

Marcin Jążyński

Świadomość w młynie

*Czego nie można powiedzieć,
tego tym bardziej nie można zagwizdać.*

Frank Ramsey

Słowa kluczowe: qualia, stany psychologiczne, stany fenomenalne, świadomość

I. Qualia i świadomość

Główny problem zagadnienia świadomości wyrażają pytania o to, czy jest ona czymś fizycznym, czy нефizycznym? Czy ma przyczynowy wpływ na zachowania i jak to możliwe, jeśli nie jest czymś fizycznym? Prowadzi to do pokrewnego problemu możliwości trzecioosobowego – fizykalistycznego lub funkcjonalnego – wyjaśnienia świadomości. Współcześnie pytania te składają się na tak zwany łatwy i trudny problem świadomości oraz problem qualiów.

Świadomość dostępu i świadomość zjawiskowa – Block i Chalmers

Podział ten jest związany z rozróżnieniem na dwa typy świadomości i odpowiednio łatwy problem ma dotyczyć świadomości dostępu, a trudny świadomości zjawiskowej. Świadomość dostępu jest zdefiniowana poprzez przyczynowe role, jakie dany świadomy stan odgrywa w procesach wewnętrznych organizmu i w jego zachowaniu. Natomiast świadomość zjawiskowa ma być określona poprzez sposób, w jaki podmioty odczuwają stany mentalne. Cechy zjawiskowe to przeżywane subiektywne cechy wrażeń, odczuć i spostrzeżeń typu: jak to jest być podmiotem, który spostrzega intensywną barwę? Przy tym własności te mają być istotne dla charakteru stanu mentalnego oraz jego identyfikacji.

Łatwy problem świadomości tworzą zagadnienia związane z funkcjonowaniem świadomych stanów umysłowych o charakterze poznawczym jak spostrzeganie, mechanizmy uwagi, pamięci, systemy percepcyjne, ale też konstruowanie sądów, wypowiedzi, czynności językowych, myślenie oraz podejmowanie decyzji. Natomiast trudny problem dotyczy subiektywnych przeżyć podmiotu i ich jakości. W jaki sposób fizyczne procesy w mózgu mogą wywoływać subiektywne stany świadomości? Czy takie stany mogą pojawić się w innych systemach niż organizmy naturalne? Wreszcie, czy subiektywne własności doznań da się wyjaśnić w terminach perspektywy trzecioosobowej, która jest typowa dla nauki?

Miłośnicy qualiów na to ostatnie pytanie odpowiadają zdecydowanie przecząco. Wedle nich funkcjonalizm nie potrafi uwzględnić jakościowego i subiektywnego charakteru stanów mentalnych, a to właśnie te cechy powodują, że stany mentalne są tym, czym są. Wprawdzie opis funkcjonalny ujmuje wzajemne związki między stanami umysłu, także świadomymi, nie mówi jednak nic o tym, jak podmiot je odczuwa. Żadna teoria funkcjonalistyczna ani tym bardziej materialistyczna nie wyjaśnia subiektywności oraz qualiów, zawsze pozostaje jakaś niewypowiedziana reszta, tym samym nie możemy stworzyć pełnego funkcjonalistycznego opisu stanów, procesów i własności umysłowych.

Z kolei stanowisko przeciwne głosi, że świadomość jakościowa daje się modelować tymi samymi środkami co świadomość dostępu. Choć wcale nie musi być tym samym. A tajemnicze qualia zależą od stanów mózgu i poddają się opisowi, wyjaśnianiu i przewidywaniu trzecioosobowemu.

Można by zapytać: dlaczego ten pogląd budzi tak silny sprzeciw? Myślę, że ma to trzy przyczyny. Pierwsza dotyczy naturalnych dla nas intuicji, które podpowiadają nam obecność subiektywnego elementu naszego świadomego życia psychicznego i jego swoisty tylko dla nas charakter. Po prostu nie chcemy rezygnować z naszego wyróżnionego miejsca w świecie, a obok tego trudno uwierzyć, że coś tak osobistego i czysto psychicznego jak jakości doznań wewnętrznych mogłoby okazać się czymś fizycznym, a przez to utracić swój wyjątkowy status. Jednak w obliczu wielu empirycznych (np. neuropsychiatrycznych) świadectw przemawiających za redukcją qualiów i umieszczeniem ich w obrębie dziedziny neurobiologicznych fenomenów nie musimy się przejmować tymi intuicjami. Inne źródło popularności antyredukjonistycznego poglądu w kwestii qualiów stanowią popularne eksperymenty myślowe, jak eksperyment odwróconego widma czy zanikających qualiów. Problem w tym, że cechą tych eksperymentów myślowych jest skrajna niewiarygodność empiryczna i wadliwa konstrukcja. Trzecia przyczyna leży w określeniu qualiów jako tzw. własności wewnętrznych, co z definicji uniemożliwia ich trzecioosobowy opis, ale oprócz tego generuje szereg poważnych wątpliwości dotyczących istnienia i roli qualiów w życiu psychicznym.

Przyjrzyjmy się konstrukcji wspomnianych eksperymentów myślowych, a potem tajemniczym własnościom wewnętrznym.

Odwrócone widmo i zanikające qualia – Chalmers

Najpierw wyróżnijmy dwie zasady na jakich opiera się Chalmers: tak zwaną zasadę niezmienności organizacyjnej – czyli możliwość wielorakiej realizacji fizycznej stanów umysłowych – wedle której liczy się tylko specyficzna struktura powiązań przyczynowych pomiędzy elementami systemu poznawczego, a nie jego fizyczna budowa. Po drugie zasadę strukturalnej koherencji, która mówi, że struktura świadomych doświadczeń wynika z treści informacji dostępnej świadomości¹.

W przypadku odwróconego widma w dwóch systemach, a w przypadku zanikających qualiów w jednym systemie, ale w różnych momentach, mimo zastą-

¹ Neurokognitywista Włodzisław Duch krytycznie analizując stanowisko Chalmersa zwraca uwagę na wysoką wątpliwość tych eksperymentów. Píše: „Zasada organizacyjnej niezmienności jest podstawą licznych eksperymentów myślowych, stanowiąc podstawę rozumowania wielu filozofów umysłu. Warto jednak pamiętać, że jest to zasada jedynie przybliżona, a jej zastosowanie do wymieniania neuronów mózgu na sztuczne elementy (jak to ma miejsce w »Conscious Mind«) jest wielce wątpliwe. Cóż to bowiem oznacza, że z funkcjonalnego punktu widzenia elementy są identyczne? Elementy krzemowe i elementy białkowe oparte na węglu są w istotny sposób różne i nie ma sposobu, by jedne zamienić drugimi, nie wpływając na pracę całości. Oddziaływania atomów i związków węgla są odmienne niż oddziaływania atomów i związków krzemu, różne są wiązania chemiczne, nie można utworzyć podobnych kanałów jonowych ani zrobić krzemowych struktur reagujących w ten sam sposób co białka na neurotransmitery obecne w szczelinach synaptycznych. Kwantowe właściwości różnych atomów nie pozwalają na ich funkcjonalnie identyczną zamianę nawet »w zasadzie«. Nie można więc eksperymentów myślowych opartych na wymianie kolejnych neuronów na krzemowe płytki traktować poważnie. (Biorąc pod uwagę »silikonowe istoty«, które prócz silikonowej struktury są prawie identyczne z ludźmi, »mózgi w kadzi«, które bez ciała wytwarzają umysł, »zombie« z doświadczeniami qualiów, to nie sposób oprzeć się przekonaniu, że ostatnimi czasy eksperymentom myślowym bliżej do science fiction niż rzeczywistości).

Możliwość zamiany wrażeń związanych z widokiem koloru czerwonego i zielonego przez zmianę połączeń w mózgu jest również czystą fantazją, albowiem konieczna byłaby całkowita reorganizacja połączeń w większej części kory mózgu, związanych z pamięcią epizodyczną zawierającą odniesienia do koloru. Takie zamiany muszą prowadzić do daleko idących zaburzeń w działaniu mózgu, a eksperymenty myślowe, które zakładają, że tak nie jest, są po prostu błędne”; i dalej o odwróconych qualiach: „Moje wątpliwości budzi samo twierdzenie, że wrażenie czerwieni i wrażenie zieleni, niezależnie od wszelkich dyspozycji pamięci, jest odmienne tak, że jest między nimi jakościowa różnica na poziomie mentalnym [...], a to zakłada problem qualiów [...] mogę jedynie powiedzieć, że mam wrażenie wzrokowe”, „nie są to wrażenia odmienne”. Rozwiązanie Ducha ma charakter neurokognitywistyczny, w którym ważną rolę gra poziom fizycznej implementacji: „Przekonanie o naszej zdolności do odmiennego postrzegania »jakości kolorów« wynika zapewne z reakcji emocjonalnych, jakie są rezultatem pobudzeń pamięci, a przez to pośrednich pobudzeń układu limbicznego, wywołującego wrażenia emocjonalne. Jeżeli odrzucę jednak takie wrażenia wtórne, nie pozostaje w mojej świadomości nic, co pozwoliłoby mi powiedzieć, że mam istotnie odmienne wrażenie. Pozostaje tylko odróżnienie wzrokowego tła. To samo dotyczy innych modalności: wszelkie wrażenia mają jakości dzięki możliwości porównania, wynikającej z istnienia pamięci poprzednich wrażeń oraz mechanizmów dyskryminacji stanów o różnych modalnościach na poziomie globalnej dynamiki mózgu”. Takie ujęcie prowadzi do wniosku, że qualia poddają się opisowi fizykalnemu i funkcjonalnemu. Wypowiedź Ducha zaczerpnąłem z M. Kasperski, *Sztuczna inteligencja*, Helion, Gliwice 2003. Patrz też www.phys.uni.torun.pl/duch/

pienia neuronów elementami krzemowymi zachodzi funkcjonalna równoważność. Systemy te (lub jeden system w różnych chwilach) są funkcjonalnie równoważne mimo zmian jakościowych stanów umysłu. Inaczej mówiąc, stany umysłu pełnią nadal tę samą funkcję mimo tego, iż zmieniła się ich charakterystyka jakościowa. Ma to znaczyć, że charakterystyka funkcjonalna „nie chwyta” jakościowego charakteru stanów umysłu, i prowadzić do wniosku, że qualiów nie można wyjaśnić na gruncie funkcjonalizmu oraz fizykalizmu. Jakościowy charakter doznań wymyka się opisowi funkcjonalistycznemu, a tym bardziej fizykalistycznemu. Oba eksperymenty myślowe opierają się na tym, że systemy są funkcjonalnie równoważne mimo odmienności ich fizycznych implementacji, co jest zgodne z funkcjonalistyczną tezą o wielorakiej realizacji oraz silną AI.

Można jednak powątpiewać w trafność tych eksperymentów. Z pewnej perspektywy – która nie wiadomo dlaczego zazwyczaj jest pomijana podczas odpierania ataków miłośników qualiów – oba są chybyne ze względu na to, jak ujmują wieloraką realizację. Bez wątpliwości akceptują funkcjonalną równoważność systemów mimo radykalnej odmienności procesów z poziomu fizycznej implementacji. Akceptuje się bez namysłu to, iż zmiany na poziomie fizycznym nie powodują zmian na poziomie funkcjonalnym. Ale i zmiany w jakościowej charakterystyce doznań następują niezależnie, gdyż nie mają konsekwencji w postaci zmian na poziomie fizycznym i funkcjonalnym. Zauważmy, że to, co się stało w eksperymentalnych sytuacjach, to zmiany na poziomie jakościowym oraz wprowadzone eksperymentalnie zmiany na poziomie fizycznym. Funkcjonalna charakterystyka systemów pozostaje niezmienna. Wniosek: funkcjonalna charakterystyka nie chwyta qualiów. Ale przynajmniej w tych eksperymentach myślowych architektura fizyczna ma znaczenie. Wprowadzenie zmian na poziomie fizycznej implementacji spowoduje różnicę na poziomie funkcjonalnym i jakościowym. Jeżeli tak, nasuwa to wniosek, że eksperymenty te są po prostu źle skonstruowane. Na ich podstawie nie uda się dowieść tego, czego rzekomo dowodzą.

Czy jednak do czegoś prowadzą? Mogą wskazywać na pewne ograniczenie możliwości wielorakiej realizacji. Kształt procesów fizycznych wpływa na przebieg procesów funkcjonalnych. Nasuwa się też wniosek, że wyjaśnienia „tajemnicy” qualiów należy szukać w procesach fizycznej implementacji, to znaczy w procesach mózgowych, które z kolei mogą posiadać swoje opisy funkcjonalne.

Eksperymenty empiryczne. Omylność pierwszoosobowego dostępu i bezpośredni dostęp trzecioosobowy

Cechy fenomenalne qualiów mają być czymś zasadniczo odmiennym od własności relacyjnych i funkcjonalnych. Między innymi ma to oznaczać, że nie poddają się one obiektywnemu badaniu i wyjaśnieniu oraz że jedyny możliwy do nich dostęp ma charakter subiektywny. Wiedza o własnościach wewnętrznych

jest możliwa tylko w perspektywie pierwszoosobowej, bowiem za pomocą opisu trzecioosobowego nie sposób uchwycić owej tajemniczej „reszty”, tj. właśnie wszystkich subiektywnych aspektów świadomego przeżycia. Innymi słowy: tylko ja mogę wiedzieć jak to jest być Marcinem Jazżyńskim, któremu smakuje czosnek, bowiem tylko ja mam pierwszoosobowy dostęp do własności wewnętrznych tego przeżycia. Co więcej, pierwszoosobowa perspektywa ze względu na swój uprzywilejowany charakter ma gwarantować mi wiedzę i nieomylną co do wewnętrznych (zjawiskowych) własności moich przeżyć.

Podsumowując: bezpośredni dostęp do wewnętrznych własności przeżyć podmiotu jest możliwy tylko z pierwszoosobowej perspektywy, to znaczy z perspektywy tego podmiotu. Uprzywilejowany dostęp jest warunkiem wiedzy o przeżyciach tego podmiotu.

Wyobraźmy sobie sytuację pod pewnym względem przeciwną, w której wiedza o przeżyciach podmiotu pochodzi z intersubiektywnych badań wyrażonych w języku trzecioosobowym. Co więcej, obserwacja pierwszoosobowa tych przeżyć czyniona przez podmiot prowadzi do fałszywych wniosków o cechach jego przeżycia. Jeżeli taka sytuacja może rzeczywiście zachodzić, co to oznacza dla problemu bezpośredniego dostępu do jakości? Po pierwsze: pierwszoosobowa perspektywa podmiotu czasem zawodzi i nie gwarantuje bezpośredniego dostępu do własności przeżyć tego podmiotu. Po drugie: własności przeżyć, których poznanie, jak sądziliśmy, było możliwe tylko w perspektywie pierwszoosobowej, okazują się dostępne intersubiektywnemu badaniu i wyjaśnieniu w trzecioosobowym języku nauki. Jeśli tak, to własności te można ująć na sposób relacyjno-funkcjonalistyczny.

Wbrew pozorom nie jest to propozycja, która ma zachęcić do kolejnego eksperymentu myślowego, ponieważ wykonano wiele eksperymentów empirycznych, które wskazują na takie wnioski, na przykład badania zjawiska tak zwanej „ślepoty na zmiany”².

Eksperment Rensika, O'Regana, Clarka i kuchnia Dennetta

W eksperymencie tym pokazuje się badanym osobom wyświetlane na zmianę niemal identyczne dwie fotografie. Czas ekspozycji to 250 milisekund. Pomiędzy nimi na czas 290 milisekund wyświetlana jest maska – pusty ekran. Badana osoba ma za zadanie dostrzec zmianę między fotografiami i zareagować naciśnięciem przycisku. Wynik eksperymentu jest taki, że ludzie biorący w nim udział dostrzegają zmianę bardzo późno, mimo iż długo patrzą na zmieniający się obraz. „Osoby badane często przypatrują się zmieniającym się obrazom przez

² Patrz Dennetta *Cousciousness Explained*, a także *Słodkie sny*, tłum. M. Miłkowski, Prószyński i s-ka, Warszawa 2007, s. 189–199, s. 122, gdzie Dennett powołuje się na badania Grimesa, Resnika, O'Regana i Simonsa.

dwadzieścia czy trzydzieści sekund, a zmiana wyświetla się dziesiątki razy przed ich oczyma, zanim ją zauważają”³.

Z kolei Dennett wykorzystując ten eksperyment pokazuje uczestnikom swojego badania na zmianę dwie bardzo podobne fotografie przedstawiające wnętrze kuchni. Różnica między nimi dotyczy barwy drzwi jednej z szafek. Jest ona dosyć trudno uchwytna, choć dostrzeżona wydaje się wyraźna. Raz drzwi są białe a raz brązowe. Podobnie jak w przypadku eksperymentu Rensika, O'Regana i Clarka mimo wielu powtórzeń ekspozycji tych dwóch odmiennych fotografii badane osoby bardzo późno świadomie zdają sobie sprawę z różnicy, mimo iż widziały ją wiele razy. Pytanie Dennetta brzmi: czy qualia badanych osób zmieniały się, a jeśli tak, to co można o tym powiedzieć na podstawie pierwszoosobowego dostępu? Należy podkreślić, że odpowiadając na to pytanie, każda z badanych osób może swobodnie korzystać z najbardziej wiarygodnego źródła, to jest ze swojego uprzywilejowanego, pierwszoosobowego dostępu.

Rozważmy najpierw dwie przeciwne odpowiedzi. Odpowiedź twierdząca oznacza, że qualia jakie ma podmiot, mogą zmieniać się niezależnie od jego o nich wiedzy. Wówczas podmiot nie będzie świadomy zmian cech własnych przeżyć. Jeśli tak, podaje to w wątpliwość podstawowe założenie nieomyślności pierwszoosobowego dostępu. Co więcej, intersubiektywny, trzecioosobowy dostęp może prowadzić do bardziej wiarygodnych wniosków dotyczących cech przeżyć podmiotu.

Odpowiedź przecząca na pytanie o zmianę qualiów ma przynajmniej jedną, za to niewątpliwą zaletę: opis zmian qualiów jest możliwy jedynie z pierwszoosobowej perspektywy podmiotu przeżywającego doznanie o pewnych jakościach. Rzetelność sądów podmiotu o jego własnych przeżyciach jest zagwarantowana przez uprzywilejowany dostęp do własnych stanów psychicznych i zostaje zabezpieczona przed zakusami i krytyką ze strony zwolenników opisu trzecioosobowego. O zmianach qualiów subiektywnych przeżyć rozstrzyga sam podmiot. Jednak wówczas, będąc konsekwentnym, trzeba by twierdzić, że qualia przeżyć podmiotu tylko wtedy się zmieniają, gdy podmiot ten uznaje, że się zmieniają. Dennett podkreśla, że takie podejście „grozi zbanalizowaniem qualiów – będą one logicznie konstituowane przez nasze sądy lub spostrzeżenia”⁴.

Zastanówmy się nad trzecią możliwą odpowiedzią. Na pytanie o to, czy qualia się zmieniły, zanim spostrzeżono zmianę, badana osoba odpowiada, że nie wie, ponieważ przyznaje, że widocznie, jak pokazują badania, nie ma do nich uprzywilejowanego dostępu, ale jednocześnie nie zgadza się na to, że istnieje trzecioosobowy dostęp do jej qualiów. To dziwna sytuacja, bo czym wtedy mają być qualia? Nie mieszczą się w obrębie intersubiektywnych badań, ale też niczego nie można o nich wiedzieć dzięki subiektywnej pierwszoosobowej perspektywie.

³ Tamże s. 123.

⁴ Tamże s. 126.

Istnieje jeszcze czwarta możliwość. Na pytanie o zmianę qualiów w eksperymentalnej sytuacji odpowiada się twierdząco, zaprzeczając nieomyślności i miarodajności subiektywnego werdyktu wydawanego na podstawie pierwszoosobowego dostępu, jednocześnie uznając możliwość trzecioosobowego opisu i wyjaśnienia tego aspektu przeżyć psychicznych oraz korekty własnego osądu dotyczącego zmian qualiów przed zdaniem sobie sprawy z różnicy w eksponowanych fotografiach.

Zajęcie takiego stanowiska oznacza zdemistyfikowanie qualiów i ich sfunkcjonalizowanie. Zdemaskowanie qualiów – ich naturalizacja – to jedyna cena, jaką musimy zapłacić za to, by mieć miarodajne empiryczne narzędzia do badania tych aspektów procesów umysłowych i móc szukać odpowiedzi na pytania o ich rzeczywistą rolę w naszym życiu psychicznym.

Te cztery stanowiska nie wyczerpują jeszcze wszystkich możliwości odpowiedzi. Pozostała ostatnia, która wcale nie jest nieprawdopodobna. Tak naprawdę nie sposób odpowiedzieć na pytanie Dennetta. Ale nie ze względu na to, że jest ono zbyt trudne. Po prostu dlatego, że nie wiadomo, co to qualia, cechy fenomenalne, a w szczególności nie wiadomo, jak można by „dobrać się” do tajemniczych własności wewnętrznych.

Własności wewnętrzne.

„Czego nie można powiedzieć, tego tym bardziej nie można zagwizdać”

Qualia określa się jako własności, jakie mają jedynie świadome przeżycia psychiczne. Co więcej, zwolennicy nieredukowalności qualiów podkreślają, że własności te przysługują stanom umysłowym ze względu na nie same. Co to może znaczyć? To, że dana własność jest wewnętrzna, znaczy, że jej występowanie jest niezależne od cegogokolwiek, co jest zewnętrzne wobec tego, czemu ta własność przysługuje. Stosując to pojęcie do przypadku qualiów, twierdzi się, że są one wewnętrznymi własnościami świadomych przeżyć psychicznych. Jeśli zaś mają one taki wewnętrzny charakter, znaczy to, że ich istnienie jest niezależne od jakichkolwiek interakcji przyczynowych umysłu (organizmu, podmiotu) z fizycznym światem zewnętrznym. Zamiast tego ich występowanie należy przypisać złożoności samego przeżycia psychicznego – strukturze doświadczenia czy stanu umysłu. Co za tym idzie, żadna nie wiadomo jak zaawansowana, a nawet kompletna teoria z dziedziny nauk poznawczych nastawionych mniej lub bardziej funkcjonalistycznie nie może wnieść poznawczego wkładu do problemu wiedzy o qualiach rozumianych jako własności wewnętrzne, ponieważ z definicji, lub z powodów czysto pojęciowych, qualia jako niemieszczące się w dziedzinie zjawisk uwikłanych w przyczynowe związki ze światem zewnętrznym nie nadają się na przedmiot badań naukowych.

Rodzi to pytania o status epistemologiczny i metafizyczny qualiów. Jak wiemy, ze względu na ich prywatny charakter mają być one bezpośrednio dostępne

jedynie subiektywnemu poznaniu, co automatycznie lokuje je poza dziedziną nauk fizykalnych. Jednocześnie zwolennicy nieredukowalności qualiów przypisują im realne istnienie. Prowadzi to do dwóch pytań. Pierwsze: jaka ma być teoria, w której można by opisywać i wyjaśniać qualia oraz czy byłaby to teoria umysłu? Drugie: czy qualia mają jakiś wpływ na nasze stany umysłowe?

Bardzo trudno odpowiedzieć na pierwsze pytanie, a raczej nie wiadomo, jak to zrobić. Bowiem uznając qualia za własności wewnętrzne, czyli za własności samego przeżycia, i jednocześnie stawiając je w roli przedmiotu badań, potrzebujemy czegoś w rodzaju teorii czystego przeżywania – czystego doświadczenia w znaczeniu przeżycia psychicznego niezależnego od interakcji przeżywającego podmiotu ze światem zewnętrznym. Nasuwa to pewną wątpliwość: czy naprawdę taka teoria mówiłaby coś wartościowego o umyśle? Czy w naszym życiu umysłowym jakkolwiek rolę grają stany umysłowe, których cechy nie są w żaden sposób powiązane ze światem zewnętrznym? Być może takie stany w ogóle się nie pojawiają. Czy na przykład własności takich doznań jak widzenie intensywnej barwy lub słyszenie głośnego dźwięku nie są powiązane interakcjami przyczynowymi ze zdarzeniami w świecie zewnętrznym? Przykłady doznań tego typu podaje się przy okazji definiowania terminu „qualia” i rzeczywiście doświadczenia takie o swoistych własnościach grają bardzo ważną rolę w życiu ludzi i zwierząt. Jednak dość trudno jest odizolować ich występowanie od przyczyn w świecie zewnętrznym.

Uznając qualia za własności wewnętrzne, płacimy olbrzymią cenę, jako że z definicji nie biorą one udziału w przyczynowych interakcjach ze światem fizycznym, pozbawiamy je możliwości przyczynowego oddziaływania w naszym życiu umysłowym. Qualia nie mogą też być epifenomenami, ponieważ te mają swoje przyczyny fizyczne. Problem w tym, że qualiów pojętych jako cechy wewnętrzne nie sposób pogodzić z obrazem umysłu zdolnego do oddziaływania przyczynowego w świecie fizycznym. Jeśli więc qualia istnieją, to nie biorą udziału w fizycznych związkach przyczynowych oraz w życiu umysłowym. Jeśli zaś uczestniczą w takich relacjach, to nie mogą być niedostępnymi dla trzecioosobowego opisu własnościami wewnętrznymi.

Rozwiązanie problemu qualiów na sposób antyredukjonistyczny i antyfunkcjonalistyczny sprawia wrażenie próby równoczesnego osiągnięcia kilku niezgodnych lub wręcz niemożliwych do połączenia celów. Z jednej strony qualiom chce się przypisać realne istnienie i jakiś udział w życiu psychicznym (mają one wszak być własnościami stanów umysłowych), a jednocześnie uznając je za własności wewnętrzne, umieszcza się je poza obrębem dziedziny fizycznej i zasięgiem przyczynowych interakcji ze światem fizycznym, do którego należy także umysł, o ile chcemy mu przypisać możliwość przyczynowego oddziaływania na przykład w postaci przyczyny zachowania.

Podsumowując: qualia istnieją, ale nie fizycznie, nie mają swoich przyczyn w fizycznym świecie, nie są więc też epifenomenami fizycznych procesów móz-

gowych. Są za to własnościami stanów umysłowych i jako takie stanowią element życia psychicznego, ale zarazem nie mogą mieć żadnego wpływu na realistycznie pojęte procesy umysłowe tj. jako rzeczywiste przyczyny fizycznego zachowania.

Z jednej strony zwolennicy qualiów z całą stanowczością podkreślają ich realną obecność i nieredukowalny charakter, a co za tym idzie – niemożliwość ujęcia w jakiejś trzecioosobowej intersubiektywnej teorii naukowej. Z drugiej zaś mówią o palącej potrzebie teorii, która jakoś jednak ujmowałaby qualia. Z tym że powinna to być taka teoria, która nie ma prawie żadnych wspólnych cech z teoriami naukowymi, jakie powstały dotychczas, ponieważ w ich intersubiektywnym języku nie można wyjaśnić subiektywnego charakteru jakości doznań. W pewnym sensie przypomina to gwizdanie, poprzez które próbuje się opowiedzieć o czymś (albo zwrócić na to uwagę), jednocześnie zaprzeczając, że cokolwiek można o tym powiedzieć. Ale „czego nie można powiedzieć, tego tym bardziej nie można zagwizdać”.

Łatwe rozwiązanie trudnego problemu

To podejście nasuwa dwie strategie rozwiązania problemu qualiów i zarazem trudnego problemu świadomości zjawiskowej. Pierwsza dotyczy sprecyzowania pojęcia nieredukowalnych qualiów tak, by wreszcie można je uchwycić i poddać analizie trudny problem. Zadanie to należy do zwolenników qualiów, skoro nie zgadzają się oni na ich kognitywistyczne ujęcie. Jednocześnie nic nie stoi na przeszkodzie, by prowadzić dalsze badania nad świadomością dostępu i łatwym problemem. A kto wie, być może istnieją duże szanse na to, że podczas rozwiązywania łatwego problemu zniknie trudny. To podejście ma jedną zasadniczą zaletę. Na gruncie łatwego problemu dobrze wiemy, czego chcemy i o czym mowa. To w znacznej mierze zapewnia sukcesy eksplanacyjne nauk o poznawaniu. Strategia ta jest naturalna dla kogoś, kto sądzi, że trudny problem to w istocie zbiór łatwych – i rozwiązywalnych metodami kognitywistycznymi – problemów. Być może świadomość nie jest większą tajemnicą niż inne cechy działania systemów poznawczych związane z „łatwym problemem” i można ją rozwikłać tymi samymi metodami i środkami, jakie nauka stosuje, badając te elementy funkcjonalnej architektury umysłu. W związku z tym proponuję jeszcze raz krótko przemysleć problem możliwości naukowego zbadania świadomości.

II. Doznania w młynie

Czym jest świadomość? Pogląd potoczny

Świadomość najczęściej określa się poprzez wskazanie na coś w rodzaju subiektywnego, prywatnego czy wewnętrznego punktu widzenia, który mogą posiadać jedynie podmioty. Określa on wiedzę o części świata zewnętrznego, w tym o położeniu czasoprzestrzennym podmiotu i jego historii, jak również o jego wewnętrznych stanach mentalnych. Obok tego świadomości przypisuje się znaczny, decydujący wpływ na działania. Wedle potocznego poglądu podejmowanie decyzji, charakter zachowania oraz doznań i uczuć zawdzięczamy właśnie subiektywnemu, pierwszo-osobowemu punktowi widzenia. Na marginesie zauważmy, że pogląd taki zakłada olbrzymią, niemal całkowitą kontrolę zachowania, oraz wręcz kompletną wiedzę o jego przyczynach – mają to wszak być świadome decyzje podmiotów.

Następnie, wedle potocznego poglądu, świadomość może być udziałem żywych podmiotów, a nie mogą jej mieć na przykład kamienie, ale też maszyny. O ile więc mózg jest specyficzną wysoce zorganizowaną maszyną biologiczną, to musi tej organizacji towarzyszyć coś więcej, co pozwala na świadomość. Przy czym to „coś więcej” nie jest jakąś emergentną cechą organizacji materii. Istnieją jednak systemy, którym można przypisać coś w rodzaju ich własnego punktu widzenia, na podstawie którego działają lub który determinuje ich zachowanie, a mimo to powszechnie nie jesteśmy skłonni przypisywać im świadomości. Te organizmy to na przykład zwierzęta, nawet rośliny czy roboty. Czego brakuje tym prostszym systemom? Z pewnością niektóre z nich są podobnie jak my systemami poznawczymi, zdolnymi jak i my do rozpoznawania różnych własności środowiska – poznawania fragmentów świata – oraz własnego w nim położenia. Co więcej, niektóre z nich zachowują się tak, jakby planowały swoje zachowania na podstawie wiedzy zgromadzonej wcześniej w doświadczeniu poprzez naukę (uczenie się) oraz przewidywanie, choć nie są w tym tak dobre jak my. Czego zatem im brak, by były świadomymi podmiotami? Kandydatem do miana brakującego elementu świadomej podmiotowości jest zdolność rozpoznawania własnych stanów mentalnych, obserwowania ich, rozróżniania, a nawet – choć to rzadkie – analizowania, co z kolei staje się dodatkową przyczyną naszych działań, jakiej brak innym systemom poznawczym. Krótko mówiąc, jest to zdolność do bieżącej obserwacji i rejestrowania nie tylko zdarzeń, procesów, jakie przebiegają w świecie zewnętrznym z naszym udziałem lub bez niego, ale też własnych percepcji, doznań, myśli, uczuć wspomnień czy przewidywań. Chodzi o to, że potrafimy identyfikować i rozróżniać percepcje przeszłe i teraźniejsze, a ponadto jesteśmy zdolni do rozróżnień dokonywanych na wyższych stopniach, na przykład możemy zidentyfikować jako dwa odmienne stany psychiczne (o różnych własnościach i charakterystyce) postrzeganie czegoś oraz wiedzę o tym, że coś postrzegamy, i dalej w górę możemy zdawać sobie sprawę z tego, że wiemy, iż coś postrzegamy.

Nasuwiają się teraz dwa blisko związane pytania: czy ta dyspozycja stanowi o jakościowej różnicy między nami – świadomymi podmiotami – a innymi systemami poznawczymi, czy też jest to różnica np. stopnia złożoności organizacji funkcjonalnej systemu poznawczego, oraz czy poddaje się ona badaniu i wyjaśnieniu naukowemu?⁵

I jeszcze jedno – powracające qualia. Można z góry powiedzieć, że co prawda możliwe jest naukowe wyjaśnienie jakichś własności świadomości, ale nie wyjaśni się najistotniejszej ich cechy tj. qualiów – własności wewnętrznych. Co można odpowiedzieć na takie postawienie sprawy? Po pierwsze, można zapytać o to, dlaczego qualia uznaje się za najistotniejszą część świadomości. Pozostaje to dla mnie zagadką. Po drugie, jak mówiłem, jeśli uzna się na wstępie qualia za własności wewnętrzne, wtedy z pewnością nie uda się ich wyjaśnić w żaden ze znanych nam sposobów. Ale także nie będzie wiadomo, jak qualia jako własności wewnętrzne mogą istnieć w przyczynowym świecie. Byłoby to dość wątpliwe zwycięstwo zwolenników nieredukcyjnego charakteru qualiów.

Wielu filozofów sądzi, że świadomość jest tajemniczym i wyjątkowym zjawiskiem umysłowym. Nawet jeżeli tak nie jest, świadomość bez wątpienia gra istotną rolę w życiu psychicznym, zatem teoria umysłu powinna ją uwzględniać i wyjaśniać. Z oczywistych względów dotyczy to także funkcjonalno-komputacyjnej strategii wyjaśniania umysłu. Jednak sama możliwość funkcjonalnego wyjaśnienia świadomości wzbudza, delikatnie mówiąc, duże kontrowersję i często jest uznawana za przedsięwzięcie z góry skazane na porażkę. Przeciw temu programowi wysuwa się zazwyczaj argumenty wskazujące na uprzywilejowanie pierwszoosobowego dostępu, a tezę o niedostępności fenomenu świadomości dla badań naukowych wspiera się eksperymentami myślowymi opartymi na znanym „młynie Leibniza”. Wydaje się jednak, że zasadność tej taktyki można podważyć. Opiera się na dwóch znanych nam filarach:

- (1) Dostęp do świadomości jest możliwy tylko z pierwszoosobowego punktu widzenia.
- (2) Znaczy to między innymi, że świadomość jest czymś нефizycznym, a jako taka nie poddaje się naukowemu badaniu, modelowaniu i wyjaśnianiu. Istotne cechy świadomości wymykają się fizykalnemu wyjaśnieniu.

⁵ Jest to dość znany problem zombie. Wyraża go hipotetyczne pytanie: czy robot lub zombie o zbliżonej do naszej pod względem złożoności funkcjonalnej organizacji byłby jak my świadomym podmiotem? Zamiast brać udział w niekończącej się wymianie zdań w sporze „Byłby czy nie byłby?”, wolę skupić się na pytaniach sformułowanych wyżej. W tym celu proponuję strategię, która wydaje mi się bardziej obiecująca. Polega ona na tym, by postawić pytanie o naukowe wyjaśnienie własności stanów świadomych, np. przy użyciu narzędzi neurobiologiczno-obliczeniowych. Naukowe badania nad świadomością pokażą przynajmniej, czy zjawisko to w ogóle poddaje się wyjaśnieniu funkcjonalnemu czy też nauka nie jest w stanie uchwycić jakichś istotnych jego własności. Wówczas należy szukać wyjaśnienia świadomości gdzie indziej.

Wątpliwości wzbudzają obie tezy (1) i (2). Fakt istnienia pierwszoosobowego dostępu do świadomości jest oczywisty. Nie znaczy to jednak, że, po pierwsze, nie istnieje trzecioosobowy dostęp do tego fenomenu. Po drugie, istnienie pierwszoosobowej perspektywy nie przesądza o niefizyczności świadomości. Eksperymenty myślowe wbrew pozorom nie pociągają takich wniosków.

Funkcja pierwszoosobowego dostępu

Jak wiadomo, tak zwana pierwszoosobowa perspektywa umożliwia swoisty tylko dla danego podmiotu dostęp do jego własnych doznań zmysłowych. Każdy organizm wyposażony w taki pierwszoosobowy dostęp poznaje własne stany umysłu w sobie tylko właściwy sposób. Jest to zatem – by tak rzec – „prywatny kanał informacyjny” zapewniający podmiotowi wiedzę o nim samym, o jego stanach, bieżących doznaniach oraz o ich własnościach – jakościach.

Istnienie pierwszoosobowego dostępu ma być podstawą dla wnioskowania o niefizycznych własnościach stanów umysłowych. Czy raczej: przez to, że istnieje pierwszoosobowy dostęp, mamy podstawy, by domniemywać istnienie niefizycznych cech stanów umysłowych, jakich nie można poznać z perspektywy trzecioosobowej, a jedynie z pierwszoosobowej. Jeżeli jednak na pierwszoosobowy dostęp spojrzeć jako na źródło informacji o stanach i własnościach stanów podmiotu, wówczas pogląd taki przedstawia się cokolwiek dziwacznie. Bo czym w gruncie rzeczy jest pierwszoosobowy dostęp? I jaką pełni funkcję? Wiadomo, że gwarantuje on coś w rodzaju wglądu w prywatne doznania podmiotu. Zgoda. Kluczowe wydaje się tu pytanie: jak to jest możliwe? „Jak on to robi?” Inaczej mówiąc: jak funkcjonuje ten swoisty, prywatny kanał informacyjny?

Transfer informacji o prywatnych doznaniach odbywa się przez funkcjonowanie wysoce złożonego systemu połączeń dróg nerwowych. Ich funkcja polega na przekazywaniu informacji o zmieniających się na bieżąco, różnorodnych prywatnych doznaniach zmysłowych organizmu. Drogi te stanowią połączenia między podsystemami sensorycznej aktywności a innymi częściami mózgu i stanowią źródło wiedzy o doznaniach. Funkcjonalnie, w olbrzymim uproszczeniu, można to przedstawić następująco: część podsystemów systemu poznawczego jest odpowiedzialna za aktywność sensoryczną, czyli za zbieranie informacji sensorycznych z modalności zmysłowych organizmu. Z kolei inne podsystemy pełnią funkcję rejestrowania tych informacji – jakby katalogowania danych o zmieniających się doznaniach zmysłowych. Między tymi dwoma typami podsystemów funkcjonalnych istnieje sieć połączeń – kanał informacyjny – przekazujący na bieżąco sygnały od pierwszego podsystemu do drugiego. Wiele wskazuje na to, że w mózgu przebiegają procesy tego typu. Takie organizmy jak ludzie, ale i zwierzęta, dysponują systemami pełniącymi scharakteryzowane przed chwilą funkcje. Dalsze badania powinny określić funkcjonowanie tych procesów, a jeśli działanie „kana-

tu informacyjnego” rzeczywiście stanowi funkcjonalny mechanizm pierwszoosobowego dostępu, zbliży to nas do wyjaśnienia tego zjawiska. W każdym razie nie znajduje się ono poza zasięgiem nauk poznawczych.

Z tej perspektywy unikatowość czy prywatność pierwszoosobowego dostępu polega na tym, że każdy organizm posiada swój neuronowy kanał informacyjny o szczególnej strukturze neuronalnej swoistej tylko dla niego, niemniej funkcjonującej na takich samych zasadach jak neuronalne kanały informacyjne organizmów o podobnej architekturze funkcjonalnej.

Nasuwa to następujący roboczy wniosek: Istnienie pierwszoosobowej perspektywy nie oznacza, że do stanów umysłowych organizmów nie ma oprócz niej także trzecioosobowego dostępu. Znaczy natomiast, że z pierwszoosobowego dostępu korzysta tylko właściciel kanału informacyjnego. Obok tego zauważamy, że samo funkcjonowanie takiego kanału daje się modelować i wyjaśniać naukowo. Czy zatem funkcjonowanie tak pojętego mechanizmu pierwszoosobowej perspektywy stanowi podstawę dla twierdzenia o zasadniczej słabości eksplanacyjnej nauk poznawczych w kwestii świadomości? Być może, ale to otwarta kwestia i na tym etapie nie należy jej przesądzać. A przynajmniej nie należy tego robić przed podjęciem próby kognitywistycznego wyjaśnienia działania procesów świadomości.

Chciałbym poczynić jeszcze pewną uwagę. Jeżeli rzeczywiście funkcjonowanie pierwszoosobowego dostępu poddaje się badaniom nauk poznawczych, daje to nadzieję na podstawy czegoś w rodzaju obiektywnego ujęcia subiektywności, co stanowi jeden z postulatów zwolenników „nowej nauki o świadomości”. Być może wiele uda się o tym zjawisku powiedzieć na gruncie „starej nauki”.

W młynie Leibniza

Ten eksperyment myślowy ma pokazywać niedostępność poznawczą świadomości dla funkcjonalno-naturalistycznego podejścia. Jest on bez wątpienia bardzo sugestywny i na pierwszy rzut oka wiarygodny. Jednak po nieco bliższym oglądzie kontrowersyjne staje się to, czy spełnia on swoje zadanie.

Krótko scharakteryzuję słowa Leibniza. W 17 paragrafie swojej *Monadologii* mówi on:

Należy wszakże przyznać, że postrzeżenie i to, co do niego należy, nie da się wytłumaczyć racjami mechanicznymi. To znaczy przez kształty i ruchy.

Twierdzenie to znaczy mniej więcej tyle, że nie można wyjaśnić takich zjawisk umysłowych jak postrzeżenia czy doznania przy użyciu narzędzi i pojęć naturalistycznych w znaczeniu np. aparatu nauk poznawczych. Zauważmy przy tym, że zjawiska, o jakich mowa, można zinterpretować jako takie, którym przysługuje aspekt świadomości. W tym miejscu w zasadzie wynik eksperymentu jest przesądzony.

Przypuściwszy zaś, że istnieje maszyna, której budowa pozwala, aby myślała, czuła, miewała postrzeżenia, będzie można pomyśleć ją, z zachowaniem tych samych proporcji, tak powiększoną, aby można do niej wejść jak do młyna.

Od tego należałoby zacząć cały eksperyment myślowy, bez zakładania wniosków z góry. Czym jest ta „maszyna” czy „młyn”? O czym mówi Leibniz? Ważne jest tu to, że rzeczona maszyna posiada elementarne cechy *świadomości*. Czy jest to coś w rodzaju sztucznego systemu poznawczego, który myśli, czuje, postrzega itp.? Stanowisko Leibniza w tej kwestii jest przecież w dużym stopniu zbieżne z krytyką idei silnej wersji sztucznej inteligencji.

A może to mózg? Dlaczego nie mielibyśmy zinterpretować Leibnizjańskiego młyna jako wnętrza mózgu? Mam nadzieję, że pod pewnym warunkiem taka zamiana może być uprawniona. Jest nim podtrzymywanie tezy o mechaniczności umysłu, a mówiąc bardziej współcześnie, o jego funkcjonalno-komputacyjnej strukturze działania. Zamieniamy więc młyn Leibniza na mózg, a towarzyszyć temu będzie teza, że jest to maszyna biologiczna. Samego zaś filozofa, który wchodzi do młyna i przygląda się działaniu jego mechanizmów, możemy zinterpretować jako zespół badaczy prowadzących badania nad poznawczymi funkcjami mózgu, którzy posługują się metodami nauk o poznawaniu ze szczególnym uwzględnieniem neurobiologii i psychologii poznawczej. Mówiąc w terminologii Leibniza, szukają oni czysto mechanicznego wyjaśnienia. Idąc za nim, powinniśmy dojść do wniosku, że nawet jeśli mózg jest maszyną i jeśli w jakiś sposób wytwarza świadomość, nie uda się nam wyjaśnić tych procesów w fizykalno-funkcjonalny sposób.

Jak dotąd chyba wszystko w porządku. Z jednej strony, jeśli nie jesteśmy dualistami, zapewne zgadzamy się co do tego, że procesy mózgowe są w jakiś sposób odpowiedzialne za „myślenie, czucie, posiadanie postrzeżeń”, a z drugiej, poprzez badanie mózgu dowiadujemy się o jego budowie i procesach neurobiologicznych, jakie w nim przebiegają. Tak jak Leibniz, który w swoim młynie obserwuje funkcjonowanie jego części. Co więcej, mówiąc nieco na wyrost, potrafimy wyjaśniać te procesy w kategoriach mechanistycznych; fizykalno-funkcjonalistycznych.

Założywszy to, odnaleźlibyśmy wewnątrz przy zwiedzaniu [...] tylko części, które popychają się wzajemnie, nigdy jednak nic, co tłumaczyłoby postrzeżenie.

W tym momencie chciałoby się wykrzyknąć: Co takiego?! Dlaczego mamy się na to zgodzić? To wcale nie jest oczywiste i wniosek taki wydaje się nieuprawniony. Bo kto wie, może byśmy znaleźli. Trzeba jednak wiedzieć, czego szukać, i dysponować odpowiednimi narzędziami. Poza tym, jeśli tego nie znajdziemy w mózgowym młynie, jak wyjaśnimy rolę procesów neurobiologicznych w świadomości? Cudem? Równie dobrze możemy postawić na taką strategię: jeżeli w młynie nie znajdujemy nic innego jak tylko jego części skonstruowane

i zorganizowane w pewien sposób, oraz widzimy działanie tego systemu i zakładamy przynajmniej możliwość, że maszyna ta w jakiś sposób „wytwarza” czy jest odpowiedzialna za powstawanie pewnego zjawiska, to rozsądne wydaje się rozważenie tego, że za powstanie tego zjawiska odpowiada budowa, organizacja i działanie tej maszyny. Być może coś przeoczyliśmy, możliwe, że czegoś nie widać na pierwszy rzut oka lub też brak nam odpowiednich instrumentów, pojęć. Zatem dobrze byłoby przyrzeć się jej bliżej i zbadać tę możliwość, zamiast ją z góry przekreślać i twierdzić, że: „[...] trzeba szukać tego właśnie w substancji prostej, a nie w rzeczy złożonej, czy też w maszynie”.

Z dwojga złego większe nadzieje na spójną teorię umysłu rokuje ponowne zbadanie maszyny młyna niż próba wyjaśnienia przyczynowego związku niefizycznej „substancji prostej” z zachowaniem, o ile chce się uniknąć epifenomenalizmu świadomości.

Przytoczone rozumowanie Leibniza można interpretować dwojako. Po pierwsze jako twierdzenie epistemologiczne. Przy tej, słabszej, interpretacji mówi ono tyle, że dysponując jedynie naturalistycznymi narzędziami nauki mechanicznej – dziś powiedzielibyśmy o naukach poznawczych – nie możemy i nie będziemy mogli wyjaśnić fenomenu świadomości i, szerzej, stanów i własności mentalnych. Przy drugiej, silniejszej, interpretacji przybiera ono postać mocnego twierdzenia metafizycznego, które można zinterpretować tak: świadomość jest czymś niefizycznym ani nie jest wytwarzana przez funkcjonowanie złożonych procesów fizycznych, na przykład neurobiologicznych. Inaczej mówiąc, świadomość nie jest skutkiem, emergentną własnością działania funkcjonalnie zorganizowanych systemów i podsystemów fizycznych – maszyny młyna. Należy ją raczej określić mianem czegoś w rodzaju niefizycznej „substancji prostej”, której istnienia i własności nie da się odkryć ani wyjaśnić, badając działanie złożonej organizacji funkcjonalnej maszyny – mózgowego młyna.

Tym dwóm interpretacjom odpowiadają dwa wspomniane wcześniej filary współczesnej wersji Leibnizjańskiego stanowiska. Jak powiedziałem, ten eksperyment myślowy wbrew pozorom nie spełnia swojego zadania, tj. nie pokazuje, że, po pierwsze, nie ma trzecioosobowego dostępu do świadomości, i po drugie, że nie jest ona czymś fizycznym, a przynajmniej czymś, czego nauki fizyczne nie mogą uczynić przedmiotem wyjaśnienia. Podobnie rzecz się ma z innymi eksperymentami tego typu jak „chiński pokój” Searle’a. Co bynajmniej nie znaczy, że teoria umysłu i świadomości, jakiej mają służyć, jest fałszywa. Jeśli jednak świadomość jest czymś w rodzaju „substancji prostej” – jak wyjaśnić jej przyczynowe własności, np. możliwość przyczynowego wpływania na procesy decyzyjne i wybory strategii zachowania czynione przez podmioty oraz na samo zachowanie organizmów w fizycznym świecie? Jak wspomniałem, alternatywę stanowi epifenomenalny charakter świadomości, ale wówczas odmawiamy jej rzeczywistego wpływu na działanie. Istnieją ważne powody, by nie iść tą drogą.

Dlaczego Leibniz nie znajdzie w mózgowym młynie świadomości? Krótko i nieco brutalnie mówiąc – nie znajdzie jej z dwóch powodów: nie wie, czego szukać, ponieważ brakuje mu odpowiednich pojęć i instrumentów, oraz dlatego, iż z góry zakłada, że w młynie nie znajdzie tego, czego szuka. Bardziej uczciwe jest inne podejście – próba wyróżnienia podstawowych, elementarnych cech świadomości i zbadania, czy jest możliwe wyjaśnianie ich metodami naukowymi przy użyciu aparatu nauk poznawczych. Możemy przyjąć tę strategię z kilku powodów. Jednym z nich jest to, że pod pewnym ważnym względem znajdujemy się w lepszym położeniu niż Leibniz i mamy nad nim tę przewagę, że dysponujemy w miarę dobrymi, rozwijającymi się pojęciami, metodami i narzędziami, które, krótko mówiąc, pozwalają na lepszą, wnikliwszą obserwację funkcjonowania młyna.

Drugi powód, dla którego nie jesteśmy zmuszeni do przyjęcia poglądu Leibniza, polega na tym, że w przedstawionym rozumowaniu, jak i w innych jego współczesnych postaciach, wniosek wydaje się nie wynikać z przesłanek. Jeśli tak jest, kwestia możliwości naukowego wyjaśnienia świadomości pozostaje otwarta i nic nie stoi na przeszkodzie, by postawić pytanie o to, jakie są elementarne, istotne cechy świadomości, by następnie podjąć próbę ich wyjaśnienia na gruncie nauk poznawczych. Na tym etapie najważniejszym zadaniem jest więc wyłonienie podstawowych własności stanów świadomych. Szukamy czegoś w rodzaju minimalnej świadomości. Inaczej mówiąc, chodzi o cechy, zdolności czy narzędzia, jakie muszą posiadać systemy, aby można im było przypisywać świadome działania czy znamiona podmiotowości. Dobrymi kandydatami do tego miana są na przykład percepcje wyższego rzędu, czyli zdolność konceptualizacji własnych stanów umysłu. Wiedza „o tym, że coś wiemy” wymaga z kolei czegoś w rodzaju kodu, językopodobnego medium, dzięki któremu możemy manipulować i przekształcać informacje. Świadomość odgrywa także jakąś rolę w lokalizacji czasoprzestrzennej podmiotu wraz z jego doznaniem (teraźniejszymi i wcześniejszymi). Tu zatem wchodzi w grę mechanizmy reprezentacji i pamięci działające na „materiale”, jakim są informacje ze świata zewnętrznego oraz informacje o wewnętrznych stanach podmiotu. Inną, ale zapewne fundamentalną funkcją konstytuującą nasze poczucie podmiotowości i tożsamości jest integracja informacji sensorycznych. Mam na myśli zdolność wykorzystywania informacji pochodzących z kilku różnych modalności zmysłowych i kształtowania z nich jednolitego wrażenia. Jednak najprawdopodobniej nasze doznania i „obrazy” rzeczy, osób itd. nie są kształtowane jedynie na podstawie informacji sensorycznych. Olbrzymią rolę grają tu nasze nastawienia, stany emocjonalne, oczekiwania czy przekonania. Zatem wyjaśnienie tej funkcji świadomości powinno obejmować zarówno neurobiologiczne mechanizmy integrowania informacji sensorycznych, jak również uwzględniać rolę nastawień itp. w kształtowaniu przedmiotu świadomości, jakim jest doznanie. Obok tego w jakiejś mierze jesteśmy zdolni do sterowania świadomością i skupiania na danym przedmiocie. Wśród

mechanizmów, jakie za tym stoją, kluczowa jest sterowalność uwagą. Zjawisko to jest zapewne pokrewne tej cesze naszych czynności umysłowych, jaką jest intencjonalność. Być może wytłumaczają ją po części te same mechanizmy.

Uderzające jest występowanie kilku typów mechanizmów we wszystkich tych zjawiskach świadomości – tj. mechanizmów reprezentacji i rejestracji reprezentacji, operacji na reprezentacjach, języka, pamięci czy uwagi. Co więcej, jak dotąd wszystkie wspomniane funkcje czy własności procesów świadomych wydają poddawać się ujęciu procesów umysłowych jako procesów przetwarzania informacji i komputacji. Kolejnym krokiem powinno być podjęcie realizacji głównego celu, tj. odkrycie neurobiologicznych i obliczeniowych mechanizmów odpowiedzialnych za funkcjonowanie procesów świadomych posiadających te cechy oraz skonstruowanie modelowych odpowiedników procesów świadomych.

Z tego punktu widzenia zachęcająca wydaje się perspektywa włączenia świadomości do arsenału środków czy narzędzi, jakimi dysponują złożone systemy poznawcze. Funkcja świadomości mogłaby polegać na wzmacnianiu zdolności poznawczych organizmów skierowanych na zdobywanie informacji – wiedzy – o świecie (środowisku, w jakim się znajdują) w celach sprawniejszego, bardziej efektywnego jej wykorzystywania. Podobnie może być z samoświadomością. Posiadanie miarodajnych informacji i wiedzy organizmu o jego własnych stanach jest równie ważne dla podejmowania decyzji i skutecznego działania, jak dysponowanie wiedzą o świecie zewnętrznym. Może to nasuwać pogląd, zgodnie z którym to wykorzystywanie podporządkowane jest celom adaptacyjnym. Być może. Ale to tylko jedna z opcji.

Oczywiście, za takim programem badawczym kryje się pewna hipoteza dotycząca tego, czym jest świadomość i jaką pełni funkcję. Ogólnie rzecz biorąc, mówi ona, że powstawanie i charakter stanów i procesów umysłowych powinny zostać wyjaśnione przez odkrycie sposobów działania mechanizmów całych podsystemów funkcjonalnych mózgu i układu nerwowego. Zgodnie z tym świadomość pojawia się jako skutek działania złożonej funkcjonalnej organizacji systemów fizycznych, np. takich jak mózg i układ nerwowy. Nawiązując do eksperymentu myślowego Leibniza: liczę na to, że odpowiedź kryje się w złożonej organizacji funkcjonalnej mózgu. Jej wystąpienia należy jednak upatrywać w funkcjonowaniu całego systemu maszynierii mózgowego młyna. Wymaga to zbadania funkcji poszczególnych podsystemów tej złożonej organizacji, ale przyglądanie się każdemu z nich z osobna i poszukiwanie w niej części świadomości jest błędem, ponieważ powstaje ona jako wynik działania całej maszynierii. Świadomość wydaje się jedną ze złożonych funkcji umysłu, na którą składa się działanie wielu podsystemów pełniących poznawcze zadania. Myślę, że argumenty z pierwszoosobowego dostępu i pokrewne im eksperymenty myślowe wbrew pozorom nie przekreślają funkcjonalistycznego programu wyjaśnienia świadomości. Kwestia ta jest otwarta. Zapewne to za mało, by w pełni wyjaśnić to zjawisko. Takie podejście ma przynajmniej jedną niewątpliwą zaletę: wiemy, czego szukamy.

Bibliografia

- Churchland P., *Mechanizm rozumu. Siedlisko duszy. Filozoficzna podróż w głąb mózgu*, tłum. Z. Karaś, Fundacja Aletheia, Warszawa 2007.
- Dennett, D.C. *Słodkie sny. Filozoficzne przeszkody na drodze do nauki o świadomości*, tłum. M. Milkowski, Prószyński i s-ka, Warszawa 2007.
- Leibniz G.W., *Monadologia*, tłum. S. Cichowicz, Wydawnictwo Comer, Toruń 1995.

Qualia and Consciousness

Key words: *qualia, consciousness, psychological states, phenomenal states, mind*

Having in mind the distinction between the easy and the hard problem of consciousness, or which is largely the same issue, the distinction between psychological states and phenomenal states in consciousness, the author argues that the essential difficulty is connected with the concept of *qualia*. They are the qualities immediately accessible to the conscious mind and incommunicable. We can form reports about their occurrence, but we cannot describe them. I can say that my hand feels cold when I hold an ice cube, I can even say how cold it feels when I hold it, but I cannot explain what it is like to hold something cold in my hand. Some philosophers call it knowledge by acquaintance. As immediate experiences are impossible to communicate the author is tempted to develop a theory which would get rid of *qualia* without questioning the fact that our experiences have some recognizable qualities in the first-person perspective. He believes that *qualia* will disappear if we develop a theory of 'integrating functions of the mind'. They may be fake representations of something else, very much like rainbows are fake representations of thick fog perceived from a certain perspective.