



doprowadzenie rozmówcy do lepszego zrozumienia. Było to wielkie odkrycie metodologiczne. Co nas popycha do zadawania dobrych pytań i szukania odpowiedzi, a co nas hamuje?

Czy starożytni faktycznie odkryli coś prawdziwego, czy też nasze myślenie nadal podąża tymi samymi torami pod ich wpływem?

**Podejście Platona** można dostrzec w fizyce teoretycznej, która dała nam obraz świata oparty na matematyce, na relacjach pomiędzy pomiarami, a nie na analizie bezpośrednich postrzeżeń zmysłowych. Fizycy teoretycy często identyfikują się z ideami Platona, szukając abstrakcyjnych struktur wyjaśniających naturę świata.

**Podejście Arystotelesa** było związane ze światem doznań zmysłowych, naukami empirycznymi. Jego pisma dotarły do Europy w XIII wieku z krajów arabskich, i pomimo początkowego potępienia przez biskupów i papieży (Heller, 2014, s. 49-50) w ciągu jednego stulecia zdominowały myślenie filozofów jak i całą ówczesną naukę.

**Każde skomplikowane zjawisko ma swoje proste wyjaśnienie, które jest błędne ... nie inaczej jest z poglądami na naturę ludzką.**

Klasyczna filozofia odkryła wiele istotnych problemów skupiając się nad spojrzeniem na rzeczywistość pod wybranym kątem. Jednakże starożytni nie mieli pojęcia o tym jak funkcjonuje organizm. **Historia medycyny** liczy sobie ponad 4000 lat, pełna jest dziwacznych wierzeń i wynikających stąd błędów medycznych. Zaburzenia psychiczne były tłumaczone zgodnie z wierzeniami religijnymi, a halucynacje uznawano za kontakt z innym światem. Do tej pory choroby psychiczne w niektórych krajach próbuje się "leczyć" przez wypędzanie złych duchów, odprawianie rytuałów takich jak egzorcyzmy.

Nie uwolniliśmy się jeszcze całkowicie od błędów przeszłości. **Arystoteles uznawał serce** za siedlisko uczuć i rozumu, inteligentne centrum kierujące wszystkimi innymi organami. W Starym Testamencie serce również uważano za siedlisko uczuć i pragnień. W efekcie do tej pory w **Katechizmie Kościoła Katolickiego** wskazuje się na "serce człowieka jako źródło, z którego wypływają uczucia" jak i źródło wszelkiego grzechu. To oczywiście można traktować jako metaforę, ale czytelnik musi sam się domyślać, co w całym tekście Katechizmu metaforą jest, a co nie jest. W 1628 roku **William Harvey** (1578-1657, Anglia) prowadząc badania na zwierzętach opisał działanie serca i krążenie krwi w zamkniętym układzie. W 1651 roku opublikował traktat o nieistnieniu samoródtwa, głosząc zasadę, że wszystko co żywe rozwija się z jaja. To oczywiście spowodowało spekulacje (Nicolas Malebranche był chyba pierwszym filozofem piszącym na ten temat), że w jaju już jest mały organizm, w pełni uformowany homunculus ("mały człowiek"), w którym jest kolejne jajo i tak w nieskończoność. Doktrynę tą nazwano **preformacjonizmem**. Ponieważ antykoncepcja i masturbacja niszczyła materiał, w którym spodziewano się małych ludzików **należało je zwalczać**. Na ile negatywne nastawienie do antykoncepcji w dzisiejszych czasach jest echem tych wierzeń?

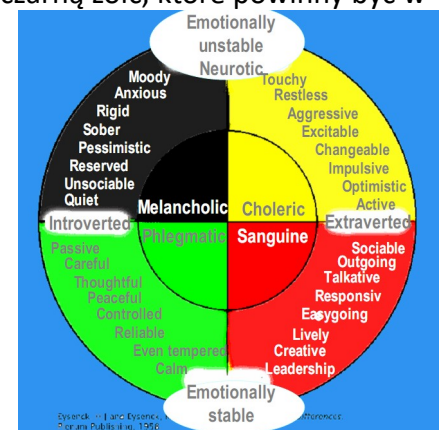
**Hipokrates** (-460,-370) był prekursorem medycyny, opierał się na empirycznych obserwacjach, do dzisiaj mamy "przysięgę Hipokratesa".

Jego teoria chorób oparta była o "humory", czyli 4 płyny: krew, śluz, żółć i czarna żółć, które powinny być w równowadze. Temperament wynika z nadmiaru jednego z tych płynów:

- sangwiczny - krwi,
- flegmatyczny - śluzu,
- choleryczny - żółci,
- melancholiczny - czarnej żółci.

Różne magiczne teorie świadomości nadal odwołują się do idei czterech żywiołów, albo kosmicznych energii, czakr i innych niewykrywalnych żadnymi metodami pomiarowymi tworów starożytnej wyobraźni.

Dopiero **Andreas Vesalius** (1514-1564, Flandria), twórca nowożytnej anatomii, ustalił znaczenie mózgu. Kościół w tym czasie uznawał, że liczba żeber u mężczyzny powinna być z jednej strony mniejsza o jedno, bo z tego żebra powstała Ewa ... Weryfikacja empiryczna takich twierdzeń nie była praktykowana (A.D. White, 1895).



**Współczesne 4 płyny to najważniejsze neurotransmitery**, ale to bardzo odległa analogia, rzeczywistość jest znacznie bardziej skomplikowana:

- acetylocholina (pamięć, uczenie się),
- dopamina (nagroda, motywacja, ruch),
- noradrenalina (uwaga, poznawanie),
- serotonina (nastrój, emocje).

**Postęp dokonywał się stopniowo dzięki empirycznym obserwacjom**, chociaż nie nadawano im przez długi okres czasu większego znaczenia niż naiwnym wyobrażeniom. Cytując jednego ze znanych komików - komu wierzyć? Mnie czy własnym oczom? Często wierzono starożytnym autorytetom.

**Analiza urazów** powypadkowych pozwoliła na obserwacje już za czasów starożytnego Egiptu: papirus sprzed 3500 lat wymienia 28 uszkodzeń czaszki. Już wtedy dokonywano trepanacji czaszki by wyciąć guzy.



Już w -3 wieku [Herofilus](#), lekarz z Aleksandrii twierdził, że mózg odpowiedzialny jest za intelekt.

[Galen z Pergamonu](#) (ok. 129–216 r) zauważył, że uszkodzenia rdzenia powodują paraliż, uszkodzenia mózdzku problemy z koordynacją ruchów, a "płyn umysłu" znajduje się w komorach mózgu.

Systematyczne badania skutków urazów rozpoczęło się dopiero w XIX wieku. W 1838 Theodore Schwann i Matthias Schleiden opisali komórkową budowę organizmu, ale do końca 19 wieku nikt nie był pewien, czy dotyczy to również mózgu. Spór pomiędzy zwolennikami ciągłej, siatkowatej organizacji materii mózgu i komórkowej, składającej się z neuronów, rozstrzygnął Santiago Ramon y Cajal (1852–1934), stosując metodę barwienia komórek opracowaną przez Golgiego.

Badania nad mózgiem mają więc dość krótką historię, więc nic dziwnego, że filozofia zainteresowała się nimi dopiero w końcowych dekadach XX wieku.



## B28.2. Potrzeba neurofilozofii



**Kognitywizm uznał umysł za system przetwarzający informację.** Wiele ważnych aspektów umysłu można zrozumieć z tej perspektywy, ale czy wszystkie?

Sztuczny umysł wymaga ogromnych bazy wiedzy, sprawnych metod reprezentacji, organizacji i korzystania z wiedzy. [Supersystem ekspertowy CYC](#), oparty na metodach reprezentacji wiedzy rozwiniętych w ramach sztucznej inteligencji, kosztował prawie 1000 osobo-lat pracy, ma miliony reguł i powinien umożliwić realizację zdrowego rozsądku, ale nadal go nie widać w naszych komputerach ... czego brakuje w takich systemach? Postęp w grach strategicznych i planszowych nie dokonał się dzięki dodawaniu coraz większej liczby reguł, tylko metodom automatycznego uczenia się w sieciach neuronowych, które działają "intuicyjnie", ucząc się podobnie jak mózgi dzięki neuroplastyczności, zmianie swojej wewnętrznej struktury.

Może mózgi są zbyt złożone by dało się stworzyć funkcjonalnie równoważne programy? **Ile neuronów ma mózg, a ile elementów ma komputer?**

Ludzkie mózgi mają pomiędzy 10 a 100 mld neuronów (ok.  $10^{11}$ ), ale współczesne mikroprocesory mają nawet miliard tranzystorów, zaś superkomputery mają nawet dziesiątki tysięcy takich procesorów, czyli



razem dziesiątki bilionów elementów ( $10^{13}$  do  $10^{14}$ ), a więc ponad 100 razy więcej niż neuronów w mózgu! Cerebras's Wafer Scale Engine (WSE) to wielki obwód scalony zawierający 1200 mld tranzystorów ( $1.2 \times 10^{12}$ ), 400.000 rdzeni procesorów, 18 GB szybkiej pamięci SRAM, oraz przesyłający 100 mln mld ( $10^{16}$ ) bitów na sekundę (dostępny od końca 2019 roku).

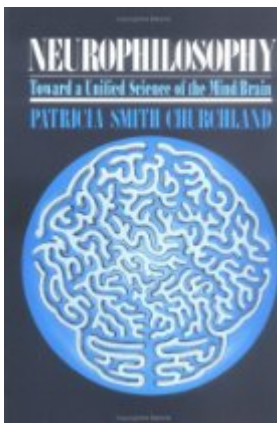
W dodatku neurony działają powoli, wysyłając zwykle poniżej 100 impulsów na sekundę (100 Hz). Nawet jeśli skupimy się na synapsach i uznamy, że każda ze 100.000 mld =  $10^{14}$  synaps działa jak bramka tranzystora, to będzie to tylko  $10^{14} \times 10^2 = 10^{16}$  elementarnych operacji (zmian w systemie) na sekundę, czyli tyle, co w jednym chipie WSE.

Superkomputery przekroczyły  $10^{16}$  operacji numerycznych na sekundę na wielocyfrowych liczbach, ale jeśli ocenimy teoretyczną liczbę elementarnych procesów dla superkomputera z zegarem 2 GHz to mamy  $5 \times 10^{12} \times 2 \times 10^9 = 10^{22}$  operacji, czyli **ponad milion razy więcej** ...

**Czemu więc superkomputery nie przewyższają jeszcze we wszystkim naszych mózgów** (choć coraz więcej funkcji realizują sprawniej)?

Zmiany strukturalne w mózgu (np. udar, lub choroba Alzheimera) powodują zanik nawet podstawowych funkcji, więc sama wielka liczba elementów nie wystarczy, by system działał w inteligentny sposób. Konieczna jest odpowiednia struktura, organizacja całości. Czy możemy taką mózgową strukturę stworzyć? Jest tu wiele pytań (Gazzaniga, 2020).

Czy systemy komputerowe **symulują jedynie inteligentne działanie, czy naprawdę mają stany poznawcze**? Czy jest sens mówić, że nigdy nie będziemy mogli stworzyć sztucznego mózgu? Na jakiej podstawie można tak sądzić? Cierpliwości, dopiero od niedawna nauczyliśmy się tworzyć systemy neuromorficzne o złożoności zbliżającej się do sieci w mózgu. Postępy w tym obszarze są bardzo szybkie i niezaprzeczalne.



Funkcjonalizm pozwala na pewno na użyteczny opis działania umysłu, ale mózgi nie przypominają komputerów, to bardzo odległa metafora dla umysłów, w opisie działania komputerów brakuje perspektywy wewnętrznej (por. Duch, 1994). Dodanie takiej perspektywy wymaga analizy treści psychicznych, a tym zajmowała się w filozofii [fenomenologia](#).

**Nie wystarczy zastosowanie funkcjonalizmu minimalistycznego**, ale też nie każdy szczegół budowy mózgu musi być istotny dla powstania umysłu. Budowa mózgu jest istotna dla zrozumienia większości funkcji umysłu, tłumaczy nie tylko nasze zdolności percepcji, ale i zdolności poznawcze: pamięć, skupianie uwagi, wyobraźnię. Budowa organizmu, w tym struktura mózgu, jest wynikiem procesów ewolucyjnych, które wytworzyły organy zdolne do działania w określonym ekosystemie. Zobaczmy w późniejszych rozdziałach, jakie zasady pozwoliły na powstanie tak złożonego narządu.

W 1986 roku [Patrycja Churchland](#) wydała książkę "Neurophilosophy: Toward a Unified Science of the Mind-Brain", rozpoczynając nowy nurt w filozofii umysłu. Podstawy neurofilozofii oparte są na całkiem innej tradycji filozoficznej niż kognitywizm.



### B28.3. [Fenomenologia](#)



[Franz Brentano](#) (1836-1916) był prekursorem fenomenologii. W książce "Psychologia z empirycznego punktu widzenia" (1874) połączył introspekcję z empiryzmem.

Źródłem wszelkiej wiedzy, w tym i psychologicznej, jest doświadczenie, w tym również doświadczenie



wewnętrzne. Zjawiska psychiczne (treści świadomości) obejmują (Brentano sklasyfikował to zgodnie z Kartezjuszem):

- przedstawienia, czyli akty działania: wrażenia zmysłowe, odczuwane stany fizyczne, widziane barwy czy słyszane dźwięki, oraz wyobrażenia;
- sądy, pozwalające oceniać przedstawienia jako prawdziwe lub fałszywe;
- emocje, czyli "akty miłości i niechęci" lub "akty zainteresowania".

Z grubsza tymi zjawiskami zajmuje się epistemologia z estetyką (poznanie i piękno), logika (prawda) i etyka (dobro).

Akty psychiczne są intencjonalne, skierowane ku czemuś (od "intendere", "kierować"). Charakteryzują się "odniesieniem do pewnej treści, skierowaniem na obiekt". Akt świadomości wykracza poza umysł, ku światu fizycznemu i tam trzeba szukać jego sensu.

Przykład: jeśli chcemy zrozumieć sens zdarzeń na ekranie analiza stanu procesora obrazu odtwarzającego sygnał z kamery, w pełni opisanego na poziomie fizycznym za pomocą napięć i stanów jego wszystkich elementów niewiele nam powie. Więcej się dowiemy wydzielając w obrazie obiekty, które ten procesor śledzi, i łącząc je z interpretacją filmowanych zdarzeń.

Podobnie dowolnie szczegółowy obraz mózgu nie wystarczy by w pełni zrozumieć stany mentalne, bo nie jest to właściwy poziom opisu. Każdy wzorzec aktywacji mózgu trzeba powiązać z percepcją, interpretacją stanów wewnętrznych w odniesieniu do świata zewnętrznego, pamięci i indywidualnej historii.

**Filozofia scholastyczna** uznawała "intencję pierwszą" (intentio prima) za akt intelektu, ujmujący rzeczy bezpośrednio; można to określić jako pierwotne pojęcia odnoszące się do bytów fizycznych.

"Intencja druga" to refleksja nad aktem intelektu, pojęcia abstrakcyjne pozwalające na porządkowanie wiedzy zdobytej za pomocą pojęć pierwotnych.

Brentano podzielił wszystkie fenomeny (zjawiska, nie poznajemy bowiem rzeczy bezpośrednio) na transcendentne (fizyczne, poza umysłem) i immanentne (umysłowe) wobec aktów intelektu.

**"Świadomość pierwotna"** istnienia obiektów fizycznych jest rezultatem aktywnej percepcji, która w odróżnieniu od biernej recepcji można traktować jako "akt intelektu".

**Świadomość refleksyjna** pozwala na intencjonalne postrzeganie stanów wewnętrznych umysłu i tworzenie pojęć opisujących relacje tych stanów.



Prowadzi to do połączenia empiryzmu (zakładającego realizm, istnienie fenomenów transcendentnych) z aktywną rolą psychiki.

Poznanie wymaga zarówno przedstawienia, jak i sądzenia (wyobraźni i oceny). Samo przedstawienie nic nie twierdzi, sądy zaś mogą być błędne lub słuszne.

Zjawiska emocjonalne, czyli akty "miłości i niechęci", obejmują zarówno uczucia jak i akty woli, wynikają z wyboru lub odrzucenia wartości rzeczy. Piękno istnieje ze względu na doskonałość przedstawienia, prawda ze względu na możliwość sądu, a dobro to właściwy stosunek do miłości i niechęci, stosunek będący podstawą intuicji moralnych.

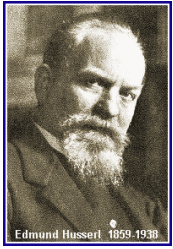
Doświadczenie wewnętrzne to fakty oczywiste, poczucie słuszności, miara prawdy i dobra.

Fenomenologia miała być "**psychognozją**", czyli psychologią opisową, opartą na obserwacjach reakcji. Czy jednak opis zewnętrzny wystarczy do zrozumienia procesów psychicznych, które powinny być podłożem tej filozofii? Czy zawsze doświadczenie wewnętrzne jest oczywiste? Czasami trudno jest sprecyzować, o co nam chodzi, przeanalizować swoje wrażenia i emocje.

Rozumienie intencjonalności nadal jest sporne: czy treść aktu intelektu to jest to samo co przedmiot tego aktu? Jak stosować pojęcie intencjonalności (ujmowanie rzeczy bezpośrednio) do odróżnienia tego co mentalne a co fizyczne? Iluzje mogą być całkiem realistyczne, zwłaszcza w przypadku chorób psychicznych i zaburzeń poznawczych.

Zagadnienia pokrewne do problemu intencjonalności nadal są dyskutowane. Jakie są relacje między semantyką i syntaktyką, czyli jak powstaje sens symboli w mózgu i w systemach formalnych, jak należy rozumieć reprezentacje mentalne? Co odróżnia zdarzenia mentalne, procesy umysłowe, od przedmiotów fizycznych lub stanów obliczeniowych procesorów komputera? Czy treść (sens) stanów mentalnych jest

cechą wewnętrzną systemu, czy też pochodzi od świata zewnętrznego? W jakim stopniu są to własności wewnętrzne, a w jakim są one rezultatem oddziaływania ze światem zewnętrznym? Czy można je zrozumieć w oderwaniu od większej całości?



**Edmund Husserl** (1859-1938) był uczniem Brentano.

Wszystkie twory umysłu muszą podlegać prawom psychologii. W "[Philosophie der Arithmetik](#)" (1891) przedstawił psychognozę matematyki, próbując połączyć matematykę, psychologię i filozofię.

Większy postęp w tej dziedzinie nastąpił dopiero w 2000 roku po opublikowaniu książki "[Where Mathematics Comes From](#)" przez lingwistę George'a Lakoffa, i psychologa Rafaela E. Núneza. Badania nad [reprezentacją pojęcia liczby](#) w mózgach ludzi i zwierząt stały się

możliwe dzięki metodom neuroobrazowania (Dehaene, 1997).

Od 1900 roku Husserl zajął się badaniem aktów i treści fenomenów świadomości, czyli **fenomenologią**. Prawa psychologii dotyczą aktów psychicznych, a nie obiektów idealnych. Psychologizm pociąga za sobą sceptycyzm i relatywizm (Hume: prawa to tylko wynik przyzwyczajień), a prawa logiki są wieczne i niezależne od psychiki.

Są więc 3 rodzaje przedmiotów: materialne, psychiczne i matematyczne. Potrzebna jest "nauka pierwsza", stwierdzająca to, co oczywiste, której szukał już Kartezjusz.

Fenomenologia miała badać samą istotę zjawisk, danych w aktach świadomości. Fenomenologiczna redukcja zjawiska do samej jego istoty wymaga, by odrzucić wszystko, co nie jest oczywiste.

Psychologia bada funkcje aktów psychicznych, a nie ich istotę. Co jest bezpośrednio dane świadomości?

Wrażenia, przedmioty idealne, rzeczy i ich stany. Intuicja jest więc najważniejszym źródłem poznania.

Poznanie to akt bierny, bezpośrednie ujęcie fenomenu, istoty zjawiska. Dzięki temu możliwe są sądy *a priori*, np. zrozumienie słowa, czy postrzeżenie barwy.

Husserl doszedł do wniosku, że istnieją [prawdy aprioryczne](#): nie można np. słyszeć zapachu lub widzieć dźwięku.

Całkiem się mylił: **wrażenia mogą się mieszać w synestezjach**, poznanie nigdy nie jest bierne, a intuicje często są całkiem mylne. Czy opieranie się na nieweryfikowalnych intuicjach prowadzi do jakiegokolwiek sensownej wiedzy? Jest to często podstawą eksperymentów myślowych w filozofii, prowadzących do dużego pomieszania (np. D. Dennett, Słodkie Sny 2005).

Z drugiej strony rola intuicji, rozumianej jako wynik nieświadomych procesów przetwarzania informacji, była przez długi czas niedoceniana. Mózg może się nauczyć prawidłowych reakcji działając intuicyjnie w złożonych sytuacjach, w których analiza logiczna jest zbyt trudna.

Dopiero od połowy lat 1990 [Gerd Gigerenzer](#), [Amos Tversky](#), [Daniel Khaneman](#), [Richard Thaler](#) i inni pokazali w niektórych sytuacjach wyższość intuicyjnych wyborów opartych na prostych, [nieuświadomionych heurystykach](#), nad racjonalnie przemyślanymi rozwiązaniami. Badania te okazały się tak istotne dla ekonomii behawioralnej, że przyznano za nie nagrodę Nobla (Khaneman dostał ją w 2002, Thaler w 2017).

**Prawa aprioryczne** pozwalają budować fenomenologię jako obiektywną filozofię. **Realizm** oznacza, że przedmioty istnieją realnie, a świadomy umysł działania intencjonalnie. Wynikiem fenomenologicznej redukcji zjawisk jest świadome wrażenie. W późniejszym okresie Husserl uznał świat za wytwór świadomości. Chociaż mylił się prawie we wszystkim fenomenologia zwróciła uwagę na istotne zagadnienia związane z intuicją.



**Max Scheler** (1874-1928) rozwinął fenomenologię wartości. Podjął próbę stworzenia obiektywnej, naturalnej etyki.

Umysł rozpoznaje *a priori* intuicyjnie, emocjonalnie, sprawiedliwość i piękno. Hierarchia wartości według Schellera, od najważniejszych do najmniej ważnych, wygląda następująco: religijne, duchowe, witalne, hedoniczne i utylitarne. Wyższe wartości nadają sens niższemu, np. przyjemności pieniądzu.

Scheler rozwinął też fenomenologię religii: istota boskości, świętości, aktów religijnych, jest też pojmowana bezpośrednio.

Wiele twierdzeń "oczywistych" dla fenomenologów okazało się fałszywych. Intuicje to wyniki wychowania w pewnej kulturze (antropologia). Sens słów wymaga odniesienia do kontekstów kulturowych. Nawet sposób patrzenia na świat jest uwarunkowany kulturowo, jak opisał to w książce [The Geography of Thought](#) psycholog społeczny Richard Nisbett.

Fenomenologia systematycznie odrzucającą konstrukcje umysłu na rzecz "bezpośredniego wglądu" to naiwna wiara, że czysto subiektywne postrzeganie daje prawdziwy obraz świata.



**Maurice Merleau-Ponty** (1907-1961).

Główne dzieło tego francuskiego filozofa to "Fenomenologia Percepcji" (fr. 1942, pol. 2001), w którym dokonał analizy roli ciała w percepcji. Przez długie lata był to bardzo niedoceniany filozof, dopiero teraz znalazł się w centrum uwagi.

Jego podejście do fenomenologii znacznie się różniło od poprzedników. Nie ma "czystych wrażeń", postrzeganie to złożona funkcja mózgu. Np. pojęcie przestrzeni zakłada możliwość doświadczania ruchu. Aktywne postrzeganie jest intencjonalne, istotna jest rola

oczekiwań pozwalających dostrzegać istotne elementy.

Dane zmysłowe są więc "pre-objektywne", wymagają wnikliwego zbadania. Percepcja zbudowana jest ze stanów świadomości, pamięci doświadczeń.

Podobnie jak buddyjscy filozofowie David Hume, opierając się na introspekcji, doszedł do wniosku, że **"... w mnóstwie wrażeń i przypomnień nie ma kogoś, kto widzi"**. Nie ma homunculusa w naszej głowie. Nie ma trwałych przedmiotów percepcji niezależnych od asocjacji z nimi związanych. Nie ma transcendentnego ego, tylko dzięki ciałom rozumiemy wzajemnie swoje przeżycia. Bez sfery cielesno-emocjonalnej nie ma świadomości.

Miał rację! Powoli weryfikujemy te twierdzenia.



**Martin Heidegger** (1889-1976) szukał filozofii stanowiącej warunek uprawiania wszelkiej nauki.

Byt, istnienie, to najbardziej fundamentalne pojęcia, opisane w najważniejszym dziele Heideggera "Bycie i Czas" (1927).

Istniejemy w świecie: nie ma do opisu tego stanu dobrego wyrażenia, dlatego trzeba było rozwinąć nowy język pojęć, Heidegger nazwał to Da-sein, "tu-bycie" lub "Jestestwo".

Jestestwo jest umiejscowione w sieci postaw wobec świata.

Świat postrzegany jest jako "poręczny", narzędzia są dopasowane do organizmu i rozszerzają możliwości naszego działania. Sednem bycia jest troska o to, czym rzeczy są lub co z nimi można zrobić. Idea [afordancji](#) jako możliwości działania rozwinęła się dopiero w latach 1970 w ramach ekologicznych teorii percepcji J. Gibsona.

Trzy tezy **metafizyki powszedniości** Heideggera:

1. Bycie w świecie implikuje jestestwo (Dasein) oraz świat, egzystencjały, składowe fenomenu bycia-w-świecie.
2. Relacja poznawcza podmiot-przedmiot ugruntowana jest przez troskę jestestwa w obliczu świata.
3. Podstawową formą troski jest "wyznawanie się Jestestwa", czyli wytwarzanie manipulatywno-przeładowych postaw wobec świata.

Świadomość powstaje z samego bycia-w-świecie, ma sens "dzięki reszcie naszych działań". Wyższe funkcje psychiczne są nabudowane na reakcjach sensomotorycznych organizmu, na "trosce i wyznawaniu się Jestestwa". Byt ma formę Czasu, bo Jestestwo aktualizuje możliwości.

Jak widać język Heideggera jest hermetyczny gdyż odnosi się do przedintelektualnych funkcji umysłu. Jednak można go zrozumieć i przełożyć na język współczesnej kognitywistyki.

Heidegger pod koniec życia przeczytał książkę D.T. Suzukiego o japońskiej filozofii Zen, która też jest skupiona na przedintelektualnym przeżywaniu istnienia. Skomentował ją: "... to jest właśnie to, co usiłowałem wyrazić we wszystkich moich pismach".

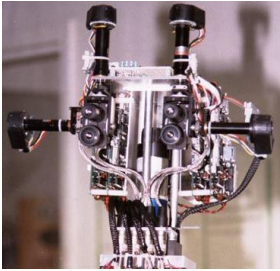
W latach 90. psychologia poznawcza dostrzegła wpływ tego podstawowego poziomu na czynności poznawcze. Troska wiążąca jestestwo z bytami obecnymi to "tematyzacja", "kategoryzacja" lub



"konceptualizacja".

**Troska i byty poręczne to rozwój zdolności manipulatywnych. Świat i jestestwo kształtują się wzajemnie.**

Metafizyka powszedniości Heideggera opisuje umysł działający w świecie, co pozwala na ugruntowanie sensu treści. Poznanie bytu, poręczności świata, wymaga sensomotorycznej integracji działań, zachodzi na poziomie pre-symbolicznym. Szczególnie wyraźnie widać to na przykładzie rozwoju niemowląt.



Paradygmat komputacyjny zajmował się wyższymi czynnościami poznawczymi, stawiając sobie inne cele, takie jak zrozumienie procesów rozumowania, rozwiązywanie problemów na podstawie werbalnych opisów. Nie sprawdzał się zupełnie w robotyce, gdzie najpierw trzeba zrozumieć percepcję, opisać działanie zmysłów (lub sensorów) w sposób symboliczny, a dopiero potem prowadzić wnioskowanie.

[Ross Ashby](#), psychiatra i pionier cybernetyki, w "Design for a Brain" (1960) zaczął podkreślać rolę środowiska, rozważając inteligencję behawioralną, związaną z zachowaniem w rzeczywistym świecie.

[Gerald Edelman](#) poszedł dalej w tym kierunku, badając ewolucję zachowania i zdolności poznawczych u robotów, które mają wbudowane tylko ogólne wartości (np. poszukiwanie większej liczby doznań). Dopiero środowisko pozwala im rozwinąć specyficzne formy zachowań.

[Rodney Brooks](#) na MIT w ramach projektu budowy [humanoidalnego robota Cog](#) (1994) próbował budować "jestestwo" robota, zgodnie z tezami Heideggera, ucząc robota zachowań sensomotorycznych, rozumienia poprzez działanie. Doprowadziło to do powstania robotyki afektywnej ([robot Kismet](#)) i robotyki rozwojowej ([developmental robotics](#)), wielkich projektów takich jak nadal rozwijany [iCube](#).



## B28.4. Umysł ucieleśniony



**Emergentyzm** to pogląd podkreślający, że nowe jakości wynikają z nowego poziomu złożoności organizacji, nie można ich wywieść z własności poszczególnych elementów. Nowe jakości substancji w oczywisty sposób wynikają z reakcji chemicznych. Silne oddziaływania między cząsteczkami prowadzą do powstania nowych struktur o zupełnie odmiennych własnościach. Jest to więc przeciwieństwo redukcjonizmu, który usiłuje zrozumieć własności całości rozkładając ją na elementarne części.

Koncepcja emergencji została rozwinięta już w 1925 r. przez [C.D. Broada](#) w książce "The Mind and its Place in Nature". Własności emergentnych nie można wydedukować z wiedzy o własnościach składników. Z mieszaniny gazów powstaje woda. Z wielokomórkowych organizmów tworzą się rośliny i zwierzęta. Z wielu komórek nerwowych tworzą się systemy sterujące ruchem zwierząt. Ewolucja organizmów zależy od środowiska, które na wiele sposobów wpływa na selekcję najbardziej przystosowanych osobników, a to zależy od ich budowy, poziomu komórkowego i sub-komórkowego, który w ten sposób również ewoluuje. W ten sposób wyższe i niższe poziomy organizacji nawzajem się warunkują.

Jak treści umysłu zależą od ciała i jak ciało zależy od umysłu? Rola procesów motorycznych w percepcji i działaniu, inicjacja ruchu i myśli, jest coraz silniej podkreślana, ale jest też oczywiste, że nasze mentalne przekonania wpływają na nasze zachowanie, dietę, ćwiczenia fizyczne. Mózg tworzy umysł, ale też umysł wpływa na mózg i całe ciało.

[Francisco Varela](#), [Evan Thompson](#) i [Eleanor Rosch](#), napisali razem książkę "The embodied mind: Cognitive Science and Human Experience". Podkreślali w niej nierozzerwalny związek pomiędzy kategoriami myślenia i odczuwania a ciałem.

Zapoczątkowało to tendencje do podkreślania wagi "ucieleśnienia" (embodiment) - ciało jest przedmiotem wrażeń i środowiskiem procesów kognitywnych. W książce podkreślono również związki pomiędzy naukami kognitywnymi a doświadczeniem ludzkim, szczególnie nawiązując do tradycji buddyjskich.



Lingwistyka również pokazuje, że korzenie wyrażen językowych leżą w czynnościach motorycznych i fizjologicznych.

Gerald Edelman pokazał, jak na skutek oddziaływań ze środowiskiem następuje selekcja połączeń neuronów w mózgu, w rezultacie pozostają precyzyjne połączenia i robot uczy się przydatnych zachowań. Procesy Darwinowskiej selekcji uważa za podstawę formowania się struktur umysłu. Fizyczna struktura i możliwości działania robota są odpowiedzialne za rozwój specyficznych form zachowania.

Rozpoznany obiekt przypisywany jest do jakiejś kategorii przez procesy poznawcze, korę mózgu, jak i nadawana jest mu określona wartość dla organizmu przez układ limbiczny. Sprzężenie pamięci wartości i kategorii prowadzi do działań intencjonalnych.



Praktycznym rezultatem tych rozważań była seria robotów "Darwin/Nomad" działających intencjonalnie. **Intencje są głęboko ukryte przed mechanizmami poznawczymi, gdyż pojemność pamięci roboczej (świadomości) jest mała.**

Wartości ogólne, np. "warto mieć bodźce zmysłowe", plus oddziaływanie ze środowiskiem, prowadzą do złożonych form zachowań.

Intencje działania są wynikiem procesów obliczeniowych. "Jestem głodny" to intencja organizmu do podniesienia poziomu cukru we krwi. Chęć działania jest wynikiem pobudzenia kory ruchowej przez procesy planowania zachodzące w płatach czołowych. Mamy tu więc kombinację przyczyn ewolucyjnych, instynktów opartych na biologii (w przypadku robota potrzeba poszukiwania pożywienia czy poszukiwanie bodźców) i środowiskowych.

Humberto Maturana stworzył pojęcie **systemu autopoietycznego, czyli systemu samoodtworzającego i samoorganizującego się, pozwalającego na osiągnięcie homeostazy**. Pojęcie to zostało wykorzystane do opisu systemów społecznych.

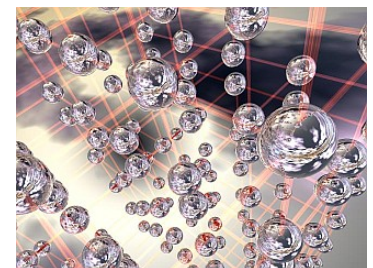
Idee te rozwinął Francisco Varela, Evan Thomson i inni badacze, podkreślając znaczenie samego systemu poznawczego, który nadaje sens odbieranym przez zmysły bodźcom, interpretując je w świetle już posiadanego modelu świata, nie mając bezpośredniego dostępu do rzeczywistości, a jedynie do efektów jej oddziaływania na system poznawczy.

W kontekście teorii systemów autopoietycznych można zdefiniować **poznawanie (kognicję) jako wszystkie działania systemów ożywionych, które wywołują w nich jakieś reakcje**. Świadomość to pewien mechanizm systemów biologicznych, który orientuje system autopoietyczny na samoobserwację i opis czy też zapamiętywanie swoich reakcji, tworząc w ten sposób obserwatora.

Koncepcja **"umysłu tylko w głowie"** trudna jest do utrzymania i była krytykowana przez licznych filozofów (H. Putnam, J.J. Gibson, D. Clark i D. Chalmers).

Hilary Putnam twierdził, że umysł jest nielokalny i nietemporalny, bo warunki prawdziwości myśli zależą od zdarzeń odległych, nie tylko mentalnych. To niepotrzebne pomieszanie filozofii umysłu i języka. Odległe zdarzenia wpływają na stany naszego umysłu. Wynika to nie tyle z "warunków prawdziwości", co z silnego sprzężenia osoby z jej otoczeniem.

**Clark i Chalmers (1998) zauważają: notesy, komputery, rozszerzają możliwości poznawcze, pamięć. Silne sprzężenie takich przedmiotów z procesami w mózgu pokazuje, że należy je traktować jako część umysłu.** Umysł może też działać w świecie wirtualnym, jest to aktywna eksternalizacja umysłu. Kiedy mówimy, że są to "moje myśli" pamiętajmy, że myśli pojawiają się w samoorganizującej się sieci, obejmującej całość oddziaływań z innymi ludźmi, są wynikiem interakcji ze światem, tego co widzimy, czytamy, o czym rozmawiamy, jakie media oglądamy. Ma to na nas większy wpływ niż przypuszczamy (John Bargh, Before You Know It: The Unconscious Reasons We Do What We Do, 2018, [Talks at Google](#)).



Umysł postrzegany jako **nielokalny, a więc zależny od stanów całego środowiska**, jest nadal przede wszystkim funkcją mózgu, ale stany dynamiczne mózgu są częścią stanów dynamicznych większego systemu i nie można ich zrozumieć badając tylko procesy w samym mózgu.

Gdzie jest granica mojego nielokalnego umysłu? To cały wszechświat! Silne oddziaływania nie pozwalają rozłożyć całości na niezależne części. Nawet zdarzenia odległe w przestrzeni i czasie wpływają na mój umysł, tak jak wybuchu supernowej miliardy lat temu, który stworzył węgiel z którego zbudowane jest moje ciało i całkowicie zmienił środowisko, w którym obecnie żyjemy. Wybuchu wulkanu zmienia moje plany, nie latają samoloty, na świecie robi się duże zamieszanie. Ludzie, z którymi się spotykam, zmieniają mój sposób myślenia, wpływają na moją przyszłość.

[Sutra Awatamsaka](#), powstała w Indiach ok. 200 roku, zawiera metaforę [sieci Indry](#), sieci peret odbijających się wzajemnie. Częścią buddyjskiej metafizyki jest **doktryna [współzależnego powstawania](#), podkreślająca powiązanie przyczyn i skutków wszystkich zdarzeń**.

Albert Einstein (list do rabina z 1950 roku, cyt. za Eves 1977): „Człowiek doświadcza siebie samego, swoich myśli i uczuć, jako czegoś odrębnego od reszty – jest to rodzaj złudzenia optycznego świadomości. To złudzenie jest dla nas rodzajem więzienia, ograniczając nas osobistych pragnień i związków uczuciowych z kilkoma najbliższymi osobami. Naszym zadaniem jest wyzwolić się z tego więzienia, poszerzając obszar naszego zrozumienia i współczucia, aż ogarnie on wszystkie żyjące istoty i całą naturę w jej pięknie. Nikt nie jest w stanie tego w pełni osiągnąć, ale wysiłek w tym kierunku jest sam w sobie częścią wyzwolenia i podstawą wewnętrznego bezpieczeństwa.”

### Semantyka i syntaktyka

[Stevan Harnad](#) zastanawiał się, jak w systemie formalnym powstać może semantyczna interpretacja symboli, semantyka wewnętrzna, a nie pasożytująca na projekcjach znaczeń dokonywanych przez człowieka? Jak manipulacja symbolami na podstawie analizy kształtów mogłaby odnosić się do czegokolwiek innego niż innych, pozbawionych wewnętrznego znaczenia symboli?

Reprezentacje symboliczne muszą być ugruntowane w niesymbolicznych reprezentacjach dwóch rodzajów:

- "Reprezentacjach ikonicznych": są to sensoryczne projekcje zdarzeń i rzeczy.
- "Reprezentacjach kategoryalnych": są to wrodzone i wyuczone detektory niezmienniczych cech projekcji sensorycznych.

[Modelowanie koneksjonistyczne](#), czyli rozproszony model kodowania informacji w postaci aktywacji sieci elementów, traktuje symbole jako elementarne reprezentacje, konfiguracje pobudzeń całych grup neuronów stanowiących węzły sieci.

W istocie w mózgu obserwujemy powtarzające się quasi-stabilne stany aktywacji, przez krótkie momenty, które można uznać za reprezentacje stanów mentalnych. Pojęcie reprezentacji mentalnych jest jednak jedynie przybliżeniem do tego, co dzieje się w takich sieciach. Zmiany stanu pobudzenia neuronów to neurodynamika. Są w niej powtarzalne wzorce pobudzeń, niezmienniki nauczone dzięki postrzeganiu stabilnych cech obiektów. Dzięki temu powstają uproszczone, ikoniczne reprezentacje obiektów w świecie. Myślenie, manipulacja mentalnymi symbolami, odnosi się do relacji pomiędzy takimi ikonicznymi reprezentacjami.

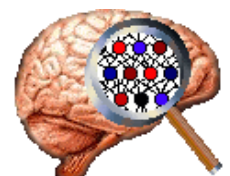


## B28.5. Neurofilozofia



Część tradycji filozoficznej jest nie tyle błędna, co o prostu nie przystaje do tego, co obecnie wiemy o rzeczywistości. To powoduje, że filozoficzna krytyka umysłu bywa bardzo dziwna.

R.H. Popkin i A.S. Stroll (Filozofia, Zysk i Ska 1994), w książce, która jest "uznana jako



jedno z najprostszych wprowadzeń do filozofii", napisali, że w mózgu nie widać obrazków, więc nasze wrażenia nie dadzą się sprowadzić do stanów mózgu. To równie dobry argument jak: "W komputerze nie widać liczb, więc komputer nie może liczyć". Nie widać też w nim pionków szachowych, więc komputer nie może grać w szachy. Wyobrażenia, jakie się nam w dzieciństwie zagnieździły w mózgu, nie jest łatwo porzucić.

Co jest potrzebne do zrozumienia percepcji i podejmowania decyzji, działania, w tym wewnętrznego, czyli pojawiania się myśli? Zdolność do dyskryminacji i kategoryzacji postrzeżeń i powiązania kategorii z działaniami, a nie szczegółowa reprezentacja. Widzimy szkice, pamiętamy zarysy, a mózg dopełnia resztę.

**Teoria identyczności typów** (Place 1956; Feigl 1958; Smart 1959) ma wyjaśnić związek między stanami mentalnymi a fizycznymi.

**E. G. Boring** w książce "The Physical Dimensions of Consciousness" (1933) napisał: "Dla tego autora doskonała korelacja oznacza identyczność. Dwa zdarzenia, które zawsze występują jednocześnie, w tym samym czasie i miejscu, nie są dwoma, lecz jednym zdarzeniem".

Błyskawica jest wyładowaniem elektrycznym: to stwierdzenie identyczności pozornie odmiennych procesów. Grzmot też wydaje się identyczny błyskawicą, ale z daleka widać, że nie ma koincydencji czasowej, bo prędkość światła jest znacznie większa niż prędkość dźwięku.

Biolog **Karl Vogt** napisał: mózg produkuje myśli tak jak wątroba żółć. Można stąd wnioskować o identyczności stanów fizycznych mózgu i stanów psychicznych w postaci myśli.

**J.C.C. Smart**: 'Widzę pomarańczowo-żółty powidok' oznacza coś takiego: "Coś się dzieje w mózgu podobnego do tego, co działo się gdy miałem otwarte oczy, byłem przytomny i patrzyłem na dobrze oświetloną pomarańcz przed sobą".

**Wniosek**: określony typ pobudzeń struktur mózgu  $\Leftrightarrow$  określony typ wrażenia.

Trzeba szukać psychoneuralnych identyczności, np. wrażenie bólu = pobudzenie C-włókien nerwowych. Ważny jest typ (token) danego stanu mózgu, który wynika z podobieństwa do utrwalonych wcześniej procesów. Typ stanu mózgu może być fizycznie realizowany na bardzo wiele sposobów (wieloraka realizowalność, multiple realizability).

**Krytyka**: nie można twierdzić, że pobudzenia nerwów są bólem; do przeżycia bólu potrzebny jest cały mózg i **procesy** interpretujące dochodzące sygnały, bez nich wrażenie bólu nie powstanie. Wrażenie bólu może natomiast powstać bez pobudzenia nerwów czuciowych, przy samym pobudzeniu neuronów w mózgu. Chociaż niektóre twierdzenia teorii identyczności były naiwne, wynikały z braku zrozumienia procesów zachodzących w mózgu, to ogólna idea znajduje coraz większe potwierdzenie: **specyficzne stany mózgu interpretowane są wewnętrznie jako specyficzne wrażenia, a podobieństwo do przeszłych stanów decyduje o tej interpretacji** (jak w przykładzie Smarta).



**Donald Hebb** napisał bardzo wpływową książkę "Organization of Behavior" (1949), był prekursorem neuronowego podejścia do wyjaśnienia percepcji, uczenia, pamięci, zaburzeń afektywnych.

**David Hubel** i **Torsten Wiesel** (1962) odkryli pola recepcyjne w układzie wzrokowym. Neurony specjalizują się w widzeniu kolorów, krawędzi, wykrywania ruchu, rozpoznawania twarzy.

**Barbara Von Eckardt-Klein** (1975) opracowała szczegółową teorię identyczności dla dotyku, rozważając neuronalne reprezentacje i fenomenologię wrażenia dotyku.

Funkcjonalizm głosił, że funkcje umysłu można zrealizować na wiele sposobów, a działanie mózgu to tylko "inżynierskie detale", mało istotne dla zrozumienia umysłu. Logika, reprezentacje symboliczne, analogie umysł-programy komputerowe powinny wystarczyć do zrozumienia działania umysłu. Jednak nadzieje na powstanie sztucznej inteligencji w oparciu o logikę i reprezentacje symboliczne znacznie zmalały po zakończeniu w 1994 roku programu komputerów **5 generacji**. Pomimo wydania miliarda dolarów nie osiągnięto zbyt wiele.

W tym czasie nastąpił burzliwy rozwój neuronauk, odkryto neurobiologiczne mechanizmy skojarzeń, pamięci i uczenia (Kandel 1976). Zrozumiano, że funkcjonalizm minimalistyczny może nie wystarczyć; jak dokładnie musimy rozumieć szczegóły działania organizmów biologicznych, by zrealizować takie same



funkcje w systemach sztucznych? Choć wieloraka realizowalność jest możliwa (zwierzęta mają bardzo różne zmysły) to biologia narzuca istotne ograniczenia na działanie systemów zdolnych do percepcji czy tworzenia skojarzeń.

Zrozumienie jak widzą zwierzęta wymaga zrozumienia nie tylko ogólnej funkcji, ale też sposobu realizacji procesów zachodzących w mózgach tych zwierząt, sposobu analizy danych sensorycznych.



[Patrycja Churchland](#) napisała książki wydane przez MIT Press: "Neurophilosophy: Toward a Unified Science of the Mind-Brain" (1986); "Neurophilosophy and Alzheimer's Disease" (1992); "The Mind-Brain Continuum" (1996); "Brain-Wise: Studies in Neurophilosophy" (2002).

Należy do nielicznej grupy filozofów zainteresowanych neurobiologią, współpracowała z [Francisem Crickiem](#). Jej mąż, [Paul Churchland](#), jest również

znany filozofem, eliminatywnym materialistą.

**Materializm redukcyjny** głosi, że wrażenia, pragnienia, myśli to typy pobudzeń mózgu.

**Materializm eliminatywny** głosi, że psychologia potoczna mówi o stanach umysłowych po prostu w błędny sposób. Wiele klas stanów mentalnych naprawdę jest tylko iluzją.

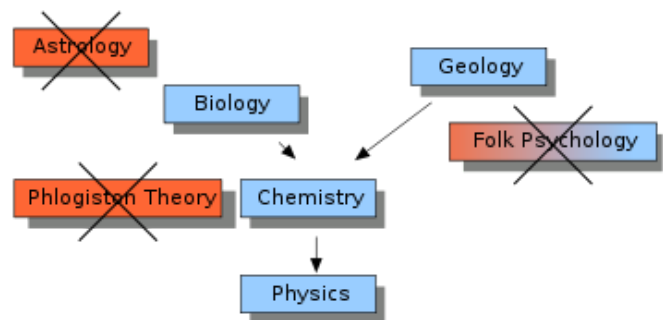
Eliminacja pojęć "siły vitalnej", pneумы, flogistonu czy eteru okazała się konieczna w nauce, te pojęcia nie przydały się do wyjaśnienia zjawisk fizycznych. Możliwa jest redukcja pojęć do całkiem innych, np. temperatura to średnia energia kinetyczna.

Wiemy, że wewnętrzne intuicje nie są źródłem wiedzy pewnej!

Rozwój teorii może doprowadzić do zmiany sensu lub nawet całkowitej eliminacji wielu pojęć "potocznej psychologii".

1. Niektóre pojęcia są mgliste np. "nastawienie, intencja, przekonanie" .
2. Inne są źle sprecyzowane, np. "śpi" to kilka różnych stanów umysłu.
3. "Pamięć, uwaga, rozumowanie" to pojęcia, które zmieniają znaczenie w wyniku eksperymentów i są obecnie rozumiane w różny sposób w zależności od kontekstu.
4. "Życzenie, pragnienie" czy "przekonanie" to kwestia dyspozycji do działania.
5. Strach jest obecnie rozumiany jako kilka różnych zjawisk: świadome uczucie strachu, reakcje behawioralne (ucieczka, zamarcie), reakcje fizjologiczne - oddech, krążenie, hormonalne, motywacja do działania (np. unikania pewnych bodźców), ocena kognitywna swojego stanu.

W efekcie potoczne pojęcia, odnoszące się do naszych stanów poznawczych i emocji są bardzo nieprecyzyjne, obejmują wiele różnych aspektów, które nie zawsze są ze sobą powiązane, bada się je za pomocą całkiem różnych eksperymentów. Nie próbuje się badać świadomości czy strachu, tylko konkretne aspekty zjawisk.



O własnych stanach umysłu tworzymy teorie i snujemy przypuszczenia równie dobre (lub równie złe) jak psychoterapeuci, którzy pomagają ludziom "zrozumieć siebie". Nasze rozumienie siebie jest pewną teorią: czy jest to dobra teoria, czy daje się ją powiązać z ugruntowaną w systematycznych obserwacjach teorią naukową?

Czy wychowanie może drastycznie zmienić psychologię potoczną? Do pewnego stopnia na pewno tak, ale na razie nikt nie próbował uczyć dzieci dokładnej analizy swoich stanów wewnętrznych.

Czy możemy wyeliminować pojęcia psychologii potocznej? Niektóre zapewne tak, ale wiele jest po prostu coraz bardziej uściślanych w powiązaniu z badaniami procesów w mózgu. Odróżniamy obecnie wiele rodzajów pamięci, uwagi, wyobraźni. Nie tyle więc zachodzi eliminacja co uściślanie nowych pojęć tak, by odpowiadały realnym procesom fizycznym zachodzącym w mózgach. Czy to się może udać? Ponieważ są to procesy silnie ze sobą powiązane, zmieniające się w ciągły sposób, dokładny opis werbalny nie będzie zapewne możliwy, ale pewnie możemy go jeszcze w znaczący sposób ulepszyć.



Co zawiera książka "Neurofilozofia" Patrycji Churchland?

Dokonuje się w niej przeglądu neuronauk i wiąże filozofię umysłu z faktami empirycznymi uważając, że filozofia umysłu i neuronauki "ko-ewoluują" nawzajem na siebie wpływając. Popiera eliminatywistów, krytykuje funkcjonalizm i argumenty dualistyczne, dyskutuje zagadnienia wzajemnej redukcji teorii.

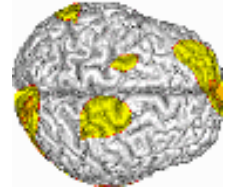
Mapy działania mózgu dają ogólne neurokorelaty stanów psychicznych.

Stany psychologiczne to pobudzenie konkretnych struktur mózgu.

Stany, które mogą przyjmować mózgi, są ściśle związane z ich konstrukcją.

Na poziomie cząsteczek neurochemicznych identyczna funkcja prowadzi do identycznej struktury.

Na poziomie sieci neuronalnych mózgu możliwa jest szybka reorganizacja, te same funkcje można realizować wykorzystując różne grupy neuronów i obszary mózgu.



Neurofilozofia stwierdza, że najbardziej płodną hipotezą jest: **umysł jest tym, co robi mózg** (dokładniej, częścią tego co robi mózg). Usuwa to problemy niefizycznej natury umysłu. Analiza działania mózgu i powstanie umysłów to kwestia empiryczna.

Redukcjonizm i eliminatywizm to dobre strategie badań.

Psychologiczne modele działania mózgu też są przydatne, np. rola hipotez wysokiego poziomu w tworzeniu się iluzji optycznych jest ważna.

Nauka rozwija się jednocześnie na wielu poziomach, psychologowie uwzględniają ograniczenia neurobiologiczne, neurobiolodzy korzystają z obserwacji psychologicznych, filozofowie umysłu próbują uściślić aparat pojęciowy.



### Krytyka neurofilozofii:

"Nie da się otrzymać świadomości z mięsa". Ból, wrażenia, uczucia nie są tylko pobudzeniami mózgu.

To naiwny punkt widzenia. Równie dobrze można zapytać: Czy życie nie może być tylko wynikiem organizacji molekuł opartych na związkach węgla? Czy siły magnetyczne mogą być wynikiem ruchu i spinu elektronów? Ludziom z dżungli pogląd, iż Ziemia jest okrągła, musi wydawać się niedorzeczny.

Błędy kategorii to kwestia teorii, w które wierzymy, brak wyobraźni lub wiedzy nie jest dobrą podstawą krytyki.

Fakt, że takie pojęcia jak "świadomość" czy "miłość" są słabo zdefiniowane i używamy ich do opisu całej kategorii różnych przeżyć nie oznacza jeszcze niemożliwości opisu, raczej konieczność uściślenia języka opisu.

Przykład: nie można leczyć "choroby", albo "raka" (ani nawet "raka piersi"), bo to nadal cała klasa różnych chorób, trzeba leczyć specyficzną chorobę. Nie można więc wyjaśniać "świadomości", trzeba najpierw dokładniej sprecyzować, co jest do wyjaśnienia, jak wygląda eksperyment, którego wyniki chcemy wyjaśnić.



**Paul Churchland** (1989, 1995) stwierdził, że neurofilozofia zmieni również filozofię nauki, a nie tylko umysłu!

Jak dochodzimy do przekonań i budujemy teorie? Tradycyjnie filozofia opierała się na logice, np. [teorii prawdy](#) zapoczątkowanej przez [Alfreda Tarskiego](#).

Czy i w jakim stopniu posługujemy się logiką? To zależy od tego, o czym myślimy. W większości przypadków logika w naszym myśleniu jest mało istotna, a czasami jej wcale nie widać.

Zachowania sensomotoryczne uznać można za przepływ aktywacji pomiędzy grupami neuronów, a nie zbiory logicznych reguł, które można zapisać symbolicznie. Logika pozwala na zbyt prosty opis, by opisać takie zachowania. Zbierając pobudzenia neuronów w tablice (położenie, pobudzenie) lub w wektory opisujące pobudzenia możemy mówić o transformacjach wektorowych lub macierzowych, opisujących procesy w sieciach neuronowych, dlatego Churchland pisze o "modelu wektorowym", lub sieciowym pojęć. Jednakże w takim modelu pojęcia początkowo nie są precyzyjnie określone.

Teoria naukowa to zbiór stwierdzeń o różnych stopniach prawdziwości.

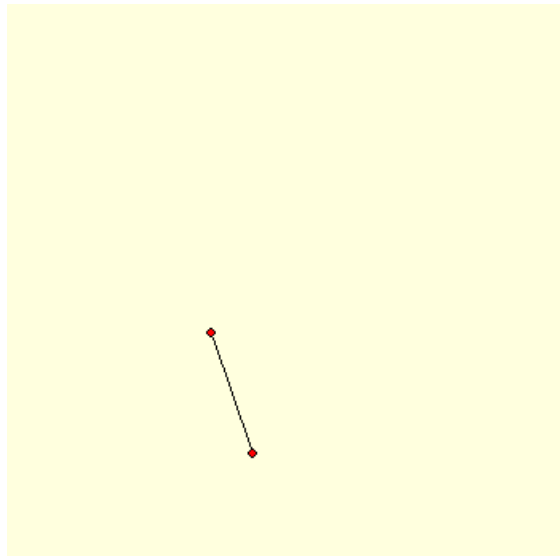
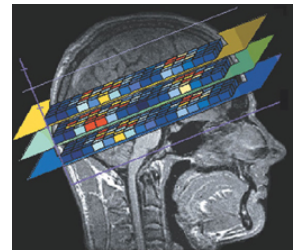
Symboliczne modele umysłu zastępowane są przez model opisujący aktywność grup neuronów w mózgu, czyli zbiór współczynników, model "sieciowy" lub "wektorowy".

Teoria jakiegos zjawiska, tworząca się stopniowo w naszych mózgach w trakcie nauki, badania i opisywania, jest wynikiem uczenia się przez sieci neuronowe w mózgu. W rezultacie powstaje zbiór stanów, wzajemnie

ze sobą skojarzonych, a niektóre z tych stanów (nie wszystkie) związane są z aktywacją fonologicznych reprezentacji pojęć językowych, interpretowanych jako symbole.

Model wewnętrzny jest rzadko w pełni werbalizowalny, na poziomie symbolicznym przekonania nie muszą więc sprowadzać się do zbioru reguł.

**Przekonania o świecie to teorie w postaci konfiguracji możliwych (potencjalnych) pobudzeń sieci neuronów**, z większością lokalnych konfiguracji można związać pewien symbol a przekonanie, że z obserwacji wynikają określone konkluzje,  $O \Rightarrow K$  jest wynikiem procesów skojarzeniowych



Jest to bardziej wierny opis rzeczywistości niż ten, który da się osiągnąć za pomocą werbalnych pojęć - neuronauki, podobnie jak fizyka, powinny się skupić na opisie rzeczywistych procesów odpowiedzialnych za stany mentalne.

Trudno to zaakceptować tradycyjnym filozofom, gdyż mają już zbyt wiele nawyków myślenia związanego z logiką i rozważaniami na poziomie wysokiej abstrakcji; historyczna spuścizna bardzo tu przeszkadza. Język logiki klasycznej jest dla nich naturalnym aparatem formułowania myśli, bo uczy się go najpierw, ale nie pozwala on opisać większości zjawisk psychologicznych, oferuje zbyt jednoznaczne rozróżnienia, nie da się za jego pomocą opisać procesów stopniowych, ciągłych.

Zastosowanie opisu sieciowego jest trudne, gdyż nie rozumiemy jeszcze szczegółów działania mózgu na tyle, by przybliżyć wszystkie aspekty jego działania za pomocą tego modelu - ale jest to jedyna droga, która prowadzi do prawdziwego postępu. Model wektorowy jest dobrym przybliżeniem tylko dla funkcji skojarzeniowych, ale trudniej go zastosować dla rozumienia języka czy myślenia (o ileż mniej jednak włożono w to wysiłku niż w rozwijanie logicznych teorii prawdy). Dopiero w XXI wieku udało się częściowo odkryć, jaki kształt mają pojęcia i wyobrażenia w mózgu.

[Modele sieci neuronowych](#) pozwalają na wyjaśnienie wielu zagadnień neuropsychologii.

Z filozoficznego punktu widzenia możliwy jest **emergentny dualizm**, głoszący pewną autonomię świata umysłu, w ramach którego można mówić o wpływie umysłu na mózg jak i odwrotnie, mózgu na umysł. Zwalczenie swoich nałogów, walka z wrodzonym lenistwem lub próba doskonalenia swojego charakteru, to walka umysłu z mózgiem.

Umysł potrzebuje substratu, jakim jest mózg, ale jest w pewnym stopniu autonomiczny, jest wynikiem genetycznych predyspozycji i społecznych uwarunkowań (Chicago Social Brain Network, 2011).

Neurofilozofia ma też coraz więcej do powiedzenia w zrozumieniu **biologicznych podstaw moralności** (dyskusji nad "naturalizacją moralności", jak to nazywają filozofowie). Patrycja Churchland (2013) w książce "Moralność mózgu" pokazuje, jak można powiązać takie pojęcia jak "sentyment moralny", używane przez filozofów, z pojęciami odnoszącymi się do wzajemnych relacji, np. przywiązaniem, troską, opiekuńczością. Takie zachowania wynikają z poznawczo-afektywnych mechanizmów, które w toku ewolucji wykształciły reakcje empatii, wstydu, poczucia winy, wynikające z rozbieżności pomiędzy zachowaniem społecznie

oczekiwanym i realnie obserwowanym. Obserwujemy je u wielu zwierząt tworzących hierarchiczne społeczności.

Tradycyjnie rozważania na temat moralności opierały się na wnioskowaniu logicznym, jednakże w większości sytuacji wymagających moralnych decyzji logika nie ma większego zastosowania. Polegamy wówczas bardziej na reakcjach emocjonalnych, które uwzględniają jednocześnie szereg czynników ograniczających dopuszczalne formy zachowania.

W terminologii uczenia maszynowego czy uczenia sieci neuronowych określa się to jako "constraint satisfaction", a więc próbę spełnienia jak największej liczby ograniczających warunków, wyboru maksymalnie dobrego rozwiązania, w sytuacji w której nie ma jednoznacznie najlepszej decyzji. Dotyczy to nie tylko dylematów moralnych typu "zderzyć się czołowo z autobusem ryzykując śmierć pasażerów czy zjechać na pobocze ryzykując przejechanie przechodnia", ale praktycznie każdej skomplikowanej sytuacji w której wybieramy np. produkty różniące się wieloma własnościami. Podążanie za intuicją, reakcją emocjonalną, prowadzi wówczas do większego zadowolenia z wyboru niż próba logicznej analizy, jak pokazał to [Roy Baumeister](#) w licznych publikacjach).



## B28.6. Enaktywizm i znaczenie ciała



Neurofilozofia skupiła się przede wszystkim na mózgu. Od "ja", zbioru procesów związanych z poczuciem tożsamości, a więc procesów postrzeganych w świadomy sposób, przeszliśmy do szerszego ja, związanego z wszystkimi czynnościami mózgu, które w większości są nieuświadomialne, ale mogą mieć wpływ na nasze decyzje.

Czy to wystarczy, by zrozumieć zachowanie ludzi i zwierząt? Trzeba jeszcze uwzględnić sprzężenie całego organizmu z otoczeniem, specyficznym ekosystemem, do którego jest dostosowany, wpływem organizacji społecznej na rozwój jednostki. Ewolucyjne podejście wyjaśnia, dlaczego organizmy są tak dziwnie zbudowane, dlaczego mają takie zmysły, jakie mają i dlaczego ich możliwości fizyczne i mentalne są ograniczone w specyficzny sposób.

Nurt **ucieleśnionego poznania** ([embodied cognition](#)), który Eleanor Rosch, Evan Thompson i Francisco J. Varela przedstawili w książce [The Embodied Mind: Cognitive Science and Human Experience](#) (MIT Press 1991), wskazywał na rolę ludzkiego doświadczenia i możliwości działania w formowaniu się mechanizmów poznawczych. Podstawą formowania się umysłu dziecka i stanów mentalnych są doznania cielesne i sensomotoryczne reakcje. Możliwości postrzegania i działania zależą od biologicznych, psychologicznych i społecznych uwarunkowań. Tak z pewnością jest, czy jednak da się określić jakie czynniki są najważniejsze i do jakiego stopnia mają wpływ na formowanie się umysłu?

**Enaktywizm** jest kontynuacją tego nurtu, skupia się nad badaniami wzajemnych oddziaływań organizm-środowisko, uwzględniając przystosowanie się organizmu do środowiska i jednocześnie wpływ organizmu na zmiany w środowisku. Słowo "enact" oznacza tu tworzenie, wzbudzenie, przywoływanie do istnienia sensu bycia w świecie, w którym percepcja łączy się z działaniem w kontekście danej sytuacji.

Chociaż enaktywiści skupiają się na dotyku i działaniu senso-motorycznym to percepcja wzrokowa czy słuchowa jest również silnie powiązana z kontrolą ruchu. W budowaniu obrazu świata i planowaniu działania współdziałają wszystkie zmysły, a mechanizmy orientacji pozwalają na maksymalizację informacji dostarczanej przez zmysły. Działając w świecie, wykorzystując możliwości ruchu całego ciała, ruchu głową, poruszania oczami, zdobywamy ważne informacje. Oczy wykonują bez przerwy ruchy sakadyczne, mózg ocenia gdzie należy patrzeć, by maksymalizować informację pozwalającą na lepsze zrozumienie tego, co widzimy.

Percepcja wymaga więc kierowanego percepcyjnie działania, a struktury poznawcze wyłaniają się z powtarzających się wzorców sensomotorycznych, które pozwalają na percepcyjne kierowanie działaniem. Zamiast patrzeć na świat i umysł jako dwa istniejące niezależnie byty podkreśla się tu aspekt dynamiczny, jaki jest współtworzenie świata i umysłu dzięki wzajemnym interakcjom.

W tym kierunku zmierza też Mark Johnson, który w książce "**Znaczenie ciała. Estetyka rozumienia ludzkiego**" skupił się nad pojęciem znaczenia, sensu wynikającego z posiadania ciała, przed-pojęciowych wzorców działania, potrzebą zrozumienia estetyki znaczenia w sztuce i muzyce. Jego wnioski można podsumować następująco:

- **Podział na umysł/ciało jest sztuczny**, te pojęcia są wynikiem powierzchownej analizy ludzkiego doświadczenia, w którym mamy ciąg interakcji między organizmem a środowiskiem. Z tych interakcji we wczesnym okresie rozwoju tworzy się w mózgu model "ja", przedmiotów fizycznych i innych ludzi.
- **Odcielesnione odczuwanie i świadomość nie jest możliwe** bo byłoby pozbawione wszelkich treści naszego doświadczenia i nie można by tego wiązać z jakimkolwiek poczuciem tożsamości.
- **Cieleśne doświadczenie** nabiera znaczenia w ciągu procesów interakcji pomiędzy organizmem i jego otoczeniem, bazując przede wszystkim na odczuciach czuciowo-ruchowych, oraz zdolnościach wyobrażeniowych tych procesów, **pozwalając zbudować na nich rozumienie pojęć metaforycznych i abstrakcyjnych**.
- **"Rozum" to proces interpretacji i skojarzeń doświadczeń cielesnych**, a więc mających głębokie znaczenie, oraz interpretacji wewnętrznych stanów mózgu prowadzących do rozważań na poziomie metafor.
- **Wyobraźnia** odpowiedzialna jest za zdolność tworzenia nowych znaczeń, nadawania sensu, za kreatywność.
- **Radykalna wolność nie jest możliwa**, "ja" nie jest abstrakcyjnym bytem kontrolującym nasze ciało i środowisko, tylko jednym z procesów wyabstrahowanym z oddziaływania całego organizmu ze środowiskiem.
- **Rozum i emocje** są ze sobą nierozdzielnie połączone, emocje leżą u podstaw wyobraźni, tworzenia pojęć, nadawania sensu doświadczeniom.
- **Ludzka duchowość jest również ucielesniona**, wynika ze zdolności transcendencji naszych doświadczeń.

Uzasadnienie powyższych tez jest bardzo ambitnym zadaniem, z którym Mark Johnson radzi sobie nieźle, odwołując się do licznych eksperymentów przeprowadzonych przez kognitywistów. Pewną słabością jego książki jest brak klarownego modelu procesów ilustrujących interakcję mózg-środowisko i samych procesów w mózgu, powiązania z neuronaukami, z użyciem dynamicznego matematycznego opisu rozwoju i sposobu poruszania się (Thelen i Smith 1994). Dlatego takie podejście może się wydawać oderwane od bardziej subtelnymi funkcji rozumu na poziomie naukowych i filozoficznych abstrakcji. Znaczna część książki Johnsona poświęcona jest metaforom i relacjom abstrakcyjnych pojęć językowych, którym ucielesnienie nadaje sens. Podsumowując, mamy cztery teorie rozwoju umysłu: w oparciu o reprezentacje wynikające z ucielesnienia i sensomotorycznych mechanizmów poznawczych, osadzenia w fizycznym, biologicznym, społecznym i kulturowym środowisku, sprzyjającym możliwościom interakcji, oraz nielokalnym stanom, wynikającym z silnego oddziaływania fizycznych artefaktów na stany umysłu. Teorie te określa się czasami jako 4E, od Embodied, Embedded, Enacted, Extended.

To całkowity odwrót od kognitywizmu, który zajmował się abstrakcyjnymi pojęciami, oderwanymi od świata fizycznego, zmysłowej percepcji. Z punktu widzenia konstrukcji programu grającego w szachy teorie 4E są mało istotne. Z punktu widzenia konstrukcji robota, który ma wykazać się inteligentnym działaniem w złożonym środowisku, lub zrozumienia jak powstały pojęcia, którymi się posługujemy, są to teorie bardzo ważne.

W drugiej części tego wykładu skupimy się na zagadnieniach związanych z świadomością i neurofilozoficznym podejściu do świadomości.



## B28.7. Współczesne dyskusje dotyczące świadomości

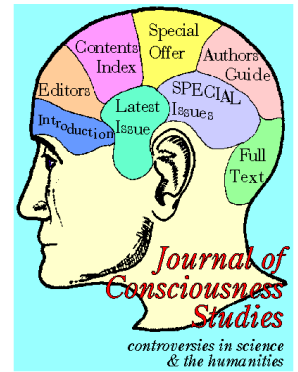




[Problem psychofizyczny](#), czyli stosunek ciała do umysłu dyskutowany w okresie przed rozwojem nauki, jak też [problem wolnej woli w filozofii](#) mają długą historię, którą warto poznać.

Badania nad świadomością znacznie przyspieszyły od założenia pisma [Journal of Consciousness Studies](#) (1994) i interdyscyplinarnej [konferencji w Tucson](#), Arizona (od 1996 r). Później powstały czasopisma "[Consciousness And Cognition](#)", "[Neuroscience Of Consciousness](#)", czy "[Anthropology Of Consciousness](#)".

Powstało wiele bardzo się od siebie różniących podejść do rozumienia świadomości, oraz kilka eksperymentów myślowych, które były i nadal są często dyskutowane. Tu nie omawiamy jeszcze mechanizmów neuronalnych świadomości. Przedyskutujemy to dopiero przy końcu drugiej części tego wykładu.



[Neomisterianie](#) to grupa filozofów, którzy sądzą, że problem umysłu przekracza ludzkie możliwości rozumienia. Najbardziej znanym przedstawicielem tej grupy jest [Colin McGinn](#), który zadaje naiwne pytania:

- Czy mózg w głowie może być źródłem świadomości nie-przestrzennej?
- Myśl lub wrażenie nie jest powyżej innej myśli czy wrażenia, dźwięk nie jest obok koloru, "gapimy się z opadniętą szczęką, nie potrafiąc zrozumieć".
- Czy umysł potrafi zrozumieć siebie? Już to omawialiśmy i nie są to realne trudności.

Funkcjonalizm, jak i neurofilozofia, pokazują, że pytania tego rodzaju nie mają sensu, nie są nawet błędne, są nieistotne. Co oznacza "zrozumieć" Czego dokładnie nie rozumiemy? Czy każdy może wszystko zrozumieć? Co musimy wiedzieć, by nabrać przekonania, że rozumiemy?

Pragmatyk John Dewey napisał: "Postęp intelektualny dokonuje się dzięki porzuceniu pytań razem z alternatywami, które proponują. Nie odpowiadamy na pytania, ale wykraczamy poza nie". Niestety brak jest dobrych metafor obrazujących aktualizację stanów mózgu i powiązanie ze stanami mentalnymi.

Umysł jest funkcją, strumieniem stanów mózgu i ciała, a nie substancją przestrzenną, w której są dźwięki i kolory. Obrazy i dźwięki na taśmie są śladami pamięci, stanów płytki CCD i membrany mikrofonu), ale nie mogą się same zmienić poprzez wzajemne interakcje lub wpływ środowiska, tak jak ludzka pamięć.

Stany elektroniki, która realizuje grę komputerową, nie mają związku z przestrzenią fizyczną, chociaż postacie gry poruszają się w wirtualnej rzeczywistości. Czy powinniśmy się zdziwić, że potwór z gry wie, że jestem na lewo od niego? Przecież "na lewo" się w moim komputerze nie zmieści ...

**Stany mózgu określone są przez aktywność neuronów, istotne są wzajemne relacje pomiędzy stanami mózgu, prowadzące do kolejnych stanów i do działania, a nie przez zdarzenia w przestrzeni fizycznej.**

Bez porządnego wykładu z neuropsychologii komputerowej nie da się tego dobrze zrozumieć, ale nie ma w tym większej tajemnicy niż w działaniu telefonu komórkowego (którego bez studiów fizyki też nie da się dobrze zrozumieć i McGinn go pewnie nie rozumie).

### [John Searle i tajemne moce neuronów.](#)

John Searle wymyślił eksperyment myślowy z "chińskim pokojem". Jest przekonany, że mózg wywołuje stany umysłu, ale nie są to stany, które można symulować obliczeniowo. Dlatego Searl sądzi, że neurony mają jakieś tajemne moce przyczynowe, które mogą wywoływać stany mentalne. Istnienie korelacji umysł-mózg nie oznacza identyczności odpowiednich stanów. Nie można tych stanów identyfikować ze stanami mózgu.

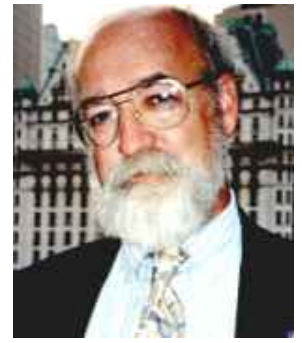
Świadomość można określić jedynie poprzez jej zawartość, nie można jej zredukować do pojęć prostszych. Czy jednak na pewno? Prąd elektryczny nie "jest wywołany" przez ruch ładunków, ale "jest" ruchem ładunków. Podobnie myśli i postrzeżenia nie są tylko wywołane zmianą aktywacji mózgu. Świadomość jest nie tylko wywołana przez stany neurofizjologiczne, lecz jest specyficznymi stanami mózgu, widzianymi niejako "od wewnątrz", bez nich nie ma świadomości. Zadaniem neuronauk jest określić dokładnie, jakie to stany. Robot wie, że widzi poszukiwany przedmiot lub osobę, na tej podstawie podejmuje działania, zwraca się do danej osoby i z nią rozmawia. Nie mamy jednak robotów, które zdolne są na tyle do autorefleksji by ich percepcja wewnętrzna była tak bogata jak świadomych ludzi czy niektórych zwierząt.



### Redukcjoniści: przykładem jest tu [Daniel Dennett](#).

Działający współbieżnie mózg symuluje wirtualną maszynę działającą szeregowo, czyli umysł. Człowiek może mówić sam do siebie, tak powstaje świadomy umysł. Umysł jest więc wynikiem działania odpowiedniego programu, realizującego maszynę wirtualną, tworzącą wirtualny świat wewnętrzny. Wewnętrzne doświadczenie jest w znacznej mierze złudzeniem (Dennett, 1992).

Komentarz: Status pojęcia "wirtualny" nie jest u Dennetta jasny, bo stoją za tym realne stany fizyczne i realne działania w fizycznym świecie. Dennett sądził np. że iluzje optyczne, takie jak "neonowe rozpytywanie się koloru" ([neon color spreading](#)) to złudzenie, a tymczasem to percepcja fizycznych stanów kory wzrokowej. Nie widzimy świata tylko reakcje swojego mózgu.



[Heterofenomenologia](#) to pomysł Dennetta, by analiza przeżyć oparta była na werbalnych raportach, które łączone są z informacjami o reakcjach, fizjologii, informacjami z neuroobrazowania. [Neurofenomenologia](#) bada korelacje pomiędzy subiektywnym opisem swoich przeżyć z perspektywy wewnętrznej i obiektywnym opisem za pomocą neuroobrazowania i innych metod neuronauk.

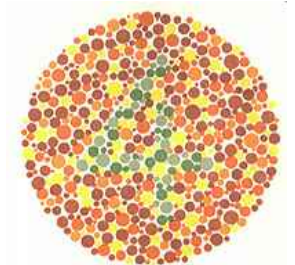
Dennett: Jesteśmy autorytetami tylko w sprawach tego, co wydaje się nam, że przeżywamy, a nie tego, co naprawdę się w naszych mózgach dzieje. To istotne rozróżnienie, ale nie do końca prawdziwe, jak pokazał w swoich książkach i pracach [Eric Schwitzgebel](#). Często bowiem nie potrafimy określić co przeżywamy i nasze interpretacje są tylko teoriami na temat siebie.

Argumenty Dennetta oparte są głównie na analizach pojęć językowych i sposobu opisywania wrażeń, ale brak języka nie oznacza braku umysłu czy świadomości. Wrażenia świadome nie redukują się do wypowiedzi, to całkiem inne stany mózgu.

Dennett: badając elementy komputera nie da się wiele powiedzieć o programie, neurobiologia nie doprowadzi więc do teorii umysłu. Tak by było, gdybyśmy nie potrafili odtwarzać funkcji za pomocą modeli komputerowych inspirowanych przez neurobiologię, co daje nam pewność, że rozumiemy jak działają mózgi.

### B28.7.1 [Mary daltonistka](#) .

[Frank Jackson](#) (1982) opisał często diskutowany eksperyment myślowy: Mary jest neurobiologiem i daltonistką; wie wszystko o procesach widzenia kolorów w mózgu. Jeśli pewnego dnia zacznie widzieć kolory, dowie się czegoś nowego. Wniosek: fakty fizyczne nie wyczerpują całej wiedzy o świecie! Fakty psychiczne dotyczące doznań są nieredukowalne do stanów fizycznych.



Dennett (1996) stwierdził: Mary wcale nie zdobyła nowej wiedzy, "wiedzieć wszystko" oznacza skorelować wszystkie wrażenia z widzianymi obrazami. Widząc niebieskiego banana Mary będzie równie zdziwiona, jak inni.

Zdziwi się, to prawda, ale jakże uboga prawda ... czy można np. na podstawie opisów utworzyć odpowiednie skojarzenia emocjonalne? Takie przeżycia nie dają się zwerbalizować, podobnie jak nie da się opisać werbalnie żadnego złożonego układu o zmieniających się w czasie aktywacjach. Zjawiska ciągłe nie można podzielić na fragmenty, którym nadamy nazwy-etykiety i opiszemy dane zjawisko.

Do pewnego stopnia można przeprowadzić rzeczywisty eksperyment tego rodzaju - istnieją [okulary Enchroma](#), dzięki którym niektórzy daltoniści mogą widzieć wybrane kolory. Co prawda takie okulary nie zrobią z nich specjalistów od widzenia kolorów, ale ich nowe odczucia mogą być dramatyczne.

Rozumienie intelektualne realizowane jest przez płaty skroniowe i płaty czołowe mózgu. Poczucie, że się **wie** dzięki doświadczeniu realizowane jest przez ośrodki podkorowe, w szczególności obszary zaangażowane w emocje i możliwości działania.

Mary wszystko rozumie intelektualnie, nie uczy się więc konceptualnie niczego nowego widząc kolory, ale jej mózg przyjmie całkiem nowe stany i wytworzy nowe doznania, nieredukowalne do poprzednio osiągalnych stanów aktywacji, a więc jest zdolna do przeżywania stanów wewnętrznych, które są całkiem odmienne niż były jej wcześniej dostępne. Nie ma tu żadnej tajemnicy.

## B28.7.2 Jak to jest być nietoperzem?

[Thomas Nagel](#) (1974) zadał pytanie: jak to jest być nietoperzem? Czy możemy to sobie wyobrazić?

Nie możemy zobaczyć świata z perspektywy wewnętrznej kogoś innego. Być nietoperzem "jest jakoś", ale być komputerem lub robotem jest nijak. Skąd mogłyby pojawić się wrażenia w mózgach robotów? Nie możemy zobaczyć świata z perspektywy wewnętrznej kogoś innego.

Zrozumienie intelektualne, jakie możemy osiągnąć w przypadku stanów mentalnych nietoperza, to korelacja stanów jego mózgu i obserwowanych zachowań. Przez analogię z naszymi doświadczeniami przypisujemy mu pragnienie, strach, przyjemność, ale potrafimy rozumieć inne istoty tylko przez analogię do swoich przeżyć.

Wiedza pełna nie jest jednak możliwa, bo struktura naszych mózgów jest dość odmienna. Wiem o sobie, bo mogę siebie obserwować od wewnątrz (dzięki wzajemnym aktywacjom różnych podsięci wiem co się w moim mózgu dzieje) i częściowo z zewnątrz, obserwując siebie i skutki swojego zachowania.

Wiedza o stanie umysłu innego człowieka możliwa jest dzięki empatii, 'rezonans' umysłów możliwy jest w przypadku w miarę podobnych mózgów. Rozumienie intencji innych osób angażuje specyficzne mechanizmy mózgowe, w tym neuronom lustrzanym. Reprezentacje wewnętrzne mogą być odmienne, mamy różne mózgi i te same funkcje realizowane mogą być w nieco różny sposób. Ważne są relacje pomiędzy różnymi stanami umysłu, bo to one decydują o interpretacji.

**Im bardziej nietypowy jest mózg, na skutek niedorozwoju, uszkodzeń, chorób, lub specyficznej kultury go kształtującej, tym trudniej zrozumieć, jakie stany przeżywa.** Członkowie grupy kibiców potrafią się wzajemnie dobrze zrozumieć, kobiety ze swoimi przyjaciółkami też, ale rozumienie osób spoza swojej grupy jest już dla wielu ludzi trudnym wyzwaniem.

Nagel twierdzi, że nie potrafimy sobie wyobrazić echolokacji, jednak [nie jest to prawdą](#), echolokacji można się w pewnym stopniu nauczyć. Niewidomi posługują się dźwiękiem stukając laską, mlaskając i nasłuchując echa, a więc uczą się echolokacji, chociaż oczywiście nie potrafią wydawać ultradźwięków a ich uszy i mózgi są dużo słabiej przystosowane do odbioru i analizy takich sygnałów. Niektórzy nawet wysyłając wysokie tony (klikli) [jeżdżą na rowerze!](#) Grający w ping-ponga nie mogą zbyt głośno tupać nogą by nie zakłócić słuchowej orientacji przeciwnika.

Liczne doświadczenia pokazały, że ludzie z opaskami na oczach dość szybko uczą się zwracać uwagę na strukturę echa wytwarzanych przez siebie dźwięków, używając echa do orientacji, rozpoznawania przeszkód, rodzaju pomieszczenia, rozróżniania kształtów obiektów spoza których wydobywa się dźwięk, a nawet odróżniania rodzaju powierzchni odbijających. Informacja tego typu jest zwykle maskowana przez bardziej przydatną informację wizualną, ale może mieć podświadomy wpływ na zachowanie.

Nawet jeśli korzystamy z takiej informacji, nie mamy związanych z tym wrażeń, trudno jest sprecyzować odczucia, ma się wrażenie zgadywania - do pewnego stopnia działamy więc jak filozoficzne zombi, albo Luke Skywalker "czujący moc".

Nie zawsze potrafimy interpretować skąd docierają do nas informacje.

Jeden z niewidomych uczestników eksperymentów (Supa i inn, 1944) upierał się, że odczuwa wibracje przeszkód rękami i policzkami, jednak zatkanie uszu nie pozwoliło mu wykrywać przeszkód. Do czasu tych eksperymentów sądzono, że prądy powietrzne i wrażliwość skóry twarzy może być rzeczywiście odpowiedzialna za takie zdolności. Jednakże wyobrażenie, które sobie skonstruował, było mylne.

Nawet jeśli nie potrafię biegać jak gepard, czy grać jak wirtuoz, to mogę sobie jednak wyobrazić bieganie, czy doskonalsze granie na instrumencie. To kwestia ekstrapolacji swoich przeżyć.

E. Schwitzgebel 2011; M. Wierchoń 2019). [podsumowanie informacji o echolokacji](#) u człowieka.

"Być kimś" to doznawać ciągu stanów mózgu, mieć subiektywny punkt widzenia, nieredukowalny do obiektywnego opisu. Dlaczego w ogóle jest to "jakoś"? To wynika z samego sposobu działania mózgu, w którym mamy do czynienia z realnie istniejącymi, fizycznymi stanami aktywacji: "jakoś" pozwala na ich odróżnienie, kategoryzację w oparciu o wiele cech i podobieństwo do poprzednich przeżyć.





Nie możemy do końca wiedzieć, jak to jest być nietoperzem, lub człowiekiem od urodzenia niewidomym czy głuchym ... Nie ma tu jednak żadnej tajemnicy.

Jak to jest być ośmiornicą lub [mątwą](#)? Żyć pod wodą, kontrolować 8 lub 10 macek, zmieniać kolory i fakturę ciała zależnie od otoczenia i komunikować się w ten sposób z innymi mątwami? To dopiero przekracza nasze wyobrażenie. Jednak nawet w przypadku ośmiornicy pewne zrozumienie jest możliwe, jak można zobaczyć na filmie "[Czego nauczyła mnie ośmiornica](#)".



### B28.7.3 [Trudny problem świadomości](#)



[David Chalmers](#) (1995) podzielił problemy świadomości na "łatwe", czyli te, które wymagają zrozumienia funkcji, i trudne, które się do tego nie sprowadzają. Np. zrozumienie sposobu generowania raportów na temat swojego stanu wewnętrznego, czy dialog wewnętrzny to problemy wymagające zrozumienia funkcji. Chociaż z punktu widzenia neurobiologii jest to skomplikowane, to z punktu widzenia filozofii łatwe.

**Problem trudny** to pytanie dlaczego istnieją jakości wrażeń, [świadomość fenomenalna](#)?

[Qualia](#), łac. qualis, to własności w oderwaniu od konkretnej rzeczy, posiadających daną własność. Czyste wrażenia to elementy nieredukowalne, nie można więc ich określić funkcjonalnie. Barwa nie da się wyrazić przez odwołanie się do prostszych własności, jest subiektywnie odczuwanym wrażeniem wzrokowym lub słuchowym. Smak czy zapach nie daje się opisać.

Jeśli wrażenia nie spełniają określonych funkcji, to jak można je symulować na komputerze? Nie da się ich sprowadzić do przetwarzania informacji. Czy jednak na pewno? Każdy stan mózgu wiąże się z asocjacjami, nawet jeśli go nie potrafimy werbalnie opisać, to jest to jeden ze stanów mózgu na trajektorii "strumienia świadomości", który zmienia tą trajektorię wpływając na nasze zachowanie. Pokazują to liczne eksperymenty pokazujące wpływ koloru (pokojów, ubrania) na zachowanie ludzi, chociaż często qualia koloru nie są nawet uświadamiane.

**Długość fali światła nie wyjaśnia wrażenia czerwieni.** Wrażenia nie dają się sprowadzić do przetwarzania informacji.

Czym są i dlaczego istnieją wrażenia?

[Filozoficzne zombi](#) to hipotetyczne istoty, które przetwarzają informację w sposób nieodróżnialny od człowieka, ale nie mają związanych z tym wrażeń, czyli świadomości fenomenalnej. Czy takie filozoficzne zombi może istnieć?

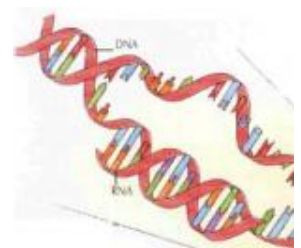
Do pewnego stopnia jesteśmy wszyscy takimi zombi, zwracamy uwagę na pewne doznania zmysłowe, ale pomijamy wiele innych, które mogą być rzeczywistą przyczyną naszych działań. Przykład: nie mamy żadnych wrażeń związanych z efektami torowania, które wpływają na nasze decyzje; bez specjalnego treningu nie zdajemy sobie sprawy do jakiego stopnia używamy echolokacji.

Wiele intuicji filozoficznych może być całkiem fałszywych, co nie przeszkadza w formułowaniu różnych myślowych eksperymentów. Czy możliwe jest np. [odwrócenie wrażeń](#) związanych z widzeniem kolorów? Widzenie koloru zielonego zamiast niebieskiego? Nie da się przeprowadzić takiego eksperymentu, ale jest to temat licznych spekulacji.

Próbując wyjaśnić trudny problem świadomości zaproponowano wiele stanowisk.

#### **Niedostrzeżenie problemu**

Konsekwentni materialści (np. Dennett) twierdzą, że wyjaśnienie funkcji wystarczy. "Czym jest życie"? Wyjaśnia to dziedziczenie, reprodukcja, wystarczy zrozumienie funkcji. Nie pytamy, czemu reprodukcji, adaptacji i innym funkcjom towarzyszy życie; czy jest sens pytać, czemu stanom mózgu towarzyszą stany świadome? To tylko zdolność komentowania własnych fizycznych stanów przez mózgi, nie istnieje "problem świadomości".





Krytyka: Dennett (1992) początkowo pisał, że wrażenia są tylko złudzeniem. Funkcja jest oczywiście konieczna, by mogło zaistnieć przeżycie. Wrażenie czerwieni zależne jest od sposobu przetwarzania bodźców wzrokowych. Oczywiście, że sama długość fali światła nie wyjaśnia wrażenia czerwieni, bo wrażenia są wynikiem reakcji złożonego układu wzrokowego. Można je wyjaśnić tylko badając reakcje mózgu. W szczególności złudzenia wzrokowe wiele nam mówią o mechanizmach powstawania wrażeń. Nawet najdokładniejszy opis funkcji realizowanej przez mózg nie wyczerpie wszystkiego, co da się powiedzieć o bezpośrednim przeżyciu. Funkcja nie istnieje w izolacji od całego systemu. Nie można uzasadnić w oparciu o opis zachowania, że wrażenia nie istnieją, jak robi to Dennett. Chcemy wyjaśnienia subiektywnego doświadczenia, a nie jego zaprzeczenia. Podstawą takiego wyjaśnienia może być zdolność komentowania własnych stanów przez mózgi, ale musimy pokazać, jak opisać ten proces tak, by w tym opisie widoczny był świat wewnętrzny, zrozumieć jak mózgi konstruują ten subiektywny świat wewnętrzny. W późniejszych pracach Dennett (2007) zmodyfikował swoje stanowisko.

## Fizykalizm i rozwiązania materialistyczne

Podstawowe pojęcia: czas, przestrzeń, ładunek nie wystarczą do zrozumienia świadomości.

Jeśli świadomość to fizyczny stan mózgu, może trzeba to po prostu przyjąć jako jeszcze jeden fundamentalny fakt natury? Chalmers proponuje **teorię podwójnego aspektu**: dwa aspekty informacji, fizyczny i fenomenologiczny, to dwie strony tego samego medalu. Uznajmy, że korelacje mózg/umysł istnieją bo odpowiednie stany mózgu skorelowane są z wrażeniami świadomymi.

Postulowanie fundamentalnych zasad jest ostatecznością. Jest to w istocie rezygnacja z poszukiwania rozwiązania!

Czy świadomość jest faktem pierwotnym, czy też da się wyjaśnić w jaki sposób powstaje? Postęp w badaniach nad świadomością i procesami zachodzącymi w mózgu pokazuje, że jest tu co wyjaśniać.

Chalmers rozważa dwie zasady:

**Zasada spójności strukturalnej**: struktura doświadczeń świadomych odpowiada własnościom strukturalnym informacji przetwarzanej przez mózg.

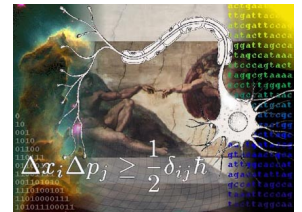
Zasada "**organizacyjnej niezmienniczości**": liczy się tylko specyficzna organizacja przetwarzania informacji, a nie sprzęt, który ją realizuje (w istocie to właśnie głosi funkcjonalizm).

Chalmers proponuje taki eksperyment myślowy: podmieńmy stopniowo neurony na identycznie funkcjonujące elektroniczne. Człowiek tego raczej nie będzie zauważał, wrażenia raczej się nie zmienią. Czy takie spekulacje mają jakieś uzasadnienie? [Implanty pamięci hipokampa](#) wydają się to potwierdzać, wymiana części głębokiej struktury mózgu nie ma widocznego wpływu na świadome wrażenia, tylko poprawia pamięć.

Z drugiej strony identyczność funkcjonalna na poziomie molekularnym oznacza identyczność substratu, muszą być te same cząsteczki, elektronika nie zastąpi w pełni neuronów. Jednakże powolne zastępowanie neuronów może być trudne do zauważenia, podobnie jak powoli zmienia się nasza percepcja z wiekiem, ale tego nie zauważamy, poczucie tożsamości jest zachowane.

Chalmers twierdzi, że granicą możliwości analizy świadomych doświadczeń może być fizykalny i fenomenalny aspekt informacji, wrażenia trzeba skorelować z czymś prostym. Fizyka bada informację "od zewnątrz", a świadomość badamy "od wewnątrz". Nie można zredukować świadomości do stanów fizycznych, mamy tylko korelacje. Dlatego uważa, że potrzebna jest nowa fundamentalna zasada, uznanie podwójnego aspektu informacji.

**Problemy**: O jaką informację jednak chodzi? Czemu tylko niektóre informacje są świadome? Znikoma część informacji, które rejestrują nasze zmysły, staje się świadoma, chociaż jest używana do utrzymywania homeostazy organizmu, sterowania napięciem mięśni w czasie chodzenia czy ćwiczeń gimnastycznych. Świadomość znika w wyniku anestezji, ale regulacja procesów życiowych nadal oparta jest na przepływie informacji. W komputerach mamy szybki przepływ informacji, ale ani śladu świadomości. Informacja mierzona za pomocą formuły Shannona to tylko miara entropii (nieporządku) sygnałów. Świadomość nie pojawia się w jakichś abstrakcyjnych przestrzeniach w czasie przepływu informacji, tylko w mózgu, które

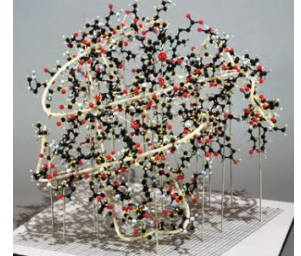


są przytomne. Widać więc, że podwójny aspekt informacji nie jest dobrym wyjaśnieniem. W jakich warunkach procesy zachodzące w mózgu stają się świadome? Jest kilka indeksów opartych na analizie sygnałów mózgu, które korelują się z subiektywnie ocenianym stopniem świadomości jak i procedurami stosowanymi przez neurologów (Seth i inn. 2006), używanymi np. w [anestezjologii](#) czy w badaniu obniżonych stanów świadomości. Na ile aktywność różnych obszarów mózgu jest zsynchronizowana, a na ile zróżnicowana? Obydwie skrajności - pełna synchronizacja jak i całkowita niezależność lokalnych procesów - prowadzą do zaniku świadomości.

### Inne próby rozwiązania trudnego zagadnienia świadomości

**Panpsychizm**, czy też w nowej wersji (Seager 1995, Rosenberg 1996) panprotopsychizm, przypisuje materii własność umożliwiającą w połączeniu z dostatecznie złożoną materią powstanie świadomości. Podstawowe elementy świadomości to "protofenomeny". Jak miałyby się wiązać z procesami fizycznymi i psychicznymi? Nie jest to jasne, ale wielu ludziom wydaje się, że to jakieś rozwiązanie.

Stwarza to więcej pytań niż odpowiedzi: czy tylko mózgi są świadome? Czemu tylko w niektórych stanach?



**Bernard Baars** uznał problem świadomości za "brutalny fakt" psychologii.

Globalnie dostępna informacja staje się świadoma. Czemu? Nie wiemy, ale trzeba przyjąć, że tak po prostu jest. To dość zbliżone podejście do podwójnego aspektu informacji, ale ze wskazaniem na globalne stany mózgu.

**Crick i Koch** (1995) uznali, że nie można opisać wrażeń świadomych, bo daje się przekazać jedynie "te różnice, które tworzą różnice".

Może istnieją specyficzne "neurony świadomości"? Jeśli nie wszystkie struktury mózgu są zaangażowane w tym procesie, to można poszukiwać neuronowych korelatów świadomości. Ludzie, którzy utracili wzrok, nadal mają wyobraźnię wzrokową. Pierwotna kora wzrokowa (V1) nie musi więc być pobudzona gdy sobie coś wyobrażamy, wystarczą wyższe piętra układu wzrokowego, ale dokładnie które obszary i dlaczego ich aktywacja wywołuje świadome wrażenia? Doprowadziło to do programu [neuronowych korelatów świadomości](#) (Crick i Koch, 1990). Po niemal trzech dekadach badań próby wyróżnienia obszarów mózgu odpowiedzialnych za świadomość dały tylko bardzo ogólne odpowiedzi (Koch i inn, 2016).

**Problem:** W mózgu nie ma słabo sprzężonych struktur, więc nie ma odrębnych neuronów świadomości, nawet aktywacja pierwotnej kory wzrokowej jest potrzebna. Kiedy widzę aktywna jest kora wzrokowa - ale czy to wyjaśnia naturę świadomych wrażeń wzrokowych?

Jeśli zrobię model, który będzie działał dokładnie tak jak mózg, wykazując takie same zdolności do rozpoznawania, robienia pomyłek, ulegania złudzeniom wzrokowym, komentowania swoich stanów, będę przekonany, że rozumiem jak to działa. Jeśli do takiego modelu dodać moduł komentujący bieżący stan aktywacji jego elementów to możemy go nauczyć tak, że będzie to przypominać komentarze świadomej osoby.

Czy telewizor, wyposażony w sieć neuronową do analizy obrazu, zdolny do komentarzy tworzonych na ekranie obrazów, uznamy za świadomy?

### Rozwiązania oparte o mechanikę kwantową.

**Mechanika kwantowa** stała się dla niektórych fizyków podstawą dla rozwijania nielokalnego modelu umysłu (Stapp 1993, Penrose 1994, Clarke 1995).

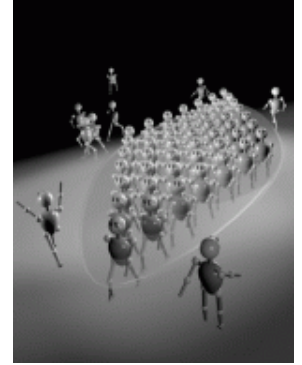
Prace te zajmują się takimi pseudoproblemami jak nieprzeźrzaną naturą umysłu. Ponieważ w mechanice kwantowej stan układu opisywany jest za pomocą jednej funkcji falowej, obejmującej cały układ, może się wydawać, że rozwiązuje to problem tożsamości, spójnego doświadczenia świadomości, pomimo zróżnicowania procesów analizy informacji w mózgu.

Podejście kwantowe prowadzi do radykalnych propozycji: umysłu jako kluczowego aspektu rzeczywistości, niealgorytmicznych teorii myślenia. Żadne eksperymenty na to nie wskazują, a sztuczna inteligencja nie napotkała na niealgorytmiczne bariery rozwoju. Głównym problemem jest złożoność systemów wymaganych do zrealizowania funkcji umożliwiających ogólną inteligencję (AGI).

Istnieją [interpretacje mechaniki kwantowej](#) wymagające świadomości, ale są też bardziej rozsądne interpretacje, w których świadomość nie gra roli, np. [teoria dekoherencji](#).

Jakości wrażeń nie dają się sprowadzić do własności fizycznych, ani klasycznych ani kwantowych. Są pewne analogie samego matematycznego opisu używanego w mechanice kwantowej, w którym mamy do czynienia zarówno z obrazem falowym, a więc oscylacji, jak i cząstek, z badaniem procesów zachodzących w mózgu, w którym zwracamy uwagę na poziom aktywności wyróżnionych obszarów i przepływ informacji pomiędzy tymi obszarami (np. w badaniach za pomocą funkcjonalnego rezonansu), oraz badań oscylacji metodami elektro lub magnetoencefalografii. Nie musimy schodzić na poziom kwantowy, by uwzględnić synchronizację aktywności różnych obszarów mózgu. Zostało to dokładniej omówione w rozdziale o [manowcach umysłu](#).

**Neurofizjolog John Eccles**, fizyk matematyczny Henry Margenau, oraz filozof Karl Popper są autorami książki: "The self and its brain" (Jaźń i jej mózg). Według nich mechanika kwantowa umożliwia wpływ ducha na materię. Wiązki dendrytów (dendrony) otoczone są polami wpływającymi na zmianę prawdopodobieństw pobudzeń synaptycznych; energia jest zachowana. "Psychony", czyli kwanty psychiki, odpowiadają za zmianę pobudzeń neuronów, nie wymaga to energii. W ten sposób mózg odbiera i wypełnia rozkazy nadchodzące ze świata umysłu.



Odsunięcie problemu umysłu w nieokreślone "zaświaty" prowadzi do nieskończonego regresu, niczego w istocie nie wyjaśniając. Takie podejście jest bezpłodne: istotne procesy decyzyjne zachodzą w sferze, w której nie można prowadzić badań, pojawiają się z zaświatów, ale mamy je uznać za najistotniejszą część naszego ja. Równie dobrze mógłbym uznać, że moje odbicie w lustrze jest takim niematerialnym światem, i zadawać pytania: gdzie się podziewa mój sobowtór gdy nie stoję przed lustrem?

Gdyby istniały psychony i odpowiadały za podejmowanie decyzji, sztuczna inteligencja, a szczególnie metody podejmowania decyzji, nie byłyby możliwe. Modele mózgu coraz lepiej wyjaśniają funkcje psychiczne i nie potrzebują psychonów. Nic nie wskazuje na istnienie takich tworów.

**Hameroff i Penrose** twierdzili, że kwantowa funkcja falowa, opisująca cały mózg jako całość, spontanicznie kolapsuje z powodu oddziaływań grawitacyjnych wytwarzając proste wrażenia. Jednak odwołanie do teorii kwantowej grawitacji, która jak na razie jest czystą spekulacją bez żadnych podstaw eksperymentalnych, niczego nie rozwiązuje. Dlaczego z miliardów elementarnych wrażeń miałyby się tworzyć świadome wrażenia? Dlaczego opis kwantowy ograniczać tylko do mózgu? Steven Hawking idzie znacznie dalej pisząc o funkcji falowej całego wszechświata.



Rozważania tego typu prowadzą nas na manowce. W mózgu nie ma miejsca na procesy kwantowe, bo:

- trwają za krótko (krócej niż pikosekundy, czyli  $10^{-15}$  sek.) by doprowadzić do synchronizacji trwającej sekundę lub dłużej;
- mózg jest za ciepły, a to oznacza szybkie fluktuacje, drgania prowadzące do dekoherencji (zaniku synchronizacji) procesów kwantowych już w obszarach wielkości pojedynczego neuronu;
- nie ma potrzeby używania wyjaśnień na poziomie kwantowym tam, gdzie wystarczą klasyczne.

[Kwantowe aspekty życia](#) to znacznie więcej niż spekulacje na temat świadomości (Abbott, Davis i Pati, 2008)

Podsumowanie kilku propozycji rozwiązania trudnego zagadnienia świadomości i wątpliwości z nimi związanych jest w tabeli poniżej.

Teoria	Proponent	Komentarz
Dualizm	Platon, Kartezjusz, Eccles, Popper ...	Stare podejście niczego nie wyjaśniło. Wielcy dualiści już wymarli? Nowe: psychony łącznikiem ze światem umysłu? Też nie pomaga.

Neomisterianie	C. McGinn	Nasze umysły nie są w stanie zrozumieć tak głębokich pytań. Skąd to wiemy? Wyjątkowo bezpłodny pogląd.
Nie ma problemu	D. Dennett, Paul Churchland	Świat umysłu zredukowany do dyspozycji, a język "psychologii potocznej" wyeliminowany na rzecz neurofizjologicznego; Nie ma mowy! Świat Ducha istnieje!
Niereduktywne	Spinoza, Chalmers	Naturalistyczny dualizm: informacja ma aspekt fizyczny i fenomenalny. To nie jest żadne wyjaśnienie!
Panpsychizm	Seager, Rosenberg	Słaby panprotopsychizm: protofenomeny tworzą świadome doświadczenia w złożonych systemach. Czym są i czemu nie jestem świadomy w komie?
Kwantowe	Stapp	Mechanika kwantowa to teoria holistyczna, umysł potrzebny do redukcji funkcji falowych; Czego można się spodziewać po takiej teorii?
Nowa fizyka	Penrose	Myślenie wymaga procesów nie-obliczalnych, potrzebna jest zupełnie nowa fizyka. 30 lat superstrun nie wyjaśniło niczego nawet na elementarnym poziomie. Powodzenia!
Redukcjonistyczne	Smart, Humprey, Searl	Stany mózgu są identyczne (lub identyczne co do typu) ze stanami umysłu, neurony mają 'moce przyczynowe' powodujące wrażenia. Nie ma identyczności, ważne są relacje! Świat umysłu jest autonomiczny, mózg to tylko substrat.
Neurobiologiczne	Crick, Koch, Taylor	Mózg 'jakoś' tworzy doświadczenia, potrzebujemy więcej szczegółów, identyfikacji 'neuronów świadomości'. Jak rozwiązanie może pojawić się z takich szczegółów?

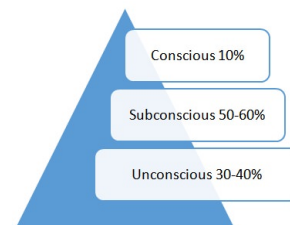


## B28.8. Świadomość potraktowana neurofilozoficznie



Pojęcia opisujące zachodzące w mózgu procesy związane ze świadomością są różnie definiowane przez różne gałęzie psychologii. Jedno z możliwych rozróżnień:

1. procesy świadome, z których zdajemy sobie sprawę;
2. procesy podświadome, wykonywane automatycznie, z których możemy sobie zdawać sprawę, jeśli zwrócimy na nie uwagę;
3. procesy nieświadome, które realizowane są przez mózg bez możliwości bezpośredniego, introspekcyjnego dostępu.



Historia odkrywania nieświadomych procesów została już częściowo przedstawiona w 2.9. Umysł, świadomość i nieświadomość. Pełny model procesów i możliwości introspekcyjnego dostępu do nich pokaże zapewne, że nie jest to proces dyskretny, tylko ma wiele stopni. Możemy całkiem jednoznacznie odczuwać i interpretować swoje stany mentalne, ale też możemy być ich mocno niepewni (np. Schwitzgebel 2011; Wierchoń i inn. 2019).



Czy wrażeniom naprawdę nie można przypisać żadnej funkcjonalnej roli? Co właściwie mamy wyjaśnić? Konieczna jest lepsza fenomenologia: czy patrząc na kolor czerwony mam istotnie inne wrażenie, niż patrząc na kolor zielony lub słysząc ton, ale skąd to wiem? Czy wie to niemowlę zaraz po narodzeniu? Oczywiście odróżniam te wrażenia, nazywając i kontrastując z innymi wrażeniami, odwołując się do pamięci. Czy mógłbym to jednak odróżniać nie mając żadnego punktu odniesienia?



**Tylko dzięki mechanizmom poznawczym potrafię rozróżnić różne stany umysłu;** gdyby nie pamięć nie mógłbym powiedzieć, czy mój stan umysłowy jest taki sam, czy się zmienił, nie mógłbym niczego porównać. Zdolność do odróżniania wrażeń wymaga więc pamięci: taki kolor morski, bardzo przyjemny, jak wody na Malediwach ... W czasie kontemplacji mogę zatrzymać swoje funkcje intelektualne i zniknie poczucie wrażeń.

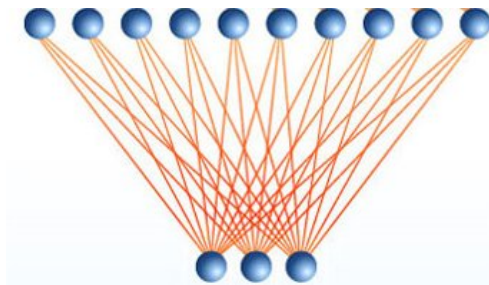
Czy podobne rozróżnienia i przekonania może mieć maszyna? Procesy obliczeniowe i maszyny Turinga nie są dobrą analogią dla funkcji mózgu. Mózgo-podobna organizacja modelu przetwarzającego informację w podobny sposób jak dynamiczna sieć neuronów prowadzi do funkcji umysło-podobnych. Czy artefakt (sztuczny intelekt) może być przekonany, że ma wrażenia podobne do naszych? Jakie warunki trzeba spełnić, by był o tym przekonany i byśmy uznali, że jest to uzasadnione przekonanie?



**Czym jest świadomość? Wynikiem zdolności do "zdawania sobie sprawy", posiadania przekonań, że się ma wrażenia.** Nie oznacza to redukcji stanów świadomości do stanów mózgu. Jakości wrażeń można wyjaśnić jedynie z punktu widzenia działania subiektywnych mechanizmów poznawczych.

Hipoteza: mechanizmy poznawcze są konieczne do odczuwania wrażeń, można je czasowo wyłączyć, np. w hipnozie czy w transie.

**Sztuczne sieci neuronowe** mogą służyć jako komputerowe modele pamięci, w których każdy element (sztuczny neuron) realizowany przez fragment oprogramowania (symbolizowany przez kulkę na rysunku poniżej) oblicza swój poziom aktywacji na podstawie informacji o aktywności innych połączonych z nim neuronów, **wykazują liczne cechy pamięci, które znamy z introspekcji:**



- Zdolność do rozpoznawania uszkodzonych wzorców.
- Przywołanie całego epizodu z fragmentarycznych wskazówek (np. samego zapachu), czyli adresowalność kontekstowa.
- Szybkość rozpoznawania wzorca nie zależy od liczby zapamiętanych wzorców, nie przeszukuje się kolejno całej pamięci.
- Częściowe uszkodzenie prowadzi do ogólnego pogorszenia się jakości rozpoznawania, ale nie zapomnienia faktów.
- Interferencja podobnych wzorców powoduje ludzko-podobne pomyłki.
- Przepiętnie pamięci prowadzi do chaotycznego zachowania.

Pamięć komputerowa ma niewiele wspólnego z biologiczną. Można wyróżnić dwa rodzaje pamięci trwałej:

- pamięć długotrwałą, wolnozmienną podstawę stabilnego obrazu świata;
- średnioterminowy podsystem pamięci, pozwalający na szybkie zapamiętywanie epizodów.



Dwa rodzaje pamięci pomagają rozwiązać dylemat stabilności-plastyczności. Zmieniamy utrwalony obraz świata bardzo powoli (wyjątki "potwierdzają regułę", zamiast ją wymazać, argumenty są ignorowane lub zapomniane), ale potrafimy zapamiętać epizody po jednym razie.

Pamięć jest własnością sieci neuronowych, pobudliwości neuronów i sprawności synaptycznych, czyli sił, z jakimi neurony mogą na siebie wzajemnie oddziaływać. Pamięć jest nieaktywna, przechowywana w strukturze sieci, dopóki elementy tej struktury nie zostaną pobudzone. Neurony zaczynają wysyłać wiele impulsów, ich działanie zaczyna się synchronizować, a ponieważ sieć zmienia swoją strukturę w wyniku intensywnej aktywności (czyli zachodzą procesy biologiczne określane jako neuroplastyczność) to pozwala na aktualizację jej przyszłych stanów. Informacja o aktywności podsieci dającej się odróżnić od przypadkowych pobudzeń staje się dostępna w większej części mózgu, staje się treścią pamięci roboczej. Dzięki temu, że nie jest zlokalizowana, możliwe są liczne skojarzenia, zapamiętanie "bogatych epizodów", różnych aspektów zdarzeń, które mogą być zapamiętane, przypominanych aspektów i sekwencji, które zakodowane są przez częściowo podobne do siebie aktywacje.

Pamięć robocza jest krótkotrwała i mało pojemna: mieści około  $7 \pm 2$  porcji informacji. Niewielka pojemność pamięci roboczej wynika z interferencji jednocześnie aktywnych procesów w mózgu. Powoduje to rozmywanie się, zanikanie aktywacji sieci neuronowych, przez co przestają się one odróżniać od neuronalnego szumu. Radzimy sobie z tym powtarzając lub zapisując informację, co pozwala odświeżać aktywację pamięci roboczej i aktualizować zawartość pamięci średniotrwałej. Proces konsolidacji tej pamięci, przeniesienie informacji do pamięci długotrwałej, trwa znacznie dłużej.



Podstawowe zadanie mózgu to **rozpoznawanie obiektów i ocena ich wartości** dla zwierzęcia.

Funkcje kognitywne można z grubsza opisać jako procesy obliczeniowe wykonywane na reprezentacjach symbolicznych; dokładniejszy opis możliwy jest na poziomie neurodynamiki; należy jednak pamiętać, że żaden opis nie wyczerpuje w pełni wszystkich własności zachodzących w mózgu procesów (podobnie jak równania hydrodynamiki nie opisują w pełni własności wody).

Rozpoznanie obiektów jest wolniejsze, konieczne jest odwołanie się do pamięci. Funkcje afektywne są związane z reakcjami fizjologicznymi całego organizmu; rozpoznawanie jest szybkie, ale niedokładne, wystarczy by zdecydować: walka czy ucieczka. Większość rozpoznanych obiektów nie wymaga natychmiastowej reakcji, ocena przyszłej wartości wymaga odniesienia do modelu wewnętrznego, reprezentacji siebie.

**Dlaczego mamy wrażenia (qualia)?** Szczur doskonale się orientuje w otoczeniu i zapewne też je ma.

- Szczur ostrożnie smakując zatrute pożywienie musi skojarzyć smak z wcześniejszymi przeżyciami by ocenić, czy jest bezpieczna.
- Reprezentacja sygnałów smakowych jest to nie-symboliczna aktywacja wielu neuronów (ten proces reprezentujemy w modelu wektorowym lub sztucznej sieci neuronowej). Można rozpoznać wiele smaków, chociaż rozróżniamy tylko cztery smaki podstawowe.
- Reprezentacja wrażeń smakowych jest rezultatem pobudzenia śladów pamięci kory smakowej i porównania z wcześniejszymi stanami mózgu. Wymaga to rozesłania informacji przez pamięć roboczą do różnych obszarów mózgu.
- Aktywacja pamięci dzięki kontekstowej adresowalności przywołuje zapamiętane informacje i udostępnia związane z nimi skojarzenia; czy po zjedzeniu podobnie pachnącego i smakującego pożywienia, w podobnej sytuacji, nie było nieprzyjemności?
- Jeśli trutka działa szybko a szczur miał z nią do czynienia w przeszłości to negatywne skojarzenie zostanie od razu przywołane i pojawi się poczucie wstrętu, reakcje fizjologiczne. Jeśli jest to trutka o powolnym działaniu, to negatywne skojarzenia się nie utworzą.
- Pobudzenia obszarów mózgu związanych z analizą danych zmysłowych i pamięcią zachodzą w różnym stopniu i są zlokalizowane w wielu miejscach, nie można więc przypisać do nich jednoznacznie symbolicznej interpretacji, ale mogą wywołać określoną reakcję organizmu.



Podsumowując szlaki pobudzeń: (węch, smak) => kora węchowa i smakowa => informacja w mózgu => szukanie skojarzeń => pobudzenie pamięci skojarzonych epizodów => aktywizacja działania i aktywizacja emocji => komentarz wewnętrzny <=> wrażenie => werbalizacja.

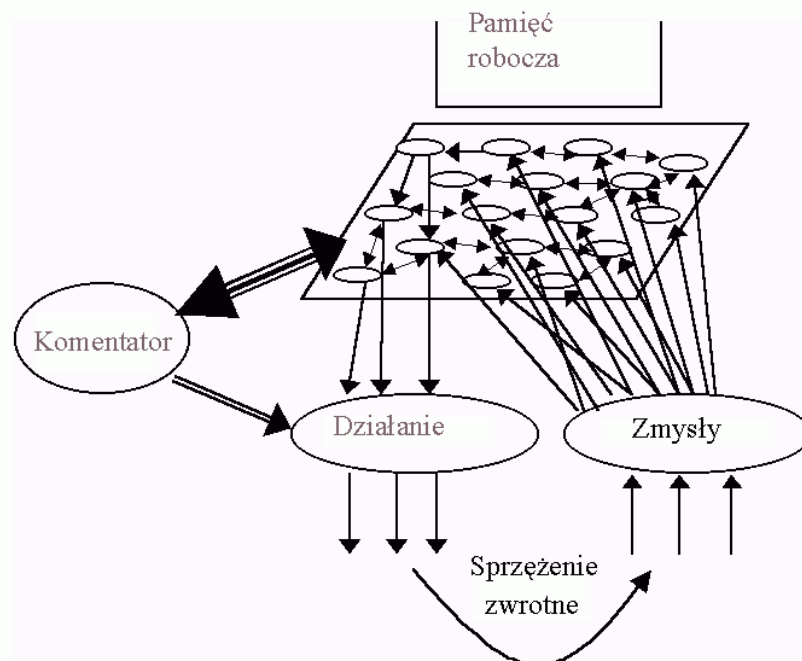
Inny przykład: [widzenie za pomocą języka](#), lub bardziej ogólnie, [substytucja zmysłów](#).

Stymulacja języka informacją z kamery prowadzi do powstania najpierw chaotycznych i niestabilnych wrażeń. Mózg uczy się pamiętając powtarzające się wzorce, porównuje zapamiętane z aktualnymi, dodaje kontekst, cała reakcja prowadzi do powstania specyficznych wrażeń. Pod wieloma względami qualia widzenia językiem są odmienne od normalnego widzenia: nie ma koloru, rozdzielczość jest niska, kontrast bardzo mały, ale orientacja przestrzenna przypomina wzrokową.

**Hipoteza:** wrażenia są zachodzącymi w mózgu procesami, które można opisać jako wewnętrzny komentarz, rozpoznanie niewerbalizowanych (w większości) stanów mózgu, reprezentowanych przez aktywacje pamięci roboczej.

Symbole wskazują na obiekty zewnętrzne lub wewnętrzne, na coś istniejącego w relacji do innych obiektów. Jakości wrażeń, oddzielone od samych wrażeń, na nic nie wskazują, są więc niewyraźne. Widząc czerwony przedmiot mam wrażenia, skojarzenia związane z kolorem wynikające z zapamiętanych wcześniej sytuacji, z wiedzy, że kolor czerwony jest podobny do różowego, ale nie do zielonego. Całkiem inne rzeczy kojarzą się z zielonym niż czerwonym (Clark, 1996). Skupiając się tylko na kolorze reagujemy na wewnętrzną reprezentację wynikającą z pobudzenia układu wzrokowego (głównie obszaru V4). Nie ma wówczas żadnego obiektu związanego z danym wrażeniem. Ten stan mózgu jest odczuwaniem "jakości wrażenia".

Na rysunku mamy najprostszą reprezentację tego procesu: zmysły po wstępnym przetworzeniu informacji z receptorów przez pierwotną korę zmysłową udostępniają ją grupom neuronów w całym mózgu, które wzajemnie się pobudzając rozsyłają informację do różnych obszarów. Informacja w pamięci roboczej może wywołać różne skojarzenia i wpłynąć na obszary kontrolujące ruch i reakcje organizmu. W przypadku ludzi jednym z działań może być werbalny komentarz, pobudzenie płatów skroniowych i czołowych, które pobudzi symboliczne reprezentacje językowe.



Postrzegane jakości wrażeń nie da się oddzielić od funkcji, którą te wrażenia spełniają. Wrażenia smakowe są konieczne by rozstrzygnąć, co jemy i na ile jest to jadalne. Mechanizmy poznawcze muszą odnieść wrażenie - ciągłą reprezentację stanu kory smakowej i węchowej - do modelu wewnętrznego świata.

Wrażenia dotykowe mają przestrzenną naturę, bo ich funkcją jest informowanie o nacisku, wzrokowe i słuchowe podobnie, węchowe w niewielkim stopniu (u ludzi), a smakowe są całkiem nieprzestrzenne.

### Wnioski:

Poprawiając dyskryminację stanów umysłu przez trening percepcyjny możemy zmienić odczuwane jakości wrażeń. Kształcenie słuchu, smaku, czy powonienia prowadzi do nowych jakości wrażeń.

Jest to wyniki trenowania pamięci wrażeń, zdolności do lepszego porównania (dyskretyzacji) ciągłych reprezentacji stanów mózgu.



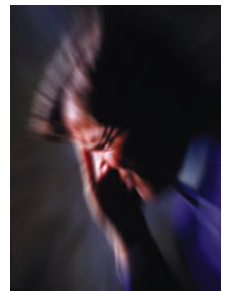
Wewnętrzne komentowanie stanów mózgu zachodzi nieustannie: proces ten wiąże się z "utrzymanie uwagi" w języku stanów mentalnych, lub synchronizacją grup neuronów na poziomie neurodynamiki. Przez pewien czas odczuwam wrażenie, bo stan mojego mózgu jest podtrzymywany przez wzajemne pobudzenia neuronów, ale brak synchronizacji, lub zbyt silna synchronizacja na innych bodźcach, powoduje rozmywanie lub zanikanie wrażeń. Wartość danego bodźca wymaga odniesienia do reprezentacji "ja". To moje wrażenie, a nie kogoś innego, więc muszę twierdzić: "(ja) słyszę". Oczywiście! Stany mózgu naprawdę istnieją, to fizyczne stany mózgu. Jedynie przekonanie o specjalnym statusie ontologicznym wrażeń jest złudzeniem. **Wrażenia są nieodłączną częścią naszej wewnętrznej rzeczywistości** (wirtualnego modelu świata).

Poczucie tożsamości jest związane z propriocepcją, stanem ciała (Damasio 2005; 2011). Kontrolowanie ciała i przewidywanie rezultatów działania mięśni wymaga wewnętrznego modelu ciała, a to prowadzi do rozwoju mózgu.

Model wewnętrzny doprowadził do powstania "jestestwa" i rozważań "co by było gdyby". Możemy trenować w wyobraźni różne reakcje. Kontrola ciała wymaga przetwarzania większej ilości informacji niż zagadnienia intelektualne: łatwiej wygrać z Kasparowem niż słabym piłkarzem. Dlatego nie mamy robotów, które wykiwają słabego piłkarza, a mamy programy szachowe, których nikt nie pobije. Samoświadomość wymaga konstrukcji reprezentacji "ja", swojej jaźni w jej licznych aspektach (Damasio 2011).

**Fenomenologia bólu** pozwala lepiej zrozumieć naturę wrażeń.

Ból nie jest tożsamy z cierpieniem. U **masochistów** ból wywołuje przyjemność. Przyjemność z jedzenia czekolady przechodzi w końcu w obrzydzenie gdy zjemy jej za wiele. Te same bodźce mogą prowadzić do całkiem odmiennych wrażeń zależnie od stanu organizmu. Pomimo intensywnego bólu można nie odczuwać cierpienia (np. w ferworze walki). Można też zapomnieć o bólu gdy mechanizmy poznawcze zaangażowane są w inne zadania. Wniosek: **odczuwanie bólu to kwestia interpretacji przez mechanizmy poznawcze**. Zwracamy uwagę na stan ciała, który wymaga szybkiej interwencji, ale jeśli coś silniej przyciąga naszą uwagę, wrażenie bólu może zostać stłumione.



Rola pamięci sensorycznej w powstawaniu wrażeń jest bardzo ważna. Dźwięk przestrzenny wymaga przechowania przez mózg sygnału porównywanego z sygnałem docierającym do drugiego ucha. Procesy kognitywne są konieczne do interpretacji wrażeń, powstania wrażenia dźwięku przestrzennego. Kognitywna interpretacja stanowi nie tylko warunek konieczny do powstania wrażenia, dzięki niej jestem przekonany, że mam wrażenie. Wynik tej oceny - wrażenie - jest częścią modelu świata, powstającego w wyniku pobudzenia neuronowych sieci, kodujących różne elementy tego modelu. Mózg pyta sam siebie "co to jest" i odpowiada: śpiew skowronka, lub pachnąca czerwona róża.

**Odczuwanie reakcji organizmu jest wynikiem interpretacji stanów mózgu**. Nie można sobie wyobrazić wrażeń oderwanych do tych reakcji (Damasio, 2005; 2011).

### Dalsze wnioski.

- Brak wewnętrznego komentarza spowoduje, że poczucie "mam dane wrażenie" osłabnie, np. na skutek habituacji, wyłączenia dominującej półkuli mózgu. Przystajemy być przekonani, że odczuwamy wrażenia.
- Trenowanie mechanizmów poznawczych prowadzi do "wyostrzenia zmysłów", a więc nowych jakości wrażeń, dostępnych również w snach - świadczy to o roli mechanizmów pamięci w procesach dyskryminacji.
- Wrażenia zależą od działania mechanizmów poznawczych. Zawieszenie działania tych mechanizmów powinno prowadzić do zniknięcia wrażeń. Dlatego możemy "patrzeć nie widząc".



- Wrażenia spełniają rolę funkcjonalną, dlatego mają różne własności: wzrokowe, dotykowe, czucie temperatury i bólu mają strukturę przestrzenną, w mniejszym stopniu słuchowe. Wrażenia smaku czy zapachu, myśli (ciche mówienie) i wyobrażenia takiej struktury nie mają.
- Stymulacja struktur mózgu odpowiedzialnych za przetwarzanie danych wrażeń powinna prowadzić do ich powstawania tylko wtedy, gdy mogą one wpłynąć na pamięć roboczą.
- Rozpoznawanie wstrętu widocznego na czyjejs twarży (realizowane przez korę przedniej części wyspy) powinno angażować te same obszary gdy wstręt wywoływany jest przez bodźce smakowe.
- Gra aktorska w stanie pełnej absorpcji powinna prowadzić do autentycznego przeżywania stanów emocjonalnych - udawane emocje mogą się stać prawdziwymi emocjami.
- Intencjonalność systemu wynika z oparcia się na oczekiwaniach, pamięci poprzednich zdarzeń.
- Pamięć proceduralna nie może być uświadamiana, gdyż obszary motoryczne nie tworzą stanów rezonansowych.

Można wyciągnąć znacznie więcej szczegółowych wniosków, brakuje jednak systematycznej fenomenologii świadomości.

Reakcje "przytomne" robota nas nie dziwią, robot postrzega i reaguje, a jego reakcje mogą być związane ze skojarzeniami, jakie podsuwa mu pamięć. Dlaczego jednak roboty nie mogłyby "zdawać sobie sprawy z tego, co się dzieje ich umysłach"? Dotychczas stosowane systemy sterujące zachowaniem robota oparte były na regułach i procedurach: po rozpoznaniu A wykonaj program B. To ogranicza kreatywność. Mózg przez większą część czasu tak nie działa. Rozpoznanie A powoduje powstanie dynamicznego, czasoprzestrzennego stanu systemu percepcji, zależnie od aktualnego stanu naszego mózgu. Stan ten jest nieprzewidywalny, zależny od wszystkiego, co się nam w życiu zdarzyło, prowadzi do różnych skojarzeń i nie da się go w jednoznaczny sposób powiązać z pożądanymi reakcjami. Takie stany mogą powstawać w wielkich sieciach neuronowych, takich jak różne wersje systemów GPT, które mają setki miliardów parametrów. Nawet stosunkowo proste roboty, sterowane przez sieci neuronowe, w których powstaje dynamiczny stan aktywacji, np. [Hector](#), [sztuczny insekt](#), mają wiele cech świadomej istoty.



Pamięć pozwala na internalizację obrazu świata, więc w mózgu takiego robota mogą zachodzić wyobrażone reakcje, z których też będzie sobie zdawał sprawę, musi więc twierdzić, że jest ich świadomy.

**Skąd będziemy wiedzieć, że sztuczny system jest już naprawdę świadomy?** Podobnie jak robi się w przypadku ludzi czy zwierząt, obserwując ich reakcje i zachowanie. Jeśli system ma odpowiednią architekturę, powstają w nim stany wewnętrzne, ciągi myśli i wrażeń, które potrafi skomentować i uzewnętrznić, jeśli te komentarze będą miały strukturę podobną do naszych, nabierzemy przekonania, że jest naprawdę świadomy.

Można przekroczyć przepaść pomiędzy światem umysłu i mózgu, psychicznym i fizycznym!

### **Czy pozostaje tu jeszcze coś istotnego do wyjaśnienia?**

Podobne poglądy wyraził Kevin O'Regan w referacie: [How to build a robot that feels](#). Szczegóły dopracowywane są w projekcie: [FEEL – A sensorimotor approach to understanding consciousness](#).

Dopiero po skonstruowaniu czującego robota zachowującego się w świadomy sposób nasze intuicje będą miały szansę ulec zmianie. Jednak filozofowie na pewno będą mieli nadal wątpliwości.

Reakcje, takie jak tropizmy, nie są jeszcze oznaką świadomości. Termostaty reagują na ciepło, ale to prosty proces fizyczny, nie ma tam interpretacji wewnętrznej tego co robią, "percepcji tego co się w nich dzieje", nie ma więc sensu mówić o świadomości. Z tego powodu robot działający w oparciu o program i reguły zachowania, nie jest jeszcze świadomy, chociaż orientuje się w otoczeniu i reaguje na różne zdarzenia, jest więc przytomny.

Dopiero model wewnętrzny świata i siebie w tym świecie, pozwalający na postrzeganie, zwracanie uwagi na wybrane aspekty zmiany modelu i komentowanie tego co się dzieje, umożliwi badania heterofenomenologiczne. Taki system, w którym utrzymuje się dynamiczny stan czasoprzestrzenny, a nie tylko realizowane są wyniki obliczeń, będzie mógł określić, w jakich warunkach jest czegoś świadomy, a w jakich nie.

Tendencje do wkraczania neuronauk na kolejne obszary, które są powiązane ze świadomością w bardziej subtelny sposób, są wyraźnie widoczne.

**Neuroteologia** zajmuje się korelatami neurofizjologicznymi stanów mistycznych (d'Aquili, Newberg, 2010). Przeżycia mistyczne są źródłem i istotnym składnikiem wielu religii (nie wszystkich). Tymi zagadnieniami interesowała się psychologia przeżyć [transpersonalnych](#), ale sama [psychologia transpersonalna](#) wyrosła z [psychologii humanistycznej](#) i na razie słabo łączy się z neuronaukami. Czy "czysta świadomość", przeżywana w pewnych stanach mistycznych, stanowi większy problem niż qualia? [James Austin](#) jest amerykańskim neurologiem, który opublikował wiele znakomitych książek na temat medytacji Zen, w których szczegółowo analizuje mechanizmy uwagi i działania mózgu w odniesieniu do przeżyć mistycznych.



**Neuroetyka** uzasadnia zachowania moralne z neurobiologicznego punktu widzenia.

[George Lakoff](#) i [Mark Johnson](#), w książkach "Metafory, którymi żyjemy" (1980), oraz "Philosophy in the flesh" (1999), przeanalizowali metafory dotyczące moralności.

Z tej analizy wnioskuje, że zasady moralnej oceny odnoszą się w pierwszym rzędzie do zagadnień, które wpływają na nasze zdrowie fizyczne i psychiczne, dobre samopoczucie.

Dobre samopoczucie psychiczne przenosi się na relacje społeczne.

Podstawą moralności jest zdolność do współ-odczuwania, empatia, a ta wynika prawdopodobnie z aktywności "[neuronów lustrzanych](#)", pozwalających na uczenie się przez imitację i współ-odczuwanie.

Wiele przekonań jest wspólnych dla wszystkich ludzi, niezależnie od kultury:

- Lepiej być zdrowym niż chorym, lepiej być silnym niż słabym.
- Lepsza jest woda, pożywienie i powietrze czyste niż skażone.
- Lepiej jest decydować samemu za siebie (być wolnym) niż być uzależnionym od cudzych decyzji (być zniewolonym).
- Lepiej posiadać zasoby, które pozwalają na niezależność, niż być na łasce innych.
- Lepiej jest czuć się bezpiecznie niż żyć w zagrożeniu.
- Lepiej być częścią grupy i wzajemnie się o siebie troszczyć, niż być izolowanym czy zaniedbywanym.



## B28.9. Podsumowanie: kompromisy ewolucyjne



Warto podsumować czego się nauczyliśmy z przeszłości i z rozważań filozoficznych. Początkowo były to naiwne próby tłumaczenia powierzchownych obserwacji i szukania wyjaśnień dalekich od empirycznych badań. Powstały liczne mity i wyobrażenia dotyczące powstania życia, człowieka, poszukiwania magicznych przyczyn zjawisk, których nie można było zrozumieć bez zbudowania naukowych podstaw fizyki, chemii, biologii czy rozwoju statystyki pozwalającej odróżnić rzeczywiste przyczyny od przypadkowych korelacji. Ceną za uniwersalne możliwości uczenia się jest utrzymywanie się wielu zabobonów, wiary autorytetom zamiast sceptycznego podejścia szukającego wiarygodnych źródeł informacji. W rezultacie nadal tylko niewielka część ludzkości ma racjonalne poglądy, wiara w magię nadal jest rozpowszechniona. To, co nam się wbija najwcześniej do mózgu, długo w nim zalega i trudno to wyplenić.

Parę tysięcy lat rozważań filozoficznych nie przyniosło wielkiego postępu. Dopiero w ostatnich dekadach filozofia kognitywna i neurofilozofia, opierając się na danych empirycznych, staje się źródłem wiedzy dobrze zweryfikowanej i silniej związanej z rzeczywistymi mechanizmami działania umysłów.

Jednym z przewijających się tematów jest konieczność kompromisów, **dylemat stabilności/plastyczności**. Stabilność prowadzi do trzymania się tradycji i dominacji poglądów oraz mechanizmów społecznych, które w obliczu współczesnej wiedzy nie mają już sensu. Nie należy psuć tego, co jeszcze dobrze działa, ale warunki się zmieniają i konieczna jest adaptacja do nowej sytuacji i dobrze ugruntowanej wiedzy. Plastyczność umożliwia zmiany, uczenie się, powstawanie nowych teorii i sposobów organizacji. Ten dylemat można

dostrzec na wielu poziomach, np. rozwoju nauki i nowych idei, które początkowo wydają się większości bezsensowne (teoria atomowa, kwanty, ruch płyt tektonicznych, ekologia itd).

Trzy teorie drastycznie zmieniły nasz sposób patrzenia na umysł, naturę człowieka i jego inteligencję, dostarczając narzędzi do weryfikacji hipotez i pokazując, jak natura poradziła sobie z dylematem stabilności/plastyczności na różnych poziomach.

Najpierw była teoria ewolucji.

Druga to rozwój neuronauk i zrozumienie sposobu działania mózgu.

Trzecia to modele komputerowe, ilustrujące jak działają mózgi i jak z takiego działania pojawia się wewnętrzna perspektywa, subiektywny umysł. Na to niestety nie jesteśmy jeszcze gotowi.

## Kompromisy ewolucyjne.

Wzajemne kształtowanie się mózgu i umysłu jest przykładem rozwiązania **kompromisu pomiędzy stabilnością a plastycznością**, a więc koniecznością nauki i przygotowaniem do działania w znanym świecie, w którym wyewoluowały mózgi.

Jednym z tematów przewodnich tego wykładu było pokazanie konieczności kompromisów, które musiała rozwiązać ewolucja. Kompromisy dotyczą budowy całego organizmu i widoczne są również w sposobie działania mózgu, pomagając osiągnąć większą sprawność działania. Każdy z tych kompromisów ma swoje zalety i wady, warto je więc dokładnie przemyśleć.

- **Przetrwanie w niebezpiecznym środowisku wymaga sprawnych zmysłów.**

Rozwój percepcji zależał od wymagań ekosystemu, w którym rozwijał się dany organizm: sposobu odżywiania, reprodukcji, konkurencji innych organizmów. Stąd potrzeba zmysłu elektrycznego, echolokacji, termowizji, widzenia przestrzennego, widzenia kolorów niezależnie od oświetlenia, widzenia w ciemności i silnym świetle, słyszenia szmerów, odczuwania słabych wibracji włosów, swędzenia, temperatury, bólu, położenia ciała, napięcia mięśni, zapachu i smaku, pragnienia, głodu, poziomu utlenienia krwi, oraz innych sygnałów płynących z wnętrza ciała.

Neurony analizujące dane zmysłowe są przystosowane do optymalnych reakcji na naturalne bodźce. Taki jest wynik licznych eksperymentów. Jest to też podstawa, by wyjaśnić szczegółową budowę układów zmysłowych, sposobu analizy dostarczanych przez nie bodźców przez wyspecjalizowane obszary mózgu (korę i jądra podkorowe).

- **Szczegółowa analiza danych ze zmysłów jest niezbędna, ale trzeba oszczędzać energię.**

Ponieważ analiza jest kosztowna, mózgi zużywają dużo energii, nie można się skupiać naraz nad wszystkim cechami dopływającej informacji.

Kompromisy potrzebne są w każdym przypadku: ostre widzenie wymaga żółtej plamki i kontroli szybkich sakadycznych ruchów oczu oraz mechanizmu oceniania istotności i przestrzennego umiejscowienia bodźców.

Skupienie się na jednej modalności zmysłowej, np. dźwiękach słyszanych w ciemności wymaga stłumienia sygnałów z innych zmysłów, do czego przydatna jest centralna struktura przekazująca informacje (wzgórze), realizująca zasadę "zwycięzca bierze większość", czyli najważniejsza informacja dominuje nad innymi.

Wymaga to skomplikowanego mechanizmu oceny istotności informacji, opartego na licznych strukturach podkorowych reagujących szybko i mało precyzyjnie, wpływając na stan całego organizmu - a więc emocjonalnie. Mamy więc kompromis między informacją powstałą dzięki szczegółowej percepcji i emocjom.

- **Dylemat stabilności-plastyczności przejawia się na wielu poziomach.**

Informacja dostarczana przez zmysły nie wystarcza, by ocenić jej wagę, trzeba ją odnieść do zapamiętanej informacji. Pamięć umożliwia rozpoznawanie obiektów, ich kategoryzację, rozpoznawanie twarzy i wyrazów twarzy. Jednakże zbyt precyzyjna pamięć nie służy sprawnemu działaniu, utrudnia powstawanie skojarzeń i tworzenie kategorii.

Pamięć i jej 7 grzechów to przykład kompromisów, rozważnych w podrozdziale na temat [niedoskonałości pamięci](#). Wynikają stąd liczne ograniczenia na sposób działania sieci neuronowych pozwalających

zapamiętać informacje.

Powolne tworzenie się pamięci semantycznej z epizodów zapamiętywanych po jednokrotnej ekspozycji wymaga dwóch odrębnych systemów w mózgu, opartego na procesach w korze mózgu i hipokampie. Kompromis pomiędzy tym, co warto zapamiętać, a tym co można łatwo odnaleźć, przejawia się szczególnie mocno w dzisiejszych czasach, gdy coraz bardziej polegamy na urządzeniach wspomagających nasze zdolności. To sprzężenie z technologią staje się coraz silniejsze i niewiele osób świadomie wykorzystuje możliwości technologii. Ma to wiele negatywnych konsekwencji:

1. Następuje alienacja człowieka od siebie, autorefleksji, spowodowana ciągłym analizowaniem nadchodzących wiadomości, graniem na telefonach czy komputerach, szukaniem w każdej chwili rozrywki. W efekcie oduczamy się być sami ze sobą, nie możemy znieść chwili ciszy, musimy włączyć telewizor lub jakąś grę.
2. Pozornie dzięki sieciom społecznościowym mamy większą łączność z wieloma ludźmi, ale są to płytkie relacje, w których tworzymy projekcję siebie, a przez to rzadziej wchodzimy w bezpośredni kontakt z innymi ludźmi.
3. Chociaż mamy dostęp do dowolnych informacji, ale nie potrafimy odróżnić faktów od opinii, prawdy od fałszu. Mnożą się więc teorie spiskowe i część załęcznionych osób kończy w zamkniętych sektach.

- **Samodzielna eksploracja i konformizm.**

Mózgi potrafią się nauczyć wszystkiego, ale na początku muszą polegać na informacji dostarczanej przez opiekunów. Wszystko, czego się dowiadujemy, zmienia sposób naszego myślenia, ale całkowity konformizm, zbytnia stabilność, uniemożliwiają potrzebne zmiany. Na poziomie pojedynczych neuronów oznacza to, że uczenie nie może być zbyt szybkie, neuron w korze mózgu zmienia swoje połączenia w sieci po wielu powtarzających się silnych seriach impulsów do niego dochodzących. Jednak potrzebny jest też mechanizm szybkiego uczenia po jednokrotnej ekspozycji, zapamiętania ważnych zdarzeń. Jest to możliwe dzięki szybkiemu zapamiętywaniu informacji o aktywacji kory przez hipokamp, szczególnie informacji, które wywołały silne emocje, oraz informacji o odwiedzanych miejscach. Na poziomie budowy całego mózgu kompromis pomiędzy stabilnością, zapewnianą przez korę mózgu, i plastycznością zapewnianą przez hipokamp, jest więc specyficzna budowa mózgow.

Dziecko musi słuchać opiekunów, ale też samo eksploruje swoje otoczenie, ucząc się podstawowych rzeczy. Potrzeba kontroli nad światem jest bardzo silna, układ nagrody silnie się aktywuje gdy można coś kontrolować zgodnie ze swoimi intencjami. Wynika stąd niezwykła atrakcyjność gier komputerowych. Konformizm potrzebny jest w okresie największej plastyczności mózgu, a więc w dzieciństwie. Doświadczenia z tego okresu kształtują sposób naszego myślenia, nasze skojarzenia, przez resztę życia. Dlatego opowiadamy nadal historie sprzed tysięcy lat, wierzymy w mity, magię, nie patrzymy krytycznie na rzekome zdolności paranormalne czy cudownych uzdrowicieli.

Kompromis potrzebny jest też na poziomie organizacji życia społecznego: raz zwycięża tradycja, konserwatywny sposób myślenia, a raz chęć zmian, postępu. Dotyczy to zarówno partii politycznych jak i obyczajowości. Zmiany w krajach, w których nie ma wolności wyrażania poglądów, są powolne. Kraje te nie dostosowują się do globalnych zmian na świecie. W krajach demokratycznych regularnie dochodzi do zmiany rządzących partii politycznych, pokazując ciągłą dynamiczną równowagę pomiędzy różnymi poglądami.



## B28.10 Perspektywy kognitywistyki



Lata powtarzania, że "nie rozumiemy, jak to działa", albo "nie możemy tego zrozumieć", argumenty oparte na twierdzeniu Gödla oraz rozważania filozoficzne skutecznie spowodowały blokadę umysłową, która



prowadzi do zadawania nieistotnych pytań i niekończących się rozważań na zupełnie nieistotne tematy. Zrozumienie ogólnych zasad działania mózgu nie jest wcale takie trudne!

Czym jest to "ja"?

Czy najpierw jest myśl, a potem aktywacja mózgu, czy odwrotnie?

Czy ja mam mózg czy mózg ma "mnie"?

Ogólnie: mózg tworzy plany działania, opiera się na nabytym doświadczeniu, czyli zachowaniach wyuczonych w przeszłości i porównaniach z zapamiętanymi, podobnymi sytuacjami.

Mózg tworzy reprezentację "ja", obserwuje swoje własne stany, klasyfikując je i pozwalając na ich werbalne komentowanie. **Mózg wpływa na umysł, umysł (przeżycia) kształtuje mózg.**

**Sztuczne umysły** będzie można realizować za pomocą uproszczonych modeli mózgu.

Co warto nazwać sztucznym umysłem? Możemy sobie wyobrazić całe spektrum modeli, od najprostszych modeli do coraz bardziej zbliżonych do umysłu ludzkiego.

- Minimalnie: pamięć długotrwała + pamięć robocza + komentator.
- Maksymalnie: model wszystkich struktur mózgu człowieka, sterujący robotem, który uczy się relacji o świecie w podobnych warunkach jak dziecko.

Model minimalny nie ma wrażeń, gdyż powstające w jego pamięci roboczej reprezentacje nie składają się z wrażeń zmysłowych, nie można ich zredukować do elementów prostszych.

### Relacja mózg-umysł .

Mózg jest substratem, koniecznym do powstania umysłu, neurony podobnie jak atomy konieczne są do powstania złożonych, dynamicznych struktur.

Mózg jest zarazem czymś więcej i czymś mniej niż umysł.

Struktury umysłu powstają w substracie mózgu, ale są odbiciem relacji w świecie, o którym zmysły dostarczają nam wiadomości, w tym relacji społecznych.

Dlatego nie można zrozumieć umysłu bez rozumienia, jakie stany mogą przyjmować mózgi.

Często mówi nam o tym więcej literatura niż nauka.

Metafory:

Do stworzenia obrazu potrzebne są barwniki, farby, struktury chemiczne, atomy różnych pierwiastków, ale piękno obrazu nie tkwi w samych farbach i płótnie.

Do stworzenia umysłu potrzebne są neurony i neuromodulatory, układ krążeniowy i komórki glejowe, ale piękno umysłu nie tkwi w neuronach.

"Ja" jest jednym z procesów realizowanych przez mózgi; nie ma co do tego żadnych wątpliwości, wynika to zarówno z psychologii rozwojowej, badań nad mózgiem (zwłaszcza pojawianiem się intencji, woli), jak i modeli komputerowych, które pokazują, jak mózgi w aktywny sposób podejmują decyzje.

Upieranie się przy starożytnych wyobrażeniach niematerialnej duszy, pociągającej za sznurki i kontrolującej zachowanie, podobne jest upieraniu się, że Ziemia jest płaska.

### Jeśli jestem tylko procesem mózgu, to czy naprawdę jestem?

Nie jestem "tylko" procesem! Identyfikacja ze swoim "ja", które jest jedynie wyobrażeniem, jakie mamy o sobie, jest błędem.

Jestem całym mózgiem, całym organizmem, a nawet więcej, całym światem, który ma wpływ na stan "mojego" umysłu, bo stany umysłu są wynikiem oddziaływania ze światem.

Potencjalnie liczba stanów umysłu jest nieskończona, ale to sytuacja, a zwłaszcza interakcje z innymi ludźmi, powodują aktualizacje wybranych stanów, które nie mogłyby powstać w izolacji.

W niektórych tradycjach to "ja" nazywane jest "małym umysłem", a ich celem jest zrozumienie, że jesteśmy „wielkim umysłem”, który jest wszystkim (por. Shunryu Suzuki, Umysł zen, umysł początkującego, Wyd. Pusty Obłok, 1997).

Ja walczę ze sobą: chciałbym być inny niż jestem, chciałbym dorównać bohaterom ...

"Ja" jest jednym z procesów, ale emergentnym, oderwanym od biologii, zrozumiałym tylko w kontekście społecznym.

Sprzęt komputerowy ogranicza do pewnego stopnia jaki rodzaj programów można na nim wykonać, ale umożliwia działanie bardzo różnych rodzajów programów.

### Jaki jest sens życia?

Szukanie sensu w świecie jest naturalną rzeczą, bo doszukiwanie się intencji i sensu działania zwierząt i ludzi jest bardzo ważne. Jednakże w przyrodzie mamy do czynienia z ewolucją, która nie ma celu, chociaż jej efekty można tak zinterpretować, jakby ten cel istniał. Ten mechanizm opisał szczegółowo Daniel Dennett w książce: *Od bakterii do Bacha. O ewolucji umysłów* (2017).

Jak można odpowiedzieć na pytanie o sens życia z punktu widzenia duchowych potrzeb człowieka, powiązanych z biologicznymi podstawami? Sens nadaje nam uczestnictwo w czymś, co nas przerasta. Problem w tym, by dobrze wybrać.

Celem życia może być pełne rozwinięcie ludzkiego potencjału, wrodzonych możliwości.

Naturalnym działaniem mózgu jest:

- poznawanie i przyjemność z tego poznawania,
- empatia, dzielenie się radościami i smutkami,
- kreatywność i rozwijanie swoich wrodzonych zdolności,
- zdolność do poświęcenia chwilowych korzyści na rzecz przyszłych.



**Ikigai** to japońskie słowo, złożone pojęcie odpowiadające z grubsza sensowi życia z uwzględnieniem wielu aspektów, które powodują, że życie jest tego warte, ma celu, kierunek rozwoju, nasze działania sprawiają nam satysfakcję, mamy poczucie misji i powołania, motywację do działania, nasze pasje i praca zawodowa się pokrywa. Pojęcie ikigai jest na razie mało znane w psychologii, ale są pewne próby by je powiązać z [badaniami nad mózgiem](#).

### Zaburzenia rozwoju mogą zniszczyć ten naturalny potencjał na różnym etapie:

- na całkiem wczesnym etapie, wpływając na rozwój mózgu płodu;
- na etapie niemowlęcia na skutek zaniedbania, braku stymulacji środowiska;
- we wczesnym dzieciństwie na skutek ciężkich urazów;
- w okresie dojrzewania i dorosłym, na skutek depresji i innych chorób;
- w życiu dorosłym na skutek niewłaściwego systemu wartości;
- na starość na skutek chorób prowadzących do otępienia.

Osoby zdrowe zastanawiają się nad sensem życia wtedy, gdy nie doszło do integracji procesów zachodzących w niedominującej półkuli ich mózgu i procesów podkorowych, pełniejszego poznania "siebie". [Abraham Maslow](#) przyczynił się do rozwoju [psychologii transpersonalnej](#), koncentrującej się na przeżyciach określanych jako "duchowe", samorealizację.

**Sumienie** = różnica pomiędzy wyobrażeniem o sobie (zwykle wyidealizowanym) a rzeczywistością: nie jestem taki, jaki myślałem, że jestem.

Sumienie gryzie (wyzwała emocje, żal), dzięki czemu możemy się zmienić, bo wzrasta neuroplastyczność.

Czy możliwe jest w pełni zintegrowane, naukowe podejście do rozwoju człowieka na wszystkich płaszczyznach? Konieczne jest do tego lepsze zrozumienie natury ludzkiej, bez tego będzie to kolejne podejście parareligijne, wiara w słabo uzasadnione modele, których stosowalność ograniczona jest do bardzo specyficznych kontekstów, które próbuje się rozciągnąć na obszary dalekie od ich stosowalności. Teorie łączące podstawowe zasady fizyki, takie jak [zasada minimalnego działania](#), z której można wyprowadzić równania opisujące ruch złożonych systemów, zasadę wolnej energii w układach biologicznych ([Karl Friston](#)), zgodnie z którą mózg optymalizuje model stanów środowiska pozwalający na przewidywanie

zmian na postawie informacji ze zmysłów i kontrolę zachowania organizmu, oraz ewolucyjne ograniczenia na konstrukcję takich systemów opisywane przez teorię systemów ewolucyjnych (Badcock, 2012). W efekcie mamy obecnie teorię łączącą fizykę (procesy neuronalne), psychologię (procesy kognitywne) oraz ewolucję (procesy biologiczne, rozwojowe). Omawianie tej teorii wykracza jednakże daleko poza wstęp do nauk kognitywnych.

**Takie pełne zrozumienie jest właśnie celem kognitywistyki.**

**Zadanie: głównym tematem tego wstępu do neurofilozofii jest znaczenie biologicznego mózgu i ciała. Zamiast rozważań filozoficznych, ujęcia koncepcyjnego, mamy osobiste doświadczenie. Spróbuj przeanalizować własne obserwacje, uzasadnić lub skrytykować tworzenie się przekonań w ujęciu emergentyzmu, enaktywizmu, lub wniosków, do których doszedł Mark Johnson w książce o znaczeniu ciała.**

Pytania, na które powinniście znać odpowiedzi po przeczytaniu notatek do tego wykładu:

1. Czym miała zajmować się fenomenologia?
2. Prawdy aprioryczne u Husserla - czy są rzeczywiście niepodważalne?
3. Znaczenie poglądów Heideggera i jego tezy metafizyki powszedniości.
4. Merlau-Ponty i filozofia percepcji.
5. Co to jest emergentyzm?
6. Na czym polega nielokalność umysłu?
7. Skąd powstaje znaczenie, intencjonalność?
8. Teoria identyczności typów.
9. Materializm eliminatywny - jakie pojęcia i dlaczego można wyeliminować?
10. Teoria Churchlanda powstawania przekonań i rozwijania nauki.
11. Jakie są wady metafory "mózg to komputer, umysł to program"?
12. Czy można powiedzieć, że umysł jest bardziej doskonały niż mózg?

---

## Literatura

### [Journal of Neurophilosophy.](#)

- [Austin J, Zen and the Brain](#). Toward an Understanding of Meditation and Consciousness. MIT Press 1999.
- Austin J, Zen-Brain Reflections. MIT Press 2006.
- Błaszak, M: Programy badawcze w naukach kognitywnych a metafizyka powszedniości. Wydawnictwo Naukowe WNS UAM 2009
- Błaszak, M: Ekotypy poznawcze człowieka. Przyczynek do kognitywnej teorii podmiotu. Poznań, Wydawnictwo Naukowe Bogucki, 2013, 203 s, ISBN 978-83-63400- 86-6
- Brentano, F: Psychologia z empirycznego punktu widzenia. Warszawa: PWN, 1999.
- Brooks R, Flesh and Machines: How Robots Will Change Us (Pantheon, 2002
- Chicago Social Brain Network, Invisible Forces and Powerful Beliefs: Gravity, Gods, and Minds. 2011
- Churchland P.S, Neurophilosophy: Toward a Unified Science of the Mind-Brain. (1986) Cambridge, Massachusetts: The MIT Press.
- Churchland P.S, Sejnowski T.J. (1992) Patricia S. Churchland and T. J. Sejnowski. Cambridge, Massachusetts: The MIT Press.
- Churchland P.S, Neurophilosophy and Alzheimer's Disease. (1992) Springer-Verlag.

- Churchland P.S, The Mind-Brain Continuum (1996). Ed R. R. Llinas and Patricia S. Churchland. The MIT Press.
- Churchland PM, Churchland PS, On the Contrary: Critical Essays 1987-1997 (1998) MIT Press.
- Churchland PS, Brain-Wise: Studies in Neurophilosophy. (2002) Cambridge, Massachusetts: The MIT Press.
- Clark Austen, Sensory Qualities. Oxford Uni. Press 1993
- Damasio Antonio, W poszukiwaniu Spinozy. Radość, smutek i czujący mózg, Rebis 2005
- Damasio A, Jak umysł zyskał jaźń. Konstruowanie świadomego mózgu. Rebis 2011
- Dehaene, S. (1997) The number sense: How the mind creates mathematics, Oxford University Press
- Duch W (1994) Życie wewnętrzne komputerów. Toruńskie Studia Dydaktyczne, rok III(6), pp.191-206.
- Edelman G, Przenikliwe powietrze, jasny ogień. O materii umysłu. PIW, Warszawa, 1999
- Eves H, Mathematical Circles. Adieu, Boston, 1979.
- Gazzaniga, M. Instynkt świadomości. Jak z mózgu wyłania się umysł? Wyd. Smak Słowa. Seria: Mistrzowie Psychologii, 2020.
- Humprey, N. [Getting the measure of consciousness.](#)
- Johnson Mark, Znaczenie ciała, Estetyka rozumienia ludzkiego. Wyd. Uniw. Łódzkiego 2015
- Lakoff, G. & R. E. Nunez (2000), Where mathematics comes from, Basic Books.
- Seth, A.K, Izhikevich, E.I, Reeke, G.N, Edelman, G.M. [Theories and measures of consciousness: An extended framework.](#) PNAS 103 (2006) 10799-10804
- Thelen E, Smith L.B. A Dynamic Systems Approach to the Development of Cognition and Action. MIT Press, 1994.
- Varela, F. J. (1996). Neurophenomenology: a methodological remedy for the hard problem. Journal of Consciousness Studies, 3(4), 330–349. Tł. polskie: [AVANT 1 \(2010\) 31-73](#)
- White A.D., [History of the Warfare of Science with Theology in Christendom](#), 1895, czyli historia pomyłek teologicznych.

[Center for Naturalism](#) | [Naturalism.org](#) |

---

---

---

Cytowanie: Włodzisław Duch, Wstęp do Kognitywistyki. Rozdz. 28: Neurofilozofia. UMK Toruń 2023.  
[Następny wykład](#): Krytyczna historia filozofii umysłu | W. Duch, [Wstęp do kognitywistyki - spis treści](#)