

Kinga Jęczyńska, Streszczenie rozprawy doktorskiej "Consciousness beyond the Cartesian Theatre: Contemporary Anti-Cartesian Theories of Consciousness"

W rozprawie doktorskiej analizuję antykartezjańskie teorie świadomości, czyli teorie, które odrzucają model teatru kartezjańskiego opisany przez Dennetta (1991): teorię globalnej przestrzeni roboczej Baarsa, model wielorakich szkiców Dennetta, teorię sensomotoryczną, którą zaproponowali O'Regan i Noë, oraz teorię kodowania predykcijnego. Teorie te odrzucają założenie o istnieniu punktu centralnego w umyśle, przez który musiałaby przejść każda świadoma informacja. Porównuję teorię globalnej przestrzeni roboczej, model wielorakich szkiców, konserwatywną teorię kodowania predykcijnego i radykalną teorię kodowania predykcijnego pod względem tego, w jaki sposób odchodzą od modelu teatru kartezjańskiego. Analizuję także moc eksplanacyjną poszczególnych teorii antykartezjańskich, skupiając się na tym, jak rozwiązują one problemy dotyczące świadomości bez uciekania się do tradycyjnego rozumienia qualiów.

Punktem wyjścia rozprawy doktorskiej był artykuł Degenaar i Keijzera (2009), w którym autorzy wskazują, że dwie teorie świadomości – teoria globalnej przestrzeni roboczej i teoria sensomotoryczna – są ze sobą kompatybilne, a ich synteza może wyjaśnić więcej zagadnień niż każda z tych teorii z osobna, ponieważ teorie te skupiają się na rozwiązywaniu innych problemów świadomości. Degenaar i Keijzer krótko zarysowują trzy możliwe sposoby połączenia tych teorii: scenariusz lokalizacji wewnętrznej, scenariusz lokalizacji zewnętrznej i scenariusz braku lokalizacji. W swojej rozprawie analizuję podstawowe założenia teorii antykartezjańskich i rozważam wszystkie trzy sposoby połączenia teorii globalnej przestrzeni roboczej z teorią sensomotoryczną. Ukazuję, że możliwe jest połączenie teorii globalnej przestrzeni roboczej z teorią sensomotoryczną według scenariusza lokalizacji zewnętrznej za pomocą radykalnej teorii kodowania predykcijnego rozwijanej przez Clarka (2015a, 2015b, 2015c, 2016).

Uzyskaną w ten sposób nową teorię, mającą postać hybrydowej architektury umysłu, wykorzystuję do wyjaśnienia podstawowych problemów związanych ze świadomością omawianych we współczesnej literaturze filozoficznej i kognitywistycznej: pojęcia świadomości; integracji świadomie przetwarzanej informacji i jej zrozumiałości dla wszystkich nieświadomych procesorów; wrażeń i qualiów; absolutnej luki eksplanacyjnej i komparatywnej luki eksplanacyjnej; różnic pomiędzy perspektywą pierwszoosobową i

trzecioosobową oraz poziomem osobowym i podosobowym; problemu łączenia (ang. *binding problem*); uwagi i penetracji poznawczej (ang. *cognitive penetrability*). Wykorzystuję także nową teorię do wyjaśnienia: zmian w procesach poznawczych występujących w chorobach psychiatrycznych i urojeniach; ślepoty na zmianę i ślepoty z nieuwagi; ślepowidzenia oraz zmian w percepcji spowodowanych soczewkami odwracającymi obraz i soczewkami dwukolorowymi.

We wstępie krótko opisuję model teatru kartezjańskiego oraz podział na słaby i silny antykartezjanizm. Omawiam tradycyjne argumenty, które podtrzymują istnienie qualiów i trudnego problemu świadomości. Prezentuję również kontrargumenty, które krytykują tradycyjne pojęcie qualiów. Wyjaśniam cel rozprawy, którym jest analiza współczesnych antykartezjańskich teorii świadomości i przedstawienie syntezy dwóch z nich w obrębie jednego z trzech scenariuszy nakreślonych przez Degenaara i Keijzera.

W rozdziale 1. przedstawiam architekturę umysłu zaproponowaną przez Baarsa jako reprezentatywny przykład dla grupy teorii globalnej przestrzeni roboczej (Baars 1988; 1997a; 1997b; 2002; 2007, Baars *et al.* 2013). Ukazuję, w jaki sposób teoria ta analizuje funkcje świadomości. Przedstawiam związki między teorią globalnej roboczej i współczesnymi badaniami z zakresu neuronauk. Wyjaśniam, w jaki sposób teoria globalnej przestrzeni roboczej odrzuca model teatru kartezjańskiego i wprowadza nową zmodyfikowaną metaforę teatru, opisującą interakcje między procesami świadomymi i nieświadomymi. Przedstawiam, w jaki sposób teoria ta wyjaśnia często dyskutowane zagadnienia związane z badaniami nad świadomością, np. strumień świadomości, eksperymenty Libeta dotyczące wolnej woli, ślepowidzenie. Zarysowuję ograniczenia tej teorii, które są związane głównie z tym, że teoria ta skupia się na absolutnej luce eksplanacyjnej, a pomija komparatywną lukę eksplanacyjną.

W rozdziale 2. opisuję model umysłu ukazany w modelu wielorakich szkiców Dennetta. Przedstawiam jego związki z badaniami empirycznymi dotyczącymi funkcjonowania umysłu. Porównuję model wielorakich szkiców z teorią globalnej przestrzeni roboczej Baarsa. Te dwie teorie są podobne ze względu na opisywanie świadomości w sposób funkcjonalistyczny, jako wynikającej ze współdziałania rozproszonych procesorów pracujących równolegle. Model wielorakich szkiców różni się od teorii globalnej przestrzeni roboczej z powodu odrzucania w nim wyraźnej granicy między procesami świadomymi a nieświadomymi. Wskazuję, jak bardzo na ten model wpłynęły poglądy filozoficzne Dennetta, w których kluczową rolę odgrywają pojęcia nastawienia intencjonalnego, heterofenomenologii i interpretywizmu. Omawiam ograniczenia modelu wielorakich szkiców, które są związane z tym, że nie jest on w stanie wyjaśnić istotnych funkcji świadomości. Do

tych problemów należą: koordynująca i integrująca rola świadomości, ważne aspekty jaźni i wpływ ucieleśnienia na stany świadome.

W rozdziale 3. przedstawiam teorię sensomotoryczną, którą rozwijali O'Regan i Noë (O'Regan, Noë 2001a, 2001b, 2001c; Noë 2004; O'Regan 2011). Ukazuję, jak odchodzi ona od tradycyjnego reprezentacjonizmu i fizykalizmu. Wyjaśniam, w jaki sposób teoria ta przyjmuje podstawowe założenia enaktywizmu. Opisuję silnie antykartezjański charakter teorii sensomotorycznej, który jest związany z odrzuceniem zarówno modelu teatru kartezjańskiego, jak i wyraźnego podziału na mózgowie i resztę ciała. Omawiam pojęcie reguł sensomotorycznych i ich rolę w wyjaśnianiu doświadczenia na poziomie osobowym. Wyjaśniam, w jaki sposób udaje się tej teorii ominąć trudny problem świadomości i zrećźnie przeformułować pewne tradycyjne problemy, takie jak luka eksplanacyjna Levine'a, dzięki przededefiniowaniu qualiów. Teoria ta opisuje różnice między modalnościami sensorycznymi i różnice między doświadczeniami w obrębie tej samej modalności sensorycznej, odwołując się do pojęcia reguł sensomotorycznych. Ponieważ teoria sensomotoryczna zamyka intermodalną i intramodalną komparatywną lukę eksplanacyjną, ale nie zajmuje się absolutną luką eksplanacyjną, korzystne jest połączenie tej teorii z teorią globalnej przestrzeni roboczej.

W rozdziale 4. omawiam najistotniejsze doświadczenia i dane empiryczne związane z teorią sensomotoryczną. Wskazuję, że teoria sensomotoryczna jest kompatybilna z wynikami badań dotyczących ruchów sakadowych, plamki ślepej, systemów substytucji sensorycznej, ślepoty na zmianę i ślepoty z nieuwagi, soczewek odwracających obraz, soczewek dwukolorowych i podejścia ewolucyjnego. Teorię sensomotoryczną można określić jako cechującą się wyższą mocą eksplanacyjną niż teorie alternatywne, np. tradycyjny reprezentacjonizm.

W rozdziale 5. opisuję teorię kodowania predykcyjnego, zwracając szczególną uwagę na jej wersję rozwijaną przez Clarka (2015a, 2015b, 2015c, 2016), którą później wykorzystuję do wykonania głównego zadania w rozprawie, polegającego na połączeniu teorii globalnej przestrzeni roboczej z teorią sensomotoryczną. Ukazuję, w jaki sposób radykalna teoria kodowania predykcyjnego wyjaśnia działanie i percepcję – wraz z interocepcją. Ta szczególna wersja teorii kodowania predykcyjnego oddaje sprawiedliwość ścisłym związkom między percepcją a działaniem podkreślanym przez enaktywizm, który stanowi inspirację dla teorii sensomotorycznej. Wskazuję, w jaki sposób radykalna teoria kodowania predykcyjnego rozwijana przez Clarka różni się od bardziej konserwatywnych teorii kodowania predykcyjnego. Ponadto przedstawiam badania Hobsona i Fristona nad snami i rozwojem

świadomości, które mają związek z analizą świadomości przez pryzmat metafory rzeczywistości wirtualnej w teorii kodowania predykcyjnego. Ukazuję, że teorię kodowania predykcyjnego można określić jako teorię antykartezjańską, gdyż odrzuca ona model teatru kartezjańskiego. Radykalna teoria kodowania predykcyjnego może być również antykartezjańska w silnym znaczeniu tego słowa, ponieważ może w jej ramach zostać także odrzucony ścisły podział na mózgowie i resztę ciała. Wyjaśniam, w jaki sposób można do analizy teorii kodowania predykcyjnego korzystać ze zmodyfikowanej metafory teatru, która jest związana z modelem rzeczywistości wirtualnej stworzonej na podstawie hipotez na temat świata.

W rozdziale 6. opisuję trzy scenariusze wyróżnione przez Degenaara i Keijzera (2009) jako możliwe sposoby połączenia teorii globalnej przestrzeni roboczej z teorią sensomotoryczną: scenariusz lokalizacji wewnętrznej, scenariusz lokalizacji zewnętrznej i scenariusz braku lokalizacji. Przedstawiam syntezę teorii globalnej przestrzeni roboczej z teorią sensomotoryczną za pomocą radykalnej teorii kodowania predykcyjnego w obrębie drugiego typu syntezy spośród trzech wyróżnionych przez Degenaara i Keijzera (2009). Synteza w obrębie scenariusza lokalizacji zewnętrznej ma wyższą moc eksplanacyjną niż każda z tych teorii z osobna. Twierdzę, że pozostałe dwa scenariusze są mniej atrakcyjne niż wybrany scenariusz lokalizacji zewnętrznej, ale nie wykluczam możliwości połączenia tych teorii w obrębie pozostałych scenariuszy przez innych badaczy.

Ponadto zwięźle omawiam zagadnienia metodologiczne związane z integracją i unifikacją teorii: problem integracji (ang. *integration challenge*) w unifikacji teorii opisany przez Bermúdeza (2010/2011), starsze podejścia do integracji (redukcja międzyteoretyczna czy analiza w terminach podejścia trzypoziomowego Marra), pojęcie architektury umysłu, a w szczególności hybrydowej architektury umysłu (Bermúdez 2010/2011) i integrację międzypoziomową (Craver, Darden 2013). Porównuję także teorię globalnej przestrzeni roboczej, teorię sensomotoryczną i teorię kodowania predykcyjnego pod względem internalizmu i eksternalizmu eksplanacyjnego. Omawiam wczesne próby połączenia teorii globalnej przestrzeni roboczej z teoriami umysłu ucieleśnionego (Shanahan 2005) oraz konserwatywnej teorii kodowania predykcyjnego z teorią globalnej neuronalnej przestrzeni roboczej (Hohwy 2013, 2015).

W rozdziale 7. porównuję teorię globalnej przestrzeni roboczej, model wielorakich szkiców, konserwatywną teorię kodowania predykcyjnego, radykalną teorię kodowania predykcyjnego i teorię sensomotoryczną pod względem tego, w jaki sposób odrzucają model

teatru kartezjańskiego. Zwięźle opisuję elementy filozofii kartezjańskiej mające związek z pojęciem teatru kartezjańskiego omawianym przez Dennetta. Wyliczam cechy charakterystyczne dla kartezjańskiego sposobu myślenia o świadomości. Wskazuję, że analizowane teorie antykartezjańskie różnią się pod względem podejścia do modelu kartezjańskiego. Proponuję analizę tych teorii za pomocą skali porządkowej, na której możliwe jest oszacowanie różnych sposobów odchodzenia przez teorie od modelu teatru kartezjańskiego. Teorie te modyfikują wyjściową metaforę teatru kartezjańskiego i wprowadzają swoje własne zmodyfikowane metafory teatru. Pokazuję, jak teoria kodowania predykcyjnego może stanowić pomost między dwoma rodzajami modeli: modelem konserwatywnym przyjętym przez teorię globalnej przestrzeni roboczej i częściowo przez model wielorakich szkiców a modelem radykalnym wspieranym przez teorię sensomotoryczną.

W rozdziale 8. przedstawiam, jak stanowisko powstałe w wyniku unifikacji teorii antykartezjańskich może wyjaśnić problemy, które były wcześniej dyskutowane z niezależnych perspektyw analizowanych teorii. Złożony opis świadomości, zawierający elementy teorii globalnej przestrzeni roboczej i teorii sensomotorycznej połączonych za pomocą radykalnej teorii kodowania predykcyjnego w obrębie scenariusza lokalizacji wewnętrznej, może dostarczyć pełniejszego wyjaśnienia problemów zazwyczaj stawianych w odniesieniu do świadomości: pojęcia świadomości; integracji świadomie przetwarzanych informacji i ich zrozumienia przez wszystkie nieświadome procesory; wrażeń i qualiów; absolutnej luki eksplanacyjnej i komparatywnej luki eksplanacyjnej; podziału na perspektywę pierwszoosobową i trzecioosobową; problemu łączenia (ang. *binding problem*); penetracji poznawczej (ang. *cognitive penetrability*); zmian poznawczych w chorobach psychiatrycznych i autyzmie; uwagi; eksperymentów dotyczących ślepoty na zmianę i ślepoty z nieuwagi; ślepowidzenia; zmian w percepcji spowodowanych soczewkami odwracającymi obraz i soczewkami dwukolorowymi. Połączenie teorii globalnej przestrzeni roboczej i teorii sensomotorycznej za pomocą radykalnej teorii kodowania predykcyjnego skutkuje pełniejszym opisem tych standardowych zagadnień.

Przedstawiona hybrydowa architektura umysłu zawiera elementy kilku współczesnych stanowisk na temat świadomości. Chociaż każde z tych stanowisk ma swoje mocne strony i może być wykorzystywane oddzielnie do analizy wybranych problemów, na ogólniejszym poziomie teoretycznym może być korzystne badanie świadomości w bardziej kompleksowy sposób.