

Newton, Natika

Problem reprezentacji w teoriach poznania ucieleśnionego

Avant 3/T, 66-82

2012

Artykuł został opracowany do udostępnienia w internecie przez Muzeum Historii Polski w ramach prac podejmowanych na rzecz zapewnienia otwartego, powszechnego i trwałego dostępu do polskiego dorobku naukowego i kulturalnego. Artykuł jest umieszczony w kolekcji cyfrowej bazhum.muzhp.pl, gromadzącej zawartość polskich czasopism humanistycznych i społecznych.

Tekst jest udostępniony do wykorzystania w ramach dozwolonego użytku.

Problem reprezentacji w teoriach poznania ucieleśnionego

Natika Newton

Department of Philosophy

Nassau County Community College, USA

przekład: Bartłomiej Świątczak

(tekst oryginalny pt. „Representation in Theories of Embodied Cognition” ukazał się w
Theoria Et Historia Scientiarum, Vol. VII, No. 1. 2003: 181–1931)

Abstract:

This paper looks at a central issue with embodiment theories in cognition: the role, if any, they provide for mental representation. Thelen and Smith (1994) hold that the concept of representations is either vacuous or misapplied in such systems. Others maintain a place for representations (e.g. Clark 1996), but are imprecise about their nature and role. It is difficult to understand what those could be if representations are understood in the same sense as that used by computationalists: fixed or long-lasting neural structures that represent the sensory stimuli that caused them (e.g. neural response patterns in the visual cortex), or whose “meaning” is fixed innately or in early development for particular functions (e.g. the body schemas for Meltzoff and Gopnik 1993). The paper proposes a distinctions between, on the one hand, neural patterns, traces of sensory activation that while not in themselves representations are available for representational activity, and on the other the act of representing, which is what gives representational content to neural patterns.

Keywords: *cognition, embodiment, neural patterns, representation, representational activity.*

¹ The translation is published with a kind permission of Holders of the copyright. / Przekład publikowany za uprzejmą zgodą właścicieli praw do tekstu.

Abstrakt:

Prezentowana praca podejmuje podstawowy problem teorii poznania ucieleśnionego. Chodzi o rolę, jaką te teorie przypisują reprezentacjom mentalnym. Thelen i Smith (1994) sądzą, że w takich koncepcjach pojęcie reprezentacji jest niewłaściwie stosowane albo nawet puste. Inni uważają, że teorie poznania ucieleśnionego przyznają pewne miejsce reprezentacji (np. Clark 1996), ale nie precyzują, jaka miałyby być ich natura i rola. Trudno zrozumieć, czym są reprezentacje, w takim sensie, w jakim mówią o nich komputacjoniści.

Uważają oni bowiem, że reprezentacje to utrwalone albo długotrwałe struktury neuronowe reprezentujące bodźce zmysłowe, które są przyczynami takich struktur (przykładem mogą być neuronowe wzorce reakcji w korze wzrokowej), albo też takie struktury neuronowe, których „znaczenie” dla poszczególnych funkcji jest wrodzone bądź wytworzone we wczesnym stadium rozwoju (np. schematy cieleśne Meltzoffa i Gopnik 1993). W niniejszym artykule proponuję dokonanie dystynkcji na: z jednej strony wzorce neuronowe, ślady aktywacji sensorycznej, które – choć nie dla samych reprezentacji – dostępne są w aktywności reprezentacyjnej, oraz z drugiej strony – na akt reprezentowania, który stanowi to, co nadaje treść reprezentacyjną wzorcom neuronowym.

Słowa kluczowe: *akt reprezentowania, poznanie, reprezentacja, ucieleśnienie, wzorce neuronalne.*

1. Poznanie ucieleśnione

W naukach kognitywnych świadomość i poznanie postrzega się coraz częściej jako samoorganizujące się systemy obejmujące emocję oraz organizację sensomotoryczną (np. Damasio 1994, 1999; Clark 1996; Glenberg, 1997; Hurley 1998). Stanowisko, według którego poznanie można najlepiej opisać jako ucieleśnione, zastępuje modele, w których przyjmuje się istnienie amodalnych systemów symbolicznych. Przykładem takich systemów są arbitralne, pozbawione znaczenia wewnętrznego manipulacje symboliczne programów komputerowych, które nie pozwalają w żaden sposób wyjaśnić zdroworozsądkowego rozumowania i świadomości. Koncepcja poznania ucieleśnionego zachowania tego rodzaju postrzega jako przedłużenie zwierzęcej interakcji ze środowiskiem. Zgodnie z modelem ucieleśnionego działania, aktywność poznawcza na wszystkich poziomach abstrakcji i złożoności, od niemowlęcego badania do kalkulacji w matematyce i logice, może być wyjaśniona w ramach jednej ramy pojęciowej. Ta rama pojęciowa, w przeciwieństwie do modelu komputacyjnego, może wyjaśnić zarówno intencjonalne, jak i funkcjonalne cechy ludzkiego poznania. Ta rama pojęciowa pozwala również sformułować intuicyjnie poprawną koncepcję znaczenia (Newton 1996; Glenberg 1997).

Model poznania ucieleśnionego jest w pełni kompatybilny ze znanymi mechanizmami mózgowymi odpowiedzialnymi za zachowanie motoryczne, poznanie i język, a na jego podstawie można sformułować predykcje weryfikowalne empirycznie. Model ten jest również kompatybilny z naturalistycznymi sposobami podejścia do świadomości, mimo że nie wszyscy teoretycy ucieleśnienia uwzględniają w swoich teoriach świadomość.

Można rozwijać pożyteczne, choć ograniczone, teorie poznania ucieleśnionego, nie podejmując jednocześnie problematyki świadomości.

Taki sposób podejścia przyjęli teoretycy systemów dynamicznych, tacy jak Thelen i Smith (1994) oraz Port i Van Gelder (1995). W swoim ostatnim artykule Thelen i inni (2001) argumentują, że pewne dobrze znane zjawisko poznawcze u niemowląt (błąd A-nie-B) może być najlepiej wyjaśnione w kategoriach dynamiki ucieleśnienia. Swoją teorię przedstawiają tak:

Podajemy formalną teorię dynamiki oraz model oparty na poznawczym ucieleśnieniu. Poznanie ucieleśnione zarówno stymuluje skutki zjawiska A-nie-B, jak i pozwala na sformułowanie nowatorskich predykcji, które pasują do nowych wyników eksperymentalnych. Ten dowód jest potwierdzeniem koncepcji ucieleśnienia dzięki temu, że zdarzenia mentalne, angażowane w procesie percepcji, planowania, podejmowania decyzji oraz przypominania, wyraża się w tym samym, dynamicznym języku, który jest używany do opisu ruchów cielesnych, tak, że mogą być one zawsze uchwycone w pewnej sieci. Uważamy, że ta sieć jest bardzo ważnym, kognitywnym aktem „poznawania”, nie tylko w okresie niemowlęctwa, ale również w codziennej aktywności, w przeciągu całego życia (Thelen i in. 2001: 1).

Chociaż przedstawiony tu model ma swoich krytyków, stanowi jednak wiarygodny opis procesu uczenia się przez niemowlęta (zob. komentarz w Thelen i in. 2001). Jednym z punktów krytycznych jest fakt, iż Thelen, w ramach poznania ucieleśnionego odmawia jakiegokolwiek roli świadomej, mentalnej reprezentacji. Chociaż będę argumentować, że takie odrzucenie świadomej, mentalnej reprezentacji jest błędne, jest to ważny sygnał przemawiający na rzecz wartości ucieleśnionego sposobu podejścia w ogóle, w tym sensie, że może być on wykorzystywany zarówno przez reprezentacjonalistów, jak i nierepresentacjonalistów. Ten sposób podejścia jest na tyle wiarygodny pod względem teoretycznym i intuicyjnym, że może być na wiele sposobów naginany i kształtowany, bez jednoczesnego popadania w jakąś niespójność. Mimo to będę argumentować, iż teorie poznania ucieleśnionego muszą uwzględnić świadome, mentalne reprezentacje, jeżeli tylko aspirują do wyjaśnienia pełnego repertuaru ludzkiego zachowania.

Jakiego rodzaju zachowanie powinna wyjaśniać pełna teoria poznania? Otóż powinna ona wyjaśniać zarówno proces konceptualnego uczenia się przez niemowłeta, jak i proces nabywania umiejętności; musi wyjaśniać źródło tradycyjnych dla „psychologii potocznej” (*folk psychology*) pojęć przekonania i pragnienia (niezależnie od tego, czy zostaną one jeszcze zachowane w takiej teorii).

Teoria taka powinna wyjaśnić również wyrafinowane umiejętności poznawcze, uwzględniając przy tym język potoczny i rozumowanie abstrakcyjne, a także – z powodów, o których mowa będzie poniżej – powinna wyjaśnić również świadomość fenomenalną. Jest rzeczą oczywistą, że świadome doświadczenie ma dla ludzi znaczenie emocjonalne. Fenomenologicznie rzecz ujmując, świadomość jest wyrazem zrozumienia naszych ucieleśnionych interakcji ze środowiskiem. Niezależnie od tego, jaka jest pełna funkcjonalna rola świadomości, spełnia ona oczywiście ważną rolę w naszym zdroworozsądkowym układzie przekonań i wartości. Wiele ważnych intuicji, które doprowadziły do teorii poznania ucieleśnionego, pochodzi z fenomenologicznych studiów nad świadomym doświadczeniem (np. Merleau-Ponty 1962, 1963; Husserl 1966). W pełni rozwinięta teoria musi uwzględnić i wyjaśnić te fakty; musi także, w taki lub inny sposób, uwzględnić i wyjaśnić naturę świadomego doświadczenia (zob. Ellis 1995, Ellis i Newton 1998, a także, aby zapoznać się z ostatnimi tego rodzaju próbami, Newton 2001). Jedną z zalet teorii poznania ucieleśnionego jest fakt, iż podejście to unifikuje różne aspekty i zdolności inteligentnych zwierząt oraz wyjaśnia ich miejsce w szeregu innych zwierząt. Pełne wyjaśnienie problemu świadomości umożliwiłoby nam zlokalizowanie jej w tym szeregu.

2. Reprezentacja

Teoria poznania ucieleśnionego ma, w swojej obecnej inkarnacji, swoich dwóch wielkich prekursorów: komputacjonizm oraz fenomenologię. Obie te koncepcje zasadniczo uwzględniają istnienie mentalnej reprezentacji. Teoria poznania ucieleśnionego zawdzięcza jednak więcej drugiej z wymienionych teorii aniżeli tej pierwszej: językiem, w którym toczy się dyskusja w ramach teorii poznania ucieleśnionego, jest język Husserla, Merleau-Ponty’ego, Sartre’a i innych. Ale teoria poznania ucieleśnionego jest również oparta na komputacjonizmie, ponieważ program badawczy dominujący w nim przez trzydzieści lat („It is the only game in town”, napisał Fodor w 1981) wyznaczył problemy, które współczesna teoria poznania ucieleśnionego stara się rozwiązać. Komputacjonizm starał się wyjaśnić mechanizmy poznawcze w taki sposób, aby (1) można było je realizować przez komputer, oraz (2) aby mogły być odwzorowywane (mapped) przez faktyczne mechanizmy mózgowie. Tradycyjna fenomenologia stawiała sobie inny cel. Starła się, dla swoich własnych potrzeb, zbadać dziedzinę zjawisk.

Opisać tę dziedzinę przy wykorzystaniu odpowiedniej terminologii oraz praw, aby rozwinąć tym samym nową naukę, naukę świadomego doświadczenia. Intuicje, które przy tej okazji rozwijano, ukształtowały teorie poznania ucieleśnionego, ale teorie te dzielą stawiane przez siebie cele raczej z komputacjonizmem aniżeli z celami właściwymi czystej, tradycyjnej fenomenologii.

Celem teorii poznania ucieleśnionego jest rozwinięcie nauki o poznaniu, która byłaby przedłużeniem nauk biologicznych oraz która mogłaby być modelowana w ramach teorii systemów dynamicznych. Pomimo takiego przeobrażenia, pojęcie intencjonalnych, mentalnych reprezentacji przetrwało w swojej oryginalnej postaci tak, jak zostało sformułowane przez Brentano (tj. jako „idee”). To Brentanowskie pojęcie funkcjonuje w ramach abstrakcyjnych systemów symbolicznych opisywanych przez komputacjonizm, a także w większości, choć nie we wszystkich, teoriach poznania ucieleśnionego.

Czym jest pojęcie intencjonalnych, mentalnych reprezentacji? Brentano wyjaśniał intencjonalność jako odniesienie albo „nakierowanie na coś”. Możemy myśleć o czymś jako znajdującym się niejako „przed umysłem” albo „obecnym dla świadomości” (Brentano 1960). Nie używał on terminu „reprezentacja”, używał on za to terminu „idee”. Miały one reprezentować albo „prezentować” rzeczy, których „dotyczą” myśli. Idee te są przedmiotami sądów (postaw propozycjonalnych) oraz stanów mentalnych, takich jak miłość czy nienawiść.

Przeskakując teraz do koncepcji Fodora, dowiadujemy się, że idee Brentanowskie są *symbolami*. Fodor swoją Reprezentacyjną Teorię Umysłu podsumowuje w sposób następujący:

- (1) Postawy propozycjonalne są relacyjne.
- (2) Wśród składników relacji znajdują się reprezentacje mentalne (często w dawniejszej literaturze nazywane „ideami”).
- (3) Reprezentacje mentalne (sic!) są symbolami: posiadają one zarówno cechy formalne, jak i semantyczne.
- (4) Reprezentacje mentalne spełniają swoje role przyczynowe na mocy ich własności formalnych.
- (5) Postawy propozycjonalne dziedziczą swoje własności semantyczne od tych reprezentacji mentalnych, które funkcjonują jako ich przedmioty. (Fodor 1981: 27)

Reprezentacje mentalne Fodora przypominają rzędy symboli opracowywane przez programy komputerowe; stąd mówi się o „komputacyjnej teorii umysłu”.

Teoria ta nie cieszy się już szeroką, bezkrytyczną akceptacją; odkryto w niej zbyt wiele nierozwiązywalnych problemów. Mimo to pojęcie reprezentacji mentalnej adaptowane przez teorię komputacyjną nie zostało – w ramach ucieleśnionego podejścia do poznania – zastąpione jakimś innym, bardziej odpowiednim pojęciem.

Problem polega na tym, że reprezentacje postrzegane są przez większość uczonych jako statyczne, niemal stałe rzędy symboli w postaci grup neuronowych, które są wzajemnie powiązane dzięki eksperymentalnemu treningowi (dzięki takim mechanizmom, jak synapsy Hebb'a) i reagują na odpowiednią daną wejściową. Takie grupy neuronowe, podobne do rzędów symboli w komputerze albo rzędów liter w słowach języka naturalnego, same w sobie niosą treść (niosą ją albo dzięki swojej „wewnętrznej intencjonalności” (Haugeland 1985), albo też, potencjalnie, czekając na wykorzystanie przez użytkownika (Dennett 1987)). Niezależnie od tego, które stanowisko zostanie przyjęte, grupy neuronowe same są reprezentacjami oraz nośnikami treści intencjonalnej myśli (jeżeli takowa istnieje). Wprawdzie Dennett wraz z Millikan (1984) uważają, że użytkownik określa treść reprezentacji, ale według nich, treść taka „określona przez użytkownika, jest [ostatecznie] ustalona przez ewolucję. Treść reprezentacji jest tym, do czego ewoluowały mechanizmy reprezentacyjne, aby osiągnąć „właściwą im funkcję”. Nowe intencje wyrastające ze spontanicznych, kreatywnych aktów użytkowników, które zajmują centralne miejsce w procesie ludzkiego poznania, nie są uwzględniane przez tego rodzaju stanowiska).

Takie pojęcie reprezentacji, jak powyższe, właściwe jest działalności neuronaukowców. Rozważmy przykład, który pojawił się ostatnio w artykule opisującym badania neuronów „kanonicznych” (czyli takich, które reagują na wizualne aspekty trójwymiarowych przedmiotów²) oraz neuronów lustrzanych (aktywnych zarówno wtedy, gdy realizowana jest pewna czynność, jak i wtedy, gdy obserwowany jest ten sam typ realizacji takiej samej czynności) w obszarze planowania ruchu (PMv, *preventral motor area*):

² Według Rizzolattiego i in. (2000) „przedmioty trójwymiarowe (...) są identyfikowane i rozróżniane nie ze względu na ich zwyczajny, fizyczny wygląd, ale relatywnie do efektu interakcji z działającym podmiotem (...). Neurony F5 [w *ventral premotor cortex*] aktywowane są tylko wtedy, gdy jakiś szczególny typ czynności (np. uchwycenie czegoś, trzymanie czegoś, itp.) jest realizowany po to, aby osiągnąć jakiegoś szczególnego typu cel (np. wejście w posiadanie pokarmu, wyrzucenie jakiegoś przedmiotu itp.)”.

Neurony kanoniczne zostały znalezione głównie w tym fragmencie PMv, który znajduje się w tylnej części bruzdy łukowatej (arcuate sulcus). W przypadku małp, neurony lustrzane są aktywne wtedy, gdy zwierzęta te wykonują określoną czynność oraz wtedy, gdy obserwują kogoś innego, kto wykonuje podobną czynność. Neurony te zostały w większości zidentyfikowane w tym fragmencie PMv, który znajduje się na powierzchni korowej w tyle bruzdy łukowatej. Obecność neuronów lustrzanych i kanonicznych wraz z odkrytym przez nas zjawiskiem kodowania zewnętrznego sugeruje, że PMv może zawierać przestrzenne reprezentacje czynności (Kakei i in. 2001).

Powyższy sposób użycia terminu „reprezentacja” zgodny jest ze stanowiskiem, według którego reprezentacje to grupy neuronów, które są bardziej lub mniej stabilnie powiązane ze sobą oraz dostępne dla organizmu do wykorzystania w różnych działaniach. Większość stanowisk głosi, że neurony, które reprezentują, są przyczynowymi pośrednikami między bodźcem a reakcją. Wpływają na nie przyczynowo zdarzenia zewnętrzne pobudzające mechanizmy zmysłowe lub zdarzenia wewnętrzne, które zostawiają ślady w ich wzorcach [pobudzeń] (tak jak ma to miejsce w przypadku reprezentacji motorycznej). Zdarzenia wewnętrzne z kolei odgrywają przyczynową rolę w zachowaniu, dostarczając wzorów, które są przewodnikami dla organizmu w planowaniu czynności oraz ich realizacji. Reprezentacje, do których odnosi się powyższy artykuł, są głęboko zakorzenione w takim kauzalnym mechanizmie, podobnie jak wiele innych reprezentacji zarówno wcześniejszych, jak i późniejszych, w sekwencji zdarzeń następujących między daną wejściową a daną wyjściową.

W jaki sposób zostaje określona treść takich reprezentacji? Według większości koncepcji: tym, co jest reprezentowane, są zewnętrzne przedmioty, których percepcja jest przyczyną powstania takich grup neuronowych. Zatem (w korze wzrokowej) różnorodne wzory neuronowe miałyby reprezentować na wiele różnych sposobów zewnętrzne, obserwowalne przedmioty. Wzory neuronowe wcześniejsze w sekwencji zdarzeń między daną wejścia a wyjścia reprezentują na przykład kształt i lokalizację bodźca, natomiast wzory późniejsze – reprezentują znane przedmioty skojarzone z innymi, nieznanymi przedmiotami. Cieleśna dana wejściowa również kształtuje reprezentacje (niektóre z nich mogą być wrodzone). Na przykład „mapy ciała” w różnych lokalizacjach, takich jak kora ciemieniowa, reprezentują różne aspekty ciała podmiotu dla celów jego orientacji, przygotowania czynności i utrzymania sprawności fizycznej. Treść jest również określona przez to, co mogą przyczynowo wywołać reprezentacje. Mówi się, że wzory ruchowe (motoryczne), które w momencie aktywacji wywołują przyczynowo określoną czynność, reprezentują taką czynność (Jeannerod 1998).

Wyjaśnienie treści reprezentacji rozumianych jako statyczne, trwałe grupy neuronowe niesie ze sobą ogromną trudność. Jednym ze słynnych problemów jest problem błędnej reprezentacji. Jeżeli wzory reprezentują swoje przyczyny, w jaki sposób aktywowane wzory mogłyby kiedykolwiek *błędnie* reprezentować swoje przyczyny? Tego rodzaju problemy doprowadziły niektórych badaczy, takich jak Thelen, do całkowitej rezygnacji z dyskusji na temat reprezentacji:

Reprezentacje w ich najmocniejszym, źródłowym i głębokim sensie są symbolami, które zastępują to, co jest reprezentowane, i są czymś różnym od mocy obliczeniowych, które na tych symbolach działają.

Zgodnie z tą podstawową definicją procesy sensomotoryczne zdecydowanie nie są reprezentacjami. Ostatnio jednak rozszerzono dziedzinę zdarzeń wewnętrznych, rozważanych jako reprezentacje. Zgodnie z tym nowym stanowiskiem każde wewnętrzne, dynamiczne zdarzenie, które przyczynowo odnosi się do zachowania, jest reprezentacją. Możemy się z tym zgodzić. Jeżeli wszyscy zgodzimy się, że istnieje tylko proces, i jeżeli odrzucimy dualistyczny podział na wiedzę (poznanie) i proces, to szczęśliwie jesteśmy reprezentacjonistami. Ale zwróćmy uwagę na konsekwencje teoretyczne takiej decyzji: zranione kolano staje się reprezentacją upadku, które dało początek temu bólowi. Nie wydaje się warte zachodu pytanie, czy jakaś teoria przyjmuje reprezentacje, czy też nie (Thelen i in. 2001).

Thelen ma rację, że niewybredne etykietowanie każdego kauzalnego komponentu sekwencji percepcyjno-behawioralnej jest bezcelowe. Ma ona również rację, odrzucając reprezentacje jako symbole, które będąc czymś różnym od procesów poznawczych, zastępują to, co jest reprezentowane. Ale mimo to nie można reprezentacjom mentalnym odmówić ważnej roli. W dalszej części niniejszego artykułu przeanalizuję dwa poważne problemy związane z traktowaniem reprezentacji jako ustalonych struktur neuronowych. Następnie zbadam pojęcie reprezentacji jako aktywności mentalnej. Termin „aktywność mentalna” jest tutaj istotny: nie jest on synonimiczny z terminem „stan mentalny”. Termin „aktywność mentalna” odnosi się do czynności intencjonalnej, podobnej do czynności ruchowej, która w przeciwieństwie do tej ostatniej znajduje się w ukryciu (jest nieujawniona), z drugiej jednak strony jest kontrolowana przez takie same struktury mózgowie (mózdzek, korę ruchową), jak te zaangażowane w czynność jawną. Myślenie, ogólnie rzecz biorąc, jest działaniem w tym sensie.

3. Reprezentacje jako nośniki o ustalonej treści.

Istnieją przynajmniej dwa problemy, które niesie ze sobą pojmowanie reprezentacji jako struktur mózgowych (nośników), których treść jest określona lokalnie przez historię i/albo relacje kauzalne tych struktur. Pierwszy problem jest taki, że teorie ucieleśnienia mają sens przede wszystkim jako metaforyczne rozszerzenie naszego pojęcia aktywności cielesnej, a nie istnieje odpowiednia cielesna analogia dla reprezentacji jako ustalonych wzorców neuronowych. Fakt ten utrudnia modelowanie. Aktywność cielesna zależna jest od potrzeb i motywów całego organizmu i nakierowuje się na te cechy środowiska, które te potrzeby zaspokajają. Zgodnie z większością teorii, te cechy są bezpośrednio dane w percepcji środowiska jako tak zwani „dostarczyciele”.

Cechy te nie mają żadnej wcześniejszej, „obiektywnie” znaczącej, treści dla ucieleśnionego podmiotu, ponieważ nabywają one treść tylko dzięki tej interakcji.

Jeżeli myślimy o poznaniu jako o przedłużeniu tego rodzaju aktywności, a jednocześnie postulujemy istnienie wzorów neuronowych jako posiadających znaczącą treść nawet wtedy, gdy wzory te „nie są używane”, to będziemy potrzebowali dwóch różnych teorii treści. Jedna teoria wyjaśni, w jaki sposób treść jest tworzona poprzez emocjonalnie kierowane czynności organizmu, kiedy organizm stara się on zaspokoić swoje potrzeby poprzez wykorzystanie dostępnych aspektów środowiska. Teoria ta będzie musiała pokazać, w jaki sposób to się dzieje. Podobnie jak w przypadku zwierzęcia zachowującego się w sposób dla niego naturalny, możemy pytać o to, co to zwierze „próbuję robić”, tak możemy pytać również w przypadku osoby, która w toku planowania czynności albo w trakcie rozwiązywania problemu tworzy obrazy zmysłowe, o to, do czego ta osoba „zmierza” poprzez takie obrazowanie. Załóżmy, że tym, co taka osoba obrazuje, jest aktywowany ślad pamięciowy oceanu zobaczony zeszłorocznego sierpnia na wyspie Fire. Nie wystarczy powiedzieć, że obraz tej osoby „wyraża” ocean, tak jak jawił się on jej zeszłego sierpnia. Ten obraz został aktywowany po to, aby mógł odegrać pewną rolę w działalności kognitywnej tej osoby: jest nią na przykład próba zdecydowania, czy wynająć domek letni nad jeziorem Michigan w przyszłym roku. Odgrywając taką rolę, ślady pamięciowe mają całkowicie nową treść, reprezentują przyszłe możliwości, które niosą teraz między innymi cechy emocjonalne, które są zupełnie obce pierwotnym wspomnieniom.

Druga teoria wyjaśni, w jaki sposób ślady neuronowe mogą posiadać treść „wewnętrznie”, nawet wtedy, gdy ślady te są zmagazynowane i nieaktywne (Karykaturą tej idei jest pojęcie „neuronu babcinego” (*“grandmother neuron”*)). Do analizy tych problemów wykorzystamy drugą teorię. Jednak nawet pomijając te problemy, należy stwierdzić, że przyjmowanie w ogólnej teorii poznania aż dwóch całkiem różnych typów reprezentacji wydaje się nieco rozrzutne i zawiłe.

Drugim problemem jest liczba wewnętrznych trudności w wyjaśnianiu treści przechowywanej we wzorach neuronowych. Trudności te uwidaczniają się w dyskusjach na temat treści językowej, ponieważ wzory neuronowe traktuje się podobnie jak wzory językowe. Co nadaje treść symbolom językowym? Teorie przyczynowe znane są z tego, że nie pozwalają wyjaśnić zjawiska błędnej reprezentacji. Poza tym teorie przyczynowe nie podają żadnego wyjaśnienia procesów twórczych, w których struktury o ustalonej treści nabywają nowych możliwości użycia, a jest to istotna i wszechobecna cecha ludzkiego poznania.

Tak jak Thelen i inni argumentują (zobacz powyżej), rozwiązywanie wspomnianych problemów w kategoriach przyczynowych prowadzi do praktycznie pustego wyjaśnienia, zgodnie z którym wszystko, co jest zaangażowane w system przyczynowy, jest reprezentacją.

Biorąc pod uwagę powyżej przedyskutowane problemy, moim zdaniem powinniśmy poszukać wyjaśnienia reprezentacji mentalnej, według której treść powstaje tylko w trakcie aktualnej aktywności poznawczej. Jeżeli pojęcie treści zarezerwujemy dla aktywnych komponentów zachowania reprezentacyjnego i jeżeli będziemy traktowali przechowywane neuronowo wzory jako ślady ich przyczyn, to te tradycyjne dla filozofii języka problemy znikną. Jeżeli tym, co nadaje symbolowi jego treść, jest sposób, w jaki ten symbol jest używany, to powinniśmy skoncentrować się na użyciu. Reprezentacje powstają w trakcie aktywnego procesu reprezentacji. Treść jest dostarczona przez kontekst i cel aktywności reprezentacyjnej. Nie jest ona w sposób gotowy upakowana w neuronowych wzorach rekrutowanych w aktywności reprezentacyjnej.

4. Reprezentacja w działaniu

Idea, według której reprezentacja powstaje w nakierowanej na cel czynności, nie jest niczym nowym. Oprócz pewnych zwiastunów w filozofii klasycznej (takich jak Arystotelesowska „entelechia”) i dziewiętnastowiecznych pragmatystów takich jak Peirce, idea ta została jasno przedstawiona między innymi przez Polanyi'ego:

Możemy powiedzieć ogólnie, że dzięki nabywaniu umiejętności, czy to mięsniowych, czy intelektualnych, osiągamy zrozumienie, którego nie możemy wyrazić w słowach i które jest przedłużeniem nieartykułowanych zdolności zwierzęcych.

To, co ja w ten sposób rozumiem, ma znaczenie dla mnie, przy czym ma to znaczenie [samo] w sobie, inaczej niż znak, który posiada znaczenie, kiedy denotuje jakiś przedmiot. Takie znaczenie wcześniej już nazwałem znaczeniem egzystencjalnym (Polanyi 1958: 90).

Realizowanie wyuczonej czynności jest zarazem ćwiczeniem zrozumienia tej czynności, to znaczy jej znaczenia. Czym jest to, co ktoś rozumie? Czym jest znaczenie? Sugerowałabym, że rozumie się to, co się próbuje robić albo to czemu służy czynność. Innymi słowy, znaczenie czynności jest celem czynności, jest tym, co ktoś przynajmniej milcząco zakłada, kiedy intencjonalnie realizuje czynność. Zatem możemy powiedzieć, że samo realizowanie czynności jest aktywnością reprezentowania określonego celu jako celu, oczekiwanego rezultatu czynności.

Cel nie jest reprezentowany odrębnie w postaci jakiegoś statycznego wzoru neuronowego, nie jest on separowalnym komponentem procesu realizacji czynności.

Podobne stanowisko przyjął Merleau-Ponty. Odrzuca on ideę „reprezentacji” jako niezależnego komponentu sekwencji zachowania; zamiast tego, postrzega on czynność w całości, jako posiadającą wewnętrzne znaczenie łączące środki i cel. Jeżeli cel i środki nie są powiązane wewnętrznie w czynności, traci ona znaczenie dla organizmu.

... dopóki świadomość jest określona przez posiadanie pewnych „reprezentacji”, (...) to świadomość aktu jest z konieczności zredukowana, z jednej strony do reprezentacji celu tego aktu, a z drugiej strony, być może, do reprezentacji cielesnych mechanizmów, które zapewniają realizację tego aktu. W tych warunkach relacja środków do celu może być tylko zewnętrzną.

Ale, jeżeli (...) świadomość reprezentacyjna jest tylko jedną z postaci świadomości i jeżeli ta ostatnia jest definiowana bardziej ogólnie przez odniesienie do przedmiotu – niezależnie od tego, czy przedmiot ten jest upragniony, oczekiwany, kochany czy reprezentowany – odczute ruchy zostaną powiązane ze sobą dzięki praktycznej intencji, która nimi porusza, która czyni je melodią, na którą świadomość nakierowała się, i staje się niemożliwe odróżnienie celu od środków jako dających się odseparować elementów, niemożliwe [staje się] traktowanie ludzkiej czynności jako alternatywnego rozwiązania problemów, które rozwiązuje instynkt. Jeżeli problemy te byłyby takie same, rozwiązania byłyby identyczne. [...] Analiza nieuchronnego znaczenia czynności i jego wewnętrznej struktury jest zastępowana analizą celów czynności i ich środków (Merleau-Ponty 1963: 173-174).

Czynność realizowana w kierunku jakiegoś celu jest czynnością reprezentowania celu *jako* celu przyświecającego środkom. Cel niezależny od środków nie byłby celem; środki niezależne od celu byłyby bezsensowne.

Nie chcę w ten sposób argumentować, że nie istnieją żadne wzory neuronowe, które mogą być używane reprezentacyjnie w czynnościach nakierowanych na cel. Na przykład Jeannerod (1998) dowodził w przekonujący sposób, że obrazowanie czynności, możliwe dzięki wcześniejszym realizacjom czynności, rekrutowane jest w ramach nowych realizacji tego typu czynności i odgrywa istotną rolę w inicjowaniu, kontrolowaniu i ocenie takiej czynności.

Można powiedzieć, że *wykorzystywane* obrazy czynności reprezentują uprzystępnianą przez nie czynność, ponieważ są one sensomotorycznymi obrazami zunifikowanej realizacji, w której środki i cele są ze sobą pomieszane i wywodzą swoje znaczenie jedne od drugich. Ale pozostając w zgodzie z tym faktem, można stwierdzić, że obrazy reprezentują czynność tylko wtedy, gdy faktycznie są wykorzystywane do tego celu. W przeciwnym razie są one zwykłymi potencjalnymi obrazami, potencjalnymi reprezentacjami czynności. Gdy są one aktualizowane w trakcie realizacji, obrazy te reprezentują egzemplarz czynności, który jednocześnie uprzystępniają.

Nie jest niczym abstrakcyjnym ani metafizycznie błędnym twierdzić, że obrazy reprezentują typy czynności. Proponuję następującą definicję reprezentacji w działaniu:

R reprezentuje przedmiot O albo stan rzeczy S wtedy i tylko wtedy, gdy za względu na pewien izomorfizm albo inną cechę korespondencji z O albo S, R odgrywa rolę O albo S w symulowanej czynności dotyczącej O albo S.

R może być dowolnym komponentem obrazu czynności, który jest w danym momencie wykorzystywany, kiedy obrazowana czynność znajduje się w interakcji z zewnętrznym przedmiotem, który spełnia [zarazem] rolę przedmiotu; przy czym „izomorfizm albo inna cecha korespondencji” jest nieskończenie różnorodna. Odegra taką rolę każda własność obrazu, która umożliwi interakcję (na przykład obraz wymaga wyłącznie odpowiednio wyobrażonych relacji przestrzennych).

Jeżeli grupy neuronowe, wykorzystywane w obrazie czynności, umożliwiają reprezentowanie przedmiotów związanych z odpowiednimi czynnościami, to można byłoby się spodziewać, że obraz czynności jako pewna całość reprezentuje czynność jako pewną całość. Tę relację można również określić w sposób następujący:

Obraz I reprezentuje egzemplarz czynności A *wtedy i tylko wtedy, gdy* I spełnia rolę A w każdej mentalnej symulacji obejmującej ocenianie, planowanie, aktywowanie, kontrolowanie albo kompletowanie A. Ale gdy obrazy znajdują się w stanie nieaktywnym, czyli reprezentują tylko potencjalnie, to faktycznie nie reprezentują one niczego, dopóki nie zostaną do tego celu aktywowane.

Zgodnie z podanymi wyżej definicjami, obrazy czynności zdolne są reprezentować uprzystępniane przez nie czynności oraz pozwalają one komponentom tych obrazów reprezentować komponenty czynności. Komponenty obrazów, w przypadku określonej symulacji czynności, reprezentują komponenty czynności w klasycznym sensie zastępowania albo spełniania ról funkcjonalnych. Nie reprezentują one jednak niczego w stanie nieaktywnym – na przykład wtedy, gdy nie są w takim celu aktywowane.

Reprezentacja jest czynnością; nie jest to zewnętrzna relacja między dwoma statycznymi przedmiotami.

Teorie poznania ucieleśnionego nie mogą kończyć analizy reprezentacji intencjonalnej na poziomie podstawowej realizacji czynności. Złożone zachowanie poznawcze obejmuje również ukrytą, abstrakcyjną myśl. W jaki sposób abstrakcyjne rozumowanie (np. logiczne i matematyczne) czyni użytek ze zdolności właściwych czynnościom cielesnym?

Wielu teoretyków argumentuje, że obrazowanie sensomotoryczne – świadomie albo nie w pełni świadomie aktywowane ślady pamięciowe doświadczeń z realizowania podstawowych czynności – funkcjonuje nie tylko w trakcie kontemplowania i planowania czynności, ale również w mentalnej manipulacji przedmiotami w abstrakcyjnym rozumowaniu. Abstrakcyjna myśl wykorzystuje podstawowe schematy czynności: cieleśne interakcje z przedmiotami w przestrzeni (np. Huttenlocher 1968; Newton 1996). Kiedy angażowane są tutaj reprezentacje mentalne, używane są one w mentalnych realizacjach wyobrażonych czynności. Ostatnia praca na temat mózdzka (Schmahmann 1997) pokazuje, że organ ten, tradycyjnie znany wyłącznie ze swojej roli w sekwencjonowaniu określonych planów czynności, angażowany jest zarówno w procesie abstrakcyjnego rozumowania, jak i – rozwiązywania problemów. Odkrycie to stanowi niezwykle potwierdzenie teorii poznania ucieleśnionego (Newton i in. 2002).

Oprócz pracy mózdzku, w pełni testowalna jest większość twierdzeń formułowanych w ramach teorii poznania ucieleśnionego, które dotyczą zjawiska reprezentacji w procesie obrazowania czynności. W stosunku do tych, którzy zaprzeczają, że istnieje takie obrazowanie można argumentować, że niekoniecznie wszystkie obrazy muszą być świadome, a dostępne obecnie badania nad obrazowaniem mózgowym mogą rozstrzygać takie sporne kwestie. Należy jednak wyjaśnić, co dokładnie rozumiemy, używając terminu „obraz”. Istnieje pewna skłonność do tego, by uważać wszystkie obrazy za wizualne. Ale pomimo [również] takiego tradycyjnego znaczenia terminu „obraz”, istnieje również poważny powód, by rozciągnąć znaczenie tego terminu, tak aby objął on wszystkie stany sensoryczne, motoryczne, proprioceptywne i afektywne, a nie tylko wizualne. Mechanizmy obrazowania wizualnego, które obejmują również wiele obszarów angażowanych w percepcji wizualnej, są całkowicie analogiczne do tych angażowanych w reaktywowanym doświadczeniu w ramach innych modalności. Na przykład obrazy ruchu związane z aktywnością w korze ruchowej okazują się identycznie do tych, które obecne są w rzeczywistym zachowaniu (Jeannerod 1998). Krótko mówiąc, możemy wyobrazić sobie wszystko, czego jesteśmy zdolni doświadczyć, co się pojawia, a słyszenie, dotykanie, smakowanie, odczuwanie emocji albo realizowanie określonych czynności motorycznych jest obrazowaniem tych czynności.

Jeżeli naprawdę możliwe jest dla nas tworzenie obrazów w dowolnej modalności, to wiele, jeżeli nie wszystkie, tradycyjnych zarzutów skierowanych przeciw stanowisku, według którego w poznaniu wykorzystuje się obrazy, upada. Na przykład Pylyshyn argumentował, że reprezentacja mentalna ma charakter sądu, a nie wyobrażenia, ponieważ możemy reprezentować relacje, których nie można pojąć obrazowo:

... podczas gdy dwa wizualne obrazy szachownicy mogą być obrazowo identyczne, to reprezentacja mentalna jednej z nich mogłaby obejmować relację pomiędzy dwiema figurami szachowymi, którą dałoby się opisać przez zwrot: „być atakowanym przez”, kiedy to reprezentacja mentalna drugiej szachownicy mogłaby takiej relacji nie obejmować. (...) W związku z tym można by oczekiwać, że reprezentacja mentalna pewnej konfiguracji figur na szachownicy byłaby o wiele bogatsza i lepiej ustrukturyzowana dla mistrza szachowego niż dla niedoświadczonego gracza w szachy (Pylyshyn 1973: 11).

Pylyshyn argumentuje, że wspomniana przez niego relacja nie jest częścią obrazu, ponieważ obraz dostępny jest tak samo mistrzowi, jak i nowicjuszowi, ale zawiera się w informacji reprezentowanej w postaci sądu, dostępnej tylko mistrzowi. Jeżeli jednak rozszerzymy pojęcie „obrazu” tak, by obejmowało ono doświadczenia we wszystkich modalnościach, łącznie z motorycznymi, proprioceptywnymi i emocjonalnymi, to możemy łatwo stworzyć obraz dwóch figur szachowych, które znajdują się w relacji „atakowania”. Jednym ze sposobów, w jaki możemy to uczynić, jest wywołanie sensomotorycznego obrazowania możliwych ruchów, tych dwóch figur, połączonego z afektywnym obrazowaniem hipotetycznych „stanów emocjonalnych” tych dwóch figur: agresji w przypadku jednej z nich, strachu albo defensywy w przypadku drugiej. Opis ten może brzmieć nienaturalnie, ale w rzeczywistości jest w pełni zgodny z językiem używanym przez Pylyshyna do opisanie tej relacji. „Bycie atakowanym przez” to relacja wewnętrznie powiązana z uczuciem. Być atakowanym to znajdować się w niebezpieczeństwie; a znajdowanie się w niebezpieczeństwie jest czymś złym. Szachy to gra abstrakcyjna, ale jeżeli abstrakcje nie są zakorzenione w konkretnych stanach, które możemy zrozumieć z perspektywy naszego własnego doświadczenia, to nie mogą mieć one dla nas żadnego znaczenia, a przez to i użyteczności.

Te konkretne stany są reprezentacjami czynności osadzonej zawsze w kontekście sytuacyjnym, który jest szerszy niż po prostu motoryczne ruchy cielesne. Czynności wymagają motywacji. Nawet subtelne czynności, takie jak utajona uwaga, zmieniają się zależnie od emocjonalnych zainteresowań organizmu.

Struktury podkorowe, takie jak migdałek, hipokamp i podwzgórze, wpływają na mechanizmy związane z wolą skupiania uwagi w przedniej obręczy. Czynności wyobrażone, ale niezrealizowane, są zarówno aktywowane, jak i hamowane w płatach czołowych i korze ruchowej. Dzięki hamowaniu kontrolowanemu w dużej części przez podwzgórze, możliwe jest świadome doświadczenie obrazów czynności (Jeannerod 1998) wraz z wartościami emocjonalnymi skojarzonymi zarówno z czynnościami, jak i z aspektami materialnego świata, od których zależy czynność.

Nie sposób tutaj w pełni obronić teorii poznania jako teorii zakorzenionej w obrazowaniu sensomotorycznym, to znaczy teorii poznania jako poznania ucieleśnionego. Skupiliśmy się na problemie reprezentacji i argumentowaliśmy, że co prawda pojęcie to jest niezbędne, ale należy je tak skorygować, aby spełniało wymagania ucieleśnienia jako podstawy poznania. Oczywiście, trzeba wykonać znacznie więcej pracy, aby rozwiązać pozostałe niejasności, które niesie ze sobą ta koncepcja, oraz umieścić ją na miejscu wciąż jeszcze silnego rywala, to znaczy stanowiska, zgodnie z którym reprezentacje są statycznymi, półtrwałymi obiektami mózgowymi o ustalonej treści³.

Bibliografia:

- Brentano, F. 1960. The distinction between mental and physical phenomena. W: *Realism and the Background of Phenomenology*, Chisolm, R.M. Red.164-176. Glencoe, IL: Free Press.
- Clark, A. 1996. *Being There*. Cambridge: MIT Press.
- Damasio, A. 1999. *The Feeling of What Happens*. New York: Harcourt Brace.
- Damasio, A. 1994. *Descartes' Error*. New York: Putnam and Sons.
- Dennett, D. 1987. *The Intentional Stance*. Cambridge, Mass.: MIT Press.
- Ellis, R. 1995. *Questioning Consciousness*. Amsterdam: John Benjamins.
- Ellis, R. i Newton, N. 1998. Three paradoxes of phenomenal consciousness: Bridging the explanatory gap. *Journal of Consciousness Studies* 5, 4: 419-42.
- Fodor, J. 1981. *Representations*. Cambridge, Mass.: MIT Press.
- Glenberg, A. 1997. What memory is for. *Behavioral and Brain Sciences*.
- Haugeland, J. 1979. Understanding natural language. *The Journal of Philosophy* 76: 619- 632.
- Hurley, S. 1998. *Consciousness in Action*. Cambridge: Oxford University Press.
- Husserl, E. 1966. *The Phenomenology of Internal Time Consciousness*. Przeł. J. Churchill. Bloomington: Indiana University Press.
- Huttenlocher, J. 1968. Constructing spatial images: A strategy in reasoning. *Psychological Review* 75, 286-298.
- Jeannerod, M. 1997. *The Cognitive Neuroscience of Action*. Oxford: Blackwell.
- Kakei, S., Hoffman, D. i Strick, P. 2001. Direction of action is represented in the ventral premotor cortex. *Nature Neuroscience* 4, 10: 1020-1025.

³ Dziękuję Ralphowi D. Ellisowi za pomocne sugestie podczas pisania tego artykułu.

- Lakoff, G. 1987. *Women, Fire and Dangerous Things*. Chicago: University of Chicago Press.
- Meltzoff, A. i Gopnik, A. 1993. The role of imitation in understanding persons and developing theories of mind: 333-366. Red. S. Baron-Cohen, H. Tager-Flusberg i D. Cohen. *Understanding Other Minds: Perspectives from Autism*. Oxford: Oxford University Press.
- Merleau-Ponty, M. 1962. *Phenomenology of Perception*. London: Routledge and Kegan Paul. Polski przekład: *Fenomenologia percepcji*. 2001. Tłum. J. Migasiński. Warszawa: Wydawnictwo Aletheia.
- Merleau-Ponty, M. 1963. *The Structure of Behavior*. Boston: Beacon Press.
- Millikan, R. 1984. *Language, Thought, and Other Biological Categories*. Cambridge, Mass.: MIT Press.
- Newton, N. 1996. *Foundations of Understanding*. Amsterdam: John Benjamins.
- Newton, N., Anderson, C., Schmahmann, J., i Ellis, R. 2002. The role of the cerebellum in cognition and affect. *Emotion and Cognition* 2, 2.
- Polanyi, M. 1958. *Personal Knowledge: Towards a Post-Critical Philosophy*. Chicago: University of Chicago Press.
- Port, R. i Van Gelder, T. 1995. *Mind as Motion: Explorations in the Dynamics of Cognition*. Cambridge, Mass.: MIT Press.
- Pylyshyn, Z. 1973. What the mind's eye tells the mind's brain. *Psychological Bulletin* 80, 1-23.
- Rizzolatti, G., Fogassi, L. i Gallese, V. 2000. Cortical mechanisms subserving object grasping and action recognition: A new view on the cortical motor functions. W: Gazzaniga 2000. *The New Cognitive Neurosciences*. Second edition. Cambridge MA: MIT Press.
- Schmahmann, J. Red. 1997. *The Cerebellum and Cognition*. Boston: Academic Press.
- Thelen, E., Schoner, G. i Scheier, C. 2001. The dynamics of embodiment: A field theory infant perseverative reaching. *Behavioral and Brain Sciences* 24, 1: 1-86.
- Thelen, E. i Smith, L.B., 1994. *A Dynamics Systems Approach to the Development of Perception and Action*. Cambridge, Mass.: MIT Press.
- Varela, F.J., Thompson, E. i Rosch, E., 1991. *The Embodied Mind: Cognitive Science and Human Experience*. Cambridge Mass.: MIT Press.