

Paweł Gładziejewski

William Ramsey o psychologii potocznej, racjonalności i pojęciu reprezentacji w naukach kognitywnych

Diametros nr 31, 33-55

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.

**WILLIAM RAMSEY O PSYCHOLOGII POTOCZNEJ,
RACJONALNOŚCI I POJĘCIU REPREZENTACJI
W NAUKACH KOGNITYWNYCH**

– Paweł Gładziejewski –

Wstęp

Przedmiotem artykułu jest rekonstrukcja oraz krytyczna dyskusja niektórych tez przedstawionych przez Williama Ramseya w jego książce *Representation Reconsidered*¹. Praca Ramseya stanowi istotny głos w toczonym obecnie na pograniczu filozofii umysłu i nauk kognitywnych sporze o istnienie oraz status eksplanacyjny reprezentacji mentalnych. Przedstawione tu rozważania koncentrować się będą na tej części wyводу autora, w której rozważa on ideę, zgodnie z którą pojęcie reprezentacji używane w ramach nauk kognitywnych powinno wzorować się na pojęciu reprezentacji wywiedzionym z psychologii potocznej, to jest pojęciu reprezentacji jako postaw propozycjonalnych. Ramsey argumentuje przeciwko takiemu ujęciu, między innymi polemizując z poglądem, zgodnie z którym wartość eksplanacyjna pojęcia reprezentacji wywiedzonego z psychologii potocznej polega na tym, że tylko ono pozwala wyjaśnić kluczową własność ludzkich procesów umysłowych – ich racjonalność. W pierwszej części artykułu przedstawiony zostanie szerszy projekt filozoficzny Ramseya, stanowiący tło dla omawianych tu szczegółowo rozważań. W części drugiej zrekonstruowana zostanie ta część jego argumentacji, która stanowi właściwy przedmiot tej pracy. Wreszcie w części trzeciej podjęta zostanie próba wprowadzenia dość istotnych uzupełnień do tez Ramseya. Będzie się tu zmierzać do pokazania, że Ramsey nie zauważył ukrytego, fałszywego założenia krytykowanego przez siebie argumentu z racjonalności. Wykazanie fałszywości tego założenia nie tylko stanowi dodatkową, silną rację za niekonkluzywnością argumentu z racjonalności, ale jednocześnie może uwolnić wywód samego Ramseya od pewnych eliminatywistycznych konsekwencji.

¹ Ramsey [2007].

1. Szkic projektu Williama Ramseya

Wydaje się, że pojęcie reprezentacji mentalnej pełni w kognitywistyce rolę fundamentalną. Słynna rewolucja kognitywna w psychologii – dająca przecież początek nie tylko psychologii poznawczej, ale też kognitywistyce jako nauce interdyscyplinarnej – polegała na wypełnieniu behawiorystycznej „czarnej skrzynki” strukturami stanowiącymi mentalne reprezentacje świata zewnętrznego, które podczas rozwiązywania określonych problemów poznawczych podlegać mogą różnorodnym manipulacjom. Niektórzy autorzy uznają używanie pojęcia reprezentacji mentalnej wręcz za cechę definicyjną kognitywistyki². Jednak zmiany teoretyczne, jakie dokonują się na przestrzeni ostatnich lat, sprawiają, że ten stan rzeczy dość szybko przestaje być aktualny. Wykorzystanie teorii systemów dynamicznych do modelowania aktywności mózgu, rozwój nurtów teoretyczno-badawczych opartych na ideach enaktywizmu czy ucieleśnionego poznania pokazały, że można owocnie uprawiać nauki kognitywne, nie posługując się pojęciem reprezentacji³. Istnienie reprezentacji mentalnych straciło status paradygmatycznego, dokonywanego często *implicite* założenia, a stało się przedmiotem coraz gorętszego sporu teoretycznego⁴. W obecnej sytuacji, na zwolennikach reprezentacyjnego ujęcia umysłu stoi obowiązek dostarczenia racji za tym, że aktywność systemu poznawczego (a w związku z tym biologicznego mózgu) wykorzystuje coś, co można poprawnie i nietrywialnie nazwać reprezentacjami świata.

Omawiana w tym artykule, wspomniana we wstępie praca Williama Ramseya stanowi być może najambitniejszy i najbardziej satysfakcjonujący, jeśli chodzi o poziom szczegółowości konceptualnej, filozoficzny głos w sporze o istnienie oraz przydatność eksplanacyjną reprezentacji mentalnych. Autor ten za swój cel przyjmuje nie tylko argumentowanie za jedną ze stron – choć trzeba zaznaczyć, że *Representation Reconsidered* ma wymowę wyraźnie antyreprezentacjonistyczną – lecz także krytyczne przemyślenie kryteriów, według których spór ten powinien być rozwiązywany. Ramseyowi chodzi o udzielenie odpowiedzi na pytanie o to, w jaki sposób czy też według jakich kryteriów możemy ocenić wartość eksplanacyjną pojęcia reprezentacji dla nauk kognitywnych. W jego propozycji pytanie to jest jednak ściśle powiązane z innym, dotyczącym tego, w jaki sposób możemy *w ogóle* stwierdzić, czy struktury (stany, procesy, mechanizmy) nazywane rutynowo przez kognitywistów reprezentacjami, *rzeczywiście* są reprezentacjami.

² Por. Thagard [1996].

³ Por. np. Chemero [2009]; van Gelder [1995].

⁴ Jeśli chodzi o głosy broniące reprezentacyjnej koncepcji umysłu, por. np. Bechtel [1998]; Grush [2003].

W bardzo ogólnym zarysie propozycja Ramseya wygląda następująco. Przede wszystkim powinniśmy oczekiwać, że struktury, które nazywamy reprezentacjami, *funkcjonują jako reprezentacje* w ramach systemu poznawczego⁵. Podchodząc więc do próby rozstrzygnięcia problemu istnienia reprezentacji mentalnych, musimy dostarczyć opis ról funkcjonalnych, jakie coś musi pełnić, by mogło ono być słusznie i nietrywialnie nazwane reprezentacją. Dopiero przejście takiego – jak je nazywa Ramsey – „wyzwania z opisu obowiązków” (*job description challenge*) może stanowić rację za uznaniem, że dana teoria posługuje się pojęciem reprezentacji w sposób uprawniony. Godna zaakceptowania teoria sformułowana z użyciem reprezentacyjnej terminologii powinna więc pozwolić na pokazanie, *w jakim sensie* postulowane w jej ramach reprezentacje rzeczywiście służą *jako reprezentacje*. Niedopuszczalna jest sytuacja teoretyczna, w której *dowolny* w zasadzie mechanizm neuronalny bądź obliczeniowy mediujący pomiędzy percepcją a zachowaniem niejako automatycznie zostaje opisywany w terminologii reprezentacyjnej. Status bycia reprezentacją musi być ugruntowany w roli, jaką dany stan, proces czy mechanizm pełni w obrębie systemu poznawczego. Ramsey wyraźnie zaznacza przy tym, że pytanie o prawomocność wyjaśnienia działania danego systemu poznawczego w kategoriach reprezentacyjnych powinno być odróżnione od pytań o to, (1) czy jest *możliwe* opisanie tego systemu w kategoriach innych niż reprezentacyjne, (2) czy jest *absolutnie konieczne* opisywanie go w kategoriach reprezentacyjnych. Przynajmniej w zasadzie, twierdzi on, zawsze możliwe jest sformułowanie opisu dowolnego systemu w kategoriach intencjonalnych (nawet kamieniowi można przypisać przekonanie, że powinien on siedzieć w miejscu), tak samo, jak z całkowitym ich pominięciem (na przykład, schodząc na poziom opisu atomów i molekuł). Zamiast tego, powinniśmy według Ramseya pytać, czy przyjęcie perspektywy reprezentacyjnej pozwala nam zrozumieć jakieś zjawisko *lepiej* niż w sytuacji, gdy porzucimy terminologię reprezentacjonistyczną, to znaczy, czy otrzymujemy dzięki pojęciu reprezentacji jakiś wymierny eksplanacyjny zysk⁶. Postulowanie reprezentacji mentalnych w ramach teorii budowanych przez kognitywistów jest według Ramseya uprawnione tylko wtedy, gdy profil funkcjonalny odpowiednich struktur spełnia „wyzwanie z opisu obowiązków” oraz gdy operowanie pojęciem reprezentacji daje nam lepszy wgląd w działanie badanego systemu niż koncepcja pozbawiona w ogóle komponentu reprezentacyjnego.

Aby projekt Ramseya się powiódł, potrzebne jest rozwiązanie kilku zasadniczych kwestii teoretycznych. Po pierwsze, należy dokładnie określić funkcjonal-

⁵ Por. Ramsey [2007] s. 24-34.

⁶ Ibidem, s. 33-34.

ne warunki bycia reprezentacją, to znaczy określić, na czym właściwie polegałoby pozytywne przejście wyzwania z zakresu obowiązków. Innymi słowy należy ustalić, na czym polega funkcjonowanie jako reprezentacja w obrębie systemu poznawczego. Według Ramseya punktem wyjścia dla rozwiązania tego problemu powinno być dokonanie analizy pojęcia reprezentacji mentalnej⁷. Zdecydowanie preferuje on tu próbę eksplikacji pojęcia reprezentacji w jego preteoretycznym sensie, a nie wprowadzenie nowej, technicznej i na ogół podatnej na kontrprzykłady definicji. Zgodnie z propozycją Ramseya, istnieją dwa preteoretyczne pojęcia reprezentacji. Jedno z nich to pojęcie reprezentacji o charakterze *stricte* umysłowym, wywodzące się z psychologii potocznej, gdzie psychologia potoczna rozumiana jest jako zestaw (przednaukowych) przekonań czy informacji na temat działania umysłu, wykorzystywany przez ludzi w celu przewidywania i wyjaśniania własnych oraz cudzych działań w toku codziennych interakcji społecznych⁸. Drugie zaś to pojęcie obiektów czy struktur reprezentacyjnych o pozamentalnym charakterze, takich jak obrazy, termometry czy mapy. W dalszej części tego artykułu istotne będzie tylko pierwsze z tych pojęć. Zgodnie z nim, prototypowymi przykładami reprezentacji są zdroworoządkowe postawy propozycjonalne, przede wszystkim przekonania i pragnienia⁹. Ich istotnymi własnościami są (1) posiadanie niewyowiedzianej treści intencjonalnej¹⁰ oraz (2) posiadanie własności przyczynowych lub inferencyjnych w oparciu o tę treść.

Po drugie, projekt Ramseya wymaga skrupulatnego pokazania, jakie rodzaje stanów, procesów czy struktur były i są powszechnie uważane przez przedstawicieli nauk kognitywnych za reprezentacje, oraz jakie rozumienie „bycia

⁷ Ibidem, s. 8-24.

⁸ Por. np. Nichols [2002]. Warto mieć jednak na uwadze, że do tak rozumianej psychologii potocznej przemykać mogą czasem treści naukowe lub aspirujące do miana naukowych, na przykład określone tezy z zakresu psychologii społecznej lub psychoanalizy. Jak się jednak wydaje, warunkiem tego jest, by tezy te były formułowane *na tym samym poziomie* opisu i wyjaśniania umysłu, co psychologia potoczna – mianowicie, na poziomie osobowym (por. dalsza część tego artykułu).

⁹ Stąd w literaturze zdarza się, że psychologia potoczna nazywana jest także „psychologią przekonań i pragnień”. Warto mieć jednak na uwadze, że w obręb psychologii potocznej wchodzić mogą także pojęcia stanów mentalnych, które są (1) niereprezentacyjne lub (2) reprezentacyjne, lecz niepropozycjonalne (autor dziękuje anonimowemu recenzentowi za zwrócenie mu uwagi na ten fakt). Do którejś z tych kategorii należeć mogą na przykład pojęcia doznań somatosensorycznych, takich jak ból, dotyk czy łaskotki. Szczegółowa dyskusja natury pojęć tego rodzaju oraz ich relacji do pojęć kognitywistycznych nie zostanie tu jednak podjęta. W dalszej części tej pracy – między innymi ze względu na rolę, jaką pełni w niej pojęcie racjonalności – omawiane będą tylko „potoczne” stany mentalne, które są reprezentacyjne i propozycjonalne.

¹⁰ Mówienie tu o „niewyowiedzianej” treści oznacza, że posiadanie treści jest pewną własnością wewnętrzną, a to znaczy przede wszystkim, że nie jest tak, iż warunkiem posiadania treści przez postawy propozycjonalne jest wchodzenie w jakiegokolwiek relacje (na przykład, relację *bycia interpretowanym*) z podmiotami.

reprezentacją” jest tu implicite zakładane. Należy więc dokonać rekonstrukcji pojęciowej praktyki kognitywistów, zbadać, jak oni sami stosują pojęcie reprezentacji. Ramsey wyróżnia łącznie pięć pojęć reprezentacji istniejących w kognitywistycznym „obiegu”: (1) potoczne pojęcie reprezentacji jako postaw propozycjonalnych, (2) pojęcie typu „wejście-wyjście” (IO-reprezentacje), (3) pojęcie reprezentacji jako symulacji (S-reprezentacje), (4) pojęcie reprezentacji jako receptorów, (5) pojęcie reprezentacji ukrytej (niejawnej)¹¹. Trzy pierwsze pojęcia są według Ramseya wykorzystywane głównie w ramach klasycznego już, symbolicznego modelu obliczeniowej architektury umysłu. Dwa pozostałe rozpowszechniły się dopiero wraz z powstaniem koneksjonizmu oraz rozwojem neuronauki poznawczej. Nie sposób tu szczegółowo omawiać, jak Ramsey dyskutuje wszystkie te sposoby ujmowania reprezentacji w ramach nauk kognitywnych. Z punktu widzenia tematyki tego artykułu istotny jest jedynie pierwszy z nich. Zgodnie z nim, naukowe pojęcie reprezentacji przyjmowane przez kognitywistów jest w zasadzie *tym samym*, co pojęcie reprezentacji obecne w psychologii potocznej, to znaczy reprezentacje odkrywane w mózgu czy systemie poznawczym przez kognitywistów powinny być w jakimś sensie utożsamiane ze zdroworozsądkowymi postawami propozycjonalnymi, takimi jak pragnienia i przekonania.

Po trzecie wreszcie, należy wykorzystać poczynione w dwóch pierwszych krokach ustalenia, aby sprawdzić, które – jeśli którekolwiek – ze struktur konceptualizowanych jako reprezentacje przez kognitywistów czynią zadość „wyzwaniu z opisu obowiązków”. Dopiero po wykonaniu tego trzeciego kroku można udzielić odpowiedzi na pytanie o to, czy użycie terminologii reprezentacyjnej jest w danym przypadku w ogóle uzasadnione. Może się okazać, że w większości przypadków proponowane przez badaczy wyjaśnienia zjawisk umysłowych w kategoriach reprezentacyjnych są kompletnie nieuprawnione. To, co w ramach tych wyjaśnień nazywa się reprezentacjami w istocie *nie* funkcjonuje jako reprezentacje. Tym samym precyzyjna analiza może pokazać, że wiele spośród dobrze ugruntowanych empirycznie teorii rozwijanych w ramach nauk kognitywnych ma reprezentacjonistyczny status jedynie pozornie, a tak naprawdę terminologia reprezentacjonistyczna jest w nich używana bezpodstawnie. Jeżeli taki stan rzeczy jest powszechny, może się nawet okazać, że już na obecnym etapie rozwoju spora część kognitywistyki ma niereprezentacjonistyczny charakter. Taka jest zresztą wymowa pracy Ramseya.

W ramach takiego szkicowego zarysu projektu Ramseya można wreszcie umiejscowić właściwy przedmiot tego artykułu. Punktem osiowym będą w nim

¹¹ Por. Ramsey [2007].

rozważania Ramseya dotyczące relacji pomiędzy pojęciem reprezentacji wywiezionym z psychologii potocznej a klasycznym, symbolicznym modelem architektury umysłu, zgodnie z którym funkcje umysłowe realizowane są przez system dokonujący czysto formalnych operacji na syntaktycznie ustrukturalizowanych symbolach¹². Zgodnie z rekonstrukcją Ramseya, istnieje przekonanie, że symboliczny, obliczeniowy model umysłu pozwala na dokonanie naturalizacji potocznego pojęcia reprezentacji¹³. Chodzi tu dokładniej o propozycję, zgodnie z którą operacje symboliczne de facto realizują opisywane przez psychologię potoczną procesy inferencyjne zachodzące na poziomie postaw propozycyjalnych. Zakłada się tu, że zachodzi paralelizm bądź izomorfizm pomiędzy symbolicznymi operacjami obliczeniowymi a wnioskowaniami (procesami inferencyjnymi) przez psychologię potoczną, dzięki czemu można powiedzieć, że scharakteryzowane syntaktycznie symbole pełnią role funkcjonalne zdroworozsądkowych przekonań, pragnień i innych postaw propozycyjalnych. Tym samym uznaje się, że operacje symboliczne stanowią bazę redukcji dla procesów inferencyjnych opisywanych przez psychologię potoczną. Opisywane przez kognitywistów, zaimplementowane w mózgu symboliczne procesy obliczeniowe powinny być zatem traktowane jako „zmechanizowana” wersja psychologii potocznej, a kognitywistyczne rozumienie reprezentacji mentalnych powinno być uznane za „legitymizujące” ich rozumienie potoczne z perspektywy filozoficznego naturalizmu.

2. Ramseya krytyka argumentu z racjonalności

Ramsey twierdzi, że stanowisko w rodzaju przedstawionego powyżej jest nieuzasadnione, ponieważ nie ma podstaw do tego, by sądzić, że struktury opisywane w ramach symbolicznego, obliczeniowego modelu umysłu spełniają „wyzwanie z opisu obowiązków” oparte na potocznym pojęciu reprezentacji¹⁴. Innymi słowy, symbole i formalne operacje na nich *nie* funkcjonują według Ramseya jak stany reprezentacyjne postulowane w ramach psychologii potocznej, w związku z czym konceptualizowanie ich jako (zdroworozsądkowych) reprezentacji jest nieuzasadnione. Ramsey przedstawia dwa argumenty za tego rodzaju tezą: jeden inspirowany jest argumentacją Searle’a przeciwko symboliczno-obliczeniowemu

¹² Ibidem, s. 38-66.

¹³ Słynnym rozwinięciem przytaczanych tu idei jest oczywiście sformułowana przez Fodora hipoteza języka myśli; por. Fodor [1975].

¹⁴ Warto przypomnieć, że zgodnie z rekonstrukcją Ramseya, zdroworozsądkowe postawy propozycyjalne posiadają dwie istotne własności: (1) posiadanie (niewyowiedzianej) treści intencjonalnej, (2) posiadanie ról przyczynowych/inferencyjnych w oparciu o tę treść. Z perspektywy projektu Ramseya szczególnie ważna jest ta druga własność, ponieważ jest ona bezpośrednio związana z *funkcjonalnymi* warunkami bycia postawą propozycyjną.

modelowi umysłu, a drugi – argumentacją Sticha za eliminatywizmem względem psychologii potocznej¹⁵.

Ramsey uznaje, że Searle swoim klasycznym eksperymentem myślowym z chińskim pokojem pokazał, iż manipulacje symbolami opisywane w ramach symboliczno-obliczeniowego modelu umysłu oparte są na czysto formalnych czy syntaktycznych regułach, a zatem ich przebieg może zostać wyjaśniony w sposób kompletny bez odwoływania się do terminologii reprezentacyjnej. Ujmując rzecz mocniej, symbole, na których przeprowadzane są czysto „syntaktyczne” obliczenia, *nie mają* po prostu żadnych własności intencjonalnych czy semantycznych, w związku z czym także ich role przyczynowe nie mogą być oparte na takich własnościach. Tym samym symbole opisywane w ramach klasycznego modelu architektury obliczeniowej umysłu nie mają żadnej z istotnych cech postaw propozycyjalnych. Standardowym sposobem odpowiadania na tego rodzaju zarzut jest propozycja uzupełnienia symboliczno-obliczeniowego modelu za pomocą serii dodatkowych twierdzeń oraz założeń. Jednak Ramsey przedstawia silne racje za tym, że nawet próba wprowadzenia stosownych uzupełnień do tego modelu nie może być wystarczająca. Można co prawda wzorem Fodora¹⁶ postulować, że model symboliczno-obliczeniowy daje możliwość pokazania, jak dany symbol *rozumiany jako nośnik treści intencjonalnej* może funkcjonować w roli przekonania, w odróżnieniu od pragnienia czy nadziei. Takie ujęcie nie stanowi jednak odpowiedzi na pytanie, dlaczego *w ogóle* mielibyśmy traktować ten symbol jako reprezentację, a to przede wszystkim znaczy – jako nośnik treści intencjonalnej. W tym celu należałoby stworzyć osobną, niezależną teorię treści. Pamiętajmy jednak, że „wyzwanie z opisu obowiązków” dotyczy takiego scharakteryzowania *funkcjonalnych* własności danych struktur, procesów czy stanów, które pozwoli nam sądzić, że są one reprezentacjami. Tymczasem funkcjonalne własności symboli są całkowicie wyczerpywane przez ich role syntaktyczne. Przypisywanie symbolom treści intencjonalnej opartej na dajmy na to relacjach kauzalno-nomologicznych z obiektami zewnętrznymi względem systemu poznawczego nie przekłada się na żadne twierdzenia o tym, jak symbole te miałyby *służyć* czy *funkcjonować* jako reprezentacje w ramach systemu poznawczego. Dlatego też nawet dostarczenie stricte naturalistycznej teorii treści intencjonalnej nie musi stanowić – i w tym przypadku nie stanowi – zadowalającej odpowiedzi na „wyzwanie z opisu obowiązków”.

¹⁵ Tu i dalej por. Ramsey [2007] s. 46-65.

¹⁶ Fodor [1975].

Drugi z argumentów Ramseya inspirowany jest z kolei pracą *From Folk Psychology to Cognitive Science* Sticha¹⁷, stanowiącej klasyczną obronę eliminatywizmu względem psychologii potocznej. Mówiąc w skrócie, Stich wychodzi w niej od obserwacji, że kryteria indywiduowania stanów mentalnych oparte o psychologię potoczną muszą z konieczności odwoływać się do treściowych czy intencjonalnych własności tych stanów. Autor ten przedstawia jednak racje za tym, że „treściowa” strategia tworzenia taksonomii stanów mentalnych nie może dać wyników, które pokrywałyby się w pełni z tymi otrzymanymi dzięki strategii posługiwania się czysto funkcjonalnymi bądź syntaktycznymi kryteriami indywiduowania tych stanów. Co więcej, ta druga strategia jest lepsza ze względu na predykcyjno-eksplanacyjne cele nauk kognitywnych niż strategia „treściowa”. Dlatego też powinniśmy odrzucić projekt zmierzający do traktowania wewnętrznych symboli jako odpowiedników postaw propozycjonalnych, a zamiast tego przyjąć czysto „syntaktyczną” teorię umysłu. Ramsey wyciąga z rozważań Sticha wniosek, że nie mamy dobrych naukowych powodów, by traktować symbole i przeprowadzane na nich operacje obliczeniowe w zdroworozsądkowych kategoriach reprezentacyjnych. Mówiąc wprost, kiedy wyjaśniamy działanie systemu poznawczego, możemy obejść się bez mówienia o zdroworozsądkowych reprezentacjach; postawy propozycjonalne są dla nauk kognitywnych eksplanacyjnie zbędne.

Biorąc wszystko to pod uwagę, wydaje się, że projekt wykorzystania potocznego pojęcia reprezentacji mentalnych w kontekście symbolicznego modelu obliczeniowej architektury umysłu nie może się powieść. Jak jednak zauważa Ramsey, istnieje pewna strategia odrzucenia tej konkluzji. Polega ona na pokazaniu, że wśród różnych własności umysłu, do których wyjaśnienia zmierza kognitywistyka, istnieje taka, która zmusza nas do tego, żebyśmy interpretowali symbole wewnętrzne i przeprowadzane na nich operacje w kategoriach psychologii potocznej. Zgodnie zatem z tą strategią, ludzkie umysły mają cechę, której nie sposób wyjaśnić rezygnując całkowicie ze zdroworozsądkowej terminologii reprezentacyjnej czy intencjonalnej. Zdaniem Ramseya, taką często przywoływaną własnością jest *racjonalność*¹⁸. Specyfiką procesów umysłowych jest to, że przynajmniej czasem egzemplifikują one cechę racjonalności. Jeśli jednak przyjmiemy czysto syntaktyczne rozumienie umysłu, racjonalność pozostaje bez wyjaśnienia. Tylko interpretowanie procesów obliczeniowych w kategoriach psychologii potocznej miałoby pozwolić naukom kognitywnym na oddanie sprawiedliwości faktowi, że

¹⁷ Stich [1985].

¹⁸ Por. np. Pylyshyn [1984].

ludzkie procesy umysłowe charakteryzują się racjonalnością. Struktura tego rozumowania, nazywanego dalej „argumentem z racjonalności”, byłaby więc następująca:

1. Teorie tworzone przez kognitywistów powinny wyjaśniać racjonalność egzemplifikowaną przez systemy poznawcze.
2. Jedynym sposobem wyjaśnienia racjonalności systemów poznawczych jest traktowanie ich stanów wewnętrznych jako reprezentacji.
3. Stąd obliczeniowe teorie poznania muszą traktować symbole wewnętrzne jako reprezentacje¹⁹.

Argument ten bez wątpienia wymaga doprecyzowania użytego w nim pojęcia racjonalności. Ramsey odrzuca na wstępie rozumienie „racjonalności” jako predykatu orzekanego o zachowaniach. Gdybyśmy uznali na przykład, że racjonalne zachowanie to zachowanie optymalne w danych okolicznościach ze względu na interesy danej osoby bądź organizmu, to nie byłoby według tego autora przeciwwskazań, by uznać, że może ono zostać w pełni wyjaśnione na poziomie czysto „syntaktycznym”, bez odwoływania się do jakichkolwiek stanów intencjonalnych. Wydaje się zatem, że lepiej jest uznać „racjonalność”, na którą powołuje się powyższy argument, za cechę samych procesów umysłowych, rozumianych jako procesy inferencyjne. W takim wypadku *same procesy obliczeniowe* mogłyby egzemplifikować w jakimś sensie cechę racjonalności; inaczej, racjonalność przysługująca procesom inferencyjnym miałaby w jakiś sposób sprawiać, że musimy w kategoriach intencjonalnych wywiedzionych z psychologii potocznej interpretować procesy obliczeniowe zachodzące wewnątrz systemu poznawczego. Nawet jednak przy takim zawężeniu, termin „racjonalność” nie jest do końca jasny, ponieważ można zadać pytanie o to, na czym polega racjonalność procesów inferencyjnych. Na przykład jak zauważa Ramsey, proste utożsamienie „racjonalnych” procesów inferencyjnych z takimi, które zgodne są z regułami rozumowania dedukcyjnego, nie byłoby strategią owocną. W takim wypadku nie byłoby przecież problematyczne uznać, że scharakteryzowane czysto formalnie czy syntaktycznie procesy obliczeniowe są racjonalne, o ile zgodne są one z określonymi regułami logicznymi. Wydaje się więc, że argument z racjonalności wspiera się na nieco innym, szerszym ujęciu racjonalności, mającym źródło w intuicji, że domena ludzkiego myślenia jest przede wszystkim domeną *racji*. Innymi słowy, procesy umysłowe są w tym sensie racjonalne o tyle, o ile uprzednie kroki motywują następne, stanowiąc dla nich racje, a więc stanowiąc dla nich pewne normatywne oparcie.

¹⁹ Za: Ramsey [2007] s. 59.

Jeżeli przyjmiemy taką koncepcję, wydaje się, że droga do konkluzywnego przeprowadzenia argumentu z racjonalności jest już otwarta. Wystarczy przyjąć dość wiarygodne założenie, że relacja „bycia racją” może zachodzić tylko i wyłącznie pomiędzy dwoma postawami propozycjonalnymi: tylko przekonania i pragnienia mogą stanowić racje sensu stricto dla innych przekonań i pragnień. Ramsey formułuje bardzo zbliżone twierdzenie, zgodne z którym bycie procesem racjonalnym oznacza bycie procesem podpadającym pod prawa czy generalizacje psychologii potocznej, czyli prawa określające z konieczności relacje pomiędzy postawami propozycjonalnymi²⁰. Łącznie wszystko to prowadzić ma zaś do wniosku, że dopóki nie będziemy traktować symboli opisywanych przez kognitywistów jako odpowiedników postaw propozycjonalnych, dopóty idea, że ludzkie myślenie jest racjonalne pozostanie bez naukowego wyjaśnienia.

Ramsey zauważa jednak, że tego rodzaju argumentacja jest formalnie niepoprawna. Fakt ten staje się jasny, kiedy rozpiszemy kolejne kroki rozumowania:

1. Teorie tworzone przez kognitywistów powinny wyjaśniać racjonalność egzemplifikowaną przez systemy poznawcze²¹.
2. Być racjonalnym w powyższym sensie, to podpadać pod prawa psychologii potocznej.
3. Podpadać pod prawa psychologii potocznej, to podpadać pod prawa dotyczące postaw propozycjonalnych.
4. Stąd teorie tworzone przez kognitywistów muszą odwoływać się do pojęć postaw propozycjonalnych, czyli potocznych czy zdroworozsądkowych stanów reprezentacyjnych takich jak pragnienia czy przekonania.
5. Stąd obliczeniowe teorie umysłu muszą traktować symbole wewnętrzne jako zdroworozsądkowe reprezentacje (jako postawy propozycjonalne)²².

²⁰ Zauważmy, że Ramsey zakłada tu, że psychologia potoczna stanowi swego rodzaju teorię specyfikującą prawa rządzące procesami umysłowymi. Należy jednak mieć na uwadze, że idea ta wcale nie jest nieproblematyczna. Po pierwsze, jest dalekie od oczywistości, że codzienna praktyka interpretowania własnych oraz cudzych działań i procesów mentalnych w kategoriach intencjonalnych jest rzeczywiście oparta na pewnego rodzaju wiedzy teoretycznej (Goldman [2006]). Po drugie, niektórzy filozofowie argumentują za tezą, że procesy umysłowe charakteryzują się swego rodzaju holizmem epistemicznym, który sprawia, że nie da się ich opisać za pomocą praw podobnych do praw naukowych (Heal [1996]). Wydaje się jednak, że „nomologiczne” ujęcie natury psychologii potocznej nie jest absolutnie niezbędne do tego, by Ramsey mógł przeprowadzić swój wywód. Warto mieć bowiem na uwadze, iż to, że (1) racjonalność konstytutywnie zakłada istnienie postaw propozycjonalnych, które mogą stanowić dla siebie nawzajem racje, *nie* zakłada z konieczności, że (2) istnieje możliwość sformułowania praw rządzących postawami propozycjonalnymi. To teza (1) wydaje się istotna dla przytoczonego dalej argumentu za koniecznością przyjęcia potocznego pojęcia reprezentacji, natomiast przyjęcie tezy (2) nie jest konieczne.

²¹ Czy inaczej: teorie te muszą traktować bądź pozwalać na traktowanie systemów poznawczych jako racjonalnych.

Ramsey zwraca uwagę na fakt, że rozumowanie to jest cyrkularne. Sprowadza się ono do stwierdzenia, że musimy – jako kognitywiści – przyjąć zdroworozsądkowe pojęcie reprezentacji mentalnej po to, by wyjaśnić racjonalność ludzkiego myślenia, jednak racjonalność jest w tej perspektywie utożsamiona z procesami obejmującymi takie zdroworozsądkowe stany reprezentacyjne. Zadanie wyjaśnienia racjonalności z *definicji* oznacza więc konieczność posługiwania się zdroworozsądkowym pojęciem reprezentacji. Jeżeli bowiem chcemy uznawać system poznawczy za racjonalny, musimy go uważać za podmiot zdroworozsądkowych postaw propozycyjalnych; bycie racjonalnym oznacza po prostu bycie podmiotem tych postaw. Mówiąc wprost, powyższy argument sprowadza się do stwierdzenia, że musimy interpretować procesy obliczeniowe w kategoriach zaczerpniętych z psychologii potocznej dokładnie po to, byśmy mogli traktować procesy obliczeniowe w kategoriach zaczerpniętych z psychologii potocznej, to jest traktować je jako procesy egzemplifikujące własność racjonalności. Próba odwołania się do pojęcia racjonalności w celu uzasadnienia postulatu o wykorzystaniu potocznego pojęcia reprezentacji w naukach kognitywnych okazuje się więc zdaniem Ramseya całkowicie niekonkluzywna. Tym samym nie mamy powodu, by uznać, że procesy obliczeniowe opisywane w ramach symbolicznego modelu architektury obliczeniowej ludzkiego umysłu powinny być konceptualizowane za pomocą terminologii zaczerpniętej z psychologii potocznej. Symbole, na których przeprowadzane są obliczenia nie spełniają „wyzwania z opisu obowiązków” opartego na potocznym pojęciu reprezentacji. Przynajmniej w takim zakresie symboliczno-obliczeniowy model umysłu powinien być uznany za w istocie niereprezentacyjny.

3. Argument z racjonalności a osobowy i subosobowy poziom opisu i wyjaśniania

Celem tej części artykułu nie jest argumentowanie przeciwko przedstawionej przez Ramseya krytyce argumentu z racjonalności. Przyjmuje się, że jest ona konkluzywna. Odwoływanie się do racjonalności w celu uzasadnienia użyteczności eksplanacyjnej potocznego pojęcia reprezentacji dla nauk kognitywnych jest zabiegiem nieskutecznym, a podstawą jego nieskuteczności jest kolistość: wymóg wyjaśnienia racjonalności jest postawiony tak, że wyjaśnienie to na mocy pojęciowej konieczności zakłada operowanie potocznym pojęciem reprezentacji mentalnej. Jednak można podjąć próbę dość istotnego uzupełnienia argumentacji Ramseya. Przede wszystkim, choć wywód tego autora pokazuje słabość argumentu z racjonalności, to nie dociera on do bardziej zasadniczej pomyłki filozoficznej, na

²² Za: Ramsey [2007] s. 62.

której argument ten się wspiera. Ponadto, Ramsey wydaje się przedstawiać raczej pesymistyczną wizję, jeśli chodzi o kwestię możliwości wyjaśnienia racjonalności ludzkiego myślenia przez kognitywistów oraz kwestię ontologicznego statusu samej racjonalności. Z jednej strony wydaje się on sugerować, że racjonalność w sensie przytoczonym wcześniej być może wcale nie należy eksplanandów nauk kognitywnych²³. Z drugiej strony autor ten jest jednym z aktywnych obrońców eliminatywnego materializmu i odrzuca istnienie stanów postulowanych przez psychologię potoczną, a co za tym idzie, racjonalności²⁴. Wydaje się jednak, że można przyjąć argumentację Ramseya przedstawioną w poprzedniej części artykułu, nie rezygnując z przekonania, że istoty ludzkie posiadają postawy propozycjonalne, a ich procesy myślowe są czy też bywają racjonalne. Aby jednak to było możliwe, trzeba wyrazić *explicite* założenie, na którym wspiera się krytykowany przez Ramseya argument z racjonalności.

Kluczem do znaczącego uzupełnienia propozycji Ramseya oraz jednoczesnego uwolnienia jej z eliminatywistycznych konsekwencji jest odwołanie się do idei, że istnieją różne *poziomy* opisywania i wyjaśniania stanów czy procesów umysłowych. Należy powołać się na wprowadzone przez Dennetta odróżnienie *osobowego* poziomu opisu i wyjaśniania od poziomu *subosobowego*. Oto jak sam Dennet wprowadza tę dystynkcję:

Gdybyśmy powiedzieli, że osoba posiada doznanie bólu, umiejscawia je i skłonna jest do reagowania w pewien sposób, powiedzielibyśmy już wszystko, co da się powiedzieć w ramach tego sposobu mówienia. Możemy domagać się dalszego wyjaśnienia, jak to się dzieje, że odsuwamy rękę od gorącego pieca, ale nie możemy domagać się dalszego wyjaśnienia w terminach „procesów mentalnych”. [...] Jeżeli się na to decydujemy, musimy porzucić ten poziom wyjaśniania, poziom w którym mówimy o ludziach, ich wrażeniach, ich działaniach i zejść na subosobowy poziom mózgu i zdarzeń zachodzących w układzie nerwowym. Lecz gdy opuszczamy poziom osobowy to zupełnie dosłownie opuszczamy również dziedzinę bólu [...] gdyż nasza alternatywna analiza nie jest wcale analizą bólu, ale czegoś zupełnie innego: ruchów ludzkiego ciała, czy też organizacji układu nerwowego²⁵.

Osobowy poziom wyjaśniania i opisywania działań oraz stanów i procesów mentalnych to ten, na którym poruszamy się w codziennej praktyce interpretowa-

²³ Por. *ibidem*.

²⁴ Por. Ramsey, Stich, Garon [1990].

²⁵ Dennett [1995] s. 106-107.

nia własnych oraz cudzych działań w kategoriach intencjonalnych. Można też powiedzieć, że jest to poziom psychologii potocznej. Procesy oraz stany wyróżniane na tym poziomie – na przykład wyobrażanie, empatyzowanie z innymi, doznawanie emocji, posiadanie pragnień i przekonań – mogą być poprawnie i dosłownie (a nie jedynie metaforycznie) przypisywane tylko i wyłącznie *osobom* czy *podmiotom*. Wyjaśnianie na poziomie personalnym ma w dużym stopniu charakter normatywny: osoby podejmują takie a nie inne działania albo przyjmują takie a nie inne przekonania, ponieważ są one w pewnych warunkach uzasadnione czy uprawnione, ponieważ dysponują one odpowiednimi ku temu *racjami*. Poziom osobowy musi być odróżniony od poziomu subosobowego, na którym porzucamy domenę podmiotów percypujących, podejmujących decyzje czy żywiących przekonania, a zaczynamy w sposób stricte mechanistyczny, nienormatywny opisywać i wyjaśniać procesy neuronalne bądź funkcjonalne czy obliczeniowe stojące u podstaw określonych zdolności poznawczych. Rozpatrywanie procesów czy mechanizmów o subosobowym charakterze w kategoriach zaczerpniętych z poziomu osobowego może co prawda pełnić pewną rolę heurystyczną, ale nie może być traktowane dosłownie pod groźbą popełnienia serii błędów kategorialnych. Mówiąc obrazowo, obszary kory wzrokowej nie percypują przedmiotów, a neurony lustrzane nie empatyzują z innymi ludźmi.

Przytoczony wyżej cytat z Dennetta wyraża także ideę, zgodnie z którą poziom osobowy charakteryzuje się swego rodzaju autonomią eksplanacyjną względem poziomu subosobowego. Wyjaśnienia formułowane na poziomie osobowym są odrębne i w pewnym sensie całkowicie niezależne od wyjaśnień na poziomie subosobowym. Można formułować wyczerpujące wyjaśnienia w kategoriach z poziomu osobowego całkowicie abstrahując od faktów stwierdzanych na poziomie subosobowym. Kiedy wyjaśniamy czyjeś odsunięcie ręki od palącego się płomienia, stwierdzając, że dana osoba doznała bólu, nasze wyjaśnienie jest już kompletne na poziomie osobowym, ponieważ nie istnieje *w ramach tego poziomu* możliwość odpowiedzi na pytanie o to, *dłaczego* odczuwamy ból w określonych okolicznościach. Wyjaśnienie takie powinno być więc uznane za kompletne i zrozumiałe nawet w sytuacji, gdy nie jest ono uzupełnione subosobowymi twierdzeniami dotyczącymi neurofizjologii bólu. Podobnie kiedy wyjaśniamy przekonanie pewnej osoby, próbując pokazać, jak wywnioskowała je ona z innych przekonań (przekonań, które stanowiły *racje* dla przyjęcia przekonania wyjaśnianego), wyjaśnienie takie jest wyczerpujące, a powołanie się na przyczynową rolę neurodynamiki kory czołowej tej osoby byłoby nieuprawnione; wykorzystanie słownika subosobowego „zmieniłoby temat” i nie mielibyśmy już w naszym wyjaśnianiu do czynienia z myślącymi osobami.

Niektórzy autorzy próbowali wykorzystać tego rodzaju autonomię eksplanacyjną jako podstawę dla przyjęcia zdecydowanie antyredukcyjnej koncepcji relacji pomiędzy poziomem osobowym a subosobowym²⁶. Przyjęcie takiej pozycji nie jest jednak uzasadnione. Aby to zrozumieć, warto wziąć pod uwagę wprowadzone przez Bermudeza rozróżnienie na wyjaśnianie o charakterze wertykalnym oraz horyzontalnym²⁷. Wyjaśnianie horyzontalne polega na wyjaśnianiu danego zjawiska poprzez odwołanie się do innych zjawisk z tego samego poziomu, które to zjawisko wywołały. Jeżeli traktujemy wyjaśnienia formułowane na poziomie osobowym jako wyjaśnienia horyzontalne (na przykład wyjaśnienie faktu, że ktoś żywi jakieś przekonanie, pokazując, jak wywiódł je z innych żywionych przez siebie przekonań), można powiedzieć, że są one autonomiczne względem wyjaśnień na poziomie subosobowym, w tym sensie, że te pierwsze mogą być uznane za kompletne i zrozumiałe bez odwoływania się do faktów wyróżnianych na poziomie subosobowym. Wyjaśnianie wertykalne polega z kolei na pokazaniu, jak zjawiska wyróżniane na niższym poziomie organizacji umożliwiają czy realizują zjawiska na wyższym poziomie organizacji. W przypadku rozróżnienia osobowe-subosobowe, wyjaśnianie wertykalne polegałoby na pokazaniu, jak kompetencje bądź własności wyróżniane na poziomie osobowym są *umożliwiane* przez leżące u ich podstaw procesy wyróżniane na poziomie subosobowym²⁸. W tym „wertykalnym” sensie wyjaśniania, poziom osobowy *nie* powinien być rozumiany jako autonomiczny eksplanacyjnie względem poziomu subosobowego. Procedura eksplanacyjna nauk kognitywnych polega na wyjaśnianiu własności lub kompetencji, które przypisywane są systemowi poznawczemu jako *całości* poprzez mechaniczną dekompozycję tego systemu, zmierzającą do pokazania, jak odpowiednio zorganizowane części (subsystemy) i wykonywane przez nie operacje umożliwiają własności lub zdolności systemowe²⁹. Przyjmuje się tu, że kognitywistyczna praktyka wertykalnego wyjaśniania tego, co osobowe poprzez to, co subosobowe, jest formą wyjaśniania mechanistycznego. Takie ujęcie pozwala też nadać ontologiczną interpretację dystynkcji osobowe-subosobowe (wcześniej zdefiniowanej w kategoriach poziomów opisu i wyjaśniania). Z perspektywy tej interpretacji, opisy sformułowane na poziomie osobowym dotyczą *systemu poznawczego* zaangażowanego w interakcje ze swoim środowiskiem. Z kolei opisy sformułowane na

²⁶ Hornsby [2000]

²⁷ Bermudez [2000]

²⁸ Por. McDowell [1994].

²⁹ Por. np. Bechtel [1994]; Bechtel [2005]; Cummins [2000].

poziomie subosobowym powinniśmy rozumieć jako charakteryzujące odpowiednio zorganizowane, interreagujące części systemu poznawczego.

Mając na uwadze te ustalenia, można już pokazać, na czym polega zasadniczy błąd argumentu z racjonalności. Otóż argument wydaje się opierać na założeniu, że wyjaśnienie racjonalności jest możliwe tylko, jeśli zachodzi *międzypoziomowa izomorfia*, to znaczy izomorfia zachodząca pomiędzy poziomem osobowym a subosobowym. Zgodnie z tym założeniem możemy ujmować systemy poznawcze jako racjonalne, to znaczy jako posiadające postawy propozycjonalne, tylko wtedy, gdy przekonaniom, pragnieniom i innym postawom propozycjonalnym wyróżnianym na poziomie *osobowym* jednoznacznie odpowiadają określone funkcjonalnie, posiadające treść intencjonalną stany na poziomie *subosobowym*. W tej perspektywie zjawiska z poziomu osobowego pozostają bez wyjaśnienia, o ile nie uznamy, że poziom postaw propozycjonalnych z ich własnościami semantycznymi i zachodzącymi między nimi relacjami inferencyjnymi nie jest „odzwierciedlony” w architekturze obliczeniowej umysłu (w znaczeniu: nie zachodzi pomiędzy nimi izomorfia). Wyjaśnienie racjonalności wymaga istnienia izomorfii międzypoziomowej, a ta możliwa jest tylko wtedy, gdy traktować będziemy „wewnętrzne symbole” jako posiadające treść, funkcjonalnie określone odpowiedniki zdroworozsądkowych postaw propozycjonalnych. Takie ujęcie w istotny sposób rozszerza argument z racjonalności w stosunku do tego, jak rekonstruuje go Ramsey. W rozszerzonej wersji, argument ten nadal wychodzi od stwierdzenia racjonalności jako eksplanandum kognitywistyki oraz od zdefiniowania racjonalności za pomocą potocznego pojęcia reprezentacji mentalnej. Jednak przejście do wniosku o konieczności przyjęcia potocznego pojęcia reprezentacji jako podstawy dla pojęcia reprezentacji w naukach kognitywnych poprzedzone jest tezą o tym, że traktowanie systemu poznawczego jako racjonalnego możliwe jest tylko wtedy, gdy obliczeniową architekturę stojącą u jego podstaw interpretować będziemy w kategoriach psychologii potocznej (a więc kategoriach stosowanych do systemu jako całości), czyli jako izomorficzną z poziomem osobowym.

Takie ujęcie usprawnia argument z racjonalności przynajmniej w tym sensie, że pozwala explicite wyrazić ukrytą do tej pory przesłankę za tym, by uznać, że konieczność traktowania systemu poznawczego jako racjonalnego, to jest jako posiadającego postawy propozycjonalne, oznacza konieczność traktowania procesów obliczeniowych w kategoriach postaw propozycjonalnych. Uzupełnienie tego argumentu o twierdzenie o międzypoziomowej izomorfii jako warunku wyjaśnienia racjonalności pozwala także na wskazanie podstaw jego niekonkluzywności, ponieważ twierdzenie to jest fałszywe. Aby to zobaczyć, należy bliżej przyjrzeć się

problemowi relacji zachodzącej pomiędzy opisami na poziomie osobowym oraz opisami na poziomie subosobowym w praktyce eksplanacyjnej kognitywistów.

Nie ulega wątpliwości, że opisy formułowane na poziomie osobowym mają w ramach nauk kognitywnych do spełnienia zasadniczą rolę. Dostarczają one bowiem przynajmniej wstępnego wyróżnienia i opisu eksplanandów, takich jak czytanie umysłów, posiadanie pojęć czy kategoryzacja percepcyjna. Wszystko to są pewne zdolności systemowe, wyróżniane na poziomie osobowym, a ich wyjaśnienie polega na dekompozycji systemu poznawczego w celu odkrycia i opisanie subosobowych mechanizmów stojących u ich podstaw czy też je umożliwiających. Jednak opis na poziomie osobowym pozostaje w zasadzie neutralny co do tego, *jaka* dokładnie postać mają przyjmować te mechanizmy. Opisy na poziomie osobowym pokazują, jakie własności, zdolności czy kompetencje systemowe wymagają wyjaśnienia, ale nie pozwalają na scharakteryzowanie subosobowej architektury umysłu wyjaśniającej te zdolności czy kompetencje. Stąd dla przykładu niektórzy autorzy słusznie stwierdzają, że na zasadniczych pomyłkach opierają się (1) próby wyjaśnienia retencjonalno-protencjonalnej struktury doświadczenia temporalnego poprzez odkrycie subosobowych mechanizmów ściśle odpowiadających określonym elementom doświadczenia fenomenalnego³⁰, (2) próby krytyki teorii symulacji i tzw. teorii teorii jako wyjaśnienia zdolności do czytania umysłów w oparciu o fakt, że naszym interakcjom z innymi ludźmi nie towarzyszy *świadome* symulowanie ich procesów umysłowych bądź przeprowadzanie *świadomych* rozumowań dotyczących ich umysłów³¹. Tak samo scharakteryzowanie na poziomie osobowym systemów poznawczych jako racjonalnych, to jest posiadających postawy propozycjonalne, nie implikuje jeszcze żadnych koniecznych ograniczeń co do subosobowej maszynerii tę racjonalność umożliwiającej, w tym ograniczenia mówiącego o tym, że musi zachodzić jakiegoś rodzaju izomorfia pomiędzy poziomem osobowym a subosobowym³². Myśl tę w następujący sposób wyraża Hurley:

³⁰ Grush [2006]

³¹ Hershsbach [2008]; por. też Hurley [2008].

³² Tego rodzaju dość mocne uniezależnienie naszego rozumienia tego, co subosobowe od tego, jak rozumiemy system poznawczy na poziomie osobowym było krytykowane. Peacocke [1983] sformułował eksperyment myślowy, w którym pewna istota behawioralnie nieodróżnialna od normalnego człowieka okazuje się istotnie marionetką sterowaną przez Marsjanina. Zachowanie tej istoty mogłoby przez nas być skutecznie interpretowane i przewidywane za pomocą kategorii psychologii potocznej. Jednak odkrycie faktu, że jest ona sterowana przez mieszkańca innej planety, a zatem jej działaniami nie kierują żadne wewnętrzne procesy neuronalne albo obliczeniowe, sprawiłoby, że odmówilibyśmy jej statusu istoty myślącej, posiadającej przekonania czy pragnienia. Tym samym opis na poziomie osobowym mimo wszystko wyznacza *pewne* kryteria, które muszą być spełnione na poziomie subosobowym. Eksperymentowi Peacocke'a daleko jednak do wykazania, że prawdziwość opisu na poziomie osobowym wymaga tego, by procesy czy mechanizmy wyróżniane na poziomie subosobowym były *izomorficzne* z tymi opisywanymi na poziomie

To, czy struktury subosobowe korespondują ze strukturami, pojęciowymi lub innymi, z poziomu osobowego, stanowi każdorazowo kwestię empiryczną. Izomorfizm międzypoziomowy nie powinien być ani wymagany, ani negowany a priori. Treść z poziomu osobowego może być dystynktywnie pojęciowa i normatywna także w sytuacji, gdy umysły wylaniają się na bazie interakcji ucieleśnionych mózgów z ich środowiskiem, w tym środowiskiem społecznym. [...] zachowanie systemu może być determinowane nieliniowymi relacjami zachodzącymi pomiędzy czynnikami i na niższym poziomie, chociaż jego struktura nie musi korespondować ze strukturą niższego poziomu³³.

Oczywiście, opis racjonalności ludzkiego myślenia na poziomie osobowym może być wykorzystany jako znakomite narzędzie heurystyczne do tworzenia hipotez dotyczących tego, jakie mechanizmy subosobowe stoją u podstaw tej racjonalności. Taką strategię obiera na przykład Davies, próbując pokazać, że opis procesów inferencyjnych zachodzących na poziomie osobowym stanowi podstawę dla przyjęcia Fodorowskiej koncepcji języka myśli³⁴. Pomijając kwestię trafności propozycji Daviesa, należy tu przede wszystkim zaznaczyć, że jego rozumowanie z poziomu osobowego do poziomu subosobowego jest w istocie rodzajem wnioskowania do najlepszego wyjaśnienia i jest ono równoznaczne z postawieniem pewnej hipotezy empirycznej, która może być sfalsyfikowana w toku rozwoju naszej naukowej wiedzy o umyśle. Davies nie wyklucza empirycznej możliwości, że hipoteza języka myśli okaże się fałszywa. Przyznaje on też, że gdyby okazało się, że hipoteza o istnieniu języka myśli jest fałszywa, to *nie* zagrażałoby to idei, zgodnie z którą ludzie są istotami racjonalnymi i posiadającymi postawy propozycjonalne (to znaczy, nie implikowałoby eliminatywizmu). Według Daviesa taki przebieg wydarzeń raczej powinien motywować nas do takiego rozszerzenia idei istot

osobowym. Jak zauważają Dennett [2008] oraz Henderson i Horgan [2004], eksperyment ten pokazuje raczej, że warunkiem koniecznym uznawania danej istoty za posiadającą postawy propozycjonalne jest to, by jej zachowania kierowane były własnymi, autonomicznymi stanami wewnętrznymi. Co ciekawe, Dennett zauważa nawet, że przy pewnych założeniach nie powinniśmy mieć żadnego problemu z nazwaniem „marsjańskiej marionetki” pełnoprawnym posiadaczem postaw propozycjonalnych. Specyfika tej istoty polegałaby na tym, że procesy przetwarzania informacji, które w normalnych okolicznościach dokonują się wewnątrz organizmu (dokładniej, wewnątrz jego czaszki), w jej przypadku odbywałyby się w sporej odległości przestrzennej w stosunku do tego organizmu, na przykład na Marsie.

³³ Hurley [2008] s. 21, tłum. P.G. Hurley używa w przytoczonym cytacie terminu „korespondencja” dla oznaczenia międzypoziomowej izomorfii.

³⁴ Davies [1998], [2000]. Przyjmuje się tu, że hipoteza języka myśli jest zgodna z twierdzeniem o zachodzeniu międzypoziomowej izomorfii. Mówiąc precyzyjniej, stanowi ona bardziej szczegółowe rozwinięcie tego twierdzenia.

ludzkich jako racjonalnych, by fakt, iż są one racjonalne *nie* zakładał tego, że ich zdolność do racjonalnego myślenia ma u swoich podstaw subosobowy język myśli³⁵.

Powyższy wniosek ma zasadnicze znaczenie dla prowadzonego tu wywo-
du. Argument z racjonalności opiera się na dużo mocniejszym niż u Daviesa uję-
ciu relacji między poprawnością opisu istot ludzkich jako racjonalnych, a teoriami
subosobowej architektury umysłu. Otóż argument ten przyjmuje zachodzenie
międzypoziomowej izomorfii jako warunek *konieczny* tego, by wyjaśnienie racjo-
nalności – dokładniej, zdolności do posiadania przekonań, pragnień i innych sta-
nów wyróżnianych przez psychologię potoczną – było w ogóle możliwe. Zgodnie
z tą perspektywą, jeżeli nie będziemy traktować subosobowych symbolicznych
procesów obliczeniowych w osobowych kategoriach psychologii potocznej (to jest,
nie będziemy traktować obydwu poziomów jako izomorficznych, na przykład
poprzez przyjęcie koncepcji języka myśli), nie będziemy mogli także wyjaśnić racjo-
nalności czy też nie będziemy w stanie nadać naukowego sensu idei, że okre-
ślane rodzaje systemów poznawczych egzemplifikują cechę bycia racjonalnym.
W tym celu należałoby uznać, że istnieje *konieczny, pojęciowy* związek pomiędzy
(1) byciem istotą posiadającą zdroworozsądkowe postawy propozycjonalne
a (2) byciem istotą, której procesy umysłowe opierają się na języku myśli czy do-
wolnej subosobowej architekturze obliczeniowej spełniającej warunek międzypo-
ziomowej izomorfii. W takiej perspektywie, (2) stanowiłoby „ugruntowany poję-
ciowo” warunek konieczny (1)³⁶. Warunek izomorfii byłby tym samym w pewien
sposób pojęciowo wpisany w psychologię potoczną, w pojęcia przekonania, pra-
gnienia i innych postaw propozycjonalnych. Taka pozycja jest jednak bardzo
problematyczna. Jest *prima facie* mało wiarygodne, by psychologia potoczna
i praktyka interpretowania siebie oraz innych w kategoriach zdroworozsądko-
wych stanów mentalnych były powiązane pojęciowo z jakimikolwiek *naukowymi*
twierdzeniami dotyczącymi subosobowej architektury umysłu. Intuicję tę wspie-
rają dodatkowo argumenty przedstawione przez Hendersona i Horgana³⁷. Po
pierwsze, kiedy wyobrażamy sobie, że zachodzi sytuacja, w której rozwój nauk
kognitywnych doprowadził do wykazania, że warunek izomorfii międzypozi-
omowej nie jest spełniony (na przykład, że hipoteza języka myśli jest fałszywa), to
nie wydaje się intuicyjnie wiarygodne, byśmy mogli powiedzieć o takim scenariu-
szu, że okazało się w nim, iż istoty ludzkie nie posiadają tak naprawdę przekonań
i pragnień. Gdyby warunek izomorfii był „wbudowany” pojęciowo w psychologię

³⁵ Davies [1998], por. też obronę podobnego poglądu zawartą w Lycan [2004].

³⁶ Por. Henderson, Horgan [2004].

³⁷ Ibidem.

potoczną, nasze intuicje semantyczne powinny być inne. Po drugie, autorzy ci zwracają uwagę na to, że wiele z pojęć o zasadniczym znaczeniu dla ludzkiego życia społecznego (stwierdzenie, posiadanie racji, uzasadnianie, działanie) presuponuje adekwatność opisu ludzkich umysłów i działań w kategoriach psychologii potocznej. Gdyby więc psychologia potoczna była rzeczywiście powiązana pojęciowo z tezą o izomorfii, mogłoby to – w sytuacji, gdyby teza o izomorfii okazała się fałszywa – pozbawić psychologię potoczną jej ważnych zastosowań. Aby psychologia potoczna mogła spełniać swoje zadania, powinna być wolna od tego rodzaju nadmiernie restryktywnych zobowiązań pojęciowych. Jednocześnie kryteria tego, jak używamy pojęć, mają przecież charakter w dużym stopniu pragmatyczny. Tym samym jest mało prawdopodobne, by psychologia potoczna była powiązana pojęciowo z tezami, które mogą w zasadniczy sposób zagrozić jej użyteczności w ludzkim życiu społecznym. Po trzecie, nie mamy problemów z wyobrażaniem sobie, że odrzucamy jakąś teorię opartą na twierdzeniu o izomorfii międzypoziomowej (na przykład, hipotezę języka myśli). Jednak wyobrażenie sobie, że odrzucamy kategorie oparte na psychologii potocznej – na przykład pojęcie celowego działania – jest bardzo trudne lub wręcz niemożliwe. Porzucenie pojęcia celowego działania *samo* musiałoby być rozumiane jako celowe działanie, podjęte na podstawie pewnych racji. Taka asymetria w możliwości wyobrażenia sobie obu wymienionych ewentualności dość naturalnie sprzyja twierdzeniu, że nie istnieje konieczny związek pojęciowy pomiędzy psychologią potoczną a tezą o izomorfii międzypoziomowej.

Wszystkie te ustalenia stanowią łącznie dodatkową, silną rację za odrzuceniem argumentu z racjonalności. Argument ten nie tylko jest, co wykazał Ramsey, cyrkularny, ale dodatkowo oparty jest na przyjętej implicite, fałszywej przesłance o międzypoziomowej izomorfii jako warunku wyjaśnienia racjonalności. Tym samym wskazano kolejną rację za tezą, że nie mamy podstaw, by w ramach symbolicznego modelu architektury obliczeniowej umysłu przyjmować potoczne pojęcie reprezentacji mentalnej. W tym też sensie, przedstawiona wyżej argumentacja Ramseya została dodatkowo wzmocniona.

Co jednak bardzo istotne, przedstawione rozważania pozwalają także uniezależnić odrzucenie argumentu z racjonalności – a wraz z nim potocznego pojęcia reprezentacji jako tego, którym powinny operować nauki kognitywne – od pewnych filozoficznych konsekwencji wyводу przeprowadzonego przez samego Ramseya. Wspomniano już, że autor ten wydaje się sądzić, iż racjonalność ludzkiego myślenia wykracza poza spektrum eksplanandów nauk kognitywnych, co może stanowić podstawę do przyjęcia względem niej pozycji eliminatywistycznej. Ujmując rzecz inaczej, jeśli w ramach teorii umysłu tworzonych przez kognitywi-

stów nie ma konceptualnego i empirycznego „miejsca” na racjonalność i postawy propozycjonalne, to tym gorzej dla ontologicznego/metafizycznego statusu racjonalności i postaw propozycjonalnych. Jeśli jednak przeprowadzone rozważania są poprawne, taką eliminatywistyczną perspektywę można oddalić. Fundamentalne znaczenie ma oczywiście odrzucenie izomorfii pomiędzy poziomem osobowym a subosobowym jako warunku koniecznego, by można było o danym systemie poznawczym orzec, że jest racjonalnym podmiotem zdroworozsądkowych stanów reprezentacyjnych. Oznacza to, że język myśli ani żadna podobna subosobowa architektura obliczeniowa nie jest konieczna do tego, by dana istota była, jak to nazwał Dennett, „naprawdę przekonana” (*true believer*)³⁸. Tym samym racjonalność może przysługiwać danemu systemowi bez względu na architekturę subosobowych mechanizmów. Chociaż argument z racjonalności nie stanowi więc dobrego uzasadnienia dla postulatu, aby subosobową maszynę umysłu rozumieć w kategoriach psychologii potocznej, to jego odrzucenie w żaden sposób *nie* oznacza odrzucenia przekonania, że określone *systemy poznawcze* mogą być całkiem dosłownie racjonalne, a w związku z tym zdolne do posiadania przekonań, pragnień, nadziei i innych postaw propozycjonalnych.

Oczywiście taka konstatacja otwiera miejsce dla pytania o to, jakie w takim razie warunki musi spełniać system, o którym moglibyśmy powiedzieć, że jest „naprawdę przekonany”. Analogicznie można pytać o to, jakie własności albo kompetencje systemowe muszą zostać wyjaśnione przez nauki kognitywne, by można było powiedzieć, że te wyjaśniły racjonalność. Udzielenie kompletnej odpowiedzi na te pytania jest tu niemożliwe. Można jednak podjąć próbę naszkicowania dwóch koniecznych warunków poprawnej atrybucji racjonalności i postaw propozycjonalnych danego systemu poznawczego. Po pierwsze, system jako całość musi wykazywać wzorce zachowania, które sprawiają, że jego zachowanie daje się w sposób optymalny interpretować i przewidywać za pomocą kategorii psychologii potocznej³⁹. Po drugie, działanie systemu musi dokonywać się na podstawie autonomicznych stanów i procesów, na przykład system taki nie może być Peacocke’owską marsjańską marionetką⁴⁰. Warunki te byłyby osobno ko-

³⁸ Dennett [2003].

³⁹ Dennett [2008], Henderson, Horgan [2004]. Zauważmy, że warunek ten można uszczegóławiać na różne sposoby. Dennett mówi o możliwości przewidywania działania systemu za pomocą normatywnej, „racjonalizującej” strategii predykcyjnej nazywanej przez niego „strategią intencjonalną” (*intentional stance*). Z kolei Henderson i Horgan (jak się wydaje, inspirowani są ideą, że psychologia potoczna jest teorią na wzór teorii naukowej) piszą, że stany intencjonalne przypisywane systemowi powinny podporządkowywać się „generalizacjom psychologii potocznej”.

⁴⁰ Por. przyp. 16 oraz Henderson, Horgan [2004]. Oczywiście, użyta tu kategoria „autonomii” jest daleka od jasności i należałoby ją w znaczący sposób doprecyzować.

nieczne do bycia racjonalnym podmiotem postaw propozycjonalnych, a łącznie wystarczające⁴¹. Są one filozoficznie nietrywialne, a jednocześnie bardzo liberalne jeśli chodzi subosobową architekturę stojącą u podstaw zdolności do racjonalnego myślenia.

Zakończenie

Przedstawiony w ostatniej części tego artykułu wywód warto jeszcze raz umieścić w szerszym kontekście. Teoretycznym celem pracy Williama Ramseya stanowiącej punkt wyjściowy przedstawionych tu rozważań, jest krytyczna analiza problemu istnienia reprezentacji mentalnych oraz eksplanacyjnego statusu pojęcia reprezentacji mentalnej w ramach nauk kognitywnych. W pierwszej części artykułu przedstawiona została ogólnie metoda Ramseya, polegająca na (1) dostarczeniu analizy pojęcia reprezentacji w kategoriach funkcjonalnych (tj. w kategoriach służenia czy funkcjonowania jako reprezentacja), (2) rekonstrukcji różnych pojęć reprezentacji używanych w obrębie nauk kognitywnych, (3) weryfikacji, czy struktury konceptualizowane jako reprezentacje w ramach nauk kognitywnych spełniają funkcjonalne kryteria bycia reprezentacją. W drugiej części artykułu przedstawiono argumentację Ramseya przeciwko tezie, że klasyczny, symboliczny model obliczeniowej architektury umysłu pozwala kognitywistom na przyjęcie zdroworozsądkowego, zaczerpniętego z psychologii potocznej pojęcia reprezentacji mentalnej. Przede wszystkim zrekonstruowano przeprowadzoną przez Ramseya krytykę argumentu, zgodnie z którym racjonalność przysługująca ludzkiemu myśleniu sprawia, że nauki kognitywne powinny interpretować symboliczne procesy obliczeniowe w kategoriach psychologii potocznej. Wreszcie w trzeciej części podjęto próbę wprowadzenia do wyводу Ramseya dość istotnych uzupełnień. Przede wszystkim chodziło o zdiagnozowanie pominiętego przez Ramseya, ukrytego założenia argumentu z racjonalności, który filozof ten krytykował. Jak się okazało, ostatecznie problem z tym argumentem nie polega tylko na tym, że racjonalność definiowana jest w nim za pomocą zdroworozsądkowych postaw propozycjonalnych, przez co argument ten jest cyrkularny. Głębokim źródłem niekonkluzywności argumentu z racjonalności jest to, że implicite zakłada się w nim, iż wyjaśnienie tak zdefiniowanej racjonalności wymaga traktowania subosobowych mechanizmów jako odpowiadających stanom i procesom wyróżnianym na poziomie osobowym. Pozbycie się tego założenia o międzypoziomowej izomorfii nie tylko stanowi kolejną podstawę do odrzucenia potocznego pojęcia re-

⁴¹ Por. Henderson, Horgan [2004].

prezentacji z konceptualnego repertuaru nauk kognitywnych⁴². Pozwala ono także pokazać, że takie odrzucenie wcale nie implikuje eliminatywizmu względem racjonalności i postaw propozycjonalnych. W dalszej perspektywie można by na takiej podstawie argumentować, że nawet przyjęcie *całkowicie* rezygnującej z pojęcia reprezentacji teorii subosobowej architektury umysłu nie oznacza jeszcze, że nie możemy całościowo traktować systemu poznawczego jako racjonalnego posiadacza przekonań i pragnień.

Bibliografia

- Bechtel [1994] – W. Bechtel, *Levels of description and explanation in cognitive science*, „Minds and Machines” (4) 1994, s. 1-25.
- Bechtel [1998] – W. Bechtel, *Representations and cognitive explanations: assessing the dynamicist's challenge in cognitive science*, „Cognitive Science” (22) 1998, s. 1-25.
- Bechtel [2005] – W. Bechtel, *The challenge of characterizing operations in the mechanisms underlying behavior*, „Journal of Experimental Analysis of Behavior” (84) 2005, s. 313-325.
- Bermudez [2000] – J.L. Bermudez, *Personal and sub-personal: a difference without a distinction*, „Philosophical Explorations” (3) 2000, s. 63-82.
- Chemero [2009] – A. Chemero, *Radical Embodied Cognitive Science*, MIT Press, Cambridge (MA) 2009.
- Cummins [2000] – R. Cummins, *'How does it work?' versus 'What are the laws?': Two conceptions of psychological explanation*, [w:] *Explanation and Cognition*, red. F. Keil, R.A. Wilson, MIT Press, Cambridge (MA) 2000.
- Davies [1998] – M. Davies, *Language, thought and the language of thought*, [w:] *Language and Thought*, red. P. Carruthers, J. Boucher, Cambridge University Press, Cambridge 1998.
- Davies [2000] – M. Davies, *Persons and their underpinnings*, „Philosophical Explorations” (3) 2000, s. 43-62.
- Dennett [1995] – D. Dennett, *Osobowy i subosobowy poziom wyjaśniania: ból*, tłum. P. Dziliński, [w:] *Filozofia umysłu*, red. B. Chwedeńczuk, Aletheia, Warszawa 1995.
- Dennett [2003] – D. Dennett, *Naprawdę przekonani: strategia intencjonalna i dlaczego ona działa*, „Przegląd Filozoficzno-Literacki” (4) 2003, s. 87-109.
- Dennett [2008] – D. Dennett, *Rzeczywiste wzorce*, (w:) *Analityczna metafizyka umysłu. Najnowsze kontrowersje*, red. M. Miłkowski, R. Poczobut, IFiS PAN, Warszawa 2008.
- Fodor [1975] – J. Fodor, *The Language of Thought*, Harvard University Press, Cambridge (MA) 1987.

⁴² Przynajmniej na subosobowym poziomie opisu i wyjaśniania działania systemu poznawczego.

- van Gelder [1995] – T. van Gelder, *What might cognition be if not computation?*, „The Journal of Philosophy” (92) 1995, s. 345-381.
- Goldman [2006] – A. Goldman, *Simulating Minds. Philosophy, Psychology and Neuroscience of Mindreading*, Oxford University Press, Oxford 2006.
- Grush [2003] – *In defense of some ‘Cartesian’ assumptions concerning the brain and its operation*, „Biology and Philosophy” (18) 2003, s. 53-93.
- Grush [2006] – R. Grush, *How to, and how not to, bridge computational cognitive neuroscience and Husserlian phenomenology of time consciousness*, „Synthese” (153) 2006, s. 417-450.
- Heal [1996] – J. Heal, *Simulation, theory and content*, [w:], *Theories of Theories of Mind*, red. P. Carruthers, P. Smith, Cambridge University Press, Cambridge 1996.
- Henderson, Horgan [2004] – D. Henderson, T. Horgan, *What does it take to be a true believer? Against the opulent ideology of eliminative materialism*, (w:) *Mind as a Scientific Object: Between Brain and Culture*, red. Ch.E. Erneling, D.M. Johnson, Oxford University Press, Oxford 2004.
- Hershbach [2008] – M. Hershbach, *Folk psychological and phenomenological accounts of social perception*, „Philosophical Explorations” (11) 2008, s. 223-235.
- Hornsby [2000] – J. Hornsby, *Personal and sub-personal: A defence of Dennett’s early distinction*, „Philosophical Explorations” (3) 2000, s. 6-24.
- Hurley [2008] – S. Hurley, *The shared circuits model. How control, mirroring, and simulation can enable deliberation, imitation and mindreading*, „Behavioral and Brain Sciences” (31) 2008, s. 1-22.
- Lycan [2004] – W. Lycan, *A particularly compelling refutation of eliminative materialism*, [w:] *Mind as a Scientific Object: Between Brain and Culture*, red. Ch.E. Erneling, D.M. Johnson, Oxford University Press, Oxford 2004.
- McDowell [1994] – J. McDowell, *The content of perceptual experience*, „Philosophical Quarterly” (44) 1994, s. 190-205.
- Nichols [2002] – S. Nichols, *Folk psychology*, [w:] *Encyclopedia of Cognitive Science*, Nature Publishing Group, London 2002.
- Peacocke [1983] – Ch. Peacocke, *Sense and Content*, Oxford University Press, Oxford 1983.
- Pylyshyn [1984] – Z. Pylyshyn, *Computation and Cognition*, MIT Press, Cambridge (MA) 1984.
- Ramsey [2007] – W. Ramsey, *Representation Reconsidered*, Cambridge University Press, Cambridge 2007.
- Ramsey, Stich, Garon [1990] – W. Ramsey, S. Stich, J. Garon, *Connectionism, eliminativism and the future of folk psychology*, „Philosophical Perspectives” (4) 1990, s. 499-533.
- Stich [1985] – S. Stich, *From Folk Psychology to Cognitive Science*, MIT Press Cambridge (MA) 1985.
- Thagard [1996] – P. Thagard, *Cognitive Science*, [w:] E.N. Zalta, *Stanford Encyclopedia of Philosophy*, dostępne na: <http://plato.stanford.edu/entries/cognitive-science/> [10.02.2011].