperzem, ponieważ nie posiadamy na przykład zdolności echolokacji i nie jesteśmy w stanie wyobrazić sobie tego, jak to jest posiadać wrażenia echolokacyjne.

W toku debaty zapoczątkowanej między innymi esejem Nagela świadome doznania, czy też subiektywną stronę bycia w jakimś stanie, zaczęto określać technicznym terminem „qualia”. Słowo to jest sztucznym tworem filozofów i zostało najprawdopodobniej po raz pierwszy użyte przez Clarence’a I. Lewisa w roku 1929<sup>17</sup>. Jako podstawowe cechy qualiów wymienia się: subiektywność, prywat-

<sup>13</sup> Polski przekład: Th. Nagel, *Jak to jest być nietoperzem?*, w: tenże, *Pytania ostateczne*, tłum. A. Romaniuk, Aletheia, Warszawa 1997, s. 203-219. Należy dodać, że sam Nagel przyznaje się do inspiracji zaczerpniętych ze znanej serii wykładów Saula Kripkego wygłoszonych w Princeton University w styczniu 1970 roku, które dwa lata później po raz pierwszy ukazały się drukiem jako *Nazywanie a konieczność* (zob. S. Kripke, *Nazywanie a konieczność*, tłum. B. Chwedeńczuk, Pax, Warszawa 1988).

<sup>14</sup> Nagel, dz. cyt., s. 204.

<sup>15</sup> Por. tamże, s. 213n.

<sup>16</sup> Por. tamże, s. 215 (przypis 11). Zdaniem Nagela źródło problemu solipsyzmu leży w tym, że nie rozumiemy, w jaki sposób subiektywność posiada obiektywną naturę (por. tamże, s. 218).

<sup>17</sup> Lewis pisze: „Istnieją rozpoznawalne jakościowe cechy tego, co dane, które mogą powtarzać się w różnych doświadczeniach, są zatem rodzajem powszechników. Nazywam je «qualia». Jednakże, choć owe qualia są powszechnikami w sensie występowania w różnych doświadczeniach,

ność, wewnętrzność (ang. *intrinsicness*), niepodważalność i niezapośredniczenie<sup>18</sup>. Rzecz jasna, debata zwolenników i przeciwników redukcjonizmu w filozofii umysłu dotyczy przede wszystkim autonomicznego istnienia tego rodzaju bytów.

Można zauważyć, że problem subiektywnych jakości doznań jest nieco inny niż wyjściowy problem Kartezjusza, któremu chodziło przede wszystkim o suwerenność woli i racjonalnego rozumu. Tymczasem, jak się wydaje, można sobie wyobrazić zbudowanie maszyny, która funkcjonalnie całkowicie odpowiada człowiekowi – a więc działa racjonalnie, podejmuje sensowne decyzje w obliczu zastanych sytuacji – lecz jest zupełnie pozbawiona doznań<sup>19</sup>. Ideę tego rodzaju maszyny można zresztą wyczytać z fundamentalnego dla komputacjonizmu tekstu Alana Turinga z roku 1950 *Maszyna licząca a inteligencja*<sup>20</sup>, w którym pojawia się pomysł „gry w udawanie”<sup>21</sup>. Ray Jackendoff, wychodząc z tradycji komputacjonistycznej kognitywistyki, zauważył wobec tego, iż rzeczywisty problem nie leży w relacji ciała do umysłu, lecz w relacji umysłu obliczeniowego (rozumianego jako przypadek maszyny Turinga) do umysłu fenomenalnego. Problem świadomości subiektywnej to zatem – według niego – *mind–mind problem*<sup>22</sup>. To bowiem między umysłem rozumianym jako zespół stanów funkcjonalnych a umysłem doświadczającym wrażeń rozwiera się „luka eksplanacyjna” (termin wprowadzony przez Josepha Levine’a<sup>23</sup>), po-

---

trzeba je odróżnić od własności przedmiotów. [...] Quale jest odczuwane bezpośrednio, jest tym co dane i nie można się co do niego myśleć, ponieważ jest czysto subiektywne.” (C.I. Lewis, *Mind and the World-Order: Outline of a Theory of Knowledge*, Courier Dover Publications, New York 1956, s. 121, tłum. fragm. – J.G.).

<sup>18</sup> Por. np. P. P r e c h t l, *Leksykon pojęć filozofii analitycznej*, tłum. J. Bremer, WAM, Kraków 2009, s. 229; R. v a n G u l i c k, *Functionalism and Qualia*, w: *The Blackwell Companion to Consciousness*, red. M. Velmans, S. Schneider, Blackwell, Oxford 2007, s. 385.

<sup>19</sup> Por. np. N a g e l, dz. cyt., s. 205. Należy jednak zauważyć, że w przypisie autor zastrzega, iż ostatecznie może się okazać, że taka maszyna jest niemożliwa.

<sup>20</sup> Zob. A. T u r i n g, *Maszyna licząca a inteligencja*, tłum. M. Szczubiałka, w: *Fragmety filozofii analitycznej*, s. 271-300.

<sup>21</sup> W „grze w udawanie” chodzi o to, by odgadnąć – wyłącznie dzięki wymianie pisanych maszynowo zdań – czy całkowicie odizolowany fizycznie partner dialogu jest człowiekiem, czy odpowiednio zaprogramowanym komputerem (przypadkiem maszyny Turinga). Procedura ta zyskała później miano testu Turinga: maszyna (ściślej mówiąc działający w niej program) zdaje test, jeśli większość osób, które się z nią komunikowały, bierze ją za człowieka. Warto wspomnieć, że w roku 2011 podczas festiwalu Techniche w indyjskim Institute of Technology Guwahati program „Cleverbot” rozwijany przez Rolla Carpentera od roku 1997 zdał test Turinga, ponieważ 59,3 procent spośród 334 ludzi go testujących uznało program ten za człowieka, przy czym jedynie 63,3 procent tych samych ludzi za człowieka uznało rzeczywistego człowieka (zob. np. J. A r o n, *Software tricks people into thinking it is human*, „New Scientist” z 6 IX 2011).

<sup>22</sup> Por. R. J a c k e n d o f f, *Consciousness and the Computational Mind*, MIT Press, Cambridge, Massachusetts, 1987, s. 20.

<sup>23</sup> Zob. J. L e v i n e, *Materialism and Qualia: The Explanatory Gap*, „Pacific Philosophical Quarterly” 64(1983) nr 3, s. 354-361.

nieważ w rzeczywistości nie wiemy, w jaki sposób jakakolwiek funkcjonalna czy fizyczna struktura może wyjaśnić powstanie subiektywnych doznań.

Zagadnienie kłopotliwej relacji mind–mind zostało podjęte przez Neda Blocka, który ujął je w dychotomii świadomość dostępu–świadomość fenomenalna (ang. access consciousness–phenomenal consciousness)<sup>24</sup>. Przez treść świadomą w sensie świadomości dostępu rozumiał on to, co wpływa na system wykonawczy, a więc może być użyte do kontroli zachowania. Świadomość dostępu może być zatem przedmiotem badań i teorii empirycznych, takich jak teoria globalnej przestrzeni roboczej Bernarda Baarsa<sup>25</sup>. Inaczej ma się sprawa ze świadomością fenomenalną, której na dobrą sprawę nie da się nawet zdefiniować bez popadania w błędne koło albo bez wskazywania na doznania zmysłowe czy świadome myśli, pragnienia i emocje. Własności świadomości fenomenalnej nie są ani poznawcze (nie odznaczają się jakimiś specyficznymi treściami), ani funkcjonalne (nie są definiowalne algorytmicznie), ani intencjonalne (nie chodzi o to, że są o czymś)<sup>26</sup>. Rozróżnienie to zostało jeszcze bardziej wyostrzone przez Davida Chalmersa, który mówił o podwójnym – fenomenalnym i psychologicznym – sensie terminów umysłowych, a w konsekwencji o podwójnym znaczeniu samego słowa świadomość<sup>27</sup>. Block, w przeciwieństwie do Chalmersa, nie jest jednak zwolennikiem podejścia, które wyklucza możliwość obserwowalnych oznak świadomości fenomenalnej – co prawda jej badanie musi rozpocząć się od introspekcji, ale nie może się na niej kończyć<sup>28</sup>.

#### NIEREDUKOWALNOŚĆ ŚWIADOMOŚCI FENOMENALNEJ ARGUMENTY ZA I PRZECIW

W filozofii umysłu można zasadniczo wyróżnić dwie formy sprzeciwu wobec koncepcji redukcjonistycznych. Pierwsza, bardziej radykalna, broni

<sup>24</sup> Zob. N. B l o c k, *On a Confusion About Function of Consciousness*, „Behavioral and Brain Sciences” 18(1995) nr. 2, s. 227-287.

<sup>25</sup> Block podaje teorię Baarsa jako przykład pomieszenia pojęciowego, z którym chce walczyć. Autor teorii globalnej przestrzeni roboczej sugeruje bowiem, że zajmuje się świadomością fenomenalną. Píše bowiem na przykład: „O czym jest teoria świadomości? Przede wszystkim [...] jest to teoria natury doświadczenia. Prywatne doświadczenie tego świata, jakie ma czytelnik, jego lub jej umysłowy obraz wczorajszego śniadania, uczucie bólu zęba – wszystko to są treści świadomości”. (B. B a a r s, *A Cognitive Theory of Consciousness*, Cambridge University Press, Cambridge 1988, s. 14; tłum. fragm. – J.G.).

<sup>26</sup> Por. B l o c k, dz. cyt., s. 234n.

<sup>27</sup> Zob. D. C h a l m e r s, *Świadomy umysł. W poszukiwaniu teorii fundamentalnej*, tłum. M. Miłkowski, PWN, Warszawa 2010, s. 49-73.

<sup>28</sup> Por. B l o c k, dz. cyt., s. 246n.



ontologicznej autonomii świadomości fenomenalnej, druga natomiast, słabsza, zakłada tylko jej autonomię w sensie poznawczym jako niemożliwość domknięcia luki ekspanacyjnej. Co ciekawe, te same argumenty mogą zostać wykorzystane zarówno przez ontologicznych, jak i epistemologicznych antymaterialistów<sup>29</sup>. Antyredukcjonizm ontologiczny może z kolei przyjmować formę mocniejszą, formę dualizmu substancjalnego, oraz słabszą – dualizmu własności. Przykładami pierwszej z nich są kartezjanizm i teoria Ecclesa (przy czym, jak już wspomniałem, trudno współcześnie wskazać na ważniejszych autorów będących jej przedstawicielami). We współczesnej analitycznej filozofii umysłu znacznie popularniejszy jest dualizm własności, głoszący, że istnieje tylko jeden rodzaj substancji, który jednak posiada dwa rodzaje własności: fizykalne i fenomenalne.

Analityczni antyredukcjoniści odwołują się do całej serii eksperymentów myślowych, z których historycznie pierwszy był argument z odwróconych qualiów. Jego korzenie sięgają wieku siedemnastego, ponieważ pojawia się już w *Rozważaniach dotyczących rozumu ludzkiego* Johna Locke'a. Ow klasyk empiryzmu brytyjskiego pisał bowiem, że możliwe jest, iż wrażenie wywołane przez rzecz błękitną (np. fiołek) w umyśle jednego człowieka może odpowiadać wrażeniu, jakie w umyśle innego wywołuje rzecz koloru żółtego (np. nagietek)<sup>30</sup>. Możemy więc wyobrazić sobie, że czyjeś fenomenalne widmo barw jest odwrócone względem naszego, że ktoś doświadcza czerwieni tak jak my błękitu czy zieleni, jak my koloru pomarańczowego, przy czym jego (jej) i nasze obserwowalne reakcje na kolory są komplementarne i w tym sensie nie występuje zasadnicza funkcjonalna różnica w doświadczaniu kolorów między nim (nią) a nami. Innymi słowy, różnica ta zachodzi w świadomości fenomenalnej, a nie zachodzi w świadomości dostępu.

Radykalizacją idei Locke'a jest eksperyment myślowy dotyczący nieobecnych qualiów. Posiada on kilka różnych postaci. Jedną z nich jest argument z wielkiego mózgu Chin sformułowany przez Neda Blocka. Autor każe nam wyobrazić sobie, że cały naród chiński połączony jest siecią łączności radiowej i przez pewien czas wspólnie symuluje przekazywanie informacji między neuronami w mózgu w ten sposób, że pojedynczy Chińczyk odgrywa rolę funkcjonalnie wyróżnionej grupy neuronów. W ten sposób skonstruowany funkcjonalny odpowiednik mózgu mógłby być podłączony do jakiegoś ciała,

---

<sup>29</sup> Por. np. J. L e v i n e, *Anti-materialist Arguments and Influential Replies*, w: *The Blackwell Companion to Consciousness*, s. 376.

<sup>30</sup> Por. J. L o c k e, *Rozważania dotyczące rozumu ludzkiego*, tłum. B. Gawęcki, PWN, Warszawa 1955, t. 2, s. 554n. Locke oczywiście nie używał pojęcia qualiów. Chalmers sugeruje, że intuicja możliwości odwróconych qualiów jest czymś powszechnym i że pomysł ten przychodzi do głowy większości ludzi niezależnie od filozoficznego wykształcenia (por. C h a l m e r s, dz. cyt., s. 435n.).

sterować jego ruchami i otrzymywać od niego informacje. Takiego rodzaju symulacja mogłaby mieć stany świadome tylko w sensie świadomości dostępu (a przynajmniej ich funkcjonalne odpowiedniki), natomiast – zdaniem Blocka – nie mogłaby oczywiście posiadać qualiów.<sup>31</sup>

Za wersję tego argumentu można by uznać również znany argument z chińskiego pokoju Searle'a opublikowany dwa lata później<sup>32</sup>, a następnie kilkakrotnie modyfikowany przez autora. W swej pierwotnej postaci opowiadał on o człowieku, który nie rozumie języka chińskiego, a który zamknięty jest w pokoju wraz z dwoma zestawami chińskich tekstów oraz instrukcjami spisаныmi po angielsku (jest to język dobrze mu znany). Dzięki tym instrukcjom człowiek ów jest w stanie powiązać ze sobą elementy obu chińskich zestawów wyłącznie na podstawie kształtu znaków. Oprócz tego człowiek w pokoju otrzymuje również trzeci zestaw chińskich napisów wraz z regułami w języku angielskim, na podstawie których jest w stanie skorelować ów zestaw z poprzednimi dwoma i wyprodukować czwarty zestaw chińskich napisów, przy czym żadna z angielskojęzycznych informacji nie mówi nic o znaczeniu znaków w żadnym z zestawów. Ludzie, który zaopatrzyli bohatera argumentu w te wszystkie symbole i informacje, nazywają pierwszy zestaw „scenariuszem”, drugi „opowieścią”, trzeci „pytaniami”, a czwarty – efekt pracy człowieka w pokoju – „odpowiedziami na pytania”. Założeniem eksperymentu myślowego Searle'a jest to, że da się stworzyć na tyle dobry zestaw instrukcji, iż władający językiem chińskim ludzie znajdujący się na zewnątrz pokoju – zapoznawszy się z pytaniami i odpowiedziami na pytania – uznałoby, że udzielił ich rodowity Chińczyk<sup>33</sup>. Człowiek z pokoju (a właściwie cały pokój wraz z instrukcjami i dwoma pierwszymi zestawami chińskich napisów) jest więc funkcjonalnym odpowiednikiem umysłu Chińczyka, w którym zachodzą procesy przetwarzania zapisanych po chińsku informacji mające wpływ na zachowanie (produkcję „odpowiedzi na pytania”), co oznacza, że owe chińskie informacje są dla systemu świadome w sensie dostępu. W całym tym procesie nie są one jednak nigdzie uświadamiane fenomenalnie, ponieważ człowiek z pokoju nie jest nigdy świadomy znaczeń napisów, które przetwarza, a byłoby jawnym absurdem sądzić, że system złożony z niego i leżących przed nim kartek papieru posiada świadomość fenomenalną bogatszą od jego własnej.

Nieco podobny do argumentu z nieobecnych qualiów jest argument z pojmomalności. Jego historia sięga Kartezjusza, a współcześnie argument ten pojawił się między innymi w pracy Saula Kripkego *Nazywanie a konieczność*,

<sup>31</sup> Por. N. B l o c k, *Troubles with Functionalism*, „Minnesota Studies in the Philosophy of Science” 1978, t. 9, s. 261-325.

<sup>32</sup> Zob. niniejszy artykuł, przypis 12.

<sup>33</sup> Por. S e a r l e, *Umysł, mózgi i programy*, s. 302-304.

lecz najbardziej znane dziś sformułowanie zawdzięcza Davidowi Chalmersowi. W sformułowaniu tym chodzi o pojmovalność możliwości tak zwanego filozoficznego zombi. Zombi filozoficzny – w przeciwieństwie do zombi hollywoodzkiego – pod względem dowolnego możliwego opisu funkcjonalnego i w ogóle zewnętrznego (zachowaniowego bądź neurofizjologicznego) jest nieodróżnialny od normalnego człowieka, nie posiada jednak świadomości fenomenalnej. Warto zwrócić uwagę na fakt, że chodzi tu o pojęciową możliwość zombi. Chalmers zakłada bowiem, że na mocy nieznanych jeszcze fundamentalnych praw psychofizycznych pozbawiony qualiów filozoficzny zombi jest niemożliwy w naszym świecie. Problem w tym, że owe prawa muszą obowiązywać, ponieważ – jak tego dowodzi możliwość wyobrażenia sobie takiej istoty – świadomość fenomenalna nie daje się pojęciowo utożsamić ani zredukować do systemu funkcjonalnego (świadomości dostępu)<sup>34</sup>. Różnica między redukcjonistą a antyredukcjonistą (w tym przypadku Chalmerssem) nie polega więc na tym, że ten pierwszy zakłada, a drugi odrzuca związek między tym, co fizyczne, a tym, co mentalne, lecz na sposobie pojmovania tego związku. Podczas gdy redukcjonista uznaje, że związek ten jest istotowy i pozwala na redukcję tego, co mentalne, do tego, co fizyczne, antyredukcjonista w typie Chalmersa zakłada związek przyczynowy pomiędzy dwoma ontologicznie różnymi poziomami. Z punktu widzenia redukcjonisty istnienie filozoficznych zombi jest logicznie niemożliwe, ponieważ wyklucza je koniecznościowy związek między fizycznymi a mentalnymi własnościami istot obdarzonych umysłem. Dlatego właśnie pojmovalność zombi jest argumentem przeciw redukcjonizmowi<sup>35</sup>.

Innym ze znanych argumentów wykorzystywanych do zwalczania redukcjonizmu jest tak zwany argument z wiedzy sformułowany przez Franka Jacksona<sup>36</sup>. W swojej najbardziej znanej postaci argument ten opiera się na eksperymencie myślowym, którego bohaterka – wybitna neurofizjolog Maria – od urodzenia zamknięta jest w czarno-białym pokoju z czarno-białym telewizorem, dzięki któremu poznaje świat. Krótko mówiąc, Maria w całym swoim życiu nie ma okazji przeżyć doświadczenia widzenia koloru. Jednakże specjalnością Marii są procesy przetwarzania bodźców wizualnych i dzięki temu posiada ona szczegółową wiedzę o tym, co zachodzi w ludzkim mózgu, gdy jego posiadacz widzi coś czerwonego czy niebieskiego. Jackson zakłada przy tym, że wiedza Marii jest neurofizjologicznie kompletna, to znaczy, że nie istnieje już nic, czego mogłaby się jeszcze dowiedzieć o funkcjonowaniu ukła-

<sup>34</sup> Por. Chalmers, dz. cyt., s. 173-181, 357-360.

<sup>35</sup> Por. Levine, dz. cyt., s. 374

<sup>36</sup> Jackson przyznaje, że formułując ten argument, inspirował się artykułem Nagela *Jak to jest być nietoperzem?*

du nerwowego reagującego na kolorowy bodziec. Pojawia się jednak pytanie: Co stanie się z Marią, gdy zostanie w końcu wypuszczona ze swojego pokoju i po raz pierwszy ujrzy barwy? Czy dowie się czegoś nowego? Dla Jacksona odpowiedź jest oczywista: Tak, dowie się, jak to jest doświadczać kolorów. Oznacza to, że cała możliwa neurofizjologiczna wiedza nie jest wiedzą kompletną – można wiedzieć coś jeszcze, a mianowicie: jak to jest mieć qualia<sup>37</sup>.

Filozofowie o nastawieniu fizykalistycznym, którzy twierdzą, że świadomość fenomenalna nie stanowi poważnego problemu dla programu redukcjonizmu, odrzucają powyższe argumenty. Najczęściej starają się oni wykazać niespójność idei leżącej u podstaw odpowiedniego eksperymentu myślowego, na przykład idei nieobecnych qualiów. Sydney Shoemaker wskazuje, że gdyby przyjąć możliwość, iż funkcjonalny odpowiednik świadomego człowieka nie posiadałby jakościowych stanów wewnętrznych, trzeba by też uznać, że owe jakościowe stany wewnętrzne nie mają żadnego wpływu ani na zachowanie, ani na to, co wiemy o sobie samych z introspekcji, ponieważ introspekcja pozwala nam posiadać przekonania na temat naszych przeżyć, a przekonania należą do opisu funkcjonalnego<sup>38</sup>. Innymi słowy, argumenty antyredukcjonistów w rodzaju wielkiego mózgu Chin, chińskiego pokoju (odczytanego jako argument na rzecz niezależności ontologicznej qualiów) czy filozoficznych zombi zakładają pojmowalność konstrukcji pojęciowej, która podważa pewność naszej wiedzy o własnych stanach świadomych.

Istnieje eksperyment myślowy, który może zostać wykorzystany do zanegowania możliwości nieobecnych qualiów. Paradoksalnie, po raz pierwszy został on wyraźnie sformułowany przez Johna Searle'a w zupełnie innym celu. Wyobraźmy sobie – proponuje Searle – że mamy robociego „kuzyna”, czyli maszynę odpowiadającą nam w sensie funkcjonalnym, lecz pozbawioną świadomości fenomenalnej, której procesy przetwarzania informacji realizowane są na „bezdusznych” układach krzemowych. Załóżmy, że struktura anatomiczno-funkcjonalna naszego mózgu całkowicie odpowiada strukturze tej maszyny. Możemy sobie teraz wyobrazić, że ktoś, dokonując serii operacji neurochirurgicznych, stopniowo podmienia owe anatomiczno-funkcjonalne części naszego mózgu na układy krzemowe robota, aż do momentu, gdy żadna z części centralnego układu nerwowego nie jest już biologiczna. Zgodnie

<sup>37</sup> Zob. F. J a c k s o n, *Epiphenomenal Qualia*, „The Philosophical Quarterly”, 32(1982), nr 127, s. 127-136; tenże, *What Mary Didn't Know?*, „The Journal of Philosophy” 83(1986) nr 5, s. 291-295.

<sup>38</sup> Zob. S. S h o e m a k e r, *Functionalism and Qualia*, „Philosophical Studies” 1975, t. 27, nr 5, s. 291-315. Kontrargument Shoemakera został poddany krytyce przez Blocka (zob. N. B l o c k, *Are Absent Qualia Impossible?*, „The Philosophical Review” 89(1980) nr 2, s. 257-274), na którą ten pierwszy odpowiedział, nieco modyfikując swój pierwotny argument (por. S. S h o e m a k e r, *Identity, Cause and Mind: Philosophical Essays*, Cambridge University Press, Cambridge 1984, s. 309-326). W niniejszym tekście nie ma jednak miejsca na szczegółowe omówienie tej debaty.

z założeniem funkcjonalnej odpowiedniości obydwu układów – naszego mózgu i maszyny – takie podmiany na żadnym etapie nie powinny wpłynąć na nasze obserwowane zachowanie. Jeśli jednak nasz roboci „kuzyn” nie posiada świadomości fenomenalnej, to i my nie powinniśmy jej posiadać na końcu serii operacji. Co jednak dzieje się na etapach pośrednich? Czy nasze qualia powoli blakną? Searle w *Umyśle na nowo odkrytym* stawia tezę, że w takiej sytuacji nasze świadome doświadczenie będzie się kurczyć. Skoro jednak nie będzie się to mogło odzwierciedlić w zachowaniu, będziemy odczuwać stopniową utratę kontroli nad tym, co robimy. Autor snuje mroczną wizję, w której podmiot takiego eksperymentu czułby się coraz bardziej zamknięty i odcięty od swojego zachowania, a przy tym nie byłby w stanie tego w żaden sposób zakomunikować<sup>39</sup>.

Inni autorzy sądzą jednak, że to, co przedstawił Searle, jest bardziej formą *reductio ad absurdum* możliwości nieobecnych qualiów niż skutecznym odparciem funkcjonalizmu. Na przykład Chalmers sądzi, że rozwiązanie proponowane przez Searle’a jest nie do przyjęcia, a treść eksperymentu myślowego sugeruje raczej, że zakładana różnica między strukturą funkcjonalną a świadomością fenomenalną nie zachodzi. Jeśli tożsamość funkcjonalna miałaby się utrzymać, w świadomości nie powinny pojawić się żadne dodatkowe przekonania (jak na przykład przekonanie, że tracę kontrolę nad własnym zachowaniem). Wynika z tego, że dopuszczenie możliwości blaknących qualiów wymaga przyjęcia albo bardzo dziwnego poglądu na naturę relacji między treścią przekonań a stanami mentalnymi, albo możliwości, że racjonalne istoty mogą się fundamentalnie mylić co do własnych świadomych doświadczeń<sup>40</sup>.

Jacksonowski argument z wiedzy sprowokował wiele różnorodnych odpowiedzi ze strony filozofów o nastawieniu redukcjonistycznym. Paul Churchland, Daniel Dennett i Kim Atkins argumentowali, że posiadanie kompletnej wiedzy neurofizjologicznej dotyczącej widzenia jest czymś, co pozwala na – przynajmniej częściowe – wyobrażenie sobie wrażenia barwy<sup>41</sup>. Według Dennetta można nawet wyobrazić sobie zupełnie inne zakończenie historii Marii, wedle którego pierwszym barwnym przedmiotem, jaki ogląda ona zaraz po wyjściu z czarno-białego pokoju, jest pomalowany na niebiesko banan przygotowany dla żartu przez pilnujących ją naukowców. Dennett utrzymuje, że wiedza o wszystkich obiektywnych faktach na temat fizjologicznych mechanizmów związanych z przeżyciem kolorów wystarczyłaby Marii do

<sup>39</sup> Por. S e a r l e, *Umysł na nowo odkryty*, s. 66n.

<sup>40</sup> Zob. D. C h a l m e r s, *Absent Qualia, Fading Qualia, Dancing Qualia*, w: *Conscious Experience*, red. Th. Metzinger, Imprint Academic, Paderborn 1995. Celem Chalmersa również nie była obrona redukcjonizmu, lecz tak zwanej zasady niezmienności organizacyjnej – elementu stanowiska, którego bronił w *Świadomym umyśle* (por. s. 411-414, 427-430).

<sup>41</sup> Por. L e v i n e, *Anti-materialist Arguments and Influential Replies* 2007, s. 377.

przejrzenia dowcipu. Innymi słowy, mogłaby ona wcześniej wiedzieć – przynajmniej w zarysie – jak to jest widzieć kolor żółty czy niebieski<sup>42</sup>.

Odmienne krytykę argumentu z wiedzy przeprowadził David Lewis. Wskazuje on, że założeniem, które doprowadziło Jacksona do stwierdzenia, że Maria z czarno-białego pokoju dowie się czegoś nowego, jest przyjmowana przezeń (i podzielana przez innych antyredukcjonistów, w tym również przez Nagela) hipoteza informacji fenomenalnej. Mówi ona, że samo doświadczenie czegoś, bycie w jakimś stanie subiektywnym, daje nam pewną informację; że istnieje coś takiego jak wiedza, jak to jest być czymś (powiedzmy, jak to jest być nietoperzem) lub jak to jest doświadczać czegoś (powiedzmy, kolorów). Lewis wskazuje jednak, że ewentualna wiedza fenomenalna musiałaby mieć bardzo nieintuicyjne własności, na przykład nie można byłoby się jej w żaden sposób nauczyć, musiałaby być całkowicie izolowana od wszystkich innych rodzajów informacji, a w konsekwencji musiałaby być przynajmniej częściowo epifenomenalna. Z tego też powodu Lewis proponuje, by zastąpić hipotezę informacji fenomenalnej hipotezą zdolności (ang. ability hypothesis). Hipoteza ta wyraża niekontrowersyjną i nieproblematyczną tezę, że gdy zaczynamy doświadczać czegoś nowego – jak Maria, gdy już wydostała się ze swego czarno-białego więzienia – nabywamy pewnych określonych zdolności, takich jak przypominanie sobie czy rozpoznawanie podobnych doświadczeń w przeszłości. Nie jest to wiedza w sensie, w jakim mówi się o wiedzy neurobiologicznej, fizycznej czy matematycznej. Mówienie o jakościach fenomenalnych w kategorii wiedzy jest, zdaniem Lewisa, pewnym zabiegiem językowym mającym zaciemnić różnicę między informacją a doświadczeniem, pewną sztuczką retoryczną mającą uderzyć w fizykalizm<sup>43</sup>.

Prawdopodobnie najbardziej znanym, a zarazem chyba najbardziej radykalnym przeciwnikiem idei qualiów jest dziś Dennett. Uważa on, że pojęcie to używane jest we współczesnej filozofii umysłu w sposób na tyle niejasny, że właściwie nie wiadomo, co oznacza. Na prostym przykładzie doświadczenia ilustrującego zjawisko ślepoty na zmiany wykazuje, że ludzie posługujący się pojęciem qualiów nie są w stanie określić, czym faktycznie są postulowane przez nich własności fenomenalne. Eksperyment ten (nie myślowy, lecz rzeczywisty) polega w dużym uproszczeniu na wyświetlaniu na przemian pary niemal identycznych fotografii różniących się drobnymi szczegółami. Badani zazwyczaj dopiero po pewnym czasie dostrzegają te zmiany. Dennett stawia pytanie, czy człowiek poddany tego rodzaju eksperymentom, zanim uświa-

<sup>42</sup> Zob. D.C. Dennett, *Consciousness Explained*, Back Bay Books–Little, Brown and Company, New York–Boston–London 1991, s. 399n.

<sup>43</sup> Zob. D. Lewis, *What Experience Teaches*, „Proceedings of the Russellian Society” (University of Sydney) 13(1988), s. 29-57.

domi sobie, że widzi dwa różne obrazy, doświadcza różnych czy identycznych qualiów? Inaczej mówiąc, czy może zachodzi nieświadoma różnica w subiektywnych własnościach fenomenalnych? Gdy odpowiadamy „tak”, rezygnujemy z idei prywatności i bezpośredniego dostępu do własnych przeżyć<sup>44</sup>. Gdy odpowiadamy „nie”, problem qualiów staje się banalny, bo okazuje się, że oznaczają one wyłącznie logiczny konstrukt stworzony na podstawie podmiotowych stwierdzeń o jakościach wrażeniowych. W tym sensie podmiot posiada quale *F* wtedy i tylko wtedy, gdy twierdzi, że posiada quale *F*. Pojęcie takie jest jednak niezbyt przydatne i żaden z obrońców qualiów nie chce się na nie zgodzić – oznaczałoby ono między innymi konieczność rezygnacji z pojmovalności możliwości filozoficznych zombi<sup>45</sup>.

W konsekwencji Dennett proponuje odrzucenie niejasnych pojęć w rodzaju „qualiów”, „surowych odczuć” (ang. raw feelings) czy „własności fenomenalnych”, a także związanych z nimi przestarzałych kartezjańskich intuicji. Zamiast nich proponuje program badawczy, który nazywa heterofenomenologią. Heterofenomenologia zakłada, że podstawowymi danymi wyjściowymi mogącymi służyć do badań stanów wewnętrznych są – oprócz danych dotyczących niewerbalnych zachowań i procesów neurofizjologicznych – wypowiedzi na temat tych stanów. Wypowiedzi te są bowiem wyrazem naszych przekonań o własnych subiektywnych doświadczeniach, a to właśnie przekonania dostępne są nam w sposób najbardziej bezpośredni. Fenomenologia błędnie zakłada, że tym, co bezpośrednio, są doświadczenia subiektywne, podczas gdy tak naprawdę nasze poglądy na temat doświadczeń i sposób ich ujmowania są nieuchronnie zapośredniczone w przyjmowanych (często nie w pełni świadomie) teoriach dotyczących natury umysłów i świadomości. Założeniem propozycji Dennetta jest więc możliwość całkowicie obiektywnych badań świadomości, ponieważ wszystkie dane, które bierzemy pod uwagę – zachowania, wypowiedzi i zjawiska neurofizjologiczne – mają charakter intersubiektywny<sup>46</sup>.

## MISTERIANIZM I METAMISTERIANIZM

Sytuacja we współczesnej analitycznej filozofii umysłu skłania pewnych autorów do przyjęcia stanowiska, które wydaje się na pozór całkiem rozsądnym kompromisem między redukcjonizmem a antyredukcjonizmem. Jest ono

<sup>44</sup> Por. D. C. 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, s. 122-126.

<sup>45</sup> Por. tamże, s. 126n.; t e n ż e, *Quining Qualia*, w: *Philosophy of Mind. Classical and Contemporary Readings*, red. D. Chalmers, Oxford University Press, Oxford 2002, s. 233.

<sup>46</sup> Zob. t e n ż e, *Who's on First? Heterophenomenology Explained*, „Journal of Consciousness Studies”, Special Issue „Trusting the Subject?” 2003, nr 9-10, s. 19-30.

nazywane misterianizmem (czy też nowym misterianizmem)<sup>47</sup>, swój początek ma w myśli Nagela, a najbardziej wyrazistą egzemplifikację znajduje w artykule Colina McGinna *Czy możemy rozwiązać problem umysł–ciało?*<sup>48</sup>. Nagel w omawianym wyżej artykule z roku 1974 zakładał, że mogą istnieć pewne fakty, których ludzie nie będą nigdy w stanie poznać ani nawet ująć w pojęcia (jeden z tych faktów jest zresztą przywołany w tytule jego artykułu – fakt dotyczący wewnętrznego stanu bycia nietoperzem)<sup>49</sup>. Nie tylko zatem konstatował aktualne istnienie luki eksplanacyjnej między świadomością fenomenalną a faktami neurofizjologicznymi i behawioralnymi, lecz także dopuszczał możliwość, że owa luka jest nieusuwalna. Z drugiej jednak strony, przynajmniej na etapie *Jak to jest być nietoperzem?*, nie wykluczał możliwości przeciwnej, mianowicie że kiedyś w „intelektualnie odległej przyszłości”<sup>50</sup> luka eksplanacyjna zostanie wypełniona i pojawi się teoria zdolna zredukować fenomen świadomości.

Główna teza McGinna wydaje mi się jednak bardzo dziwaczna. Zakłada ona bowiem, że teorie naukowe istnieją niezależnie i uprzednio względem naszych umysłów, i to nawet w tak mocnym sensie, że mogą być z istoty przez te umysły niepoznawalne. Wraz z Nagelem proponuje on zatem (bardzo) naiwną wersję realizmu pojęciowego<sup>51</sup>, którą można określić nieco żartobliwym mianem naturalizmu an sich.

Tę drugą możliwość jednoznacznie zanegował McGinn. Wedle niego problem świadomości i jej relacji do świata fizycznego należy do zagadnień, na które nasze umysły – ze względu na pewne cechy swojej architektury – są zamknięte poznawczo. Zamknięcie poznawcze (ang. cognitive closure) w ujęciu autora dotyczy nie tyle własności świadomości fenomenalnej, ile teorii wyjaśniającej

<sup>47</sup> Określenie to pochodzi od Owena Flanagana. Według tego autora „starzy misterianie” to dualiści, którzy nie potrafili podać rozwiązania problemu psychofizycznego. Flanagan uważa, że stanowisko prezentowane przez „nowych misterian” – choć uznaje je za reakcyjne i złośliwe – jest jednak z punktu widzenia naturalizmu spójne. Por. O. J. Flanagan, *The Science of the Mind*, MIT Press, Cambridge, Massachusetts, 1991, s. 393.

<sup>48</sup> Zob. C. McGinn, *Czy możemy rozwiązać problem umysł–ciało?*, tłum. M. Iwanicki, S. Judycki, w: *Analityczna metafizyka umysłu*, s. 360-383 (oryginał ukazał się w lipcu 1989 roku). W poczet nowych misterian zaliczany jest również Noam Chomsky ze względu na swoje rozróżnienie problemów (których rozwiązanie znajduje się w zasięgu naszych możliwości poznawczych) i tajemnic (których nie jesteśmy w stanie poznać). Chomsky jednak za tajemnicę uważa przede wszystkim zagadnienie wolności i kreatywności (por. N. Chomsky, *Reflections on Language*, Pantheon Books, New York 1975, rozdz. 4, w szczególności s. 137n.). Chalmers do grona prekursorów misterianizmu oprócz Nagela zalicza także Franka Jacksona, przywołując jego artykuł *Epiphenomenal Qualia* z roku 1982 (por. Chalmers, *Świadomy umysł*, s. 292).

<sup>49</sup> Nagel, dz. cyt., s. 209n.

<sup>50</sup> Por. tamże, s. 204.

<sup>51</sup> Trzeba przyznać, że Nagel jest tego całkowicie świadomy i przyznaje się do poglądów „celowo reakcyjnych” wobec rozpowszechnionych współcześnie stanowisk, które nazywa idealistycznymi. Por. Th. Nagel, *Widok znikąd*, tłum. C. Cieśliński, Aletheia, Warszawa 1997, s. 14.



związek tych własności z rzeczywistością fizyczną<sup>52</sup>. Trzeba tu podkreślić, że zarówno Nagel, jak i McGinn uważają siebie za naturalistów, przy czym ten pierwszy postuluje rozszerzoną, niefizykalną i nieredukcyjną formę naturalizmu<sup>53</sup>, McGinn natomiast utrzymuje fizykalizm i przyznaje, „że mózgi są de facto przyczynową podstawą świadomości”<sup>54</sup>. Jest to jednak fizykalizm dość osobliwy. Głosi on, że co prawda pewne fizyczne własności mózgu stanowią przyczynę świadomego umysłu, że istnieje teoria wyjaśniająca tę relację przyczynową i że nie musi to być nawet bardzo skomplikowana teoria, ale że właśnie owa teoria jest czymś, czego nasz umysł ze swej istoty nie może pojąć<sup>55</sup>.

Nie zgadzam się zasadniczo z tezą McGinna (i za chwilę to uzasadnię), jednakże przyznaję, że argumentacja, jaką rozwija na jej rzecz, jest częściowo przekonująca. Autor wymienia dwa możliwe kierunki badań, dzięki którym moglibyśmy próbować dotrzeć do naturalistycznej teorii świadomości. Pierwszym z nich jest fenomenologia w sensie bezpośredniego badania świadomości (notabene tą właśnie drogą próbował podążać Nagel<sup>56</sup>), drugim jest neuronauka. Żadna z tych dróg nie doprowadzi nas jednak do celu, ponieważ w obu przypadkach mamy zawsze do czynienia tylko z jednym członem relacji świadomość–ciało. Zarówno poznanie fenomenologiczne, dzięki któremu odsłania się nam realność przeżycia subiektywnego, jest zamknięte na intersubiektywną, „trzecioosobową” rzeczywistość, jak i wszelkie możliwe metody obiektywnego badania zjawisk neurofizjologicznych są ostatecznie zamknięte na podmiotowy aspekt przeżyć świadomych. Z tego też względu fizykalna teoria świadomości nie jest i nigdy nie będzie nam dostępna<sup>57</sup>. Argument ten w części dotyczącej neuronauki jest do pewnego stopnia zbieżny ze znanym stanowiskiem Chalmersa, według którego problem świadomości fenomenalnej należy do problemów „trudnych”. Nie mamy bowiem pojęcia, jak mogłoby ewentualnie wyglądać jego naukowe rozwiązanie, ponieważ nauka przyrodnicza – taka, jaką znamy – składa się zasadniczo z wyjaśnień w kategoriach struktur, funkcji i procesów dynamicznych, a subiektywność zdaje się wymykać tego rodzaju wyjaśnieniom<sup>58</sup>.

Jeśli odrzucenie „opowieści transcendentalnej”, a więc zrozumiałości pojęcia „rzeczy samej w sobie” jest idealizmem, to sądzę, że należy przyjąć

<sup>52</sup> Zob. M c G i n n, dz. cyt., s. 362.

<sup>53</sup> Por. N a g e l, *Jak to jest być nietoperzem?*, s. 214; t e n ż e, *Widok znikąd*, s. 12, 23.

<sup>54</sup> M c G i n n, dz. cyt., s. 360.

<sup>55</sup> Flanagan trafnie zauważa, że nie jest jasne, czy McGinn twierdzi, że nie możemy poznać fizycznej teorii świadomości, czy też że nie możemy zrozumieć, jak pewna teoria fizyczna może uczynić świadomość zrozumiałą (por. F l a n a g a n, dz. cyt., s. 314, przypis 15).

<sup>56</sup> Por. N a g e l, *Jak to jest być nietoperzem?*, s. 218n.

<sup>57</sup> Por. M c G i n n, dz. cyt., s. 367-371.

<sup>58</sup> Zob. D. C h a l m e r s, *The hard problem of consciousness*, w: *The Blackwell Companion to Consciousness*, s. 225-235.

idealizm, podobnie jak uczynili to Ludwig Wittgenstein, Peter F. Strawson, Hilary Putnam, czy John McDowell. Z perspektywy tak szeroko pojętego idealizmu misterianizm jest obarczony błędem podobnym do błędu Kanta; autor *Krytyki czystego rozumu* usiłował ujmować relację między sferą fenomenalną a sferą noumenalną w kategoriach, które sam rezerwował dla obszaru możliwego doświadczenia. Przedstawiciel tak szeroko pojętego idealizmu nie może oczywiście zaakceptować propozycji McGinna. Zarazem trzeba przyznać, że jest w misterianizmie coś pociągającego. Jest on bowiem próbą odnalezienia „trzeciej drogi” pomiędzy redukcjonizmem a antyredukcjonizmem.

Dlaczego w filozofii umysłu potrzebujemy dziś „trzeciej drogi”? Sądzę, że wskazuje na to pewna część argumentacji Chalmersa. Gdy mówi on o bezpłodności rozważań fizykalnych dotyczących „trudnego” problemu świadomości, ma również na myśli wszelkie próby zbudowania kwantowej teorii umysłu, czyli odwoływanie się do nowej fizyki, wyjaśnienia bowiem będące ich wynikiem ostatecznie nie mogą wyjść poza opisy struktur i procesów<sup>59</sup>. Uważam, że tę argumentację można uogólnić i odnieść do każdego możliwego rodzaju ontologii świadomości fenomenalnej bez względu na to, czy jest ona materialistyczna, idealistyczna, dualistyczna (substancjalnie bądź własnościowo), panpsychiczna, panprotopsychiczna, anomalna, czy też stanowi trudny do klasyfikacji przypadek w rodzaju niejasnej teorii Searle’a. Podawanie opisów w kategoriach struktur i funkcji nie jest bowiem cechą szczególną wyjaśnienia należącego do nauk przyrodniczych – jest cechą wszelkiego wyjaśnienia, jakie nas satysfakcjonuje. Chalmers ma jednak rację – żadne takie wyjaśnienie nie jest przydatne w przypadku problemu świadomości fenomenalnej. Jest to więc prawdziwie „trudny” problem, a nawet – używając rozróżnienia Chomsky’ego – tajemnica<sup>60</sup>.

To właśnie, jak zresztą wskazuje jego nazwa, ma nam do przekazania „nowy misterianizm” McGinna i do tego miejsca, jak sądzę, jest to stanowisko słuszne. McGinn idzie jednak dalej i buduje pewną wersję „opowieści transcendentalnej”. Według niego za tajemnicą kryje się „rzecz sama w sobie”, rozumiana jako fizyczna teoria świadomości, do której nie mamy dostępu z powodu konstrukcji naszego systemu poznawczego. Proponuję zrobić krok w tył i pozostać przy konstatacji tajemnicy świadomości bez jakichkolwiek prób określania tego, co za nią stoi. Być może nawet nie stoi za nią nic. Ścisłej mówiąc, uważam, że nie mamy i nie będziemy mieć intelektualnych narzędzi pozwalających ustalić, czy w ogóle coś za nią stoi, czy też nie. Innymi słowy,

<sup>59</sup> Por. Chalmers, *Świadomy umysł*, s. 214.

<sup>60</sup> Mówiąc precyzyjniej, rozróżnienie problemów i tajemnic w sensie, w jakim przedstawia je Chomsky, pojawiło się w filozofii znacznie wcześniej w myśli francuskiego filozofa religii Gabriela Marcela.

rozwiązanie, a raczej brak rozwiązania problemu świadomości, lokuję o poziom wyżej, niż robi to McGinn, dlatego też proponuję nazwać je metamisterianizmem.

Pogląd, który właśnie przedstawiłem, może się na pierwszy rzut oka wydawać jeszcze bardziej dziwaczny niż naturalizm an sich postulowany przez McGinna. Faktycznie wynika on jednak z dość nieskomplikowanego zabiegu polegającego na uznaniu krytycznych argumentów wysuwanych przez obie strony sporu o redukcję przy jednoczesnym zastąpieniu absurdalnie silnego realizmu pojęciowego McGinna i Nagela czymś, co bliskie jest putnamowskiemu realizmowi wewnętrznemu<sup>61</sup>. Tezę metamisterianizmu można ująć także w ten prosty sposób – stosując pojęcia Johna McDowella z pracy *Mind and World* – że schemat pojęciowy charakterystyczny dla „przestrzeni racji” (ang. space of reasons) jest niesprowadzalny do schematu pojęciowego „królestwa praw” (ang. realm of laws)<sup>62</sup> i nie powinniśmy się łudzić, że jesteśmy (czy będziemy) w stanie opisać relację między „przestrzenią racji” a „królestwem praw” w kategoriach któregoś z tych schematów.

Moja propozycja zakłada, że problem świadomości fenomenalnej ma charakter podobny do charakteru hipotezy continuum w sensie, w jakim myślał o niej Georg Cantor. Jak wiadomo, ten pierwszy z dwudziestu trzech problemów Hilberta znalazł swoje zaskakujące rozwiązanie w postaci dowodu na swą niezależność od aksjomatyki teorii mnogości Zermelo-Fraenkela. Okazało się więc, że nie można podać odpowiedzi na (pozornie) proste pytanie Cantora, czy istnieje zbiór, którego moc jest większa od mocy zbioru liczb naturalnych, a zarazem mniejsza od mocy zbioru liczb rzeczywistych<sup>63</sup>. Metamisterianizm (w jednej ze swych możliwych eksplikacji) głosi natomiast, że nie można podać odpowiedzi na pytanie o to, czy istnieje „trudny” problem świadomości.

Metamisterianizm jest formą „trzeciej drogi” w filozofii umysłu, która nie jest jednakże obciążona absurdalnie silnym realizmem charakterystycznym dla misterianizmu. Stanowi więc dostępną dla krytyków tego rodzaju realizmu trzecią opcję, sytuującą się pomiędzy fizykalistycznym redukcjonizmem a antyredukcjonizmem w jego różnorodnych odmianach. Z tym pierwszym zgadza się w twierdzeniu, że wszystkie możliwe do jasnego wyrażenia zagadnienia w filozofii umysłu można ostatecznie rozwiązać standardowymi metodami nauk przyrodniczych. Temu drugiemu przyznaje natomiast rację w tym, że

<sup>61</sup> Zob. np. Putnam, *Wiele twarzy realizmu*, s. 342n. Osłabienie realizmu jest założeniem metamisterianizmu, analogicznie jak silny realizm pojęciowy jest założeniem misterianizmu McGinna.

<sup>62</sup> Zob. McDowell, *Lecture IV*, w: tenże, *Mind and World*, Harvard University Press, Cambridge, Massachusetts, 1994, por. w szczególności s. 71 (przypis 2).

<sup>63</sup> Gwoli ścisłości, dowód niezależności hipotezy continuum mówi o tym, że można podać dowolną odpowiedź i potraktować ją jako dodatkowy aksjomat teorii. Cantor sądził jednak, że stawia pytanie o rzeczywistość, a nie o konwencję matematyczną.

rozwiązanie wszelkich możliwych „łatwych” (a więc dających się w ogóle zrozumiale sformułować) problemów naukowych nie usunie tajemnicy subiektywności. Tajemnica ta zawsze pozostanie pewną „resztą”, z którą nie będzie wiadomo, co robić. Owej „reszty” nie ma jednak sensu hipostazować, ponieważ powrócilibyśmy w ten sposób w stare koleiny „opowieści transcendentalnej”. Powinna ona pozostać tym, czym jest obecnie: niespójnością naszego schematu pojęciowego.

\*

Tekst ten rozpocząłem od przywołania postulatu metodologicznego Michała Hellera dotyczącego filozofowania w kontekście nauki. W pierwszych dwóch jego częściach zaprezentowałem różnorodne poglądy i argumenty na temat świadomości fenomenalnej, jakie pojawiły się w analitycznej filozofii umysłu. Większość tych poglądów jest całkowicie zgodna z założeniem metodologicznym, wedle którego filozofia nie może rozwijać się bez ścisłego związku z osiągnięciami naukowymi. Niektórzy z omawianych autorów – na przykład Dennett – oczekują wręcz, że odpowiednie rozwinięcie naszej wiedzy neurofizjologicznej ostatecznie rozwiąże wielowiekowe filozoficzne problemy wynikające z niejasności relacji między umysłem (i świadomością) a światem fizycznym. Innymi słowy, że neuronauka zastąpi filozofię umysłu.

Można zauważyć, że stosunek do zagadnienia redukcjonizmu w kwestii psychofizycznej ma swoje konsekwencje w rozumieniu roli filozofii w poznaniu. Program fizykalnego redukcjonizmu jest bowiem zarazem programem usunięcia filozofii, ewentualnie programem jej redukcji do czegoś w rodzaju nauki pomocniczej. Z kolei propozycje antyredukcyjne można budować wyłącznie na gruncie filozofii – bądź w dziedzinie ontologii wykraczającej poza założenia scjentyzmu, bądź w dziedzinie epistemologii wykraczającej poza programy badawcze psychologii poznawczej.

Koncepcja metamisterianizmu, chociaż, jak się wydaje, we wszystkim ustępuje redukcjonizmowi, to jednak wbrew temu, co mogliby sądzić jej krytycy<sup>64</sup>, pozwala zachować autonomię rozważań filozoficznych o subiektywności i wszystkich związanych z nią zagadnieniach (w tym na przykład problematyce etycznej). Metamisterianizm nie rezygnuje z mówienia o tajemnicy świadomości, nawet jeśli jedyną możliwą w jego obrębie wykładnią owej tajemnicy jest niedająca się usunąć niespójność schematów pojęciowych. Eksplikacja roli tej tajemnicy możliwa jest tylko na gruncie filozofii, przy czym nie może być to filozofia mająca na celu wyjaśnianie tajemnicy w kategoriach któregoś ze

<sup>64</sup> Uwagę tę sprowokowała moja prywatna dyskusja z Adamem Workowskim, który wobec metamisterianizmu przedstawił zarzut, że uśmierca on filozofię.

schematów, lecz objaśnianie jej konsekwencji dla konkretnych sfer ludzkiego życia<sup>65</sup>.

---

<sup>65</sup> Filozofia ta musi więc spełniać postulat metodologiczny późnego Wittgensteina (zob. np. L. Wittgenstein, *Dociekania filozoficzne*, §109, tłum. B. Wolniewicz, PWN, Warszawa 2000, s. 72).