

Włodzisław Duch

Uniwersytet Mikołaja Kopernika w Toruniu

NEUROKOGNITYWNA TEORIA ŚWIADOMOŚCI

John Searl w artykule o przyszłości filozofii, napisanym dla milenijnego numeru *Philosophical Transactions of Royal Society* przypomina, iż wiele problemów uważanych za filozoficzne udało się sprowadzić na grunt nauk szczegółowych. Debaty na temat różnic pomiędzy materią „bezwładną” a „ożywioną” na początku XX wieku toczono z wielką intensywnością, lecz dopiero rozwój genetyki i biologii molekularnej zmienił zupełnie meritum tej dyskusji. Podobnie obecne debaty dotyczące filozofii umysłu mogą po latach odejść w niepamięć dzięki osiągnięciom nauk kognitywnych, a w szczególności neurobiologii. Za największą przeszkodę w znalezieniu poprawnego rozwiązania problemu psychofizycznego Searl uznał „zbiór tradycyjnych, lecz przestarzałych, pojęć takich jak ciało i umysł, materia i duch, mentalne i fizyczne”. Zamiast martwić się, w jaki sposób mózg może stworzyć umysł należy według niego przyjąć, że tak po prostu jest. Umysły, kwarki, polityka czy kultura są w różny sposób częścią jednego świata.

Spróbuję tu pokazać, że tak jest w istocie. Użyję przy tym argumentu, który wydaje mi się dość oryginalny. Przedstawię mianowicie prosty schemat mózgowopodobnego układu – nazwijmy go artelektem – który może znajdować się w różnych fizycznych stanach. Relacje pomiędzy stanami artektu będą podobne do relacji pomiędzy stanami umysłu (ma to być podobieństwo relacji między stanami, a nie samych stanów). Artekt interpretując w symboliczny sposób swoje stany dynamiczne będzie musiał twierdzić, że jest świadomy i ma wrażenia. Nasze mózgi są takim właśnie układem. Pozwala to spojrzeć na zagadnienie świadomości z nowej strony i prowadzi do wielu wniosków dotyczących natury wrażeń świadomych. Najpierw jednak przedstawię bardzo krótko trudny problem świadomości i dotychczasowe próby jego rozwiązania.

1. Trudny problem świadomości

Świadomość stała się w ostatnim dziesięcioleciu modnym zagadnieniem. Ponieważ jest to pojęcie często stosowane w języku potocznym, używane w bardzo wielu kontekstach, wywołujące wiele skojarzeń, nie da się go ująć w ramach teorii naukowej, która pokryłaby wszystkie jego znaczenia. W najprostszym sensie problem świadomości sprowadza się do znalezienia neurofizjologicznych korelatów odróżniających sytuacje, w których jesteśmy jakiegoś bodźca świadomi od sytuacji, w których nie jest on uświadamiany. Są realne szanse na znalezienie takich korelatów dzięki analizie czynności bioelektrycznej i biomagnetycznej mózgu. Globalna dynamika mózgu jest prawdopodobnie związana z aktualizacją, w danej chwili czasu, niektórych śladów pamięci, skojarzeń przez nie wywołanych, stanu obszarów sensorycznych i motorycznych mózgu. Następujące po sobie stany charakteryzują się pewną czasoprzestrzenną strukturą, która zapewne odpowiada stanom pamięci roboczej, strumieniowi świadomych wrażeń, czyli treści umysłu. Nie mamy jeszcze na to dowodów, jednakże taka możliwość jest wielce prawdopodobna i nie budzi kontrowersji.

Z filozoficznego punktu widzenia najtrudniejszym zagadnieniem jest problem subiektywnych wrażeń lub ściślej rzecz biorąc problem specyficznych, subiektywnych jakości, jakie wiążą się z wrażeniami. Jest to stary problem przedstawiony w nowym ujęciu przez Davida Chalmersa (1995,96) i szczegółowo omawiany na łamach „Journal of Consciousness Studies”. Łacińskie słówko *qualis* oznacza własność w oderwaniu od rzeczy, posiadających daną własność, stąd mówi się potocznie o „problemie kwaliów”. Większość zagadnień dotyczących działania umysłu sprowadza się do wskazania na specyficzne funkcje mózgu, które są za nie odpowiedzialne. Możemy wyjaśnić powstawanie halucynacji pod wpływem środków halucynogennych wskazując na załamanie się mechanizmów pamięci asocjacyjnej w wyniku zaburzeń procesów neurochemicznych, koniecznych do jego prawidłowego działania, nie pomoże to nam jednak zrozumieć, jakiego rodzaju wrażenia ma osoba doznająca takich halucynacji. W tym przypadku nie wystarczy wyjaśnienie funkcji (np. problem integracji wrażeń o różnych

modalnościach w jedną całość sprowadza się do wyjaśnienia funkcji), opisanie dostępności informacji, sposobu tworzenia raportów na temat stanu wewnętrznego, czy powstawania dialogu wewnętrznego itp. Wrażenia subiektywne nie wydają się zredukować do funkcji. Wrażenie barwy nie daje się rozłożyć na czynniki prostsze, stąd wyjaśnienia redukcjonistyczne nie są możliwe.

Jest to zagadnienie ważne, gdyż bez jego zrozumienia nie potrafimy odpowiedzieć na pytanie „czy komputer lub inny sztuczny system może mieć wrażenia?” Zaprogramować można jedynie „różnice, które tworzą różnice” (differences, that make a difference), a więc stany informacyjne. Istnienie wrażeń nie jest obserwowalną różnicą – tak się przynajmniej twierdzi. Przetwarzanie informacji, przynajmniej w niektórych warunkach, może zachodzić bez wrażeń wzrokowych (np. ślepotą korowa, por. Dennett 1991). Z drugiej strony wrażenia wzrokowe mogą pojawiać się pomimo utraty wzroku, jeśli ta utrata związana jest z uszkodzeniami na poziomie pierwotnej kory wzrokowej lub szlaków do niej dochodzących. Często dyskutowana jest logiczna możliwość przetwarzania informacji „w ciemności” i istnienia zombi, a więc istot, które zachowując się zewnętrznie podobnie jak istoty świadome nie mają żadnych wrażeń związanych z przetwarzaniem informacji. Trudno jest jednak wyprowadzić takie wnioski z przypadków ślepoty korowej. Przypadki te pokazują raczej, że uszkodzenia układu wzrokowego bardzo mocno redukują możliwości posługiwania się sygnałami docierającymi do mózgu ze strony siatkówki. Wszelkie uszkodzenia mózgu prowadzą do obniżenia kompetencji behawioralnej. Możliwość działania pomimo braku wrażeń wzrokowych nie oznacza braku wszelkich wrażeń, a jedynie nieumiejętność interpretacji tych wrażeń. Po pewnym czasie pojawiają się nowe wrażenia świadome, odmienne od zwykłych wrażeń wzrokowych. Ślepotą korowa może być argumentem za koniecznością nauczenia się właściwej interpretacji stanów mózgu do powstania świadomych wrażeń – do zagadnienia tego jeszcze powrócę.

W dyskusji po artykule Chalmersa formułującym trudny problem świadomości, czyli problem zrozumienia jakości wrażeń świadomych, zaproponowano kilka rozwiązań, o których warto wspomnieć.

2. Czy jest tu jakiś realny problem?

Daniel Dennett (1996) i kilku innych filozofów uznało ten problem, podobnie jak problem wyjaśnienia istoty życia czy pojęcia takie jak „fajność” za problem sztuczny. Czy jest sens dyskutować o wyjaśnieniu istoty „fajności” poza opisem, co uznać można za fajne, a co nie? Według niego wrażenia subiektywne są kwestią mentalnych dyspozycji. Ta odpowiedź wskazuje na trudności w dostrzeżeniu sedna „trudnego” problemu. Próba wyobrażenia sobie, jak można by zaprogramować wrażenia w sztucznym systemie pokazuje dobitnie, że nie jest to tylko kwestia dyspozycji systemu do produkowania pisku po naciśnięciu klawisza. Życie czy fajność to kategorie abstrakcyjne, wrażenia są przeżywane w bezpośredni sposób. Czy sztuczne systemy mogą mieć wrażenia? W miarę budowania coraz bardziej złożonych systemów pytanie to stanie się bardzo ważne.

3. Rozwiązanie nie-redukcyjne

Chalmers (1995) powrócił do starego pomysłu, który można wywieść jeszcze od Spinozy. Skoro przetwarzanie informacji przez mózgi wywołuje w tych mózgach wrażenia to może istnieje podwójny aspekt informacji – fizyczny i fenomenalny? Otrzymamy w ten sposób naturalistyczny dualizm, który dużo łatwiej zaakceptować niż tradycyjny dualizm Kartezjański. Ten pomysł wydaje się mieć wielu zwolenników, ale w istocie nie jest to rozwiązanie problemu, a raczej rezygnacja z próby rozwiązania: doświadczenia subiektywne istnieją, bo tak jest. Nadal nie rozumiemy, dlaczego tak się dzieje, jakie systemy będą miały wrażenia, a jakie nie, dlaczego pewne bodźce przetwarzane przez mózg wywołują w nas wrażenia a na inne reagujemy automatycznie, bez wrażeń świadomych. Pomysł ten wydaje się bezpłodny, nie pozwala zrozumieć specyfiki wrażeń o różnych modalnościach, odpowiedzieć na pytania „dlaczego wrażenia są takie a nie inne”?

W istocie nie jest to lepsze rozwiązanie niż jawna rezygnacja, głoszona przez McGinna (1990). Być może ludzki umysł nie może zgłębić wszystkich

tajemnic natury, ale nie oznacza to, że należy szybko zrezygnować z poszukiwań dobrych rozwiązań.

4. Panpsychizm

Seager (1995) i Rosenberg (1996) opowiadają się za słabą wersją panpsychizmu, który można określić jako panprotopsychizm. Materia ma własność X, która w połączeniu ze złożoną materią tworzy świadomość. Jeśli za złożoną materię uznać taką, która przetwarza informację jest to postawa bliska podwójnemu aspektowi informacji Chalmersa. Świadomość nie pojawia się we wszystkich złożonych systemach, a tylko w takich, które przetwarzają informację w specyficzny sposób. Co odróżnia superszybkie, niezwykle złożone komputery, które nie wydają się mieć świadomości, od mózgów, które ją mają? Jest tu więcej pytań, niż odpowiedzi.

5. Protofenomeny

Podstawowe elementy świadomości, które można by nazwać „protofenomenami”, w złożonych systemach wywołują doświadczenia świadome. Jak jednak te protofenomeny wiążą procesy fizyczne z umysłowymi? Czym są i jak wiążą się z neuronami? Przypomina to próbę „kwantyzacji pola świadomości”. Eccles sądził (1985), że niematerialny umysł steruje mózgiem zmieniając prawdopodobieństwa transmisji synaptycznych za pomocą *psychonów*, pełniących rolę kwantów pola świadomości. Nie ma jednak żadnych poszlak świadczących o tym, że coś takiego istnieje i jest do czegoś przydatne. Im dokładniejsze są czysto klasyczne modele sieci neuronowych, tym więcej funkcji mózgu udaje się w nich odtworzyć. Istnienie protofenomenów lub psychonów wydaje się wielce nieprawdopodobne.

6. Świadomość kwantowa

Od ponad 40 lat próbuje się połączyć problemy związane z interpretacją podstaw mechaniki kwantowej, w szczególności tzw. problem obserwatora i kolapsu funkcji falowej, z zagadnieniem świadomości. Świat kwantowy opisany jest przez funkcję falową jako świat potencjalnie istniejących możliwości, które aktualizują się dopiero w wyniku pomiaru, stając się konkretnymi zjawiskami. Czy potrzebna jest do tego świadomość obserwatora, czy wystarczy nieodwracalny zapis wyniku? Jest wiele alternatywnych teorii pomiaru w mechanice kwantowej, które nie odwołują się do świadomości.

Jak na razie ten kierunek myślenia nie przyniósł żadnych rezultatów. Czy nielokalny model umysłu ma coś wspólnego z nielokalnością w mechanice kwantowej, jak sądzi Stapp (1993) czy Clarke (1995)? Czy do rozwiązania problemu świadomości potrzebna jest kwantowa teoria grawitacji, jak sądzi Penrose (1994)? Nie ma żadnych poszlak doświadczalnych popierających takie tezy. Efekty kwantowe w mózgu musiałyby ulec zniszczeniu z powodu jego wysokiej temperatury lub w polach magnetycznych tomografów komputerowych, ludzie nie odczuwają jednak żadnych zaburzeń świadomości w takich warunkach. Rozwiązania kwantowe niczego nie tłumaczą!

7. Neurobiologiczna teoria znaczenia

Crick i Koch (1995) próbowali podkreślić, że w przetwarzaniu informacji przez mózg tylko te bodźce, które mają dla organizmu istotne znaczenie, stają się świadome. Jest to zapewne stwierdzenie prawdziwe, ale z filozoficznego punktu widzenia mało przydatne. Trudno jest bowiem zdefiniować pojęcie znaczenia i zrozumieć, dlaczego znaczeniu miałyby towarzyszyć doświadczenia świadome. Na podobnej zasadzie, co przetwarzaniu informacji? Nie wszystkie zachodzące w mózgu procesy wydają się być konieczne dla powstawania wrażeń świadomych. Sprzężenia z niektórymi grupami neuronów są na tyle powolne, że nie mogą one wpływać na pojawianie się wrażeń świadomych, stąd pomysł poszukiwania „neuronów świadomości”. Ponieważ wszystkie sprzężenia w mózgu są dość silne jest całkiem prawdopodobne, że nie da się go podzielić na rozłączne obszary, że stany umysłu odpowiadają

globalnym stanom dynamiki a obszary reagujące wolniej mają początkowo słaby, a potem narastający, wkład do powstających wrażeń.

8. Teoria Baarsa

Bernard Baars, twórca psychologicznej kognitywnej teorii świadomości (1988), proponuje uznać istnienie świadomości jako „brutalny fakt”, z jakimi często mają do czynienia psycholodzy. Globalnie dostępna informacja staje się w jakiś sposób świadoma. Zaletą tego podejścia jest możliwość odróżnienia, kiedy informacja może być uświadamiana (gdy jest globalnie dostępna), a kiedy nie. Decydują o tym procesy skojarzeniowe zachodzące w mózgu, w szczególności szybka ocena informacji przez ośrodki podkorowe, zaangażowane w powstawanie emocji. Jest to więc krok w dobrym kierunku. Jednakże wiele pytań pozostaje bez odpowiedzi. Czy sztuczne systemy mogą mieć wrażenia i jakie warunki muszą spełnić? W jaki sposób powstają wrażenia w oparciu o informację, która jest globalnie dostępna w systemie?

Żaden z tych punktów widzenia nie prowadzi do realnego postępu w zrozumieniu trudnego problemu świadomości.

9. Sztuczne umysły

Komputery i maszyny Turinga nie są właściwą metaforą dla działania umysłu. Architektura maszyny Turinga nie przypomina sposobu działania mózgu. Analogia oparta jest jedynie na przetwarzaniu informacji i nie należy jej rozciągać na inne aspekty działania umysłu, takie jak wrażenia. Chociaż rezultaty działania umysłu i programów komputerowych mogą być podobne, wewnętrzne stany różnią się w obu przypadkach w zasadniczy sposób.

W danej chwili czasu rejestry procesora komputera mają określone wartości binarne, decydujące o dalszym działaniu systemu. Do jego opisu niepotrzebne są relacje przestrzenne czy czasowe, wystarczy abstrakcyjny opis typu maszyny Turinga. Tymczasem stan bioelektrycznej aktywności mózgu ma

określony rozkład przestrzenny, rozciągłą, zmieniającą się w czasie strukturę. Zawiera ona w każdym momencie wkład od wielu procesów peryferyjnych (nieuświadamianych), skojarzenia z przeszłymi zdarzeniami, pobudzenia sensoryczne i wiele innych elementów. Fizycznego stanu mózgu nie można utożsamiać ze zmianą wartości binarnych rejestrów procesora. Działanie komputera nie przypomina działania mózgu i dlatego za pomocą komputerów można w dużym przybliżeniu oddać tylko niektóre aspekty działania mózgu, związane z przetwarzaniem informacji. Dotyczy to np. rozwiązywania abstrakcyjnych problemów, budowania modeli mentalnych, procesów wymagających przetwarzania informacji symbolicznej, ale nie dotyczy tak podstawowych funkcji jak czucie i świadomość.

Z drugiej strony nawet najprostsze neuronowe modele pamięci asocjacyjnej wykazują zdolność do rozpoznawania uszkodzonych wzorców, czas rozpoznania nie zależy w nich od liczby zapamiętanych wzorców, uszkodzenie części macierzy połączeń nie prowadzi do zapominania konkretnych wzorców a jedynie ogólnego pogorszenia pamięci, pomyłki dla podobnych wzorców są częstsze niż dla wyraźnie różnych, przepełnienie pamięci prowadzi do chaotycznego zachowania. Są to znane z psychologii cechy charakterystyczne ludzkiej pamięci. W ostatnich latach bardzo intensywnie rozwija się komputerowa psychiatria (por. Duch 2000), dziedzina wykorzystująca proste modele oparte na sieciach neuronowych do wyjaśnienia wielu aspektów syndromów neuropsychologicznych i chorób psychicznych, począwszy od różnego rodzaju amnezji (w tym dziecięcej), procesów degeneracji funkcji poznawczych, natury schizofrenicznych halucynacji jak i pojawiania się odczuć kończyn fantomatycznych. Nie ma wątpliwości, że im dokładniejsze modele mózgu będziemy tworzyć, tym więcej funkcji umysłu uda się za ich pomocą uchwycić. Czy proces ten zbiega się do „prawdziwego”, świadomego umysłu?

Spróbuję tu uzasadnić, że mózgowopodobna organizacja przetwarzania informacji musi doprowadzić do pojawienia się „przestrzeni wewnętrznej”, a w niej umysłu i wirtualnej świadomości. Sztuczny umysł, w miarę doskonalenia modelu obliczeń opartego na mózgowopodobnej organizacji, stanie się coraz bardziej podobny do prawdziwego umysłu. Będzie nie tylko twierdzić, iż ma

wrażenia, ale też będzie sam o tym przekonany, interpretując jako wrażenia stan swojej pamięci roboczej.

10. Czym są wrażenia?

Umysł jest układem kontrolnym systemu (organizmu) o specyficznej architekturze, pozwalającej tworzyć model świata (por. Piłat 1999). Najbardziej istotne części tej architektury to pamięć robocza (WM), pamięć trwała, automatyczna aktywacja skojarzeń przez stany pamięci roboczej, oraz oparta na niej zdolność do rozróżniania zmieniających się w sposób ciągły stanów pamięci roboczej. Bardziej rozwinięte umysły, takie jak ludzkie, posiadają również reprezentację swojego jestestwa (w postaci zespołu struktur stałej pamięci oraz mechanizmu filtracji informacji zagrażającej stabilności tego zespołu) oraz możliwość symbolicznego komentowania stanu WM.

Podstawowe operacje wykonywane przez takie umysły to kategoryzacja, wartościowanie i odnoszenie percepcji do modelu wewnętrznego jestestwa. Konieczne jest szybkie porównanie np. postrzeganego smaku ze smakami zapamiętanymi w pamięci długotrwałej. Szczury, jak i wszystkie ssaki, wykształciły tę umiejętność w wysokim stopniu, gdyż ich organizm jest bardziej delikatny niż organizm gadów, a środowisko wymusiło na nich szybkie podejmowanie decyzji. Pamięć roboczą można uznać za centralny system dystrybucji informacji do wszystkich części mózgu (por. teoria Baarsa 1988). Pamięć robocza (WM) jest niewielka, oparta na dynamicznych stanach mózgu, przechowuje tylko informacje konieczne do podjęcia bieżących decyzji. Pamięć długotrwała jest ogromna, pozwala odnieść się do zgromadzonej w ciągu całego życia wiedzy. Jej konstrukcja umożliwia pojawienie się szybkich skojarzeń poprzez interferencję ze stanami pamięci roboczej. Szczur smakując nowe pożywienie wprowadza informację o wyniku przetwarzania bodźców węchowych i smakowych do swojej pamięci roboczej. Informacja ta udostępniana jest pamięci trwałej i pojawia się skojarzenie, oparte na wspomnieniu, które można np. zinterpretować jako: ostatnim razem po zjedzeniu czegoś podobnego ledwo przeżyłem. Pojawia się reakcja wstrętu i związane z nią reakcje pochodne – wrażenia.

Adekwatna reprezentacja symboliczna zmieniających się w sposób ciągły wrażeń nie jest możliwa. Zwierzęta i niemowlęta nie używają języka, a jednak mają podobne wrażenia jak dorośli ludzie. Czym więc są wrażenia i dlaczego istnieją? Dyskryminacja sygnałów z różnych zmysłów i różnych sygnałów tej samej modalności ułatwia porównania. Rezultaty tej dyskryminacji pojawiają się jako stany umysłu – może pojawić się myśl (reprezentacja dźwięku, symbolu, *common sensorum*), lub wrażenie, reprezentacja niewerbalna. Powstanie wrażenia wymaga stałej, niewerbalnej aktualizacji stanu pamięci roboczej. Wrażenia można więc traktować jako „niewerbalne nazwy” pozwalające na porównanie ciągłych reprezentacji stanów zmysłów na poziomie globalnej dynamiki mózgu. Mechanizm tej dyskryminacji oparty jest na „adaptacyjnych stanach rezonansowych”, zgodnie z modelem ART Grossberga (1995).

Wrażenia różnią się między sobą jakościowo, gdyż są to różnego rodzaju pobudzenia (globalne stany dynamiczne) mózgu, obejmują aktywację różnych struktur mózgu w różnym stopniu. Do powstania wrażeń konieczne są prawidłowo przebiegające funkcje mózgu, ale wrażeniom powstającym na skutek tych funkcji w mózgu skłonny jestem przypisać status ontologiczny – są, istnieją nie mniej realnie niż jakiegokolwiek inne stany przedmiotów fizycznych. Stany mózgu pełnią funkcję w relacyjnych strukturach charakteryzujących obiekty umysłu, podlegając ciągłej „interpretacji” przez wszystkie struktury mózgu, które w wyniku tej „interpretacji” przyczyniają się do powstawania nowych stanów. Czerwony kolor wywołuje we mnie specyficzne wrażenie nie tylko z powodu moich dyspozycji mentalnych (w tym skojarzeń wpływających na stan pamięci roboczej), lecz również realnych stanów mózgu/umysłu, które powstają w czasie percepcji. Nie chodzi tu o redukcję stanów świadomości do stanów mózgu – raczej o teorię relacyjną, gdyż wszelka interpretacja tych stanów odwołuje się do ich wzajemnych powiązań, decydujących o powstającym ciągu stanów. Zrozumienie struktury wzajemnych relacji pomiędzy stanami mózgu musi uwzględniać rolę środowiska w kształtowaniu się tych relacji.

Takie ujęcie natury wrażeń pozwala zrozumieć wiele aspektów dotyczących ich przeżywania. Wrażenia są ściśle związane z mechanizmami

uwagi, które decydują o tym, co dla organizmu powinno być najważniejsze do podjęcia decyzji kontrolujących zachowanie. Większość bodźców nie wywołuje w nas wrażeń gdyż nie dociera do pamięci roboczej, przetwarzane są na poziomie lokalnych struktur mózgu. Czyste doświadczenie fenomenalne to filozoficzna fikcja. Wrażenia spełniają rolę funkcjonalną, taki przynajmniej jest ich sens ewolucyjny. „Odczuwanie jakości wrażeń” to nic innego jak proces interpretacji stanów WM przez obszary mózgu odpowiedzialne za percepcję i mechanizmy poznawcze. Jeśli te mechanizmy nie działają odczuwanie jakości wrażeń powinno zniknąć. Obserwujemy to np. w stanach absorpcji kontemplacyjnej (Austin 1999).

W tym ujęciu pytanie „czym jest świadomość” nie znika, tak jak to się dzieje u Dennetta (1991), który sprowadza wszystko do roli dyspozycji. Świadomość pierwotna jest zdolnością do interpretacji stanów pamięci roboczej. Świadomość refleksyjna wymaga modelu jestestwa (Dasein). Poczucie tożsamości wiąże się z wewnętrznym modelem ciała i propriocepcją, przewidywaniem skutków swoich ruchów. Wrażenie istnienia, poczucie bytu, jest prawdopodobnie związane z procesami dyskryminacji reprezentacji sensomotorycznych podstawowych procesów wynikających z posiadania ciała (por. Damasio 2001).

11. Rola mechanizmów poznawczych w powstawaniu wrażeń

Z powyższych rozważań można wyciągnąć wiele wniosków i przewidywać dotyczących sposobu odczuwania wrażeń.

Wrażenia świadome nie mają specjalnego statusu wśród innych zjawisk umysłowych. Za poczucie jakości wrażeń odpowiedzialne są skomplikowane mechanizmy poznawcze. Zawieszenie działania tych mechanizmów prowadzi do zaniku wrażeń (np. habituacja, brak uwagi). Brak interpretacji docierających do mózgu sygnałów jest równoznaczny z brakiem wrażeń. Można „patrzeć nie widząc”, nie zauważając nawet wyraźnych zmian w środowisku, np. nowych mebli w domu, pomimo prawidłowego działania, np. omijania tych mebli (por. O'Regan, Nöe 2001). Dopiero pojawienie się faktu w pamięci roboczej powoduje, że stajemy się go świadomi i pojawiają się

związane z nim wrażenia. Silna koncentracja na wykonywanym zadaniu modyfikuje lub całkowicie blokuje wrażenia (np. bólu w czasie walki), absorpcja w stanach kontemplacji może nawet spowodować całkowity zanik wrażeń. Z drugiej strony gra aktorska, udawanie emocji przy pełnej identyfikacji z odgrywaną rolą, może doprowadzić do przeżywania prawdziwych emocji.

Wyłączenie dominującej półkuli mózgu np. za pomocą próby Wada (Churchland 1986) lub drażnienia prądem zaburza mechanizmy interpretacyjne i powoduje zanik rozpoznawalnych wrażeń. Aktywność pozostałej półkuli nie wystarcza do precyzyjnej interpretacji, stąd wrażenia stają się nieokreślone, rozmyte. Brak aktywności mózgu znosi oczywiście wszelkie wrażenia.

Trenowanie mechanizmów poznawczych prowadzi do „wyostrzenia zmysłów”, nowych jakości wrażeń słuchu, smaku, wzroku. Np. już po pół godzinie pracy nad dużą układanką (puzzle) można zauważyć subtelne zmiany w postrzeganiu kształtów i kolorów. Pamięć wrażeń konieczna jest do dobrej dyskryminacji.

Nowe wrażenia dostępne są również w snach, gdyż opierają się na zmianach w pamięci długotrwałej. Zmiany te powodują powstawanie innego rodzaju skojarzeń z informacją zawartą w pamięci roboczej. Odczuwanie wrażeń zachowane jest w trakcie snów, w fazie REM, w której widać dużą aktywność kory (świadcząca o działaniu pamięci roboczej). Możliwa jest kontynuacja wrażeń sennych na jawie, np. pod wpływem *chlordiazepoxidu*.

Znikające wrażenia pociągają za sobą zredukowaną kompetencję behawioralną, co widać np. w przypadku ślepoty korowej. Funkcjonalna rola wrażeń pociąga za sobą różne własności strukturalne odczuwania wrażeń. Wrażenia wzrokowe, dotykowe, czucie temperatury i bólu mają strukturę przestrzenną, w mniejszym stopniu mają ją wrażenia słuchowe, za to wrażenia smaku, zapachu, myśli czy wyobrażenia nie mają jej wcale. Modalności, za pomocą których odbieramy te wrażenia pełnią rolę drugorzędną, najważniejsza jest ich struktura. Aparatura dotykowa, przekształcająca sygnały z kamery na pobudzenia skóry na plecach, po okresie przyzwyczajenia się do jej używania dostarcza osobom niewidomym wrażeń podobnej natury co wzrokowe, z

możliwością oceny położenia przestrzennego, wielkości, kształtu i ruchu obiektu.

Stymulacja struktur mózgu odpowiedzialnych za przetwarzanie danych wrażeń powinna prowadzić do ich powstawania tylko wtedy, gdy mogą one wpłynąć na globalną dynamikę działania mózgu. Dlatego zbyt słaba stymulacja (podprogowa) nie wywołuje wrażeń. Powstanie wrażenia wymaga dostarczenia dostatecznie dużej ilości energii, na tyle silnej pobudzającej neurony mózgu, by wpłynęło to na globalną aktywność mózgu, a nie tylko na lokalne obszary (cf. Libet 1993).

Poczucie wstrętu może powstać w wyniku reakcji na bodźce wzrokowe, słuchowe i smakowe. Czym uzasadnione jest używanie tego samego określenia „wstręt” na wrażenia wywołane w różny sposób? Z introspekcji wiemy, że chodzi o ten sam stan. Powinniśmy więc oczekiwać, że pewne wrażenia sensoryczne obecne w pamięci roboczej wywołują interferencję z jakimiś obszarami mózgu, które zmieniają stan WM na stan interpretowany jako wstręt, stan który możemy ukryć, a który u zwierząt pociąga za sobą określone zachowania motoryczne. Dyskryminacja stanów pamięci roboczej, stanowiąca podstawę ich interpretacji, może polegać na prostym procesie interferencji z wyspecjalizowanym obszarem mózgu służącym do rozpoznawaniu jakiegoś stanu. Tak się istotnie dzieje – pobudzenia różnej modalności prowadzą do tego samego stanu mózgu, pobudzenia przedniej części kory wyspy, rozpoznawanego jako wstręt. Być może pobudzenia tego obszaru przez nerwy czuciowe dochodzące od różnych kończyn są odpowiedzialne za dziwaczny syndrom neuropsychologiczny jakim jest dysmorfia ciała. Pacjenci cierpiący na ten syndrom odczuwają bardzo silny wstręt do swoich kończyn decydując się czasami na ich amputację. Błędna interpretacja braku sygnałów w mózgu może też być odpowiedzialna za zaniedbanie jednostronne, czyli brak wrażeń dotyczący połowy postrzeganego pola widzenia. Syndrom ten pojawia się najczęściej po wylewie w obszarze prawego płata ciemieniowego i nie pozwala pacjentom wyobrazić sobie lewej połowy przestrzeni.

Z badań nad mózgiem wiadomo, że uszkodzenia pierwotnych obszarów sensorycznych lub szlaków doprowadzających do nich sygnały powodują zanik wrażeń. Uszkodzenia wtórnych obszarów sensorycznych wywołują

asymbolię (Ramachandran 1999), np. asymbolię bólu, zjawisko polegające na niezdolności do interpretacji odbieranych wrażeń jako bólu. Ramachandran opisuje fascynujące eksperymenty z osobami cierpiącymi z powodu paraliżu kończyn fantomowych. Osoby te czują swoje amputowane kończyny, cierpiąc okropne bóle, np. z powodu paznokci wrastających w nieistniejącą dłoń, której nie potrafią otworzyć. Jeśli pacjentom wydaje się, że amputowana ręka jest widziana – iluzję taką można stworzyć pokazując zdrową rękę w odpowiednio ustawionym lustrze – zaczynają czuć jej ruch i w końcu udaje im się rękę otworzyć, co prowadzi do zaniku bólu. Wrażenie bólu jest więc związane z interpretacją pojawiających się w mózgu pobudzeń.

Fenomenologia bólu wydaje się być bardzo obiecującym terenem szczegółowych badań nad naturą wrażeń. Już w okresie wojennym chirurdzy zauważyli, że bez interpretacji poznawczej ból nie zawsze prowadzi do cierpienia, pozostając czystym wrażeniem (np. Beecher, 1946). Causalgia, piekący ból pourazowy, pojawić się może pod wpływem dowolnego bodźca. Pełna koncentracja na przeżyciu bólu może doprowadzić do zaniku cierpienia. Ból zwykle wywołuje cierpienie, ale u masochistów przyjemność. Próg wrażliwości na ból zależy od oczekiwań. Wszystkie te zjawiska potwierdzają rolę mechanizmów interpretujących stany mózgu w odczuwaniu bólu. Analiza zjawisk związanych z przeżyciem bólu jest dobrym testem przedstawionej tu hipotezy dotyczącej natury wrażeń i świadomości.

12. Zarzuty

Wszystkie wypowiedzi dotyczące wrażeń wymagają działania mechanizmów poznawczych. Jestem przekonany, że mam wrażenia, które nie są złudzeniem, ale realnymi stanami mojego mózgu, stanami, które mogę skomentować (niekoniecznie werbalnie), a więc stanami umysłu. Czy podobne przekonania mieć może maszyna? Jeśli tak, należy uznać, że komputery o odpowiedniej architekturze mogą być świadome i mieć wrażenia, mogą przeżywać ich specyficzne jakości, podobnie jak my przeżywamy nasze wrażenia. Mózgopodobne przetwarzanie informacji musi prowadzić w granicy coraz

bardziej złożonych modeli do powstania systemów, które komentując stany swojej pamięci roboczej będą miały przekonanie, że mają wrażenia. Czy możemy uznać, iż będzie to przekonanie fałszywe? Nie widzę dobrych argumentów przeciw takiej możliwości. Eksperymenty myślowe mające świadczyć przeciw możliwości istnienia sztucznych umysłów nie są przekonywujące. Oczywiście wrażenia identyczne z ludzkimi wymagają identycznych mózgow, ale wrażenia ludzkopodobne tylko podobnych relacji pomiędzy wzorcami pobudzeń pamięci roboczej. Te pobudzenia tworzą tkanę wewnętrzną przestrzeni, w której mieści się nasz model świata, nazywany umysłem. Iluzja, że to właśnie „duch w maszynie” rządzi naszym zachowaniem jest bardzo silna (por. Eccles 1985). Brakuje dobrej fenomenologii świadomości, potrzeba neurofenomenologii (cf. F. Varela 1996), głębokiej introspekcji, a nie powierzchownych obserwacji.

Rozważania filozoficzne dotyczące umysłu uwikłane są w problemy lingwistyczne. Używając języka potocznego do opisu umysłu skazani jesteśmy na brak precyzji. Próba wprowadzenia pośredniego języka, opisującego zdarzenia umysłowe, języka na poziomie pośrednim pomiędzy zdarzeniami neurofizjologicznymi a psychologicznymi, przedstawiona została w (Duch 1994, 1997) oraz inne ujęcie w (Churchland 1995). W istocie nie potrafimy do końca zdefiniować żadnego obiektu czy pojęcia fizycznego, możemy natomiast mówić precyzyjnie o relacjach pomiędzy nimi. Wprowadzenie precyzyjnego języka powinno się przyczynić do rozpoznania pozornych problemów dotyczących natury umysłu, np. jego nielokalności czy nieprzestrzennej natury, dyskutowane przez zwolenników teorii kwantowych.

Czy mamy dobry test, odróżniający systemy świadome, posiadające wrażenia, od systemów przetwarzających informację? Jeśli kiedyś dojdzie do spotkania dziwacznych form inteligentnego życia czy będziemy mogli rozróżnić pomiędzy mechaniczną inteligencją a świadomą istotą? Słynny eksperyment z Chińskim Pokojem (Searl, 1980; Kloch 1996) nie jest tu szczególnie przydatny. Nie jest to test, gdyż jego wynik jest zawsze negatywny, z góry przesądzony. Czy zagłębimy do wnętrza mózgu, czy też Chińskiego Pokoju, nie zobaczymy tam zrozumienia. Należy się raczej

zapytać, jakie spełnione muszą być warunki by obserwator rozumiał, co robi dany system?

Alison Gopnik (1998) napisała wielce interesujący artykuł pod tytułem „Zrozumienie jako orgazm”. Zbudowanie modelu, zrozumienie czegoś wymaga z ewolucyjnego punktu widzenia u inteligentnych istot sygnału nagrody, zachęcającego do tego typu działania. „Narrator”, czyli ta część mózgu, która usiłuje tworzyć symboliczne, łatwe do zapamiętania komentarze zachowania organizmu, daje sygnał zakończenia pracy, interpretowany jako przyjemne zdarzenie. Narrator rozwinął się z obszarów zaangażowanych w tworzenie teorii o świecie, aktywnych w dzieciństwie (i u naukowców). Ten sam mechanizm działa na poziomie rozumienia zdania jak i złożonej teorii.

Zrozumienie w sensie emocjonalnego poczucia „rozumiem”, oraz zrozumienie intelektualne w sensie integracji nowej wiedzy i odpowiedzi na pytania z nią związane, to dwie różne rzeczy. Tłumacze kabinowi poprawnie tłumaczący zdania nie potrafią czasami odpowiedzieć na najprostsze pytania dotyczące tłumaczonej rozmowy. Koncentracja na tłumaczeniu i poczucie „rozumiem”, nie gwarantuje zrozumienia intelektualnego. Wewnętrzne poczucie zrozumienia jest tylko sygnałem mózgu: zrozumiałem, mogę działać dalej. Dysocjacja pomiędzy poczuciem rozumienia a rozumieniem intelektualnym pojawia się po zażyciu niektórych środków halucynogennych (np. LSD). Skrajnym przypadkiem braku poczucia zrozumienia pomimo zrozumienia czy rozpoznania intelektualnego jest syndrom Capgrasa.

Nie warto więc odwoływać się do tajemnych mocy biologicznych neuronów by uzasadnić, że w naszych mózgach zachodzi prawdziwe rozumienie, a w systemach sztucznych nie jest ono możliwe (Searl 1999a). Nie potrafimy tego rozpoznać inaczej, niż poprzez testy behawioralne, a więc test Turinga lub jego odmiany. Jedynie doprowadzając system sztuczny i nasze mózgi do rezonansu, mapując i przekształcając struktury pobudzeń systemu sztucznego na struktury mózgu, moglibyśmy się przekonać czy jest to rzeczywiście system rozumiejący w naszym sensie. Trudno sobie jednak wyobrazić tego typu rezonans. Wymaganie, by system działał na tyle podobnie do naszego mózgu, by dało się wprowadzić pomiędzy nimi jakies jednoznaczne odwzorowanie, jest niepotrzebnym ograniczeniem na

architekturę systemu, który mógłby mieć całkiem odmienne poczucie rozumienia od naszego. Najważniejsze są relacje pomiędzy stanami wewnętrznymi takiego systemu i jego działaniami sensomotorycznymi. Dzięki tym relacjom system taki mógłby wykazywać jakąś formę samoświadomości, „wiedzieć, że wie”.

Hilary Putnam (1998) doszedł do przekonania, że za całe zamieszanie w filozofii umysłu odpowiada przekonanie, iż postrzeżeniom towarzyszą jakieś wewnętrzne stany fenomenalne, identyczne za każdym razem, gdy mamy takie same wrażenia. Chociaż argument wysuwany przez Putnana na poparcie tej tezy jest zły to sama teza wydaje się być poprawna. Putnam rozważa eksperyment, w którym mamy 100 kart, kolejne karty pomalowane białą farbą, do której przed malowaniem kolejnej karty dodajemy kroplę czerwonej farby. Karty porównywane parami (1,2), (2,3) aż do (99,100), wydają się mieć identyczną barwę, a więc wrażeniom związanym z ich postrzeganiem powinny odpowiadać identyczne stany wewnętrzne. Ponieważ relacja „bycia w określonym stanie” jest przechodnia, pisze Putnam, więc dla karty 1 i 100 stany wewnętrzne są identyczne, ponieważ dla każdej pary były identyczne. Założenie, że oglądając dwie karty jesteśmy w jednym ze stanów wewnętrznych typu „barwa nr x ” jest oczywiście błędne, bo mamy tu proces ciągły (x jest ciągłe), a nie stany dyskretne. Mózg nie jest urządzeniem dyskretnym, nie ma w nim identycznych stanów za każdym razem, gdy rozpoznajemy ten sam obiekt. Skłonni jesteśmy przypisać do jednej kategorii nieskończenie wiele wrażeń, wywołujących nieznacznie się od siebie różniące stany. Prawdopodobnie nawet to nie jest potrzebne. Istotne są relacje pomiędzy stanami, a nie same stany wewnętrzne. Walter Freeman (1995) badając zachowania królików rozpoznających ten sam zapach i reagujących na niego w taki sam sposób stwierdził, że stan kory węchowej jest za każdym razem odmienny, nie ma prostych „reprezentacji wewnętrznych”. Stan kory węchowej zależny jest od wielu procesów zachodzących w mózgu królika. Jedynie relacje pomiędzy stanem „rozpoznano określony zapach”, a wywołanym tym rozpoznaniem zachowaniem, są istotne. Nie istnieje bezpośrednia reprezentacja postrzeganych obiektów, do działania wystarczy tylko zachowanie relacji pomiędzy stanami mózgu podobnych do relacji

zachodzących w obserwowanym środowisku, a więc podobieństwo drugiego rzędu.

Zarzut matematyczny – twierdzenia Gödla, Churcha, Turinga o zdaniach nierozstrzygalnych (por. Penrose 1994) – nie wydaje mi się istotny. Pokazuje on jedynie, że nie można stworzyć maszyny wszechwiedzącej, która rozstrzygnie wszystko za pomocą aksjomatów. Ograniczenia umysłu człowieka są znacznie poważniejsze niż ograniczenia maszyny Turinga. Wynika to choćby z niewielkiej pamięci roboczej, jaką dysponujemy, i skończonego czasu życia. Sztuczna inteligencja największe sukcesy odnosi właśnie w symulowaniu myślenia, grach planszowych, systemach ekspertowych, dowodzeniu twierdzeń matematycznych. Nie ma wątpliwości, że potężniejsze moce obliczeniowe, pamięci skojarzeniowe oraz reprezentacje wiedzy stosowane w takich badaniach pozwolą na rozwiązywanie problemów daleko przekraczających możliwości naszych mózgów.

Intencjonalność również nie wydaje się być przeszkoda na drodze do sztucznego umysłu. Tworzone obecnie roboty, takie jak roboty z serii Darwin (Edelman 1999) czy projekt Cog rozwijany na MIT, są przykładem inteligencji behawioralnej, rozwijającej się na skutek kontaktów z otoczeniem przez ugruntowanie sensu rozpoznawanych obiektów w reprezentacjach sensomotorycznych. Bardzo ogólne potrzeby, takie jak potrzeba doznań, prowadzą w wyniku ewolucji umysłu pod wpływem środowiska do złożonych zachowań i potrzeb. Na poziomie działania świadomego nasze głębsze motywy nie są widoczne, lecz są one wbudowane przez ewolucję w stare części mózgu.

Nagel (1974) w słynnym artykule „Jak to jest być nietoperzem” pokazuje naszą bezradność wobec problemu cudzych wrażeń, zwłaszcza wrażeń tak odmiennych istot jak nietoperze. Jedyna perspektywa, z jakiej potrafimy oglądać świat, to nasza perspektywa wewnętrzna. To samo dotyczy sztucznych umysłów: będą mogły komentować i przeżywać jedynie swoje stany pamięci roboczej. Możemy oczywiście prowadzić intelektualne dyskusje na temat bycia nietoperzem czy daltonistą, ale zrozumienie intelektualne nigdy nie będzie tym samym, co bezpośrednie przeżycie. W obu przypadkach mamy bowiem do czynienia z zupełnie innymi stanami mózgu: w pierwszym mamy model symboliczny, narracyjny, realizowany głównie przez korę czołową i

skroniową, a w drugim zaangażowane są wszystkie obszary mózgu. Mary daltonistka (Jackson, 1982) może o kolorach wszystko wiedzieć intelektualnie, ale nie zastąpi to bezpośredniego przeżycia. Wrażenia to realne stany mózgu, nowe wrażenia pociągają za sobą nowe stany, a braku odpowiednich stanów kory wzrokowej (wrażeń koloru) nie sposób zastąpić stanami kory czołowej (wiedzy intelektualnej). Modele narracyjne nie wystarczą choćby z powodu niewerbalnej natury prawej półkuli, Mary nie może więc wiedzieć „wszystkiego”.

13. Podsumowanie

Poszukiwanie rozwiązań w teoriach fundamentalnych jest błędne. Nic nie wskazuje na konieczność odwołania się do efektów kwantowych, protofenomenów, psychonów, fenomenalnych aspektów informacji ani innych koncepcji fundamentalnych. Świadomość i wrażenia pojawiają się jedynie w normalnie działających systemach o mózgowopodobnej organizacji. Przedstawione tu idee po raz pierwszy pozwalają na wyciągnięcie wniosków dotyczących natury wrażeń świadomych, przełamując impas w dyskusjach nad trudnym problem świadomości. Wydaje mi się, że żadna alternatywna teoria nie potrafi wyjaśnić, dlaczego wspominając smak ostatnio zjedzonego lodu mam wrażenia, a wspomnieniom o zawiązywaniu krawata żadne jakości wrażeń nie towarzyszą. W pierwszym z tych przypadków ślady pamięci pobudzają się poprzez rezonans adaptacyjny z pamięcią roboczą. W drugim mamy do czynienia z pamięcią typu umiejętności, działaniem sensomotorycznym, które nie może modyfikować pamięci roboczej, a może jedynie być wykonane.

Prawidłowo działające funkcje mózgu są konieczne by być świadomym wrażeń, ale wrażenia nie sprowadzają się do funkcji – tu Chalmers (1996) ma rację. Wrażenia istnieją w postaci procesów fizycznych, którym – poprzez relacje z innymi procesami i działaniami sensomotorycznymi – nadajemy sens. Kognitywna teoria świadomości wymaga zdefiniowania i określenia relacji pomiędzy pojęciami: świadomość, pamięć robocza, uwaga, nieświadomość, umysł. Na umysł składa się zbiór stanów interpretowalnych wewnątrznie,

struktura i możliwe relacje między tymi stanami. Struktura ta jest odbiciem pewnych relacji pomiędzy stanami pamięci roboczej, które powstają na skutek pobudzeń sensorycznych i własnych działań motorycznych. Zadaniem umysłu jest kontrolować organizm na podstawie wewnętrznego modelu świata.

Wydaje mi się, że są realne szanse zbudowania kognitywnej teorii świadomości w oparciu o zarysowane tu idee. Przy odpowiedniej konstrukcji sztuczne mózgi będą twierdzić, że mają wrażenia i są świadome. Jednak konstrukcja sztucznych mózgów, kiedy już zaczniemy je tworzyć, jeszcze długo będzie się różniła od mózgów prawdziwych, a tym samym wrażenia sztucznych umysłów będą odmienne od naszych.

Literatura:

- Austin, J. H. (1999): *Zen and the Brain. Toward an Understanding of Meditation and Consciousness*. MIT Press.
- Baars, B. J. (1988), *A cognitive theory of consciousness*, Cambridge Univ. Press
- Beecher, H.K. (1946), Pain in men wounded in battle. *Annals of Surgery*, 123(1), 95-105
- D.J Chalmers (1995) Facing up to the problem of consciousness. *J. of Consciousness Studies* 2: 200-219
- D.J. Chalmers (1996) *The conscious mind: In search of a fundamental theory*. New York, NY, USA: Oxford University Press
- Churchland, P. S. (1986), *Neurophilosophy. Toward a unified science of the mind/brain*, MIT Press, Bradford Book
- Churchland, P. (1995), *Engine of reason*, MIT Press
- Clarke C.J.S. (1995), The nonlocality of mind. *Journal of Consciousness Studies* 2:231-40
- Crick, F, Koch, C. (1995), Why neuroscience may be able to explain consciousness. *Scientific American* 273(6): 84-85
- Damasio A. (1999), *The Feeling of What Happens*, tłum. *Tajemnica świadomości*, Poznań, Rebis 2000.

- Dennett, D. C. (1991), *Consciousness explained*, Boston, Toronto, London: Little, Brown and Co.
- Dennett, D.C. (1996), Facing backwards on the problem of consciousness. *Journal of Consciousness Studies* 3:4-6
- Duch, W. (1994), A solution to the fundamental problems of cognitive sciences. International Philosophical Preprint Exchange (1994)
- Duch, W. (1997), Platonic model of mind as an approximation to neurodynamics. W: *Brain-like computing and intelligent information systems*. red. S-i. Amari, N. Kasabov. Springer, Singapore, rozdz. 20, str. 491-512
- Duch, W. (2000), Sieci neuronowe w modelowaniu zaburzeń neuropsychologicznych i chorób psychicznych. W: *Biocybernetyka, Tom 6: Sieci neuronowe* (red. W. Duch, J. Korbicz, L. Rutkowski i R. Tadeusiewicz), rozdz. II.18. Warszawa, Akademicka Oficyna Wydawnicza-Exit.
- Eccles, J. (1985), *The Brain and the unity of conscious experience*, Cambridge University Press
- Edelman, G. (1999), *Przenikliwe powietrze, jasny ogień. O materii umysłu*, Warszawa: PIW
- Freeman, W.J. (1995), *Societies of Brains: A study in the neuroscience of love and hate*. Lawrence Erlbaum Associates
- Gopnik, A. (1998), Explanation as orgasm, *Minds and Machines*, 8, s. 101–118
- Grossberg, S. (1995), The attentive brain, *American Scientist*, 83, s. 483-449
- Jackson, F. (1982), Epiphenomenal qualia, *Philosophical Quarterly*, 32, s. 127-136
- Kloch, J. (1996), *Świadomość komputerów?*, Biblios Tarnów i OBI Kraków
- Libet, B. (1993), *Neurophysiology of Consciousness. Collected papers and new essays*, Boston, Basel, Berlin: Birkhäuser
- McGinn, C. (1990), *The Problem of Consciousness*. Oxford: Blackwells.
- Nagel, T. (1974), What is it like to be a bat?, *Philosophical Review*, 4, s. 435-50
- O' Regan, J.K, Noë, A. (2001), A sensorimotor account of vision and visual consciousness. *Behavioral and Brain Sciences* 24 (5): 883-917

- Penrose, R. (1994), *Shadows of the mind*, Oxford University Press; tłum: *Cienie umysłu*. Poznań, Wyd. Zysk i Ska, 2001
- Putnam, H. (1998), dodatek do "The Royce Lectures in the Philosophy of Mind", wykładów wygłoszonych na Brown University, 1998 (dziękuję autorowi za przesłanie mi manuskryptu)
- Piłat, R. (1999), *Umysł jako model świata*, Wyd. IFiS PAN
- Ramachandran, V. S. (1999), *Consciousness and body image: lessons from phantom limbs, Capgras syndrome and pain asymbolia*, Phil. Trans. R. Soc. Lond. B, 353, s. 1851-1859
- Rosenberg G.H. (1996), Rethinking nature: A hard problem within the hard problem. *Journal of Consciousness Studies* 3:76-88.
- Stapp H.P. (1993), *Mind, matter and quantum mechanics*. Berlin: Springer.
- Seager W. (1995), Consciousness, information, and panpsychism. *Journal of Consciousness Studies* 2:272-88.
- Searle, J.R. (1980), Minds, Brains and Programs. *Behavioral and Brain Sciences* 3: 417-457
- Searle, J. R. (1999), The future of philosophy, *Philosophical Transactions of Royal Society*, B 354, s. 2069-2080
- Searle, J. R. (1999a), *Umysł na nowo odkryty*, Warszawa: PIW.
- Varela, F. (1996), Neurophenomenology: A methodological remedy for the hard problem, *Journal of Consciousness Studies*, 3, s. 330-349