
Hipoteza umysłu normatywnego

BARTOSZ BROŻEK

*Katedra Filozofii Prawa i Etyki Prawniczej UJ
Centrum Kopernika Badań Interdyscyplinarnych*

Streszczenie. *Głównym celem artykułu jest analiza filozoficznych konsekwencji niektórych teorii wypracowanych w ostatnich latach w kognitywistyce i neuronaukach. Uznanie, iż umysł jest ucieleśniony (embodied), osadzony w interakcjach społecznych (embedded) oraz zorientowany na działanie (enacted) prowadzi do odrzucenia trzech tez charakterystycznych dla nowożytnej wizji umysłu: dualizmów umysł-ciało, umysł-inne umysły, oraz percepcja-wola. Odrzucenie tych dualizmów wspiera z kolei tezę, że umysł jest u swych korzeni normatywny, tj. że jego pierwotną funkcją jest działanie nakierowane na rozwiązywanie napotykanym przez nas problemów, zaś takie deskryptywne funkcje umysłu jak konstruowanie teorii czy snucie historii mają charakter wtórny — są nadbudowane nad naszymi normatywnymi zdolnościami.*

Słowa kluczowe: *umysł, normatywność, ucieleśnienie, osadzenie w kulturze, zorientowanie na działanie*

Moim celem jest zarysowanie hipotezy¹ umysłu normatywnego — obrona poglądu, że umysł jest, u swych korzeni, normatywny, nawet jeśli jego najbardziej wyrafinowane wytwory, takie jak matematyka, fizyka czy literatura są, przynajmniej w pewnym sensie, deskryptywne. W pierwszej części zarysuję trzy dogmaty kartezjanizmu: dualizm umysł-ciało, umysł-inne umysły, oraz percepcja-wola. Pokazuję następnie, że ustalenia współczesnych nauk kognitywnych i ewolucyjnych kwestionują wszystkie te dogmaty, zastępując je wizją umysłu, który jest

¹Używając słowa „hipoteza” nie chcę sugerować, że broniona przeze mnie koncepcja ma charakter naukowy i że, w związku z tym, można ją zweryfikować lub sfalsyfikować. Jest to koncepcja filozoficzna. Uważam przy tym jednak, że nie istnieje metodologiczna przepaść między teoriami konstruowanymi na gruncie filozofii i nauk kognitywnych. W szczególności, twierdzenia filozoficzne także mogą mieć charakter hipotez, czyli sądów, które podtrzymujemy dopóki nie przemawiają przeciw nim wystarczająco mocne argumenty (por. Brożek 2009).

ucieleśniony (*embodied*), osadzony w praktykach społecznych (*embedded*) i zorientowany na działanie (*enacted*). Twierdzą dalej, że ta nowa perspektywa współgra z hipotezą umysłu normatywnego, wedle której pierwotną funkcją umysłu jest rozwiązywanie problemów, a nie posiadanie przekonań.

1. Trzy dogmaty kartezjanizmu

Upraszczając nieco, można powiedzieć, że wizja ludzkiego umysłu zarysowana przez Kartezjusza, a przyjmowana przez wielu myślicieli od siedemnastego wieku do dzisiaj, opiera się na trzech twierdzeniach, czy wręcz dogmatach, które sformułować można w następujący sposób:

(Dogmat 1) *Dualizm umysł-ciało*: umysł jest ontologicznie odrębny od ciała. Stanowiska tego Kartezjusz broni *explicite* w *Medytacjach*, gdzie pisze:

Na podstawie samego faktu, że poznaję w sposób pewny, iż istnieję, oraz że nie stwierdzam wcale, aby cokolwiek innego musiało z koniecznością należeć do mej natury czy istoty, poza tym, że jestem czymś myślącym, wnioskuję z przekonaniem, że moja istota polega wyłącznie na tym, że jest czymś myślącym, a więc substancją, której całą istotą czy naturą jest myśleć. I chociaż być może, czy raczej na pewno, jak to powiem wkrótce, mam ciało, z którym jestem bardzo ściśle związany, to jednak, jako że z jednej strony mam jasną i wyraźną ideę siebie samego jako czegoś będącego wyłącznie czymś myślącym, a nie przestrzennym, a z drugiej posiadam dokładną ideę ciała jako czegoś będącego wyłącznie rzeczą przestrzenną, a nie myślącą, pewne jest, że ja, to znaczy moja dusza, przez którą jestem tym, czym jestem, jest całkowicie i prawdziwie różna od mojego ciała i że może ona być czy istnieć bez niego (Kartezjusz 2004, s. 73).

Oczywiście, Kartezjusz opisuje także związek pomiędzy umysłem i ciałem na inne sposoby, stwierdzając nawet w jednym z listów do Regiusa, że są one realnie złączone: „I gdziekolwiek nadarzy się okazja, powinienes oświadczać, zarówno prywatnie, jak publicznie, że jesteś przekonany, iż człowiek *to prawdziwy byt przez się, nie zaś przez przypadłość*, a umysł jest złączony z ciałem realnie i substancjalnie” (Kartezjusz 1996, s. 29). Co więcej, w kontekście niektórych problemów — na przykład natury emocji — autor *Medytacji* zmuszony jest podkreślać, jak ścisły jest to związek:

Ujmowanie, chcenie i wszystkie modyfikacje tak ujmowania, jak i chcenia, odnoszą się do substancji myślącej; do rozciąglej natomiast — wielkość (...), kształt, ruch, położenie, podzielność samych części i in-

ne takie rzeczy. Ale doświadczamy w sobie i jakichś innych, których nie można odnieść ani do samego umysłu, ani do samego ciała. Te (...) wypływają ze ścisłego i najtajniejszego związku naszego umysłu z ciałem; mianowicie pożądanie głodu, pragnienia, itd. A tak samo poruszenia, czyli doznania duszy, które nie polegają na samym myśleniu, jak poruszenie gniewu, wesołości, smutku, miłości itd. A wreszcie wszystkie wrażenia zmysłowe, jak wrażenie bólu, laskotania, światła i barw, dźwięków, zapachów, smaków, ciepła, twardości i innych jakości dotykowych (Kartezjusz 2001, s. 42–43).

Wszystko to nie zmienia jednak podstawowego założenia kartezjanizmu w odniesieniu do umysłu: jest on ontologicznie różny od ciała, nawet jeśli funkcjonuje w ścisłej z nim współpracy.

(Dogmat 2) *Dualizm umysł-inne umysły: umysł każdej osoby jest oddzielony od umysłów innych osób.* Realizując swój wielki projekt ugruntowania wiedzy, Kartezjusz zaczyna od potwierdzenia własnego istnienia jako *res cogitans*, umysłu (powiedzielibyśmy dziś), który „jest całkowicie niepodzielny. Wobec tego, gdy zastanawiam się nad nim, to znaczy, gdy zastanawiam się nad samym sobą jako czymś, co jedynie myśli, to nie mogę wyróżnić w sobie żadnych części, lecz poznaję i pojmuję nader jasno, że jestem rzeczą absolutnie jedną i całkowitą” (Kartezjusz 2004, s. 93). Wewnątrz tego „metafizycznego punktu” Kartezjusz odnajduje ideę nieskończonego bytu, która pozwala „dowieść” istnienia Boga, co z kolei gwarantuje istnienie świata poza umysłem. W świecie tym istnieją inni ludzie:

Gdybym oto przypadkiem ujrzał za oknem ludzi idących ulicą, a widząc to zwykłem mówić, że widzę ludzi, podobnie jak też i mówię, że widzę wosk, to przecież cóż innego byłbym widział, niż kapelusze i płaszcze, które mogłyby okrywać sztuczne roboty, poruszające się dzięki mechanizmowi? A jednak sądzę, że są to ludzie, a znaczy to, że rozumiem, dzięki samej tylko zdolności pojmowania, która tkwi w mojej duszy, to, co — jak mi się wydawało — widziałem za pomocą oczu (Kartezjusz 2004, s. 46).

Zatem Kartezjusz twierdzi, że istnienie innych umysłów stwierdza się w następujący sposób: potwierdziwszy istnienie *ego cogitans* możemy następnie dowieść istnienia Boga, który gwarantuje istnienie innych ciał, a wśród tych ciał są takie, którym przypisać można „posiadanie” umysłu. To skomplikowane rozumowanie zakłada, oczywiście, że umysł każdego z nas jest ściśle oddzielony od umysłów innych ludzi.

(Dogmat 3) *Dualizm percepcja-wola: umysł ma dwie odrębne zdolności, percepcję i wolę.* W *Zasadach filozofii* Kartezjusz stwierdza:

W istocie wszystkie modyfikacje myślenia, których w sobie doświadczamy, można sprowadzić do dwóch; z tych jedna to ujmowanie (*perceptio*), czyli operacja intelektu, druga zaś to chcenie, czyli operacja woli. Bo odczuwanie, wyobrażanie sobie i czyste rozumienie to są tylko różne modyfikacje odczuwania, tak jak pożądanie, wstręt, twierdzenie, przeczenie, wątpienie są różnymi modyfikacjami chcenia (Kartezjusz 2001, s. 38).

W kontekście tego fragmentu trzeba zwrócić uwagę na dwie kwestie. Po pierwsze, dla Kartezjusza percepcja obejmuje kontemplację dowolnych — propozycjonalnych i niepropozycjonalnych — treści psychicznych: danych zmysłowych, wytworów wyobraźni i rzeczy „czysto myślonych”. Po drugie, autor *Medytacji*, zgodnie zresztą ze współczesnymi tendencjami, uznaje, że myślenie jest analogiczne do widzenia: myśleć to tyle, co „widzieć oczami duszy”. Mówiąc precyzyjniej, Kartezjusz zdaje się twierdzić, że myślenie i widzenie to dwie formy tej samej zdolności postrzegania (por. Hacking 1975). Trzeba podkreślić, że w takim ujęciu myślenie jest procesem pasywnym, a staje się elementem bardziej aktywnych funkcji umysłowych dopiero wtedy, gdy na scenę wkracza wola, pozwalająca czemuś przyświadczyć, coś odrzucić lub w coś zwątpić. Związek między percepcją i wolą zarysowuje Kartezjusz w *Zasadach filozofii*:

Tak więc do sądzenia potrzebny jest wprawdzie intelekt, ponieważ o rzeczy, której w żaden sposób nie ujęliśmy, nie możemy niczego sądzić, ale potrzebna jest do tego także i wola, aby rzeczy w pewien sposób ujętej udzielić zgody. Nie jest natomiast potrzebne (co najmniej dla jakiegokolwiek sądzenia) całkowite i wszechstronne ujęcie rzeczy; możemy bowiem zgadzać się na wiele rzeczy, które ujmujemy tylko bardzo niejasno i niewyraźnie (Kartezjusz 2001, s. 39).

Pokazuje to, że jakaś forma percepcji musi poprzedzać działanie woli — jest zatem uprzednia wobec sądzenia i działania. W tej wizji sądzenie i działanie są zależne od (pasywnej) percepcji, która — warto to raz jeszcze podkreślić — obejmuje nie tylko doświadczenie zmysłowe, ale także wyobraźnię i kontemplację treści propozycjonalnych.

Opisane wyżej trzy kartezjańskie dogmaty tworzą wizję *umysłu deskryptywnego*, która zakłada, że umysł jest oddzielony od ciała i od innych umysłów, a jego główną, a przynajmniej pierwotną funkcją jest pasywne rejestrowanie bodźców ze środowiska i „obserwacja” idei pojawiających się na scenie „teatru myśli”. Tylko w oparciu o tak szeroko pojęte zdolności percepcyjne umysł — korzystając z woli — może formować sądy i podejmować działania. Mówiąc inaczej, podstawową funkcją umysłu deskryptywnego jest tworzenie adekwatnej reprezentacji świata, a podejmowanie działań jest czymś wobec niej wtórnym.

Wizji tej przeciwstawić można ideę umysłu normatywnego, czyli takiego, w którym reprezentowanie świata jest pochodne wobec pierwotnej jego zdolności, jaką jest podejmowanie działań zmierzających do rozwiązania problemów napotykanym przez organizm².

Nie twierdzę, że wszelka filozofia od czasów Kartezjusza opierała się na wizji umysłu deskryptywnego — istnieje wiele wyjątków³. Co więcej, można bronić tezy, że sam Kartezjusz nie był ślepym wyznawcą zaprezentowanych powyżej dogmatów, a jego teorie były bardziej subtelne; twierdzi się na przykład, że wcale nie uznawał pełnokrwistego dualizmu umysłu i ciała oraz percepcji i woli (por. Nelson 1997). Zaprezentowana przeze mnie rekonstrukcja nie ma jednak stanowić egzegezy stanowiska tego czy innego filozofa, ale podkreślać tendencje obecne w dominującym od siedemnastego wieku obrazie świata. Może i Kartezjusz myślał o związku umysłu i ciała jako o „niemal doskonałej unii”, a interakcje percepcji i woli przedstawił w pełni ich złożoności. Jednakże to, co z filozofii kartezjańskiej i pokartezjańskiej przeniknęło do nowożytnego obrazu świata, to tezy o trzech dualizmach: umysłu i ciała, umysłu i innych umysłów oraz percepcji i woli — czyli koncepcja umysłu deskryptywnego.

2. Umysł 3E

Różne aspekty koncepcji umysłu deskryptywnego krytykowano i modyfikowano w wieku dziewiętnastym i dwudziestym. Wydaje się jednak, że nie podjęto systematycznej próby przededefiniowania układu pojęciowego związanego z naszym rozumieniem umysłu. Impulsu dla takiej paradygmatycznej zmiany dostarczyła dopiero w ostatnich latach neurobiologia i teoria ewolucji.

Dualizm umysłu i ciała odrzucony został, między innymi, przez zwolenników idei *ucieleśnienia* (*embodiment*). Wedle tej koncepcji, „poznanie usytuowane jest w interakcji ciała ze światem; dynamiczne procesy

²Być może sformułowania „umysł deskryptywny” i „umysł normatywny” nie są najbardziej fortunne i lepiej byłoby mówić o „umyśle pasywnym” i „umyśle aktywnym” lub „kontemplacyjnym” i „proaktywnym”, czy nawet — odwołując się do Kanta — o „teoretycznym” i „praktycznym”. Pozostają jednak przy tej kontrowersyjnej terminologii, bo dobrze podkreśla ona główną tezę tego eseju — że funkcje umysłu pozwalające na *opis rzeczywistości* są pochodne względem jego funkcji umożliwiających *działanie w świecie*.

³Wśród tych, którzy kwestionują choć jeden z trzech kartezjańskich dogmatów, znajdują się na pewno kantyści i neokantyści, amerykańscy pragmatyści, niektórzy filozofowie analityczni (np. Gilbert Ryle i P. F. Strawson) czy przedstawiciele hermeneutyki.

cielesne, takie jak działania motoryczne, mogą być częścią procesu rozumowania, a poznanie ‘offline’ także bazuje na ciele” (Hutchins 2010, s. 428). Innymi słowy, teza, że poznanie jest ucieleśnione, sprowadza się do podkreślania, że nasze myślenie jest w sposób zasadniczy kształtowane przez interakcje naszych ciał ze środowiskiem. Niektóre z ostatnich odkryć biologicznych sugerują, że nasz aparat pojęciowy opiera się na architekturze neuronalnej odpowiedzialnej pierwotnie za motorykę. Dla przykładu: twierdzi się, że ludzki język bazuje na schematach, które są pochodną neuronalnych programów kontroli motorycznej (por. Brożek, Hohol 2014). Lakoff i jego zwolennicy zauważają, że tworzone przez nas abstrakcyjne pojęcia są wynikiem „metaforyzacji” pojęć bardziej konkretnych, wyrażających relacje przestrzenne i bezpośrednio zakorzenionych w schematach motorycznych (por. Brożek, Hohol 2014). Uważa się także, że język nie opiera się na żadnej wrodzonej gramatyce uniwersalnej, ale wyewoluował z gestu poprzez proces konwencjonalizacji (por. Tomasello 1995).

Dualizm *umysł-inne umysły* odrzucany jest, w pewnym przynajmniej sensie, przez zwolenników wizji umysłu osadzonego w kulturze (*embedded mind*). Teoria osadzenia — tak jak ją tu rozumiem — głosi, że interakcje z innymi ludźmi są warunkiem koniecznym rozwinięcia się w pełni funkcjonującego ludzkiego umysłu. Nie jest to trywialne stwierdzenie; w szczególności, nie chodzi o to, że prawidłowy rozwój umysłu wymaga odpowiedniego środowiska społecznego, a raczej o to, że bez systemu interakcji społecznych w ogóle nie byłoby ludzkiego umysłu. Tezę tę można lepiej zrozumieć w kontekście dyskusji wokół problemu przypisywania innym stanów umysłowych (*mindreading*). Wspominałem o tej kwestii analizując kartezjańskie rozumowanie, prowadzące do ustalenia istnienia innych umysłów. Przebiega ono — przypomnijmy — tak, że uznajemy istnienie innych umysłów wtedy, gdy wnioskujemy *per analogiam*, opierając się na naszym „prywatnym” doświadczeniu istnienia jaźni. Podobny pogląd wyrażają zwolennicy tzw. inferencyjnej koncepcji „odczytywania umysłu” — zdolność do przypisywania innym stanów psychicznych opiera się na „teorii” skonstruowanej z perspektywy pierwszoosobowej, a mówiącej, co to znaczy mieć stany umysłu („Zachowuję się w sposób *x*, kiedy czuję ból; zatem, jeśli on zachowuje się w sposób *x*, to znaczy, że czuje ból”). W ujęciu tym poczucie tożsamości osobowej jest czymś, co wyprzedza interakcje z innymi umysłami (Hohol, Urbańczyk 2013). Istnieje jednak ujęcie alternatywne, które głosi, że rudymetarna forma poznania społecznego jest uprzednia wobec poczucia tożsamości. Na przykład Vittorio Gallese formułuje *hipotezę wspólnej rozmaitości* (*shared manifold hypothesis*) (por. Gallese 2005). „Rozmaitość” ta stanowi prostą, intersubiektywną przestrzeń

informacji, bazującą na mechanizmie neuronów lustrzanych. Sposób działania tych neuronów sugeruje, że umysł koduje „własne” zachowania tak samo jak zachowania „cudze”. Pozwala to na pojawienie się rudymen tarnej formy intersubiektywności:

Ta wspólna przestrzeń umożliwia społeczne zapoczątkowanie rozwoju kognitywnego i afektywnego. Skoro tylko kluczowe więzi ze światem innych ludzi zostają nawiązane, przestrzeń ta stanowi podstawę w pełni rozwiniętej zdolności do rozumienia społecznej tożsamości i różnicy („Jestem oddzielnym podmiotem”). (...) Wspólna przestrzeń stanowi [zatem] niezwykle potężne narzędzie wykrywania i wprowadzania spójności, regularności i przewidywalności do indywidualnych interakcji ze środowiskiem. (...) Równoczesny rozwój języka wpływa zapewne na wydzielenie pojedynczych aspektów czy modalności doświadczenia z pierwotnego wieloaspektowego postrzegania świata, ale wspólna intersubiektywna przestrzeń nie znika, lecz stopniowo przyjmuje inną rolę: pozwala nam na doświadczanie różnic i podobieństw w relacji „ja-inny”. (...) Twierdzę, że odczucie „swojności” (*selfness*), którą przypisujemy też innym, wewnętrzne poczucie „bycia jak ja” wzbudzone kontaktami z innymi, jest wynikiem istnienia tej „wymieszanej” intersubiektywnej przestrzeni. Fizyczne i epistemiczne interakcje „ja-inny” są kształtowane i warunkowane przez podobieństwo organizmów i ograniczeń środowiskowych (Gallese 2005, s. 106).

Gallese twierdzi zatem, że „wspólna różnorodność” jest uprzednia wobec odróżnienia „ja-inny”, i to zarazem w wymiarze filogenetycznym, jak i ontogenetycznym. Ta teza, jeśli tylko jest choćby z grubsza poprawna, wskazuje, że rudymen tar na, biologicznie warunkowana forma intersubiektywności pojawiła się wcześniej w filogenezie i pojawia się wcześniej w ontogenezie. W konsekwencji można bronić twierdzenia, że zarówno w ewolucji naszego gatunku, jak i w rozwoju osobniczym, społeczny wymiar poznania jest warunkiem pojawienia się umysłu.

Wreszcie, dualizm *percepcja-wola* kwestionują ci, którzy opowiadają się za koncepcją umysłu zorientowanego na działanie (*enacted mind*)⁴ „Idea zorientowania na działanie głosi, że organizmy tworzą swe doświadczenia podejmując działania. Organizmy nie są pasywnymi odbiorcami danych ze środowiska, ale aktorami w tym środowisku, a to,

⁴Pogląd zwany *enaktywizmem* rozpropagowali Varela, Thompson i Rosch. Dziś ten paradygmat analizowania umysłu, stanowiący zdaniem niektórych konkurencję dla kognitywizmu, obejmuje wiele różnorodnych, czasem wzajemnie wykluczających się tez. Dodać też trzeba, że — w większości ujęć — enaktywizm obejmuje teorię ucieleśnienia. W niniejszym szkicu nie zajmuję się jednak tymi, często subtelnymi, rozróżnieniami. Interesuje mnie jedynie ten aspekt enaktywizmu, który związany jest z poglądem, że poznanie jest ściśle powiązane z naszym działaniem w świecie.

czego doświadczają, kształtowane jest przez to, jak działają” (Hutchins 2010, s. 428). Z tej perspektywy, myślenia nie można — jak chciał Kartezjusz — wyjaśniać, odwołując się do metafory widzenia. Trzeba raczej powiedzieć, że zarówno myślenie, jak i widzenie, są dwoma typami działania. Alva Noë zauważa:

Postrzeganie nabiera treści dzięki zdolnościom naszych ciał. To, co postrzegamy, jest określone przez to, co robimy (i co kiedykolwiek potrafimy zrobić); jest kształtowane przez to, co jesteśmy gotowi zrobić (...). Wcielamy w życie nasze doświadczenia percepcyjne. Odgrywamy je (*we enact our perceptual experience; we act it out*) (Noe 2004, s. 1).

I dodaje:

Bez wątplenia percepcja zależy od tego, co dzieje się w naszych mózgach, w których z pewnością tworzą się reprezentacje [świata zewnętrznego]. Jednak percepcja nie jest procesem toczącym się w mózgu, ale aktywnością całego zwierzęcego organizmu (Noe 2004, s. 2).

Noë podkreśla tu, że nie postrzegamy świata jako czegoś danego niezależnie od naszych zachowań, całego naraz, ale stopniowo budujemy jego umysłową reprezentację, zawsze w kontekście — i na potrzeby — podejmowanych przez nas działań.

Mówiąc ogólniej, wszelkie formy postrzegania i myślenia są rodzajem działania, a zatem są — pierwotnie — nakierowane na coś lub na kogoś. Myśląc, zaczynamy od pewnych oczekiwań (filozof powiedziałby zapewne: hipotez lub przedrozumienia), które można w trakcie tego procesu modyfikować lub odrzucać. Zatem wedle koncepcji umysłu zorientowanego na działanie, widzenie i myślenie nie są pasywnymi stanami poznawczymi. Trzeba odrzucić nieadekwatną wizję, wedle której widzenie sprowadza się do rejestracji bodźców docierających do nas z otoczenia, a co za tym idzie — także Kartezjańską metaforę myślenia jako widzenia.

Tak się jednak składa, że inny historycznie doniosły model myślenia — jako rozmowy duszy z samą sobą — także nie sprawdza się w interesującym nas kontekście. Konieczna jest nowa metafora, a jedną z możliwości jest uznanie, że myślenie przypomina *taniec* (por. Sheets-Johnstone 1981). Nie jest to wcale prowokacja, czy próba podjęta *pour épater la bourgeoisie*, uważam bowiem, że analogia pomiędzy tańcem a myśleniem zawiera w sobie ważne treści. Po pierwsze, mówiąc, że myślenie przypomina taniec, podkreślamy fakt, że procesy myślowe są głęboko zakorzenione w schematach motorycznych naszych mózgów. Po drugie, ta nowa metafora wychwytyje aspekty myślenia, na które zwracają naszą uwagę zarówno model myślenia jako widzenia, jak

i koncepcja myślenia jako rozmowy duszy z samą sobą. Z jednej bowiem strony, metafora myślenia jako tańca wskazuje na przestrzenny wymiar procesów myślowych (tak jak model myślenia jako widzenia); z drugiej — podkreśla dynamiczny charakter myślenia (co czyni także koncepcja myślenia jako rozmowy). Metafora myślenia jako tańca może zatem — w pewnym przynajmniej zakresie — posłużyć do wyjaśnienia, czemu dwie wzajemnie sprzeczne wizje myślenia — oparte na modelach widzenia i rozmowy — tak długo dominowały w filozoficznej tradycji Zachodu, a także w myśleniu potocznym. Po trzecie, metafora myślenia jako tańca wydaje się lepsza niż teza, że myślenie jest jakimś nieokreślonym rodzajem ruchu, gdyż taniec — w przeciwieństwie do „jakiegokolwiek” ruchu — jest działaniem celowym, nawet jeśli cel, o którym mowa, nie jest w pełni uświadamiany. Po czwarte wreszcie, taniec, a przynajmniej typowe jego formy, jest wspólnym działaniem więcej niż jednej osoby, wymaga zatem tego, co psychologowie określają mianem wspólnej bądź dzielonej intencjonalności (*shared intentionality*). W ten sposób metafora myślenia jako tańca wychwytuje kluczowe aspekty trzech idei: że umysł jest ucieleśniony (bo myślenie — tak jak taniec — bazuje na schematach motorycznych); że umysł jest osadzony w interakcjach społecznych (bo nie ma myślenia, podobnie jak nie ma tańca, bez udziału innych); i że umysł jest nakierowany na działanie (bo myślenie, tak jak taniec, jest pewnym *działaniem*).

Trzeba też podkreślić, że we wszystkich trzech przypadkach — ucieleśnienia, osadzenia i nakierowania na działanie — kluczową rolę zdają się odgrywać systemy neuronów lustrzanych, tj. takich, które uaktywniają się zarówno wtedy, gdy wykonujemy jakąś czynność, jak i wtedy, gdy ją jedynie obserwujemy. Przypomnijmy, że idea ucieleśnienia głosi, że nasz aparat pojęciowy kształtowany jest poprzez interakcje naszych ciał ze środowiskiem. W tym kontekście Lakoff i Gallese twierdzą, że pojęcia konkretne (np. „chwycić”) tworzone są na drodze symulacji, wykorzystującej te same obwody neuronalne, które stanowią podstawę dla działania i obserwacji działań innych związanych z danym pojęciem (innymi słowy: pojęcie „chwycić” powstaje na drodze symulacji realizowanej przez te same systemy neuronalne, które aktywne są w trakcie wykonywania czynności chwytania, jak i obserwowania, jak ktoś chwyci jakiś przedmiot) (Gallese, Lakoff 2005). Skoro tak, to system lustrzany jest niezbędny dla wytworzenia pojęć konkretnych — jest zatem podstawowym mechanizmem ucieleśnienia. Podobnie, osadzenie umysłu w interakcjach społecznych nie może się obyć bez systemu lustrzanego: wytworzenie się „wspólnej rozmaitości”, o której mówi Gallese, czyli intersubiektywnej sfery doświadczeń uprzedniej wobec odróżnienia „ja-inny”, musi wykorzystywać zdolność do symulowania stanów psy-

chicznych innych osób, a system lustrzany wydaje się najprostszym mechanizmem, który umiejętność taką zapewnia. Wreszcie, nakierowanie na działanie można wyjaśnić najłatwiej odwołując się do mechanizmu lustrzanego, który wszak bezpośrednio łączy obserwację z działaniem.

3. Hipoteza umysłu normatywnego

Teza, że umysł jest ucieleśniony, osadzony w interakcjach społecznych i nakierowany na działanie prowadzi, w moim przekonaniu, do ważnej zmiany w sposobie, w jaki konceptualizujemy zjawiska psychiczne. Twierdząc jednak, że konsekwencje tej zmiany idą znacznie dalej niż się zwykle przyjmuje. Jest więcej rzeczy w tej „nowej filozofii umysłu” (Rowlands 2010) niż to się śniło naszym filozofom. Można bowiem przekonywać, że w tradycyjnym ujęciu umysł jest „posiadaczem przekonań”. Próbując przedstawić formalny model umysłu, powiedzielibyśmy zapewne, że utożsamiać go można z „siecią przekonań”, które są powiązane ze sobą na rozmaite sposoby. Powiedzielibyśmy także, że niektóre z tych przekonań odnoszą się do faktów (i w tym przypadku możemy twierdzić, że są prawdziwe lub fałszywe), a inne nakazują nam działać w taki a nie inny sposób (i tym nie przypiszemy prawdziwości ani fałszywości, ale powiemy, że są właściwe, uzasadnione, sprawiedliwe, itd.). Co ważne, w tym ujęciu przekonania odnoszące się do faktów są w istotnym sensie paradygmatyczne: gdy spojrzymy na logikę klasyczną, dotyczy ona prawdziwych lub fałszywych sądów; gdy rozważymy najważniejsze koncepcje w epistemologii, okaże się, że przykładają one największą wagę do kryteriów uzasadniania twierdzeń dotyczących faktów. I w logice, i w epistemologii, sfera normatywna ma drugorzędne znaczenie — dysponujemy powszechnie akceptowaną logiką sądów, którym przypisać można prawdziwość lub fałszywość, ale mamy problem z logiką norm; zgadzamy się — mniej lub bardziej entuzjastycznie — na pewne minimalne wymogi uzasadniania w dyskursie teoretycznym, czego nie można powiedzieć o dyskursie praktycznym (Brożek 2007). Trudno się temu dziwić, jeśli się uwzględni się fakt, że „pracujemy” w paradygmacie umysłu deskryptywnego — stwierdzamy, że pierwotna funkcja umysłu polega na opisywaniu rzeczywistości, a pytanie o działanie w świecie uznajemy za problem drugorzędny, którym można się zająć dopiero wtedy, gdy już opiszemy odpowiednie aspekty świata.

Uważam, że warto — szczególnie w kontekście ostatnich odkryć w neurobiologii i teorii ewolucji — rozważyć porządek odwrotny. W tym ujęciu, które określam mianem *hipotezy umysłu normatywnego*, umysł nie jest traktowany jako „posiadacz przekonań”, ale jako narzędzie „rozwiązywania problemów”. Pierwotną funkcją umysłu jest kierowa-

nie działaniem w odpowiedzi na napotkane okoliczności. Funkcję tę dobrze wychwytyją idee ucieleśnienia (*embodiment*), osadzenia w interakcjach społecznych (*embedment*) i nakierowania na działanie (*enactment*). Ucieleśniony umysł oddziałuje ze środowiskiem, kształtując w ten sposób aparat pojęciowy, który stanowi potężne narzędzie w próbach rozwiązywania problemów generowanych przez to środowisko. To, co pozwala nam działać, samo jest współtworzone przez działanie. To, że umysł jest osadzony w interakcjach społecznych, oznacza, że inni są ważni dla nas i naszych działań — naśladując ich i współpracując z nimi zwiększamy swoje szanse na przetrwanie. Wreszcie, umysł jest nakierowany na działanie w tym sensie, iż poznanie zawsze współkształtowane jest przez to, co robimy: nie sposób oddzielić percepcji od akcji.

Jeśli tak, to trzeba zapytać, jak nasza architektura poznawcza zdołała wytworzyć takie „deskryptywne” osiągnięcia, jak teorie fizyczne, matematykę, prozę, poezję czy filozofię. Innymi słowy, zastanowić się wypada, jaka droga prowadzi od ucieleśnionego, osadzonego w interakcjach społecznych i nakierowanego na działanie umysłu do naszych najbardziej wyrafinowanych kulturowych wytworów. Jak umysł normatywny mógł stworzyć geometrie nieprzemienne, ogólną teorię względności, *Tractatus logico-philosophicus*, teorię sekwencjonowania DNA czy *Braci Karamazow*? Twierdzę, że byliśmy zdolni do tych osiągnięć, bo przyjęliśmy specyficzny sposób rozwiązywania problemów: jesteśmy gatunkiem zdolnym do *kierowania się regułami*⁵. Uważam, że z filo- i ontogenetycznego punktu widzenia umiejętność kierowania się regułami dotyczyć musiała najpierw tego, co nazywam *regułami rudymenarnymi*. Reguły takie są: (a) przedjęzykowe; (b) proste; (c) konkretne; (d) wieloaspektowe; i (e) niemodalne. Język nie jest potrzebny do wytworzenia się praktyki kierowania się regułą; przeciwnie, to istnienie jej rudymenarnej formy było warunkiem koniecznym pojawienia się języka (oczywiście wymawianie pewnych „słów” czy wykonywanie odpowiednich gestów może — i musiało — być aspektem niektórych reguł rudymenarnych). Reguły te odnoszą się do relatywnie prostych i konkretnych wzorców zachowań. Dla przykładu: nie istnieje reguła rudymenarna, która nakazywałaby złożenie zeznania podatkowego, gdyż taki wzorec zachowania jest złożony i ogólny. Co najważniejsze — reguły rudymenarne są wieloaspektowe: dotyczą całych zachowań, a nie poszczególnych ich aspektów. Z tego względu reguł rudymenarnych nie można dzielić na rodzaje (moralne, językowe, prawne, matematyczne, prudencjalne, itd.). Wreszcie, reguły rudymenarne określają, jak

⁵Zarysowywaną tu jedynie koncepcję kierowania się regułą rozwijam w pełni w innej książce (por. Brożek 2013).

pewne rzeczy „się robi” w danej społeczności, ale nie ma tu mowy o modalnościach deontycznych, takich jak nakaz, zakaz czy dozwolenie. Można też twierdzić, że reguły rudymen tarne charakteryzują się ubogą, rudymen tar ną formą normatywności⁶. Przyjmijmy, że pewna reguła postępowania jest normatywna (w pełnym sensie tego słowa), gdy może służyć do uzasadnienia działania. Reguły rudymen tarne nie mają tej cechy; równocześnie jednak stanowią obiektywne wzorce zachowań i mogą służyć za kryterium korekty czyj egoś zachowania. Dla przykładu: wyobraźmy sobie, że w pewnej wspólnotie istnieje reguła rudymen tar na, „nakazująca” bieganie za królikiem i nazywanie go „quine”. Jeśli ktoś się pomyli, i zamiast „quine” krzyknie „gavagai”, naruszy pewien obiektywny standard. Może to być podstawą dla negatywnej reakcji wspólnoty i, w konsekwencji, doprowadzić do tego, że następnym razem ów ktoś krzyknie poprawnie „quine”. Trudno byłoby jednak utrzymywać, że taka reguła rudymen tar na uzasadnia bieganie za królikami i nazywanie ich „quine”. Pomimo to, reguła ta posiada pewne charakterystyczne cechy takich reguł, które rzeczywiście generują uzasadnienie; w szczególności, odnosi się do pewnego wspólnego, obiektywnego (czyli niezależnego od przekonań jednostki) wzorca zachowania.

Co sprawia, że rudymen tar na forma kierowania się regułą jest możliwa? Jedną z lepiej uzasadnionych odpowiedzi głosi, że postępowanie zgodne z regułami możliwe jest dzięki naszej zdolności i tendencji do imitowania zachowań innych. Co ciekawe, wiele argumentów przemawia za tym, że imitacja *sensu stricte*, czyli zdolność do kopiowania zarówno sposobu działania, jak i jego celu (czyli zmiany w środowisku, do której działanie to prowadzi), jest warunkowana istnieniem systemu neuronów lustrzanych. Dla przykładu: Giacomo Rizzolatti rozróżnia dwa typy reakcji neuronów lustrzanych, a w związku z tym dwa typy imitacji. Reakcja wysokiego poziomu polega na kopiowaniu celu zachowania, zaś niskiego poziomu — sposobu działania. Wedle Rizzolattiego, tylko ludzki mózg wykorzystuje oba mechanizmy, co umożliwia imitację *sensu stricte*. Inne naczeln e posiadają systemy lustrzane zdolne jedynie do reakcji wysokiego poziomu, co wyjaśnia ich umiejętność uczenia się przez emulację (kopiowanie samego celu działania). Wykorzystanie równocześnie reakcji wysokiego i niskiego poziomu tłumaczy elastyczność i stabilność ludzkich zachowań społecznych, jako że umożliwia — poprzez swoistą rekombinację — użycie tych samych, wyuczonych przez

⁶Naszkicowaną tu jedynie koncepcję normatywności rudymen tarnej i abstrakcyjnej rozwijam szerzej w trzech książkach (por. Brożek 2013, Brożek 2012, Brożek, Brożek, Stelmach 2013).

imitację wzorców zachowania do realizacji różnych celów, bądź realizację tego samego celu różnymi środkami (por. Rizolatti 2005; Brożek 2013).

Jak jednak normatywny umysł może budować teorie, snuć historie, dowodzić twierdzeń czy ważyć racje? Twierdzę, że rozwój systemu reguł rudymenarnych umożliwi powstanie języka. Język, rozumiany jako izolowany system reguł syntaktycznych i semantycznych, jest naszą teoretyczną konstrukcją: powstaje przez wyabstrahowanie niektórych aspektów rzeczywistej praktyki kierowania się regułami rudymenarnymi. A zatem „język w księgach” jest tym, co opisujemy w podręcznikach do gramatyki i słownikach, podczas gdy „język w działaniu” osadzony jest w szerszym kontekście interakcji społecznych. Rozwój języka prowadzi do wytworzenia się tego, co Karl Popper nazywa trzecim światem: sfery naszych abstrakcyjnych teorii, matematyki, opowieści, wartości i reguł abstrakcyjnych. Ujmując rzecz inaczej: opowiadanie historii, budowanie teorii, dowodzenie twierdzeń czy ważenie racji to pewne aktywności, które toczą się wedle określonych reguł; bez istnienia reguł rudymenarnych nie mogłyby istnieć. Jednakże wytwory tych procesów — opowieści, teorie, twierdzenia, tak jak normy prawne i moralne — istnieją w trzecim świecie, a oceniamy je podług różnych kryteriów: prawdziwości, poprawności, bezstronności, efektywności, spójności, itd. Co istotne, wszystkie te kryteria oparte są na rudymenarnym rozumieniu normatywności — stanowią one rozmaite probierze uzasadnienia, ale samo pojęcie uzasadnienia byłoby niezrozumiałe, gdyby nie istniało prymitywne pojęcie normatywności.

Więcej kontrowersji budzi pytanie, w jaki sposób umysł staje się „posiadaczem przekonania”. Wedle szeroko rozpowszechnionej definicji, przekonanie to rodzaj nastawienia wobec treści propozycjonalnych — osoba *O* jest przekonana, że *p*, jeśli uważa *p* za prawdziwe (ewentualnie bardzo prawdopodobne) (Schwitzgebel 2011). Wydaje się jednak, że używanie pojęcia przekonania może prowadzić do nieporozumień. Po pierwsze, sformułowana powyżej definicja głosi, że o przekonaniu możemy mówić wtedy, gdy ktoś uznaje pewne zdanie (sąd) za prawdziwe. Rzecz w tym, że zapewne większość żywionych przez nas przekonań nie może być ocenianych w kategoriach prawdy lub fałszu. Jeśli uważam, że powinienem zażywać witaminy, że nie powinno się zabijać innych ludzi, czy że trzeba być ostrożnym, atakując tradycyjne doktryny filozoficzne, nie myślę, że przekonania te są prawdziwe lub fałszywe. Uznanie przekonań, którym można przypisać prawdziwość lub fałszywość, za paradygmatyczne przypadki przekonań, zakłada deskryptywną wizję umysłu. Po drugie, kategoria przekonań jest bardzo heterogeniczna. Rozumiana w sposób permissywny, obejmuje rozmaite rodzaje „nastawień

propozycjonalnych” — od uznawania dobrze uzasadnionych i przemyślanych poglądów, po twierdzenia, które przyjmujemy w oparciu o przeczuca czy — psychologicznie rozumianą — intuicję. Tezę, że umysł jest „posiadaczem przekonań” można, w związku z tym, rozumieć na dwa przeciwstawne sposoby. Jeśli przyjmiemy, że podstawową funkcją umysłu jest wytworzenie sieci przekonań, które rysują prawdziwy obraz świata, to jest to koncepcja z gruntu fałszywa — umieszcza w centrum uwagi to, co jest pochodną bardziej podstawowych funkcji umysłu. Jeśli zaś uznamy, że „przekonanie” obejmuje jakiekolwiek „nastawienia propozycjonalne”, od świadomie zaakceptowanych i dobrze uzasadnionych tez, po twierdzenia oparte o przeczuca i intuicję, które niekoniecznie oceniamy w kategoriach prawdy lub fałszu, to koncepcja „umysłu jako posiadacza przekonań” staje się trywialna, gdyż „przekonanie” jest tu synonimem „treści psychicznych”. Jeśli jednak jesteśmy skazani na jeden z tych dwóch sposobów rozumienia pojęcia „przekonanie”, to pojęcie to staje się bezużyteczne, a w niektórych przypadkach — wręcz szkodliwe.

Broniona przeze mnie hipoteza umysłu normatywnego głosi, że pierwotną funkcją umysłu jest działanie nakierowane na rozwiązywanie problemów, przy czym w toku ewolucji naszego gatunku wykształcił się służący temu szczególnie sposób działania: postępowanie zgodnie z regułami. To istnienie systemu reguł rudymenarnych pozwoliło nam wytworzyć język i, szerzej, kulturę: prozę, poezję, matematykę, fizykę i prawo. W tym kontekście kluczowa zdaje się następująca obserwacja: jeśli chcemy wyjaśnić, skąd bierze się w pełni funkcjonujący ludzki umysł, nie powinniśmy skupiać się na jego najsubtelniejszych wytworach, ale na tym, co je warunkuje — a to coś jest normatywne, przynajmniej w rudymenarnym sensie tego słowa.

Literatura

Brożek, A., Brożek, B., Stelmach, J. (2013). *Fenomen normatywności*. Kraków: Copernicus Center Press.

Brożek, B. (2007). *Rationality and Discourse*. Warszawa: Wolters Kluwer.

Brożek, B. (2009). Philosophy in Neuroscience. W: Brożek B., Mączka, J., Grygiel W. (red.) *Philosophy in Science*. Kraków: Copernicus Center Press, s. 163-188.

Brożek, B. (2012). *Normatywność prawa*. Warszawa: Wolters Kluwer.

Brożek, B. (2013). *Rule-Following. From Imitation to the Normative Mind*. Kraków: Copernicus Center Press.

- Brożek, B. (2014). The Normative Mind. In Defence of a Heresy. W: Brożek, B., di Pellegrino, G., Stelmach, J., Winkielman, P. *The Normative Mind. Dimensions of Decision-Making*. Kraków: Copernicus Center Press.
- Brożek, B., Hohol M. (2014). *Umysł matematyczny*. Kraków: Copernicus Center Press.
- Gallese, V. (2005). "Being Like Me": Self — Other Identity, Mirror Neurons, and Empathy. W: Hurley, S., Chater, N. (red.), *Perspectives on Imitation vol. 1: Mechanisms of Imitation and Imitation in Animals*, Cambridge, Massachusetts: MIT Press, s. 101–118.
- Gallese, V., Lakoff, G. (2005). The Brain's Concepts: The Role of the Sensory-Motor System in Conceptual Knowledge, *Cognitive Neuropsychology* 22 (3–4), 455–479.
- Hacking, I. (1975). *Why Does Language Matter to Philosophy*. Cambridge: Cambridge University Press.
- Hohol, M., Urbańczyk, P. (2013). Self-Deception. Between Philosophy and Cognitive Neuroscience. W: Stelmach, J., Brożek, B., Kurek Ł. (red.), *Philosophy in Neuroscience*. Kraków: Copernicus Center Press.
- Hutchins, E. (2010). Enaction, Imagination, and Insight. W: Stewart, J, Gapenne, O., Di Paolo, E. A. (red.) *Enaction*. Cambridge, Massachusetts: MIT Press.
- Kartezjusz (1996). *Listy do Regiusa. Uwagi o pewnym piśmie*, Warszawa: PWN.
- Kartezjusz (2004), *Medytacje o pierwszej filozofii*. Kraków: Zielona Sowa.
- Kartezjusz (2001). *Zasady filozofii*. Kęty: Antyk.
- Nelson, A. (1997). Descartes's Ontology of Thought. *Topoi* 16, 163–178.
- Noe, A. (2004). *Action in Perception*. Cambridge, Massachusetts: MIT Press.
- Rizolatti, G. (2005). The Mirror Neuron System and Imitation. W: Hurley, S., Chater, N. (red.), *Perspectives on Imitation vol. 1: Mechanisms of Imitation and Imitation in Animals*, Cambridge, Massachusetts: MIT Press, s. 55–76.
- Rowlands, M. (2010). *The New Science of the Mind*, Cambridge, Massachusetts: MIT Press.
- Schwitzgebel, E. (2011). Belief. Zalta, E. (red.) *The Stanford Encyclopedia of Philosophy* (Winter 2011 Edition), URL = <<http://plato.stanford.edu/archives/win2011/entries/belief/>>.
- Sheets-Johnstone, M. (1981). Thinking in Movement. *Journal of Aesthetics and Arts Criticism* 39 (4), 399–407.
- Tomasello, M. (1995) Language is Not an Instinct. *Cognitive development* 10, 131–156.

The Hypothesis of the Normative Mind

BARTOSZ BROŻEK

Department for the Philosophy of Law and Legal Ethics, Jagiellonian University Copernicus Center for Interdisciplinary Studies

Abstract. *The main goal of the paper is the philosophical reflection over some consequences of the theories, which have recently been developed in cognitive science and neuroscience. The claim that the mind is embodied, embedded and enacted leads to the rejection of three theses characteristic of the modern replies view of the mind: the mind-body dualism, the mind-other minds dualism, and the perception-volition dualism. The rejection of those dualisms support the thesis that, at its roots, the mind is normative, i.e. that its primary function is to act in order to solve the encountered problems. In turn, such descriptive faculties of the mind as theory-construction or story-telling have a derivative character — they are built upon our normative abilities.*

Keywords: *mind, normativity, embodiment, embedment, enactment*