

MICHAŁ WYRWA

Uniwersytet im. Adama Mickiewicza w Poznaniu

Instytut Psychologii

WŁASNA A CUDZA ŚWIADOMOŚĆ. PRÓBA WYJŚCIA Z NAUKOWEGO LABIRYNTU ŚWIADOMOŚCI NA PRZYKŁADZIE FILOZOFII EKSPERYMENTALNEJ

Wstęp

Choć o znaczeniu badania zjawiska świadomości mówiło się od początku rozwoju psychologii naukowej, to zainteresowanie nią było zmienne. Z jednej strony obiekt badań psychologii według jej twórców (James, 2002; Wundt, 1921), a z drugiej – jedna z wielu *personae non gratae* behawiorystów (Blanshard i Skinner, 1967). Mimo że zainteresowanie świadomością powróciło pod koniec XX wieku, do dziś badacze nie są zgodni co do tego, jak należy sformułować naukowy problem świadomości.

Istotny w tym względzie jest diagnozowany przez Davida Chalmersa (1995) trudny problem świadomości, postulowana niemożność wyjaśnienia świadomości fenomenalnej – zorientowanej wokół pierwszoosobowych doznań – w kategoriach czysto fizykalnych. Jak nietrudno się przekonać, sięgając po artykuły nawiązujące do tej tematyki z ostatnich lat (Carruthers i Schier, 2017), do argumentacji Chalmersa odwołują się najczęściej jego oponenti.

Nie można powiedzieć, że antyredukjonistyczne idee pokroju trudnego problemu świadomości wpłynęły w znaczący sposób na rozwój empirycznej nauki o świadomości. Mówiąc szerzej, problem ten dotyczy całej części filozofii umysłu, która rozważa zagadnienie poznania z perspektywy analitycznej. Niektórzy badacze pochyłili się nad problemem *qualiów* (Ramachandran i Hirstein, 1997), mierzalności świadomości fenomenalnej (Tononi, 2008), ale są oni w mniejszości. Jeśli świadomość ma niejako z definicji wykraczać poza wyjaśnienie empiryczne albo też samo rozumienie świadomości wyrażane jest przez miłą język analitycznej filozofii umysłu, to trudno oczekiwać, że odbije się on większym echem w świecie nauki. Jak zauważa Andrzej Klawiter (2005), wiedza o umyśle wyrażona jest w ję-

zyku nauki (empirycznej) i jeśli filozof chce coś do tego korpusu wiedzy wnieść, powinien robić to w sposób zrozumiały dla pozostałych badaczy.

Po latach obecności trudnego problemu świadomości w każdej dyskusji wyrażającej z filozofii umysłu, w sensie jego obrony bądź krytyki, zaczęły się ostatnio pojawiać nowe, interesujące wariacje na temat problemu świadomości. Jedno z takich pytań brzmi: Czy i w jaki sposób dokonujemy atrybucji stanów świadomych.

Najogólniej mówiąc, hipoteza mechanizmu przypisywania stanów świadomych głosi, że potrafimy przypisywać innym osobom świadome stany mentalne oraz że warto wyróżnić tego typu zdolność w arsenale naszych kompetencji poznawczych. Może się bowiem okazać, że choć faktycznie potrafimy dokonywać rzeczzonej atrybucji, to nie różni się ona w żaden sposób od „zwykłej” atrybucji stanów mentalnych.

To, że potrafimy szacować świadomość innych, jest dość jasne, pod warunkiem stosownego rozumienia świadomości. Diagnostyka takich stanów, jak śpiączka, stan wegetatywny czy minimalny stan świadomości (Bruno, Vanhaudenhuyse, Thibaut, Moonen i Laureys, 2011), polega najczęściej na korzystaniu z odpowiednich kwestionariuszy behawioralnych (Giacino, Kalmar i Whyte, 2004). Są to narzędzia służące właśnie do szacowania poziomu świadomości pacjentów. Przy czym szacuje się nie zdolność do doznawania, ale wymiar świadomości znany w literaturze filozoficznej pod pojęciem świadomości dostępu (Block, 1995). Skoro potrafimy wytworzyć takie narzędzia, to zasadne jest pytanie, czy potrafimy przypisywać świadomość także bez wiedzy eksperckiej, na poziomie psychologii potocznej.

Niniejszy tekst prezentuje koncepcję nastawienia fenomenalnego, czyli mechanizmu przypisywania świadomych stanów mentalnych. Wyrasta ona z programu badawczego filozofii eksperymentalnej.

1. Charakterystyka filozofii eksperymentalnej

Jednym z celów stawianych sobie przez filozofię eksperymentalną jest empiryczne testowanie intuicji dotyczących zagadnień filozoficznych. Korzystająca z analizy pojęciowej filozofia, a taką właśnie jest filozofia umysłu, opiera swe ustalenia ostatecznie na intuicji filozofów (Williamson, 2007). Akceptuje zatem jakąś wersję tezy o zgodności intuicji (Machery, Mallon, Nichols i Stich, 2004) lub twierdzi, że możemy ustalić poprawne intuicje drogą refleksji epistemologicznej (Weinberg, Nichols i Stich, 2001). W wersji słabej teza o zgodności mówi o podzieleniu tych samych, trafnych, intuicji przez określoną grupę ekspertów (filozofów), zaś w wersji silnej – że każdy podziela takie same intuicje. Filozofia eksperymentalna poprzez sprawdzanie powszechności intuicji jest projektem weryfikacyjnym względem

intuicji filozoficznych, a zatem stoi w opozycji do tezy zgodności (Nadelhoffer i Nahmias, 2007). Nietrudno napotkać argumenty starające się bronić tradycyjne podejście filozoficzne (Kauppinen, 2007), jednak zostaną one tutaj pominięte.

Badania filozofii eksperymentalnej dotyczą także tego, w jakim stopniu różnicowane są intuicje w populacji. Przykładem są wyniki prac zespołu Jonathana M. Weinberga (2001), które wskazują na występowanie różnych intuicji epistemicznych w zależności od kultury. Ludzie wychowani w kulturze azjatyckiej są bardziej skłonni przypisywać wiedzę, a nie tylko przekonania, niż ludzie z kultury zachodniej. Podobną zależność obserwuje się w przypadku statusu socjoekonomicznego. Ludzie uzyskujący niski dochód częściej przypisują innym stan wiedzy niż uzyskujący wysoki dochód. Z kolei wyniki zespołu Shauna Nicholasa (Machery i O'Neil, 2004; Mallon, MacHery, Nichols i Stich, 2009) wskazują na występowanie różnic na tle kulturowym w zakresie intuicji semantycznych, a dokładniej pojęcia odniesienia przedmiotowego (*reference*). Te intuicje u ludzi wychowanych w kulturze zachodniej częściej mają charakter przyczynowo-historyczny, jej przedstawiciele biorą bowiem pod uwagę to, co ktoś inny wie na temat odniesienia danej nazwy. Osoby wychowane w kulturach wschodnich skłaniają się natomiast ku deskrytywistycznemu rozumieniu odniesienia przedmiotowego w sensie Gottloba Fregego (1977).

Ze względu zarówno na selekcję testowanych intuicji co do pojęć, jak i dobieraną metodologię granica między filozofią eksperymentalną a naukami poznawczymi jest mglista. Badacze nie ograniczają się tylko do roli sprawozdawczej, próbują także modelować mechanizmy leżące u podstaw generowania danego typu intuicji (Fiala, Arico i Nichols, 2012) oraz poszukiwać ich neurokorelatów (Greene, 2001; Jack, Dawson, i Norr, 2013). Powszechnie znanym w psychologii rezultatem badań z zakresu filozofii eksperymentalnej jest efekt Knobe'a (2003; Nichols i Knobe, 2007), mówiący o wpływie moralnych konsekwencji działań na ich percypowaną intencjonalność. Możemy zatem pojmować filozofię eksperymentalną nie jako dziedzinę autonomiczną, gałąź filozofii, a jako interdyscyplinarne przedsięwzięcie.

Oprócz intuicji związanych z filozofią języka (Haukioja, 2015; Machery i O'Neill, 2014), metafizologią (Fischer i Collins, 2015) oraz moralnością swoistym leitmotivem filozofii eksperymentalnej jest atrybucja stanów mentalnych. Wynika to z dominującej metodologii, opierającej się głównie na kwestionariuszach wzorowanych na scenariuszach Edmunda L. Gettier'a (1963). Nie zawsze dotyczą one bezpośrednio wiedzy jako uzasadnionego prawdziwego przekonania, jak w przypadku oryginalnych badań Gettier'a. Jednak zawsze w przedstawianych historiach dowiadujemy się o pewnych faktach na temat świata oraz działaniach jednego lub kilku bohaterów nie zawsze znających owe fakty. Osoby te muszą się ustosunkować do prezentowanej treści, określając, jakiego typu stany mentalne przypisałiby bohaterowi historii. Zmiennymi niezależnymi mogą być w tym para-

dygmacie badawczym zarówno fakty na temat świata, jak i działania występujących postaci. Przykładowo, w oryginalnym badaniu Knobe'a badani dowiadywali się o kierowniku pewnej firmy, który mógł wdrożyć nowy program rozwoju, pozwalając znacząco zwiększyć zysk. W zależności od wersji skutki uboczne zmiany mogły mieć pozytywny lub negatywny wpływ na środowisko. Dyrektora firmy te konsekwencje nie obchodziły. Skutki uboczne wystąpiły. Czy zatem dyrektor wywołał je intencjonalnie? By odpowiedzieć na to pytanie, osoby badane musiały ocenić, czy dyrektor wiedział o programie rozwoju, o tym, jakie niósł ze sobą konsekwencje, i na którym przekonaniu polegał, gdy podejmował ostateczną decyzję.

Filozofia eksperymentalna, korzystając z metodologii obecnej w innych naukach, dziedziczy też jej ograniczenia i problemy. Uwagi o konieczności większej roli replikacji w badaniach psychologicznych (Yong, 2012) spotkały się w ostatnich latach z odpowiedzią w postaci m.in. *Reproducibility Project: Psychology*. Niemal dwóch trzecich spośród wybranych ważnych dla psychologii badań nie udało się zreplikować (Open Science Collaboration, 2015). Na gruncie filozofii eksperymentalnej podjęta została podobna inicjatywa – baza danych zbierająca replikacje eksperymentów¹. Powiązanym z kwestią replikacji problemem jest pojawiający się zarzut o *fishing expedition*, czyli poszukiwanie w danych jakichkolwiek zależności statystycznych i formułowanie odpowiednio ogólnych hipotez, które na to pozwalają (Alfano i Loeb, 2016). Oczywiście, ocena kondycji całej dyscypliny jest czymś innym od analizy konkretnych badań. Przy okazji krytyki koncepcji nastawienia fenomenalnego zwrócę uwagę na problemy bardziej szczegółowe.

2. Świadomość w badaniach filozofii eksperymentalnej

Filozofia eksperymentalna ma też sporo do powiedzenia na temat intuicji co do pojęcia świadomości, szczególnie zaś świadomości fenomenalnej. Bez trudu można znaleźć zarówno proponentów (Knobe i Prinz, 2008, Buckwalter i Phelan, 2013), argumentujących na rzecz posiadania takowego pojęcia na poziomie psychologii potocznej, jak i oponentów (Sytsma i Machery, 2009; 2010). Rzeczona dyskusja, choć porusza zagadnienia posiadania i posługiwania się pojęciem świadomości także w stosunku do innych ludzi niż my sami, to nie odnosi się wprost do kwestii tego, jak miałby działać i kiedy miałby być wykorzystywany taki mechanizm atrybucji. Dlatego też nie traktuję jej jako dyskusji nad jakąś konkretną koncepcją. Niemniej, by przedstawić pełniejszy obraz nurtu filozofii eksperymentalnej w odniesieniu do problematyki świadomości, pokrótce ją przybliżam.

¹ <http://experimental-philosophy.yale.edu/xphipage/Experimental%20Philosophy-Repliations.html> [30.01.2018].

W badaniu Johna Knobe i Jessego Prinza (2008) zakładali, że atrybucja stanów fenomenalnych jest odróżnialna od atrybucji innych stanów mentalnych oraz że pewne wypowiedzi niosą ze sobą atrybucję świadomości fenomenalnej. Plan eksperymentalny polegał na przedstawieniu szeregu stwierdzeń na temat stanów mentalnych (fenomenalnych i niefenomenalnych) oraz pytaniu, jak bardzo naturalnie brzmią one dla badanych. Kolejne etapy postępowania naukowego pozwoliły ustalić, że dla osób badanych wyrażenia typu: „szczegółowo wyobraża sobie”, „odczuwa wielką radość” brzmią nienaturalnie, gdy przypisać je agentom grupowym (np. firmom)². W przypadku, gdy jako agenta przedstawiono mebel, wyposażony w zdolność myślenia, planowania, komunikowania i emocje, badani chętnie dokonywali atrybucji świadomości. Badacze uznali, że kryterium atrybucji jest budowa ciała agenta, a nie po prostu podobieństwo do budowy człowieka. Mimo to wydaje się, że można ich traktować jako zwolenników tezy ucieleśnienia (*embodiment hypothesis*). Wedle niej kryterium dokonywania atrybucji doznań stanowi struktura biologiczna ciała agenta (Buckwalter i Phelan, 2014).

W odpowiedzi na wnioski sformułowane przez Knobe’a i Prinza Justin M. Sytsma i Edouard Machery (2009) opublikowali artykuł krytyczny. Ich zdaniem, aby stwierdzić, czy na poziomie psychologii potocznej posiadamy pojęcie świadomości fenomenalnej, należy wykazać prawdziwość dwóch tez. Pierwsza z nich mówi o tym, że faktycznie ludzie dokonują atrybucji stanów, które dla filozofów są stanami fenomenalnymi względem niektórych agentów. Natomiast druga uściśla, że kryterium kategoryzacji na posiadających i nieposiadających stany fenomenalne nie ogranicza się do różnic behawioralnych i funkcjonalnych pomiędzy agentami. Autorzy, zarzucając Knobe’owi i Prinzowi niespełnienie warunku drugiej tezy, uzasadniają jej istnienie faktem, że atrybucja stanów mentalnych mogłaby się odbywać w różny sposób względem różnych agentów nawet wtedy, gdy pojęcie świadomości fenomenalnej w psychologii potocznej (*folk psychology*) nie występuje. Innymi słowy, Knobe i Prinz popełniają błąd, gdy porównują agentów grupowych z jednostkowymi. Ci pierwsi funkcjonalnie i behawioralnie zasadniczo różnią się od drugich. Sytsma i Machery argumentują, że skoro nie sposób sobie wyobrazić firmy, która mogłaby się uśmiechać albo śmiać (behawioralne wskaźniki odczuwania radości), to nic dziwnego, że nienaturalnym wydaje się sformułowanie przypisujące jej odczuwanie radości. Przeprowadzone przez nich badanie wydaje się potwierdzać tę interpretację. Badani uznali za mniej naturalne wyrażenia, które przypisywały agentom grupowym zachowanie kojarzone z działaniem jednostkowym, niż takie, które w ten sposób się nie kojarzyły.

² Knobe i Prinz uważają, że atrybucja tego rodzaju wypowiedzi zakłada przypisanie świadomości fenomenalnej. Inaczej jest z – wydawałoby się, podobnie brzmiącymi – wyrażeniami typu: „Firma A rozgniewała się”, „Firma A cieszy się z”. Te ostatnie nie zawierają *explicite* zwrotów opisujących doznawania (np. „odczuwać”).

Część wyników Knobe'a i Prinza tłumaczy niefortunny dobór bodźców w eksperymentach. Chcąc ocenić naturalność brzmienia wyrażen zawierających, wedle badaczy, ukrytą atrybucję świadomości fenomenalnej, do porównania używali wyrażen rzekomo jej niezawierających. Sytsma i Machery zwrócili uwagę, że zdania z obu warunków różniło więcej niż tylko opis stanu świadomego bądź jego brak. Te drugie zawierały także wyrażenia przyimkowe, wyjaśniające wywołanie danego stanu u agenta grupowego. Przykładowo, „Firma A była zdenerwowana z powodu orzeczenia sądu” nie jest pozbawionym odwołania do doznawania odpowiednikiem zdania „Firma A odczuwała złość”. Po ponownym doborze bodźców i przeprowadzeniu badania efekt uzyskany przez Knobe'a i Prinza zniknął, a dla badanych równie naturalnie brzmiały obie grupy wyrażen.

Dalsze wyniki Sytsmy i Machery (2010) sugerują, że intuicje potoczne i filozoficzne co do subiektywnego doświadczenia są na tyle różne, że nietrafne byłoby szukanie pojęcia świadomości fenomenalnej na poziomie psychologii potocznej. Niefilozoficzne kryterium przypisywania stanów, które filozofowie nazywają fenomenalnymi, opiera się na ocenie wartości emocjonalnej (*valence*) tych stanów. Te, które jej nie posiadają, np. widzenie koloru czerwonego, są przypisywane zarówno ludziom, jak i prostym konstrukcyjnie robotom. Nie jest tak natomiast w przypadku tych stanów, które oceniamy jako przyjemne/nieprzyjemne (odczuwanie bólu, przyjemności).

Wesley Buckwalter i Mark Phelan (2013) otrzymali wyniki odmienne od Sytsmy i Machery (2010), stawiając poprzedniemu badaniu podobne zarzuty jak te wobec badań Knobe'a i Prinza, czyli niewłaściwy dobór bodźców. Wedle Buckwaltera i Phelana tym, co kieruje atrybucją subiektywnych doznań, nie jest walencja emocjonalna, ale charakterystyka funkcjonalna. Co więcej, w kolejnych badaniach występują oni przeciw tezie o ucieleśnieniu (Buckwalter i Phelan, 2014). Choć nie twierdzą, że struktura ciała nie stanowi w ogóle kryterium, to fakt swobodnej atrybucji przeżyć emocjonalnych agentom bezcielesnym przez badanych uznają oni za wynik na rzecz nadrzędnej roli w tym procesie charakterystyki funkcjonalnej.

Brak konsensusu co do występowania pojęcia świadomości fenomenalnego na poziomie psychologii potocznej pokazuje, jak znaczącym problemem w badaniach z zakresu filozofii eksperymentalnej jest kwestia właściwego dobrania bodźców. Badacze muszą zdecydować, jakiego rodzaju intuicje faktycznie testują: co do świadomości fenomenalnej czy też atrybucji stanów mentalnych jako takich. Jak wykazał zespół Heather M. Gray (2007), ludzie w różnym stopniu dokonują atrybucji stanów mentalnych różnym agentom. Niektórym rodzajom podmiotów przypisujemy raczej stany związane z doznawaniem, emocjami i osobowością, a innym ze sprawczością. Właśnie na tym rozróżnieniu zbudowana jest teoria atrybucji świadomości, która wyłoniła się z badań filozofii eksperymentalnej.

3. Nastawienie fenomenalne

Koncepcja Jacka i Robbinsa, zaproponowana w 2006 r. w artykule *Phenomenal Stance*, nawiązuje do podziału nastawień dokonanych przez Daniela Dennetta (1971; 1987). Przypomnę, że nastawienie intencjonalne to strategia opisu świata polegająca na przypisywaniu ludziom i innym wystarczająco złożonym obiektom przekonań i pragnień w celu wyjaśnienia i przewidywania ich zachowania. Postulowana przez Jacka i Robbinsa kategoria nastawienia fenomenalnego ma być sposobem widzenia rzeczywistości, oceny działań jej wyróżnionych elementów zwanych agentami. Niemniej nie jest to rozszerzenie idei amerykańskiego filozofa w ścisłym sensie. Propozycję Jacka i Robbinsa można zatem uznać co najwyżej za inspirowaną Dennettem.

Przedstawienie koncepcji warto zacząć od jej ogólnej charakterystyki, przechodząc następnie do poszczególnych tez stawianych w jej ramach oraz analizy eksperymentów, mających te tezy empirycznie koroborować.

Przymiowanie wobec X-a nastawienia fenomenalnego to rozumieć X-a jako „system fenomenalny”, czyli traktować X-a jako ośrodek doświadczenia fenomenalnego (Jack i Robbins, 2006, s. 69).

Traktowanie agenta jako systemu fenomenalnego oznacza przypisywanie mu stanów fenomenalnych w rodzaju emocji czy doznań percepcyjnych. Atrybucja w ramach nastawienia fenomenalnego jest ponadto obwarowana szeregiem obostrzeń:

- przypisać można tylko te spośród stanów świadomych, które się samemu doświadczyło,
- stany fenomenalne mają wartość emocjonalną,
- konsekwencją identyfikacji jako systemu fenomenalnego jest rozpatrywanie agenta na płaszczyźnie moralnej.

Ze względu na drugie i trzecie założenie nastawienie fenomenalne odgrywa istotną rolę w interakcjach społecznych. Jack i Robbins twierdzą, że pełni tę rolę w równie zasadniczy sposób, w jaki fizyka potoczna (*folk physics*), utożsamiana przez autorów z nastawieniem fizykalnym Dennetta, pomaga nam widzieć świat w kategoriach przyczynowo-skutkowych (Jack i Robbins, 2006, s. 61). Podobnie jak fizyka potoczna, nastawienie fenomenalne ma być względnie niezależne od zdolności do *mindreadingu* (Baron-Cohen, 1995) – zdolności do przypisywania stanów mentalnych innym ludziom i wnioskowania na ich temat, często utożsamianej z nastawieniem intencjonalnym na gruncie psychologii³. Niezależność nie

³ Jako że nie znajduje się w centrum zainteresowania niniejszej pracy, ewentualne różnice semantyczne pomiędzy pojęciami, takimi jak: mentalizacja, *mindreading*, teoria umysłu, zostają pominięte. Wszystkie są traktowane synonimicznie, chyba że zostanie zaznaczone inaczej.

oznacza przy tym braku kompatybilności, co zobaczymy przy okazji powoływania się przez autorów na badania neurokorelatów tych trzech nastawień.

Tak silnie zaakcentowana rola nastawienia fenomenalnego może dziwić, jeśli przywykło się do myślenia o atrybucji świadomości w kategoriach rozumienia czyichś działań. Zdaniem Jacka i Robbinsa rola procesu przypisywania stanów fenomenalnych jest jednak zupełnie inna. Nastawienie fenomenalne służy poznaniu moralnemu, a zatem nie rozumieniu, lecz ocenianiu działań obserwowanych agentów.

Empirycznym uzasadnieniem tak pojmowanej roli nastawienia fenomenalnego, a co za tym idzie – rozpoznawania świadomości fenomenalnej, mają być rezultaty badań zaburzeniami poznawczymi (Jack i Robbins, 2006, s. 64-69). Autorzy wskazują, że psychopatia wiąże się z niezdolnością do współodczuwania. Mają tu na myśli prostą formę empatii, polegającą na automatycznej reakcji emocjonalnej na obserwowany stan drugiej osoby (Frith, 2003). Jest ona niezależna od kompetencji do przypisywania stanów mentalnych, w przeciwieństwie do intencjonalnej empatii. Ta ostatnia polega na rozumieniu, dlaczego druga osoba jest w takim, a nie innym stanie, i na odpowiedniej reakcji. Niektóre badania pokazują, że psychopaci radzą sobie lepiej od grupy kontrolnej w zadaniach na mentalizację (Blair, 1996). Zatem osoby o skłonnościach psychopatycznych, z zachowaną zdolnością do myślenia w kategoriach intencjonalnych, nie wykazują stosownych do sytuacji automatycznych reakcji emocjonalnych. Odwrotnie dzieje się w przypadku autyzmu, gdzie występuje deficyt mentalizacji, przy jednoczesnym zachowaniu reaktywności emocjonalnej, mierzonej stopniem pobudzenia fizjologicznego na dystres drugiej osoby (Blair, 1999). Rezultatem deficytów w psychopatii może być także brak wrażliwości moralnej. Osoby diagnozowane jako wysoce psychopatyczne nie odróżniają transgresji moralnych od konwencjonalnych, w przeciwieństwie do osób cierpiących na autyzm (Blair, 1995; 1996)⁴. Argumentując w powyższy sposób, Jack i Robbins wskazują zatem na związek między sferą emocjonalną człowieka a poznaniem moralnym.

Również u Dennetta pojawia się wzmianka na temat nastawienia, mającego służyć ocenie moralnej (Dennett, 1981, s. 240). Nastawienie personalne, jak je nazywa, stałoby ponad nastawieniem intencjonalnym. Wyróżniałoby zatem jeszcze bardziej złożone systemy niż intencjonalne. By przyjąć perspektywę moralną względem agenta, należy go postrzegać jako osobę, a to z kolei oznacza przypisanie zobowiązań moralnych temu, co już zostało rozpoznane jako intencjonalne. Propozycja Jacka i Robbinsa różni się od intuicji Dennetta w kilku aspektach:

⁴ Transgresja moralna (złamanie normy moralnej) jest w przypadku osób zdrowych oceniana surowiej niż transgresja konwencjonalna. Nawet gdy nie istnieje reguła zakazująca, to przekroczenie normy moralnej jest traktowane jako niedopuszczalne.

- pojęcie nastawienia fenomenalnego jest bezpośrednio związane z pojęciem świadomości (a dokładnie: świadomości fenomenalnej). W przypadku nastawienia personalnego tak nie jest. Nie znajdziemy w uwagach amerykańskiego filozofa niczego, co pozwalałoby przypuszczać, że w nastawieniu personalnym przypisuje się stany fenomenalne;
- nastawienie personalne jest, w przeciwieństwie do nastawienia fenomenalnego, bezpośrednio związane z pojęciem intencjonalności. Dennett argumentuje za takim rozwiązaniem, stwierdzając, że nie jest w stanie wyobrazić sobie zobowiązań moralnych w sytuacji, gdy brakuje intencjonalności. Przywołuje tu postępowanie osoby nadmiernie oszczędnej. Pieniądz nie jest systemem intencjonalnym, nie można więc zachowania skąpca oceniać w kategoriach zobowiązania moralnego;
- cała konstrukcja nastawień Dennetta ma charakter instrumentalny. Przyjęcie takiej czy innej postawy jest wyznaczone pragmatycznością. Nastawienie personalne musiałoby być podobnym środkiem do celu jak nastawienie intencjonalne. Nastawienie fenomenalne kieruje się jednak inną logiką – nie celem predykcji dalszych zachowań, lecz oceną moralną.

Przeprowadzone przez zespół Jacka badania nad neurokorelatami nastawień wiążą nastawienie fenomenalne z większą aktywacją sieci domyślnej w mózgu, zaś nastawienie fizyczne – z siecią aktywowaną zadaniem (*task positive network*; Jack i in., 2013). Gdy obserwujemy większą aktywność jednej sieci, zwykle idzie za tym spadek aktywności w drugiej (i na odwrót), natomiast w sytuacji korzystania z nastawienia intencjonalnego jest inaczej. Wydaje się, że nastawienie intencjonalne korzysta z rejonów należących do obu sieci. Dlatego też autorzy uznają myślenie w kategoriach przyczynowo-skutkowych za przeciwne do nastawienia fenomenalnego, a nastawienie intencjonalne za niezależne, ale kompatybilne (Jack, 2014, s. 130).

Autorzy przeprowadzili wiele badań testujących ideę atrybucji stanów fenomenalnych dla celów oceny moralnej. W artykule z 2012 r. przedstawili wyniki czterech testów przeprowadzonych metodą kwestionariuszową (Jack i Robbins, 2012). Kontrolowanymi zmiennymi były – zgodnie z drugim założeniem ich koncepcji oraz wynikami zespołu Gray z 2007 r. – poziom agencyjności (sprawczości) oraz doznawania (*experience*). W przypadku dwóch pierwszych eksperymentów badani zapoznawali się z historią o połowie homarów. Rybacy zastawiają pułapki na dnie basenu wodnego i co jakiś czas sprawdzają, czy nie wpadły do nich skorupiaki. Zdarza się, że homary umierają z głodu, zanim zostaną wyłowione. Osoby biorące udział w eksperymencie musiały odpowiedzieć, jak bardzo przejmują się dobrostanem skorupiaków, jak czuliby się, wiedząc, że przez zbyt rzadkie sprawdzanie pułapek część homarów ginie, a także jak duża kara powinna być prawnie przewidziana za zbyt rzadkie sprawdzanie pułapek. Manipulacji

podlegała, dostarczana wraz z historią o polowie, informacja „naukowa” na temat homarów. W zależności od warunku skorupiaki były przedstawiane bądź jako pozbawione emocji, inteligentne, planujące i rozumujące podobnie do ludzi, bądź jako posiadające bogate życie emocjonalne, zdolne do popadania w depresję, ale niespecjalnie inteligentne. Po przedstawieniu tej dodatkowej informacji badani byli proszeni o ponowną odpowiedź na pytania. W drugim eksperymencie badacze zamiast dwóch użyli czterech warunków, manipulując dostarczaną badanym informacją. Przypisywano tym samym homarom albo wysoki poziom agencyjności i niski doznawania, niski poziom agencyjności i wysoki doznawania, wysoki poziom doznawania i wysoki agencyjności, albo niski poziom doznawania i niski agencyjności. Wyniki Jacka i Robbinsa wskazują na istotną statystycznie różnicę pomiędzy przedstawianymi sytuacjami. W przypadku narracji o inteligentnych homarach badani nie byli skłonni myśleć o nich w kategoriach moralnych. Z kolei narracja o homarach przeżywających stany emocjonalne tym bardziej wyczulała badanych na moralne zachowanie względem skorupiaków, im więcej kompetencji w zakresie doświadczania czy doznawania świata im przypisywano.

Kolejne dwa eksperymenty dotyczyły nowo odkrytego podwodnego gatunku, którego przedstawiciela (zwanego ARTIC23) naukowcy zabrali do laboratorium, rekonstruując warunki w akwarium na wzór jego naturalnego habitatu. Najpierw manipulowano sposobem przedstawienia istoty: albo była ona w wieku dziecięcym i miała trudność z przystosowaniem się do życia w laboratorium, albo też była dorosła, a przystosowanie do nowych warunków przyszło jej z łatwością. W obu przypadkach istota ostatecznie się zadomowiła. Badani mieli stwierdzić, jak bardzo zgadzają się z wyrażeniami: ARTIC23 odczuwa ból, gdy się go zrani; ARTIC23 jest zdolny do autorefleksji, może odczuwać radość, jest inteligentny, wyczuwa uczucia innych i po skończeniu badań niemoralne byłoby uśmiercenie go. Czwarte wyrażenie służyło rozpoznaniu poziomu atrybuowanej agencyjności, reszta zaś poziomu doznawania. Wyniki wskazują na silniejszą atrybucję doznań w przypadku warunku z przystosowującego się z trudnością ARTIC23 w wieku dziecięcym. W kolejnym eksperymencie modyfikacji uległy zarówno stwierdzenia dla osób badanych, jak i pierwszy z warunków. Tym razem ARTIC23 był niewielkich rozmiarów dziecięcym przedstawicielem swego gatunku, raniącym się w procesie przystosowywania do akwarium. Pytania brzmiały następująco: ARTIC23 doświadcza podstawowych emocji, takich jak ból i przyjemność, doświadcza świata wokół siebie, posiada wewnętrzne życie mentalne, intencje, własny umysł i wolną wolę, jest organizmem biologicznym, został stworzony przez człowieka. Ostatnie dwa stwierdzenia służyły eliminacji tych badanych, którzy bez należytej uwagi podchodzili do całego badania. Tym razem poziom agencyjności i doznawania zyskały równomierną reprezentację w odpowiedziach badanych – trzy pierwsze stwierdzenia odnosiły się do przypisywanych doznań, trzy

kolejne do przypisywanej agencyjności. Dodatkowo badani otrzymali na początku eksperymentu do wypełnienia trzy kwestionariusze: część Interpersonal Reactivity Index dotyczącą empatii (IRI-EC, Davis, 1983), afektywną część Self-Reported Psychopathy Scale IV (Paulhus, Neumann i Hare, 2009), a także Three-Item Loneliness Scale (Hughes, Waite, Hawkey i Cacioppo, 2004). Jack i Robbins odnotowali korelację między odpowiedziami na temat atrybucji doznawania a wynikami badanych w IRI-EC. Podobnie jak w poprzednim eksperymencie, w przypadku gdy ARTIC23 był opisywany jako niewielki i młody osobnik, atrybucja doznań była większa.

Rezultaty badawcze, zdaniem autorów pojęcia nastawienia fenomenalnego, koroborują tezę o uwikłaniu atrybucji stanów fenomenalnych w moralną ocenę obserwowanych działań⁵. Fundamentalne znaczenie ma w tym względzie związek świadomości ze sferą emocjonalną człowieka (założenie drugie). Najprostszą drogą falsyfikacji projektu Jacka i Robbinsa byłoby zatem podanie przykładu stanu świadomego, w sensie świadomości fenomenalnej, niemającego komponentu emocjonalnego. Nastawienie fenomenalne jest w tej koncepcji rozumiane jako powszechnie wykorzystywany proces poznawczy; autorzy stawiają je obok nastawienia intencjonalnego.

4. Uwagi krytyczne wobec koncepcji nastawienia fenomenalnego

Propozycja Jacka i Robbinsa jest z pewnością kusząca. Oto za pomocą jednego mechanizmu wyjaśnia się, przynajmniej częściowo, wiele zjawisk. Wskazuje się tym samym, że świadomość fenomenalna nie jest pozbawiona funkcji. Nie chodzi o interpretację czyjegoś zachowania celem predykcji i dostosowania własnych działań, jak w przypadku nastawienia intencjonalnego, ale o ocenę, wartościowanie obserwowanego behawioru.

Wobec propozycji Jacka i Robbinsa można przedstawić wiele zarzutów. Część z nich odnosi się do sposobu przeprowadzenia eksperymentów, inne do kwestii natury teoretycznej. Zanim je rozpatrzę, przywołam kryteria, które moim zdaniem koncepcja atrybucji stanów świadomych powinna spełniać.

⁵ Nie zostały omówione wszystkie badania kwestionariuszowe (Jack i in., 2014). W obecnej wersji koncepcji autorzy bardziej skupiają się na zależnościach między emocjami a poznaniem moralnym, a mniej na faktycznym rozpoznawaniu świadomości fenomenalnej.

5. Kryteria

Taka koncepcja powinna przede wszystkim rozstrzygać, jak powszechne jest przypisywanie świadomych stanów mentalnych komuś drugiemu. Czy jest to umiejętność ekspercka, którą możemy nabyć w toku specjalnego treningu, czy też kompetencja powszechna. Jeżeli jest to kompetencja powszechna, to w jaki sposób sytuuje się względem innych, ogólniejszych teorii atrybucji stanów mentalnych, takich jak teoria umysłu (Frith i Frith, 2005) czy nastawienie intencjonalne (Dennett, 1971).

W koncepcji powinien też znaleźć się zrozumiały dla czytelnika model procesu przypisywania świadomości. Taka propozycja będzie zatem odpowiadać na pytania: Czy atrybucja opiera się na przesłankach behawioralnych? Czy jest jedno- czy wieloetapowa? Czy możemy wyróżnić typowe sytuacje, w których korzystalibyśmy z takiej atrybucji? Jaki pożytek dla jednostki ma stosowanie tej atrybucji? Nieuchronne dla każdej takiej koncepcji jest także wybór i uzasadnienie sposobu rozumienia świadomości. Mówiąc o atrybucji świadomych stanów mentalnych, autorzy z pewnością mają na myśli konkretny sposób rozumienia zjawiska świadomości, np. jako świadomości fenomenalnej.

Autorzy powinni również dążyć do takiego jej formułowania, aby była możliwa konfirmacja lub falsyfikacja empiryczna ich propozycji teoretycznych. Oznacza to również dbałość o standardy metodologiczne.

5.1. Charakter atrybucji stanów świadomych

W przypadku koncepcji Jacka i Robbinsa mamy do czynienia z kompetencją powszechnie przez ludzi posiadaną. Mamy tu wyróżnione nastawienie, wydawałoby się, w rozumieniu Dennetta, a to implikuje powszechność kompetencji. Warto jednak odnotować, że mimo powszechnego charakteru nastawienia fenomenalnego Jack i Robbins dopuszczają możliwość szkolenia się w atrybucji stanów świadomych, ale nie precyzują, na czym takie szkolenie miałoby polegać. Traktują jednak nastawienie fenomenalne jako rodzaj rozumowania związanego z refleksją (Jack, 2014, s. 133). Nie sposób więc traktować koncepcji nastawienia fenomenalnego jako zwykłego rozszerzenia koncepcji Dennetta.

Zdaniem autorów nastawienie fenomenalne nie sprowadza się do teorii umysłu (Jack i Robbins, 2006, s. 61). Podczas gdy teoria ta ma na celu wyjaśnianie i predykcję zachowań, nastawienie fenomenalne ma na celu ocenę moralną czyjegoś zachowania.

Relacja między atrybucją dotyczącą wyjaśniania zachowania i atrybucją dotyczącą potrzeb przetwarzania moralnego nie jest zatem w świetle omawianej koncepcji tak prosta.

5.2. Model mechanizmu atrybucji

Nastawienie fenomenalne polega na przypisywaniu stanów fenomenalnych, co czyni świadomość fenomenalną jednym z głównych pojęć tej koncepcji. Czy jednak pojęcie to jest dobrą propozycją wyjaśnienia potrzeby powszechnej kompetencji do przypisywania stanów świadomych? Pojęcie świadomości fenomenalnej nie zostało przeniesione do dyskursu filozoficznego z języka naturalnego, ale jest efektem wielowiekowych przemian w myśli filozoficznej. Jego źródła można poszukiwać u Immanuela Kanta (1787/1957) i empiryków brytyjskich (Hume, 1739/2005). Należy jednak pamiętać, że były to konstrukcje filozoficzne znacznie szersze, systemy epistemologiczne, nierzadko z aspektem ontologicznym. Współczesne ujęcie jest efektem dwóch procesów w historii nauki: rozwoju myśli fenomenologicznej (Husserl, 1913/1967) oraz wycofywania się w drugiej połowie XX wieku klasycznego behawioryzmu w naukach poznawczych (Feigl, 1958; Campbell, 1970; Nagel, 1974)⁶. Jako wytwór działalności akademickiej, nie jest oczywiste, że to intuicji dotyczących właśnie tego pojęcia powinniśmy szukać na poziomie potocznym.

Jack i Robbins postulują silny związek stanów fenomenalnych z emocjami. Nawet jeśli pominiemy brak dookreślenia, co autorzy rozumieją pod pojęciem emocji, związek ten jest problematyczny. Standardowe ujęcie świadomości fenomenalnej nie ogranicza jej wyłącznie do stanów emocjonalnych, a stwierdza jedynie, że należą do niej takie stany, dla których możemy wyróżnić, „jak to jest w nich być” (Block, 1995). Autorzy nie zauważają nawet, że zawężają rozumienie tego pojęcia. Stwierdzają, że „dla ich posiadacza stany fenomenalne mają zwykłą wartość hedonistyczną, pozytywną (przyjemność, radość) albo negatywną (ból, smutek)” (Jack i Robbins, 2006, s. 70; tłum. własne).

Ta niezwykle mocna teza jest trudna do obrony. O ile podawane przez autorów przykłady: doznawania bólu, przyjemności czy złości stosunkowo łatwo powiązać ze sferą emocjonalną, to taka intuicja niekoniecznie pojawia się w przypadku percepcji wzrokowej. Nietrudno wyobrazić sobie sytuację, w której percypując zwykły przedmiot codziennego użytku i znajdując się w stanie „jak to jest widzieć taki przedmiot”, nie odczuwa się żadnej walencji emocjonalnej. Nie zawsze widzimy w zachowaniu innych, czy ich działanie sprawia im przyjemność, czy też przy-

⁶ Warto zaznaczyć, że pierwszym, który użył terminu *qualé* (najprostszych jednostek doznań) we współczesnym rozumieniu był Clarence I. Lewis (1929).

krość. Jeżeli nawet takie wartościowanie emocjonalne ma miejsce, to w większości przypadków będzie ono poniżej progu naszej wykrywalności. Co więcej, oznacza to, że nastawienie fenomenalne działa tylko wtedy, gdy nie dość, że dostrzegamy u kogoś dany stan fenomenalny, to widzimy również, że towarzyszy mu silne przeżycie natury emocjonalnej. Tylko ten drugi fakt sprawia, że czyjeś zachowanie jest przez nas rozpatrywane w nastawieniu fenomenalnym.

Być może zamiast mówić o nastawieniu fenomenalnym, należałoby mówić o nastawieniu emocjonalnym. Autorzy podkreślają różnicę między zdolnością do wydawania ocen moralnych i rozpoznawaniem emocji a kompetencjami mentalizacyjnymi. Trudno dostrzec powód, dla którego rozpoznawanie stanów fenomenalnych miałyby być pierwszym etapem dokonywania ocen moralnych. Także wyniki badań z nurtu filozofii eksperymentalnej nie pomagają udzielić jasnej odpowiedzi na pytanie o związek między świadomością fenomenalną a emocjonalną walencją doznań. Z jednej strony mamy proponentów (Sytsma i Machery, 2010), z drugiej oponentów (Buckwalter i Phelan, 2013). Podobnie jest w przypadku przywoływanych przez zespół Jacka rezultatów badań nad neurokorelatami (Jack i in., 2013; Jack, 2014). Choć ukazują one różnicę między wzorcem aktywacji dla myślenia zgodnie z nastawieniem fizykalnym a myśleniem w kategoriach społecznych, to nie wskazują na wiele więcej. Zadania na percepcję społeczną, aktywujące sieć domyslną, są tak różnorodne, że i one nie dostarczają przekonującego argumentu, dlaczego komponent poznawczy pomagający nam dokonywać ocen moralnych miałyby być określane właśnie nastawieniem **fenomenalnym**.

Autorzy nie precyzują również, na jakich przesłankach opiera się atrybucja w nastawieniu fenomenalnym. Rozpoznawane są stany fenomenalne bądź też – zgodnie ze zgłoszonymi wyżej wątpliwościami – nie wszystkie stany fenomenalne, a tylko takie, które mają wartość emocjonalną. Jednak niejasne jest, co pomaga nam ocenić, czy dana osoba spełnia warunki bycia rozpoznaną w nastawieniu fenomenalnym. Czy jest to dostrzegalny aspekt jej zachowania, czy może posiadana wiedza na temat tej osoby?

5.3. Ewidencja empiryczna

Autorzy nie stosowali jednakowej metody werbowania osób badanych (Jack i Robbins, 2012, s. 386-396). W przypadku pierwszego eksperymentu wszyscy uczestnicy byli studentami pierwszego stopnia, biorącymi udział w zajęciach wprowadzających z psychologii. Osoby badane w drugim eksperymencie były rekrutowane na kampusie uniwersyteckim i przez Internet. Informacja o trzecim eksperymencie przekazywana była drogą ustną. W przypadku czwartego zastosowano selekcję na podstawie odsetku wypełnionych zadań na platformie

Amazon's Mechanical Turk, służącej do poszukiwania drobnej pracy, zaś wybrani otrzymywali rekompensatę finansową. Sam sposób werbowania badanych nie byłby może wart odnotowania, gdyby nie fakt wyróżnienia zupełnie odmiennych grup. Przeciętny wiek uczestnika czwartego eksperymentu to 32 lata, a trzeciego – 23. W przypadku pierwszego eksperymentu danych na temat wieku i płci nawet nie zebrano, można co najwyżej przypuszczać na podstawie uczestnictwa w przedmiocie akademickim, że były to osoby w wieku studenckim. Ponadto tylko w czwartym i pierwszym eksperymencie badani otrzymywali gratyfikację (pieniądze lub punkty z przedmiotu) za udział.

Wskutek takiej selekcji badanych osób pojawia się wątpliwość, czy rzeczywistość mierzono to, co było deklarowane. Fundamentem badań z zakresu filozofii eksperymentalnej jest testowanie intuicji potocznych, niezwiązanych z posiadaniem wiedzy naukowej na dany temat. W przypadku koncepcji nastawienia fenomenalnego badane były intuicje, zgodnie z kolejnymi pozycjami testowymi w użytych kwestionariuszach, odnoszące się do tego, czym jest świadomość, doznanie, intencjonalność, inteligencja, planowanie, czy chociażby to, kiedy mamy do czynienia z zachowaniem wartościowanym moralnie. Wszystkie te kategorie są związane z płaszczyzną psychologiczną człowieka, trudno więc zakładać, że grupa składająca się ze studentów przedmiotów psychologicznych nie będzie posiadać intuicji zmodyfikowanych przez proces edukacji. W przypadku pozostałych eksperymentów nie zbierano informacji na temat wykształcenia osób badanych, więc trudno ocenić, czy wyłoniona z populacji próba była wystarczająco reprezentatywna. Ze względu chociażby na średnią wieku można co najwyżej przypuszczać, że były to osoby uczące się.

Autorzy popełniają zatem błąd doboru próby (*selection bias*) (Rosenthal i Rosnow, 2009, s. 671-672). Błąd ten nie jest zjawiskiem nowym, studia na jego temat powstają już przynajmniej od kilkudziesięciu lat (Smart, 1966; Berk, 1983). Wskutek procesu selekcyjnego wyłoniona została grupa, której przebadanie nie sprawdza założeń postulowanej przez autorów teorii.

Popełniono również błąd uczestnika (*volunteer bias*), raz zachęcając do udziału w badaniu pieniędzmi, a raz ogłaszając jedynie, że poszukują osób do badania (Rosenthal, 1965). Poświęcone temu błędowi eksperymenty wskazują z kolei, że można spodziewać się różnych poziomów motywacji wykonywania zadania, w zależności od tego, w jaki sposób osoby są zachęcane do wzięcia udziału w badaniu (Rosenthal i Rosnow, 2009, s. 672-742).

Kwestia wyróżnionych grup badanych wiąże się też z metodą wypełniania kwestionariuszy. W połowie przypadków badani wypełniali je online. Problem pojawia się w czwartym eksperymencie, gdy eksperymentatorzy nie zadbali o to, by dostęp był możliwy z więcej niż jednej platformy (byłoby to zgodne z zaleceniami Reips, 2002). Nie wiadomo, jak specyficzna jest społeczność użytkowników

platformy Amazon's Mechanical Turk, ale ponieważ umożliwia ona znajdowanie prac drobnych, nie jest raczej odwiedzana przez przedstawicieli wszystkich warstw społecznych. Skutkiem ponownie jest prawdopodobne wyłonienie nieodpowiedniej grupy do testowania założeń postulowanej teorii.

Powyższe zarzuty nie oznaczają, że wyniki raportowane przez Jacka i Robbinsa są nieprawdziwe. Każą natomiast wątpić, czy mają jakąkolwiek siłę argumentacyjną na korzyść postulowanej przez nich teorii⁷.

Słowo końcowe

Uwagi krytyczne dotyczące omawianej koncepcji można streścić następująco: pisząc o atrybucji świadomości, musimy przyjąć pewne rozumienie zjawiska świadomości. Jeżeli staramy się pokazać, że zdolność do takiej atrybucji jest kompetencją powszechnie przez ludzi posiadaną, to prawdopodobnie pojęcie świadomości fenomenalnej nie jest dobrym kandydatem. Nie odnajduje się ono w koncepcji nastawienia fenomenalnego, zaś przywołane wyniki badań z zakresu filozofii eksperymentalnej wskazują, że jego status na gruncie psychologii potocznej jest niejasny.

Być może jest tak, że propozycja Jacka i Robbinsa jest lepszą teorią atrybucji stanów mentalnych na potrzeby dokonywania ocen moralnych aniżeli teorią atrybucji stanów świadomych. Za tą oceną nie powinno jednak iść porzucenie badań nad przypisywaniem świadomości innym ludziom. Plany eksperymentalne można poprawić, przedstawić inne podstawy teoretyczne, dobierając taką eksplikację pojęcia świadomości, które łatwiej będzie pogodzić z psychologią potoczną.

Innymi słowy, potrzeba większej liczby badań. Dobra wiadomość płynąca z zaprezentowanych wyników i koncepcji dla filozofów o skłonnościach kognitywistycznych jest taka, że zajmowanie się zjawiskiem świadomości nie musi oznaczać zamknięcia się na obszar empirycznej rewidowalności.

Literatura

Alfano M., Loeb D. (2016), *Experimental Moral Philosophy*, w: E.N. Zalta (red.), *The Stanford Encyclopedia of Philosophy*, Stanford: Metaphysics Research Lab, Stanford University.

⁷ W badaniach kwestionariuszowych (Jack i in., 2014) popełniono te same błędy. Nie mamy pewności, czy grupa osób badanych była reprezentatywna na tyle, by uznać, że faktycznie testowano intuicje potoczne.

- Baron-Cohen S. (1995), *Mindblindness: An Essay on Autism and Theory of Mind*, Cambridge, Mass.: The MIT Press.
- Berk R.A. (1983), An Introduction to Sample Selection Bias in Sociological Data, *American Sociological Review*, 48(3), 386-398, <https://doi.org/10.2307/2095230>.
- Blair R. (1995), A cognitive developmental approach to morality: Investigating the psychopath, *Cognition*, 57(1), 1-29.
- Blair R. (1996), Brief Report: Morality in the Autistic Child, *Journal of Autism and Developmental Disorders*, 26, 571-579.
- Blair R. (1999), Psychophysiological Responsiveness to the Distress of Others in Children in Autism? *Personality and Individual Differences*, 26, 477-485.
- Blanshard B., Skinner B.F. (1967), The problem of consciousness – a debate, *Philosophy and Phenomenological Research*, 27(3), 317-333.
- Block N. (1995), On the Confusion About a Function of Consciousness, *Behavioral and Brain Sciences*, 18, 227-247.
- Bruno M.-A., Vanhaudenhuyse A., Thibaut A., Moonen G., Laureys, S. (2011), From unresponsive wakefulness to minimally conscious PLUS and functional locked-in syndromes: recent advances in our understanding of disorders of consciousness, *Journal of Neurology*, 258(7), 1373-1384, <https://doi.org/10.1007/s00415-011-6114-x>.
- Buckwalter W., Phelan M. (2013), Function and feeling machines: A defense of the philosophical conception of subjective experience, *Philosophical Studies*, 166(2), 349-361, <https://doi.org/10.1007/s11098-012-0039-9>.
- Buckwalter W., Phelan M. (2014), Phenomenal Consciousness Disembodied, w: J.M. Sytsma (red.), *Advances in Experimental Philosophy of Mind*, New York: Bloomsbury.
- Campbell J. (1970), *Body and Mind*, New York: Doubleday.
- Carruthers G., Schier E. (2017), Why are We Still Being Hornswoggled? Dissolving the Hard Problem of Consciousness, *Topoi*, 36(1), 67-79, <https://doi.org/10.1007/s11245-014-9266-3>.
- Chalmers D.J. (1995), Facing Up to the Problem of Consciousness, *Journal of Consciousness Studies*, 2(3), 200-219, <http://cogprints.org/316/1/consciousness.html>.
- Davis M.H. (1983), Measuring individual differences in empathy: Evidence for a multi-dimensional approach, *Journal of Personality and Social Psychology*, 44(1), 113-126.
- Dennett D. (1981), *Brainstorms. Philosophical Essays on Mind and Psychology*. *Brainstorms: philosophical essays on mind and psychology*, Hanover: Bradford Books, <https://doi.org/10.1017/CBO9781107415324.004>.
- Dennett D.C. (1987), *Intentional Stance*, Cambridge, Mass.: The MIT Press, <https://doi.org/10.1017/S0140525X00058611>.
- Dennett D.C. (1971), Intentional Systems, *The Journal of Philosophy*, 68(4), 87-106.
- Feigl H. (1958), The “Mental” and the “Physical”, w: H. Feigl, M. Scriven, G. Maxwell (red.), *Concepts, Theories and the Mind-Body Problem*, Minneapolis: University of Minnesota Press.
- Fiala B., Arico A., Nichols S. (2012), On the Psychological Origins of Dualism: Dual-Process Cognition and the Explanatory Gap, w: E. Slingerland, M. Collard (red.), *Creating Consilience: Integrating the Sciences and the Humanities*, New York: Oxford University Press, <https://doi.org/10.1093/acprof:oso/9780199794393.003.0005>.

- Fischer E., Collins J. (red.) (2015), *Experimental philosophy, rationalism, and naturalism: rethinking philosophical method*, London: Routledge.
- Frege G. (1977), *Pisma semantyczne*, Warszawa: PWN.
- Frith C., Frith U. (2005), Theory of mind, *Current Biology*, 15(17), R644-R645, <https://doi.org/10.1016/j.cub.2005.08.041>.
- Frith U. (2003), *Autism: Explaining the Enigma*, Oxford: Blackwell.
- Gettier E.L. (1963), Is Justified True Belief Knowledge? *Analysis*, 23(6), 121-123, <https://doi.org/10.2307/3326922>.
- Giacino J. T., Kalmar K., Whyte J. (2004), The JFK Coma Recovery Scale-Revised: Measurement Characteristics and Diagnostic Utility, *Archives of Physical Medicine and Rehabilitation*, 85(12), 2020-2029, <https://doi.org/10.1016/j.apmr.2004.02.033>.
- Gray H.M., Gray K., Wegner D.M. (2007), Dimensions of Mind Perception, *Science*, 315, 619, <https://doi.org/10.1126/science.1134475>.
- Greene J.D. (2001), An fMRI Investigation of Emotional Engagement in Moral Judgment, *Science*, 293, 2105-2108, <https://doi.org/10.1126/science.1062872>.
- Haukioja J. (red.) (2015), *Advances in Experimental Philosophy of Language*, London: Bloomsbury.
- Hughes M., Waite L., Hawkey L., Cacioppo J. (2004), A short scale for measuring loneliness in large surveys: Results from two population-based studies, *Research on Aging*, 26(6), 655-672.
- Hume D. (2005), *Traktat o naturze ludzkiej*, Warszawa: Aletheia.
- Husserl E. (1967), *Idee czystej fenomenologii i fenomenologicznej filozofii*, Warszawa: PWN.
- Jack A.I. (2014), A Scientific Case for Conceptual Dualism: The Problem of Consciousness and the Opposing Domains Hypothesis, w: J. Knobe, T. Lombrozo, S. Nichols (red.), *Oxford Studies in Experimental Philosophy*, t. 1, Oxford: Oxford University Press, <https://doi.org/10.1093/acprof:oso/9780198718765.003.0008>.
- Jack A.I., Dawson A.J., Norr M.E. (2013), Seeing human: Distinct and overlapping neural signatures associated with two forms of dehumanization, *NeuroImage*, 79, 313-328, <https://doi.org/10.1016/j.neuroimage.2013.04.109>.
- Jack A.I., Robbins P. (2006), The Phenomenal Stance, *Philosophical Studies*, 127, 59-85, <https://doi.org/10.1007/s11098-005-1730-x>.
- Jack A.I., Robbins P. (2012), The Phenomenal Stance Revisited. *Review of Philosophy and Psychology*, 3(3), 383-403, <https://doi.org/10.1007/s13164-012-0104-5>.
- Jack A.I., Robbins P.A., Friedman J.P., Meyers C.D., Sytsma J. (2014), More than a feeling: counterintuitive effects of compassion on moral judgment, w: J. Sytsma (red.), *Advances in experimental philosophy of mind*, London: Bloomsbury.
- James W. (2002), *Psychologia. Kurs skrócony*, Warszawa: Wyd. Naukowe PWN.
- Kant I. (1957), *Krytyka czystego rozumu*, Warszawa: PWN.
- Kauppinen A. (2007), The Rise and Fall of Experimental Philosophy, *Philosophical Explorations*, 10(2), 95-118, <https://doi.org/10.1080/13869790701305871>.
- Klawiter A. (2005), Myśleć jak filozof, argumentować jak kognitywista, *Diametros*, 3, 176-181.
- Knobe J. (2003), Intentional action in folk psychology: An experimental investigation, *Philosophical Psychology*, 16(2), 309-324, <https://doi.org/10.1080/09515080320103373>.

- Knobe J., Prinz J. (2008), Intuitions about Consciousness: Experimental Studies, *Phenomenology and the Cognitive Sciences*, 7(1), 67-83.
- Lewis C.I. (1929), *Mind and the World Order*, New York: Dover.
- Machery E., Mallon R., Nichols S., Stich S.P. (2004), Semantics, Cross-Cultural Style, *Cognition*, 92(3), B1-B12, <https://doi.org/10.1016/j.cognition.2003.10.003>.
- Machery E., O'Neill E. (red.) (2014), *Current Controversies in Experimental Philosophy*. New York: Routledge, <https://doi.org/10.4324/9780203122884>.
- Mallon R., MacHery E., Nichols S., Stich S. (2009), Against arguments from reference, *Philosophy and Phenomenological Research*, 79(2), 332-356, <https://doi.org/10.1111/j.1933-1592.2009.00281.x>.
- Nadelhoffer T., Nahmias E. (2007), The Past and Future of Experimental Philosophy, *Philosophical Explorations*, 10(2), 123-149, <https://doi.org/10.1080/13869790701305921>.
- Nagel T. (1974), What Is It Like to Be a Bat? *The Philosophical Review*, 83(4), 435-450, <https://doi.org/10.2307/2183914>.
- Nichols S., Knobe J. (2007), Moral responsibility and determinism: The cognitive science of folk intuitions, *Nous*, 41(4), 663-685.
- Open Science Collaboration (2015), Estimating the reproducibility of psychological science, *Science*, 349(6251), aac4716.
- Paulhus D., Neumann C., Hare R. (2009), *Manual for the self-report psychopathy scale*, Toronto: Multi-Health Systems.
- Ramachandran V.S., Hirstein W. (1997), Three Laws of Qualia, *Journal of Consciousness Studies*, 4(5), 429-458.
- Reips U.-D. (2002), Standards for Internet-Based Experimenting, *Experimental Psychology*, 49(4), 243-256, <https://doi.org/10.1027//1618-3169.49.4.243>.
- Rosenthal R. (1965), The Volunteer Subject, *Human Relations*, 18, 389-406.
- Rosenthal R., Rosnow R.L. (2009), *Artifacts in Behavioral Research*, New York: Oxford University Press, <https://doi.org/10.1093/acprof:oso/9780195385540.001.0001>.
- Smart R.G. (1966), Subject selection bias in psychological research, *Canadian Psychologist/Psychologie Canadienne*, 7a(2), 115-121, <https://doi.org/10.1037/h0083096>.
- Sytsma J.M., Machery E. (2009), How to study folk intuitions about phenomenal consciousness, *Philosophical Psychology*, 22(1), 21-35, <https://doi.org/10.1080/09515080802703653>.
- Sytsma J., Machery E. (2010), Two conceptions of subjective experience, *Philosophical Studies*, 151(2), 299-327, <https://doi.org/10.1007/s11098-009-9439-x>.
- Tononi G. (2008), Consciousness as integrated information: a provisional manifesto, *The Biological Bulletin*, 215(3), 216-42, <http://www.ncbi.nlm.nih.gov/pubmed/19098144>.
- Weinberg J.M., Nichols S., Stich S. (2001), Normativity and Epistemic Intuitions, *Philosophical Topics*, 29(1/2), 429-460, <https://doi.org/10.5840/philtopics2001291/217>.
- Williamson T. (2007), *The Philosophy of Philosophy*, Oxford: Blackwell.
- Wundt W.M. (1921), Logik. Eine Untersuchung der Principien der Erkenntnis und der Methoden Wissenschaftlicher Forschung, t. 3, Stuttgart: Verlag von Ferdinand Enke.
- Yong E. (2012), Nobel laureate challenges psychologists to clean up their act: Social-priming research needs "daisy chain" of replication, *Nature News*, <https://doi.org/doi:10.1038/nature.2012.11535>.