

TOMASZ WYSOCKI

FILOZOFIA EKSPERYMENTALNA
JAKO METODOLOGIA FILOZOFICZNA

WROCŁAW 2011

Kopiowanie i rozpowszechnianie całości lub fragmentu książki jest dozwolone.

Książkę w postaci elektronicznej można pobrać ze strony:

www.xphi-europe.org

Copyright © by Tomasz Wysocki, Wrocław 2011

Korekta: Katarzyna Szubert

Skład i projekt okładki: Tomasz Wysocki

ISBN 978-83-932208-0-9

spis treści

wstęp.....	5
pochodzenie i założenia filozofii eksperymentalnej.....	9
filozofia analityczna	9
argumentacja z intuicji filozoficznej	20
filozofia eksperymentalna: cele i metody	30
trzy klasyczne eksperymenty filozoficzne	39
deskryptywna a kauzalna teoria referencji.....	39
badania nad pojęciem wiedzy.....	46
działanie intencjonalne.....	53
eksperymenty Wrocławskiego Koła Xφ.....	59
sprawiedliwość i zasada maksyminu u Rawlsa	59
eksperyment w filozofii religii.....	67
zarzuty do filozofii eksperymentalnej.....	89
filozofowie nie odwołują się do intuicji	90
zarzut z refleksji i zarzut z intuicji austyńskich	91
model dialogu i model refleksji.....	93
jeszcze jeden eksperyment.....	100
filozof jako ekspert.....	103
zakończenie.....	113
bibliografia	115

Dwie osoby klóćą się o dopuszczalność aborcji. W pewnej chwili jedna z nich pyta: *wyobraź sobie, że ktoś podłączył do ciebie chorego z niewydolnością układu trawienia. Chory wyzdrowieje tylko wtedy, jeżeli przez dziewięć miesięcy będzie podtrzymywany przy życiu przez twój organizm – przez ten czas zdasz wyzdrowieć. Czy masz prawo odłączyć się do chorego?* Druga osoba po chwili zastanowienia odpowiada, że oczywiście, ma takie prawo. Wtedy spostrzega, że skoro tak, to powinna także uznać prawo kobiety do przerywania ciąży, ponieważ te sytuacje nie różnią się niczym istotnym. Zmienia zdanie w kwestii aborcji – została przekonana.

Jak taka sytuacja jest możliwa? Co powoduje, że możemy się przekonywać bez odwoływania do doświadczenia zmysłowego lub do (samego) znaczenia spójników i kwantyfikatorów logicznych? W tej rozmowie jeden z rozmówców przekonał drugiego dlatego, że obaj podzielali tę samą intuicję filozoficzną. Co jednak, gdyby druga osoba uważała, że w opisanej sytuacji nie miałaby prawa odłączyć się od chorego? Czy można jej intuicję uznać za gorszą?

Intuicja filozoficzna jest jednym z najważniejszych narzędzi filozofa. Używał jej Sokrates w swoich rozmowach z Ateńczykami, obecnie odwołuje się do niej prawie cała filozofia analityczna. Jednak wysiłek filozofów poświęcony uprawomocnieniu tego narzędzia jest nieproporcjonalnie mały w stosunku do roli, jaką intuicja pełni we współczesnej filozofii. Pytania o to, jak funkcjonuje intuicja filozoficzna i czy intuicjom można ufać (a jeżeli tak, to którym), legły u podstaw nowego ruchu filozoficznego, filozofii eksperymentalnej.

Filozofia eksperymentalna jest wyjątkowo młodą dziedziną – pierwsze artykuły pochodzą z roku 2001.¹ Dziesięć lat to niewiele jak na ruch w naukach szczegółowych, a co dopiero w filozofii. Obecny dorobek filozofii eksperymentalnej to kilkadziesiąt artykułów w rozpoznawalnych czasopismach filozoficznych i psychologicznych oraz jedna książka.² W Polsce nie ukazał się jeszcze żaden artykuł, który prezentowałby ten kierunek.³ Dlatego właśnie podstawowym celem tej książki jest przedstawienie głównych założeń filozofii eksperymentalnej, jej metod badawczych i nurtów wewnątrz kierunku. Ten cel staram się zrealizować w pierwszym rozdziale. Sam opis założeń teoretycznych nie wystarcza oczywiście do opisania żadnego kierunku badawczego, dlatego w drugim rozdziale przytaczam trzy najważniejsze – jedno z najwcześniejszych i jednocześnie najbardziej dyskutowane – eksperymenty filozoficzne, opisując ich metodologię, wyniki i filozoficzne konsekwencje. Przeciwno filozofii eksperymentalnej jako programowi, to znaczy przeciwko jej podstawowym założeniom, tradycyjnie zorientowani filozofowie wysunęli kilka zarzutów. Najpoważniejsze z nich przytaczam w czwartym rozdziale, jednocześnie przedstawiając kontrargumenty, którymi filozofowie eksperymentalni starają się bronić przed tymi zarzutami.

W trzecim rozdziale przedstawiam dwa eksperymenty, które przeprowadziłem razem z pozostałymi członkami Koła Naukowego Filozofii Eksperymentalnej, działającego przy Uniwersytecie Wrocławskim, w szczególności z Katarzyną Szubert i Dominikiem Dziedzicem.⁴ Stąd też w tym rozdziale w niektórych miejscach używam liczby mnogiej – *zaprojektowaliśmy, przeprowadziliśmy, zbadaliśmy*. Pierwszy eksperyment dotyczy filozofii politycznej i korzysta z nowej w filozofii eksperymentalnej metody dyskusji. Drugi to eksperyment z filozofii religii. W rozdziale czwartym przedstawiam wyniki jeszcze jednego

¹ J. Weinberg, S. Nichols, S. Stich, *Normativity and Epistemic Intuitions* (pełny adres bibliograficzny znajduje się na końcu książki w bibliografii).

² J. Knobe, S. Nichols, *Experimental Philosophy*, Być może należałoby także uwzględnić książkę: K. Appiah, *Experiments in Ethics*.

³ W *Polish Journal of Philosophy* (vol. III, wiosna 2009, s. 131-134) ukazała się recenzja książki Knobe'a i Nicholasa, *Experimental Philosophy*, autorstwa A. Feltza. Był to artykuł anglojęzyczny i bardzo krótki, dlatego trudno mówić w tym wypadku o prezentacji kierunku.

⁴ Szczegóły nt. działalności koła: www.xphi-europe.org.

eksperymentu, stanowiące kontrargument dla niektórych z przedstawionych tam zarzutów.

Ta książka jest nieznacznie zmienioną pracą magisterską, obronioną na Uniwersytecie Wrocławskim w styczniu 2011 roku. Chciałbym w tym miejscu podziękować mojemu promotorowi, dr. hab. Markowi Magdziakowi, za bardzo cenne uwagi do tekstu, prof. Stephenowi Stichowi za wskazówki dotyczące rozdziałów pierwszego i ostatniego oraz dr. Łukaszowi Nyslerowi za opiekę nad Kolem Filozofii Eksperymentalnej.

W tym rozdziale przedstawiam podstawowe założenia filozofii eksperymentalnej, jej metody i trzy główne nurty badawcze: analizę eksperymentalną, deskrytywizm eksperymentalny i restrykcjonizm eksperymentalny. Najpierw opisuję rozpowszechnioną w filozofii analitycznej metodę eksperymentów myślowych i odwołania do intuicji, ponieważ filozofia eksperymentalna powstała właśnie jako reakcja na nieskrępowane użycie intuicji w argumentacji filozoficznej.

filozofia analityczna

Filozofia eksperymentalna powstała na gruncie filozofii analitycznej: filozofowie eksperymentalni przeważnie zajmują się problemami, którymi zajmują się filozofowie analityczni, a niektórzy z nich są sławni z powodu swoich dokonań na polu filozofii analitycznej. To, że różnice między tradycją analityczną a badaniami eksperymentalnymi są istotne, nie zmienia faktu, że filozofia eksperymentalna jest nurtem wewnątrz tej tradycji lub, jeżeli ktoś woli traktować ją jako osobny kierunek, jest w tej tradycji głęboko zakorzeniona. Dlatego, zanim przedstawię podstawowe założenia filozofii eksperymentalnej, krótko opiszę te cechy filozofii analitycznej, które są ważne z punktu widzenia eksperymentalistów – przedmiot, istotną rolę przypisaną językowi i metodę argumentacji w oparciu o intuicję.

W książce *Philosophy's Second Revolution* David Clarke wyróżnia trzy fazy w historii filozofii, wyznaczone przez różne przedmioty i metodologie.⁵ Faza kończyła się, kiedy właściwe jej problemy okazywały się nie

⁵ D. Clarke, *Philosophy's Second Revolution: Early and Recent Analytic Philosophy*, s. 10 i nast.

do rozwiązania przy pomocy używanych narzędzi. Część z nich filozofowie uznawali za pseudo-problemy i jako takie po prostu porzucali, a część przejmowały nauki szczegółowe, które w trakcie rewolucji odrywały się od filozofii. Filozofia zmieniała swój zbiór problemów – dokonywała się rewolucja.

Pierwsza faza – filozofia klasyczna – rozpoczęła się pytaniem fizyków greckich o tworywo świata, a swój najdojrzały wyraz znalazła w *Metafizyce* Arystotelesa, której koncepcje w średniowieczu przyswoiła i przetworzyła scholastyka. Praca filozofa polegała na dostarczaniu wyjaśnień na podstawie abstrakcyjnego, bardzo ogólnego modelu świata.⁶ Według Arystotelesa dąb rośnie, uaktualniając swoją potencję bycia potężnym drzewem, według Platona konie są podobne, ponieważ wszystkie uczestniczą w jednej idei, a dla Epikura ludzie są wolni dzięki temu, że atomy odchylają się od właściwej sobie trajektorii. W *Księdze I Metafizyki* przedmiot filozofii (jako filozofii pierwszej) został zdefiniowany bardzo wyraźnie:

Istnieje nauka, która bada Byt jako taki i przysługujące mu atrybuty istotne. Nie jest ona żadną z tzw. nauk szczegółowych, bo żadna z tych nauk nie bada ogólnie Bytu jako takiego, lecz wyodrębnia pewną część Bytu i bada jej własności, jak na przykład nauki matematyczne. [...] Przedmiotem naszych badań są zasady i przyczyny Bytu.⁷

Przedmiotem filozofii klasycznej był *byt jako byt* i rządzące nim zasady. Metodą – obserwacja zmysłowa i spekulacja.⁸ Szczególnie w późniejszych teoriach istniał podział filozofii na dyscypliny, przeważnie na fizykę, etykę i logikę, ale etyka i logika miały swoje oparcie i uzasadnienie w fizyce.⁹ Modele metafizyczne opisywały i wyjaśniały wiele zjawisk, ale zdecydowanie nie spełniały dzisiejszych kryteriów porządnej teorii, bo nie przewidywały prawie nic.

⁶ D. Clarke, *Philosophy's Second Revolution*, s. 13.

⁷ Arystoteles, *Metafizyka*, *Księga I*, 1003a.

⁸ D. Clarke, *Philosophy's Second Revolution*, s. 17.

⁹ Gdzie fizyka w ówczesnym znaczeniu to filozofia przyrody, a logika to teoria poznania i psychologia deskryptywna.

Ten sposób uprawiania filozofii przetrwał do siedemnastego wieku, kiedy, przede wszystkim za sprawą Galileusza, Tycho Brahe i Kartezjusza, dokonała się rewolucja. Użycie przez Galileusza i Kartezjusza modeli matematycznych oraz zastosowanie przez Tycho wyjątkowo zaawansowanej (jak na XVII wiek) techniki przy prowadzeniu obserwacji rozpoczęło proces wypierania spekulacji z fizyki i zmieniło sposób wyjaśniania zjawisk: zamiast ogólnych metafizycznych modeli, tłumaczących świat przy pomocy przyczyn celowych, powstawały mechaniczne modele opisujące interakcję matematycznych obiektów. Nowe modele sprawdziły się lepiej – i pod względem teoretycznym jako źródła wyjaśnień i predykcji, i pod względem praktycznym jako podstawa dla wynalazków – przez co filozofia utraciła legitymację do wyjaśniania zjawisk naturalnych za pomocą swoich metafizycznych pojęć substancji, natury konstytutywnej i przyczyn celowych.

Odlączając się od filozofii, fizyka, chemia i fizjologia zagarnęły jej dotychczasowy przedmiot – przyrodę. Tymczasem, niezależnie od tego, filozofia natrafiła na problemy wewnętrzne: dyskutowane kwestie, takie jak istnienie i natura uniwersaliów lub zasada jednostkownienia, nie znajdowały rozwiązania. Spekulacja nie dostarczała argumentów wystarczających do opowiedzenia się za którymkolwiek z proponowanych rozwiązań. Jednocześnie w naukach szczegółowych, pomimo ich bardzo młodego wieku, zdążyła zapanować o wiele większa zgoda niż ta, która była udziałem filozofii. Ten stan rzeczy zawdzięczały posiadaniu własnej metody, dającej wyniki dużo bardziej jednoznaczne niż spekulacja.

W tych okolicznościach Kartezjusz we Francji, a za nim Locke w Anglii zaproponowali nowy przedmiot filozofii – operacje umysłowe i ich treści, nazwane przez Locke'a ideami – i nową metodę, introspekcję. Dotychczasowe problemy filozofii zostały, przynajmniej częściowo, zarzucone, a nowa metoda miała pozwolić na osiągnięcie konsensusu podobnego do tego w naukach szczegółowych. Takie kwestie, jak zadane na gruncie filozofii Kartezjusza pytanie o zależność między nierozciąglym umysłem i rozciąglą ciałem, były spadkiem po scholastycznym aparacie pojęciowym i zostały porzucone przez najbardziej dojrzałych filozofów nowej fazy, Hume'a i Kanta. I tak *Krytyka czystego rozumu* ma opisywać, jak

umysł organizuje bodźce zmysłowe – fenomeny – przy pomocy apriorycznych form zmysłowości oraz kategorii; nie mówi natomiast niczego o naturze noumenów, przedmiotów niedostępnych bezpośredniemu poznaniu.

Oprócz opisu funkcjonowania umysłu systemy drugiej fazy dostarczały narzędzi do krytyki zastanych pojęć. Według Locke'a należało porzucić te idee złożone, którym nie odpowiadały żadne idee proste – takie idee jak anioł lub jednorożec nie powstały z idei prostych, ale wzięły się ze sztucznego połączenia idei człowieka z ideą skrzydeł lub idei konia z ideą rogu. Nieporuszony poruszyciel i materia pierwsza także nie oparły się krytyce – te arystotelesowskie pojęcia nie pochodzą z żadnych idei prostych. Na gruncie swojej psychologii poznania Hume przeprowadził krytykę tak podstawowych dla dotychczasowej metafizyki pojęć jak substancja, związek przyczynowo-skutkowy, a nawet pojęcie *ja*. Według niego źródłem tych pojęć jest jedynie nawyk obserwatora do grupowania impresji w pewien szczególny sposób. Podobnie Kant krytykował metafizyczne dowody na istnienie Boga, ponieważ wykraczają poza dziedzinę, w której człowiek ma prawo formułować sądy.

Tak jak w przypadku pierwszej rewolucji, z końcem filozofii opartej na introspekcji Clarke również wiąże powstanie nowej nauki szczegółowej – tym razem jest to psychologia. Przejęła ona od filozofii nie tylko jej przedmiot, zjawiska psychiczne, ale i metodę. Jednak w psychologii od początku introspekcja została połączona z metodą eksperymentalną: w pierwszym laboratorium psychologicznym, założonym przez Wundta na Uniwersytecie Lipskim, subiektywna rejestracja przeżyć psychicznych przebiegała w kontrolowanych warunkach, celem powtarzalnych badań było potwierdzenie lub obalenie wysuniętej hipotezy, a samo badanie przeprowadzano przy pomocy zaprojektowanej w tym celu aparatury.¹⁰

Jednocześnie filozofia oparta o introspekcję natrafiła na problemy metodologiczne. Używając terminu *czerwony* albo *krzesło* wcale nie muszę przywoływać w świadomości wrażenia czerwieni lub wyobrażenia krzesła. Dwoje ludzi, którzy skutecznie porozumiewają się między sobą przy

¹⁰ B. Hergenhahn, *An Introduction to the History of Psychology*, s. 230.

użyciu języka, mogą wiązać z pojęciami odmienne wrażenia lub nawet nie doświadczać żadnych wrażeń bezpośrednio z tymi pojęciami powiązanych (potwierdzają to przypadki, w których – całkiem sensownie – wypowiadamy jedno zdanie, a myślimy o czymś zupełnie innym). Dlatego, aby badać takie filozoficznie istotne terminy jak *materia*, *umysł*, *sprawiedliwość*, wcale nie trzeba, a właściwie nawet nie należy odwoływać się do wrażeń. Ta obserwacja zapewniła filozofii nowy przedmiot, autonomiczny względem wrażeń: pojęcia. Powstała filozofia analityczna – produkt drugiej rewolucji filozoficznej, żeby jeszcze raz przywołać tytuł książki Clarke’a.

W istocie filozofia od swoich początków zajmowała się badaniem pojęć, jednak nigdy wcześniej filozofowie nie uznawali analizy pojęciowej za swoje podstawowe zadanie. Jedne z pierwszych, a jednocześnie najslawniejszych analiz pojęciowych w historii filozofii zawierają dialogi Platona. Inspiracją do snucia wizji doskonałego ustroju jest w *Państwie* analiza pojęcia sprawiedliwości, zaproponowanego przez Simonidesa.¹¹ W *Eutyfronie* Platon analizuje pojęcie dobra – czy coś jest dobre i dlatego podoba się bogom, czy też podoba się bogom i dlatego jest dobre. *Sofistę* i *Politykę* poświęca poszukiwaniu definicji tytułowych zawodów. Te dwa dialogi mają nawet własną metodę definiowania, diairezę. Sokrates dopytujący się swoich rozmówców *ale czym są sprawiedliwość, dobro, odwaga same przez się?* został patronem dwudziestowiecznej analizy pojęciowej.¹²

Arystoteles nie zarzucił badań nad pojęciami, uważnie definiując m.in. w *Księdze Δ Metafizyki* używane przez siebie terminy. W filozofii nowożytnej dokonana przez Hume’a krytyka pojęcia substancji lub związku przyczynowo-skutkowego nie byłaby możliwa bez analizy pojęcia substancji lub przyczyny, podobnie jak krytyka pojęcia materii z dialogów Berkeley’a bez analizy terminu *materia*. Jednak u Arystotelesa wyjaśnianie terminów jest nierozzerwalnie związane z jego fizyczno-metafizycznym modelem świata i nie ma w jego koncepcji wyraźnego podziału na język

¹¹ Platon, *Państwo*, 331D.

¹² Podręcznik z serii *Blackwell Companions To Philosophy*, poświęcony filozofii analitycznej, wręcz rozpoczyna się słowami: *Chociaż filozofia analityczna była praktykowana przez Platona [...].* A. Martinich, D. Sosa, *A Companion to Analytic Philosophy*.

i rzeczywistość.¹³ Podobnie u Hume'a, Locke'a lub Berkeleya analiza pojęć opiera się na rozstrzygnięciach przyjętej przez nich psychologii poznania. Pojęcia stały się samodzielnym przedmiotem filozofii dopiero w dwudziestowiecznej filozofii analitycznej.

Oczywiście schemat zaproponowany przez Clarke'a upraszcza historię. Przede wszystkim należy pamiętać, że rewolucje były udziałem niewielu, a nowy prąd przyjmował się powoli. Na przykład po pierwszej rewolucji scholastyka długo jeszcze była dominującym nurtem na wielu uniwersytetach, a w postaci neotomizmu dotrwała w murach niektórych katolickich uczelni nawet do czasów współczesnych.

Po drugie, w swojej opowieści Clarke zupełnie ignoruje współczesną filozofię kontynentalną. Celem nurtów zapoczątkowanych pracami Brentano – a więc przede wszystkim Husserlowskiej fenomenologii, ale także filozofii Kazimierza Twardowskiego i niektórych jego uczniów ze Szkoły Lwowsko-Warszawskiej – było opisanie struktury świadomości. Ich narzędzie stanowiła introspekcja, która przez fenomenologów została przekuta w analizę ejdetyczną. Zresztą fenomenologia bardzo dobrze wpisuje się w schemat Clarke'a. Podobnie jak filozofia analityczna, fenomenologia przyjęła nowy przedmiot – znaczenie – ale w przeciwieństwie do wczesnej filozofii analitycznej fenomenolodzy traktują znaczenie jako zjawisko z natury ejdetyczne, idealne, a nie językowe.

Po trzecie, być może opisując początki filozofii analitycznej należałoby raczej wskazać na brytyjską (Moore, Russell) i austriacką (Kolo Wiedeńskie, Brentano i jego uczniowie) reakcję przeciwko heglizmowi, kontynuację myśli empiriokrytycznej (Kolo), inspirację dokonaniem w logice (Frege, Russell) i naukach szczegółowych (członkowie Koła Wiedeńskiego mieli wykształcenie ściśle, przeważnie fizyczne). Następnie należałoby przedstawić dwa podejścia we wczesnej filozofii analitycznej – podejście Koła Wiedeńskiego, nazywane filozofią języka idealnego lub

¹³ Na przykład Arystoteles w tym samym fragmencie *Księgi* Δ definiuje substancję jednocześnie jako to, co istnieje samo przez się, i jako to, co nie może być orzekane o czymś innym. Arystoteles, *Metafizyka*, 1017b.

analizą logiczną, oraz wczesne nurty brytyjskiej filozofii języka potocznego.

Jednak do wyjaśnienia powstania filozofii eksperymentalnej wystarczy zdać sprawę z obecnego stanu filozofii analitycznej, a w zasadzie jedynie z pewnej rozpowszechnionej wśród analityków metody uzasadniania sądów filozoficznych. Pierwotne pomysły z początków tego ruchu – zarówno od strony filozofii języka idealnego, jak filozofii języka potocznego – zostały przez filozofów zasadniczo zarzucone, a w każdym razie bardzo istotnie przekształcone. Chciałbym natomiast uważnie przyjrzeć się stwierdzeniu, że istotnym (bo nie jedynym) przedmiotem współczesnej filozofii analitycznej są pojęcia.

W rozdziale *Analytical Philosophy: Two Analogies* z książki *Analysis and Metaphysics* P.F. Strawson jednoznacznie utożsamia filozofię analityczną z analizą pojęciową:

Jakie sobie stawia zadania [filozof analityczny]? Czym się zajmuje? Cóż, na pewno zajmuje się ideami lub pojęciami. Toteż samozwańczy tytuł *filozofa analitycznego* sugeruje, że ulubionym określeniem jego ulubionego zajęcia będzie *analiza pojęciowa*.¹⁴

Analizę pojęciową Strawson porównuje do gramatyki. Mówiąc, buduje zdania zgodnie z zasadami polskiej składni – przypadek przymiotnika zgadza się z przypadkiem rzeczownika, z którym ten przymiotnik się łączy, liczba i rodzaj podmiotu zgadzają się z liczbą i rodzajem orzeczenia. W ten sposób postępuje większość osób, o których powiedzielibyśmy (lub które same by o sobie powiedziały), że mówią w języku polskim. Jednak przeważająca liczba tych osób miałaby problem ze sformulowaniem reguł, zgodnie z którymi postępują, gdy konstruują poprawne wypowiedzi. Zdolność gramatycznego wypowiedziania się to jedno, zdolność sformułowania reguł gramatycznych to zupełnie co innego – pierwsze wymaga obycia z językiem i jest w większym lub mniejszym stopniu

¹⁴ P. Strawson, *Analiza i metafizyka*, tłum. A. Grobler, s. 8.

udziałem każdego, drugie wymaga poświęcenia się studiom nad strukturą wypowiedzi.¹⁵

Podobnie jest z pojęciami. Ludzie stosują pojęcia, ale nie umieją podać ich znaczenia lub podają definicje nieprzystające do sposobu, w jaki tych pojęć używają. W tym wypadku rolę gramatyków spełniają filozofowie, przed którymi stoi zadanie opisanie regul rządzących użyciem pojęć. Zgodnie z przykładem Strawsona, wiemy, co znaczy *wiedzieć*, bo potrafimy stwierdzić, czy w danych warunkach osoba *O* wie, że *p*, a nie na przykład jedynie zgadła, że *p*. To jednak nie pociąga za sobą umiejętności wysłowienia znaczenia terminu *wiedzieć*.¹⁶

Michael Dummett przedstawia cel filozofii jako badanie struktury myśli, przy czym struktura myśli objawia się w języku. Dummett nie wspomina bezpośrednio o pojęciach, ale nie odbiega od analogii gramatyki i filozofii. Według niego metodą filozofii pozostaje badanie użycia języka, bo nawet jeżeli *język jest lustrem, które zniekształca obraz, to jest to jedyne lustro, którym dysponujemy*.¹⁷

Strawsonowi i Clarke'owi można zarzucić, że sympatyzują z filozofią języka potocznego, a tym samym ich wizja filozofii analitycznej może być jednostronna. W szczególności filozofowie z pewnością nie zarzucili ambicji poprzednich okresów i nadal formułują metafizyczne modele świata i człowieka. W filozofii (już nie tylko w filozofii analitycznej, ale filozofii w ogóle) można próbować wskazać zbiór pytań przewijających się – mimo rewolucji – przez całą jej historię: jakie są podstawowe składniki świata, jak funkcjonuje umysł, jak człowiek wydaje sądy etyczne, czy istnieje wiedza wrodzona, czy jesteśmy wolni... Próba odpowiedzi

¹⁵ Można uprawiać gramatykę deskryptywną, czyli badać, jak ludzie rzeczywiście mówią, lub gramatykę normatywną, czyli jak ludzie powinni mówić lub jak mówi się właściwie. Jeżeli ktoś używa sobie właściwych reguł gramatycznych, gramatyki deskryptywnej można najwyżej opisać jego idiolekt, ale nie może ocenić, czy ten człowiek mówi poprawnie, czy nie. To rozróżnienie ma przełożenie na analizę pojęciową: można opisywać, jak ludzie używają pojęć, i można próbować formułować poprawne znaczenie pojęcia. To rozróżnienie dokładniej poruszam w ostatnim rozdziale.

¹⁶ P. Strawson, *Analysis and Metaphysics*, s. 7. Rawls posługuje się taką samą analogią, opisując zadanie skonstruowania teorii etycznej: J. Rawls, *A Theory of Justice, Revised Edition*, s. 40.

¹⁷ M. Dummett, *Origins of Analytical Philosophy*, s. 6. Tłum. własne. W kolejnych przypisach do cytatów niewymienienie tłumacza oznacza tłumaczenie własne.

na część z nich wymaga co najmniej pewnego składnika wiedzy empirycznej. Filozofia analityczna w swoich wiedeńskich początkach była zainteresowana naukami szczegółowymi, ale dotyczyło to wyłącznie ich języka. Obecnie części filozofów chodzi o coś zupełnie innego – o wykorzystanie wyników nauk szczegółowych w argumentacji filozoficznej. Tym zajmuje się, jak ją określa Jesse Prinz, filozofia empiryczna:

Filozofowie empiryczni to teoretycy. Nie zbierają danych, ale raczej konstruują argumenty i teorie, próbując sformułować implikacje wyników badań empirycznych. Filozofowie empiryczni raczej łączą wyniki z różnych dziedzin, niż komentują to, czego dokonano przy pomocy podobnych technik i paradygmatów.¹⁸

Metafizycy korzystają z rezultatów fizyki, filozofowie umysłu i etycy z wyników psychologii, neurobiologii i nauk kognitywnych, filozofowie polityczni z wyników ekonomii i socjologii i tak dalej. Kiedyś Arystoteles wyciągał metafizyczne wnioski na podstawie własnych obserwacji, których dokonywał z perspektywy lekarza – a więc z perspektywy nauki szczegółowej (medycyny), którą uprawiał. Teraz to filozofowie czekają, aż nauki szczegółowe same dostarczą materiału do interpretacji. Ale cel jest wciąż ten sam.

Mimo że w projektach określanych mianem filozofii empirycznej głównym zadaniem nie jest rozjaśnianie pojęć, analiza pojęciowa nadal odgrywa w nich istotną rolę. Jeżeli pytanie brzmi, czy jesteśmy wolni (a jest to pytanie jednocześnie filozoficzne i, przynajmniej częściowo, empiryczne), nie można na nie odpowiedzieć bez wyjaśnienia, czym jest wolność. Jeżeli pytanie brzmi, jaka jest natura czasu, interpretując wyniki fizyki trzeba jednocześnie ostrożnie podchodzić do sposobu, w jaki terminem *czas* posługują się fizycy.

Argumentację w podobnym duchu przedstawia Frank Jackson w książce *From Metaphysics to Ethics: a defense of conceptual analysis*. Jackson wychodzi od czegoś bardzo odległego od analizy pojęciowej – od metafizyki

¹⁸ J. Prinz, *Empirical Philosophy and Experimental Philosophy* [w:] J. Knobe, S. Nichols, *Experimental Philosophy*, s. 205.

i to od metafizyki pewnego konkretnego rodzaju: opartej na osiągnięciach fizyki.

Metafizyka zajmuje się tym, co jest, i jakie to jest. [...] Metafizycy starają się wyczerpująco opisać pewne przedmioty lub zjawiska – umysł, semantykę lub, ambitniej, wszystko – przy użyciu ograniczonej liczby mniej lub bardziej podstawowych terminów.¹⁹

Fizykalizm jest bardzo oszczędny. Stwierdza, że pełny opis naszego świata [...] może być w zasadzie stworzony przy użyciu względnie małej liczby własności i relacji fizycznych.²⁰

Metafizyka fizykalistyczna [*physicalism*] wymaga znajomości osiągnięć fizyki i neurobiologii i należy do filozofii empirycznej, nie mając przy tym nic wspólnego z analizą pojęć potocznych [*folk concepts*]. Kiedy taka metafizyka zostaje zbudowana — lista zawierająca kwarki, elektrony, siły elektromagnetyczne i prawa rządzące nimi wszystkimi — metafizyk chce wiedzieć, co zrobić z takimi terminami jak myśl, ból, odpowiedzialność, tożsamość osobowa. Możliwości są dwie: takie terminy można albo wyeliminować, uznać za bezprzedmiotowe, albo stwierdzić, że oznaczają pewne rodzaje przedmiotów zbudowanych z elementów sporządzonej uprzednio listy. To zadanie wcale nie jest proste, bo na liście nie ma żadnych potocznych terminów opisujących stany umysłu, żadnej potocznej psychologii [*folk psychology*].

Tak więc metafizyk dysponuje z jednej strony słownikiem terminów współczesnej fizyki, a z drugiej słownikiem języka potocznego. Pytanie brzmi, które terminy z drugiego słownika dają się zdefiniować w oparciu o pierwszy, a tym samym zostaną zachowane, a które nie mają tłumaczenia, a tym samym zostaną wyeliminowane. Odpowiedzi może udzielić tylko analiza pojęciowa.²¹

Jackson wskazuje na dwie role, które może pełnić analiza pojęciowa: skromną [*modest*] i nieskromną [*immodest*]. Opisanie świata w czterowymiarowej przestrzeni (trzy wymiary przestrzenne plus czas)

¹⁹ F. Jackson, *From Metaphysics to Ethics*, s. 3-4.

²⁰ F. Jackson, *From Metaphysics to Ethics*, s. 6.

²¹ F. Jackson, *From Metaphysics to Ethics*, s. 41.

powoduje, że zamiast mówić o zmianie, na przykład temperatury, należy mówić o różnych temperaturach w różnych miejscach na osi czasu. Świat rozpada się na nieskończoną liczbę zastygłych trójwymiarowych kostek indeksowanych czasem. Uprawiając analizę pojęciową skromnie, filozof powie, że ten obraz i potoczne rozumienie zmiany są niezgodne. Uprawiając ją nieskromnie powie, że ten obraz jest fałszywy, ponieważ zmiana w potocznym rozumieniu jest faktem, a nie da się jej ufundować w tym modelu. Skromna analiza pojęciowa dotyczy języka, nieskromna analiza pojęciowa ma większe ambicje i stara się powiedzieć, jaki świat jest. Jackson jest podejrzliwy w stosunku do tej drugiej roli i nie uważa, aby analiza pojęciowa mogła wnieść cokolwiek do wiedzy o rzeczywistości pozajęzykowej.²²

Podobny podział stosuje Stephen Stich, który wskazuje na trzy rodziny projektów filozoficznych wewnątrz współczesnej filozofii analitycznej: analizę pojęć [*conceptual analysis*], projekty neoplatonickie [*neo-platonic projects*] i projekty normatywne [*normative projects*].²³ Przedmiotem analizy pojęć są pojęcia jako pewne obiekty językowe lub stany psychiczne, przedmiotem projektów neoplatonickich są istotne pod względem filozoficznym zjawiska, celem projektów normatywnych jest ustalenie reguł, jakimi powinny kierować się jednostki w swoim postępowaniu. Jeżeli projekt z pierwszej grupy dotyczy pojęcia wiedzy używanego przez pewną wspólnotę językową, ambicją projektu z drugiej grupy jest opisanie wiedzy niezależnie od czyjegokolwiek pojęcia wiedzy, a celem projektu z trzeciej grupy jest podanie reguł, do których powinno stosować się poznanie, aby jego rezultatem była wiedza. Jeżeli projekt z pierwszej grupy dotyczy pojęcia sprawiedliwości, celem projektu z drugiej grupy jest opisanie sprawiedliwości jako takiej, a projekt z trzeciej grupy polega na wskazaniu zasad, którym powinno czynić zadość społeczeństwo, aby było sprawiedliwe. Trzecia rodzina projektów silnie zależy od drugiej – opisanie istotnego filozoficznie zjawiska implikuje pewne normatywne rozwiązania, tak samo jak trudno jest poszukiwać normatywnych rozwiązań, jeżeli zjawisko nie zostało wystarczająco dobrze opisane. Natomiast pierwsza

²² F. Jackson, *From Metaphysics to Ethics*, s. 43.

²³ S. Stich, wystąpienie na konferencji *How and why economists and philosophers do experiments: dialogue between experimentaleconomics and experimental philosophy*, Kioto, 27 marca 2010.

rodzina projektów nie ma ambicji do opisywania zjawisk innych niż psychiczne lub językowe, a tym samym jest niezależna od dwóch pozostałych. Mimo tych różnic, we wszystkich projektach filozofowie często używają tych samych narzędzi, tego samego rodzaju argumentacji, właściwego analizie pojęciowej. Dlatego pierwszą rodzinę można utożsamić z analizą pojęciową w skromnej roli, a drugą i trzecią, o ile stosuje się w nich te same narzędzia, co w pierwszej, z nieskromną analizą pojęciową.

Niektórzy autorzy w ogóle nie zgodziliby się na nazywanie projektów neoplatońskich analizą pojęciową, ponieważ są one po prostu teoriami na temat obiektywnego świata, a nie na temat pojęć, które służą opisywaniu tego świata.²⁴ Niektóre z projektów neoplatońskich rzeczywiście nie mają wiele wspólnego z analizą pojęciową – na przykład kiedy projekt polega na sformułowaniu fizykalistycznej metafizyki w oparciu o wyniki fizyki (aż do momentu, kiedy filozof chce w tej metafizyce zidentyfikować takie obiekty jak przekonania lub emocje). Inne jednak korzystają z tej samej metody, co analiza pojęciowa, niezależnie od tego, czy zajmujący się nimi filozofowie nazywają siebie analitykami pojęciowymi, czy nie.

argumentacja z intuicji filozoficznej

Filozofia klasyczna miała swój przedmiot – zjawiska naturalne – i metodę – spontaniczną obserwację i spekulację. Przedmiotem filozofii nowożytnej były zjawiska psychiczne, a metodą – introspekcja. Skoro pojęcia są istotnym przedmiotem filozofii analitycznej, jaka metoda służy ich badaniu?

Przedstawię trzy przykłady argumentacji filozoficznej, jeden z filozofii starożytnej, dwa ze współczesnej filozofii analitycznej. Te dwa współczesne przykłady odegrały podstawową rolę w filozofii analitycznej: pierwszy z nich spowodował odejście od deskryptywnej teorii znaczenia nazw własnych na rzecz teorii kauzalnej, w wyniku drugiego filozofowie porzucili definicję wiedzy jako prawdziwego uzasadnionego przekonania.

Pierwszy argument pochodzi od Platona:

²⁴ D. Papineau, *What is X-Phi Good For?*

– Bardzo pięknie mówisz, Kefalosisie – powiedziałem. – A co do tego właśnie, jeżeli mowa o sprawiedliwości, to czy powiemy, że sprawiedliwość to jest tyle co: prawdomówność – tak po prostu – i tyle co: oddawanie, jeżeli ktoś coś od kogoś weźmie, czy też i to nawet raz będzie działaniem sprawiedliwym, a innym razem niesprawiedliwym? Ja na przykład coś takiego mam na myśli: przecież każdy przyzna, że gdyby ktoś wziął od przyjaciela, od człowieka przy zdrowych zmysłach, broń, a ten by potem oszalał i zażądał jej z powrotem, to nie trzeba jej oddawać, i nie byłby sprawiedliwy ten, który by ją oddał; ani ten, który by człowiekowi zostającemu w takim stanie chciał mówić wszystko zgodnie z prawdą.

– Słusznie mówisz – powiada.²⁵

Drugi pochodzi od Kripkego:

Wyobraźmy sobie następującą, jaskrawo fikcyjną sytuację. (Mam nadzieję, że nie ma tu profesora Gödla.) Założmy, że Gödel nie był faktycznie autorem owego twierdzenia. Faktycznie dokonał tego dzieła człowiek o nazwisku *Schmidt*, którego ciało znaleziono w tajemniczych okolicznościach przed wieloma laty w Wiedniu. Jego przyjaciel Gödel dostał jakoś ów rękopis, który później przypisano Gödlowi. A więc na gruncie poglądu, który rozpatrujemy, gdy zwykły człowiek [*ordinary man*] posługuje się nazwą *Gödel*, w zamyśle odnosi się do Schmidta [*really means to refer*], ten bowiem jest jedyną osobą spełniającą opis: *człowiek, który odkrył niezupełność arytmetyki*. (...) A więc ponieważ człowiekiem, który odkrył niezupełność arytmetyki, jest faktycznie Schmidt, gdy mówimy o *Gödlu*, faktycznie zawsze odnosimy się do Schmidta. Wydaje mi się jednak, że się nie odnosimy. Po prostu nie odnosimy się.²⁶

Autorem trzeciego przykładu jest Gettier:

Przypuśćmy, że Smith i Jones zaczęli się starać o pewną posadę. Ponadto przyjmijmy, że dla Smitha jest zupełnie oczywista następująca koniunkcja:

²⁵ Platon, *Państwo*, 331C. Tłum. W. Stróżewski. Podkreślenie moje.

²⁶ S. Kripke, *Nazywanie a konieczność*, s.117. Tłum. B. Chwedeńczuk. Podkreślenie moje.

(d) *Jones jest tym, który dostanie posadę i Jones ma dziesięć monet w swej kieszeni.*

Dla Smitha może być oczywiste (d), na przykład dlatego, że prezes spółki zapewnił go, że w końcu zostanie wybrany Jones; jest też dla niego oczywiste, że on, Smith, policzył monety będące w kieszeni Jonesa przed dziesięcioma minutami. Zdanie (d) pociąga zdanie:

(e) *Człowiek, który otrzyma posadę, ma dziesięć monet w kieszeni.*

Przypuśćmy teraz, że Smith dostrzeżę wynikanie (e) z (d) i uznaje (e) na podstawie (d), które jest dla niego zupełnie oczywiste. W tym przypadku przekonanie Smitha, że (e) jest prawdziwe, istotnie jest uzasadnione. Lecz wyobraźmy sobie dalej, że to właśnie Smith (a nie Jones) dostanie posadę, choć nie będzie o tym wiedział, oraz że to on będzie miał dziesięć monet w kieszeni i tego również nie będzie wiedział. Wówczas zdanie (e) jest prawdziwe, chociaż zdanie (d), z którego Smith wyprowadził (e), jest fałszywe. W naszym przykładzie wobec tego, to wszystko jest prawdziwe: (i) (e) jest prawdziwe, (ii) Smith jest przekonany, że (e) jest prawdziwe oraz (iii) przekonanie Smitha, że (e), jest uzasadnione. Jednakże jasne jest też, że Smith nie wie, że (e) jest prawdziwe; (e) jest bowiem prawdziwe z racji liczby monet w kieszeni Smitha, gdy tymczasem Smith nie wie ile monet jest w kieszeni Smitha, i opiera swoje przekonanie, że (e), na obliczeniu monet w kieszeni Jonesa, co do którego jest fałszywie przekonany, że jest tym człowiekiem, który otrzyma posadę.²⁷

Przykłady mają wspólny schemat. W dialogu Platona Sokrates sprawdza trafność definicji sprawiedliwości: *mówić prawdę i oddać każdemu, co się należy*.²⁸ Przedstawia swojemu rozmówcy hipotetyczny przypadek osoby, która pożyczyła przyjacielowi broń, a następnie oszalała – jednym słowem, Sokrates przeprowadza pewien eksperyment myślowy. Następnie pyta, czy sprawiedliwe byłoby zwrócenie broni jej właścicielowi, i stwierdza, że *przecież każdy przyzna, iż nie byłoby to sprawiedliwe*. Kefalos

²⁷ E. Gettier, *Czy uzasadnione i prawdziwe przekonanie jest wiedzą?*, tłum. J. Hartman, J. Rabus. Podkreślenie moje.

²⁸ Platon, *Państwo*, 331D.

się z nim zgadza – definicja Simonidesa zawiodła, bo nie ujmuje koniecznych i wystarczających warunków, w których zachowanie jest sprawiedliwe. Sokrates wskazał na sytuację, w której zwrócenie komuś jego własności nie jest sprawiedliwe.

Historia opisana przez Kripkego służy mu jako argument przeciwko deskryptywnej teorii znaczenia, a za teorią kauzalną. W swoim eksperymencie myślowym Kripke wyobraża sobie, że osoba nazywana Kurtem Gödlem wcale nie sformułowała twierdzenia o niezupełności. Następnie zastanawia się, czy wypowiadając nazwisko Gödel ktoś, kto kojarzy z tym nazwiskiem jedynie autorstwo twierdzenia o niezupełności, odnosi się do Kurta Gödla, czy do Schmidta, prawdziwego autora twierdzenia. To Schmidt spełnia konieczne i wystarczające warunki bycia referentem nazwy *Gödel*, sformułowane w teorii deskryptywnej. Ale według Kripkego dla zwykłego człowieka referentem jest nadal Kurt Gödel, niezależnie od dziejów twierdzenia o niezupełności. Teoria deskryptywna zawodzi, a stosunek nazwy do referenta lepiej opisuje teoria kauzalna.

Przykład Gettier'a służy wykazaniu, że wbrew obowiązującej przez długi czas definicji nie wystarczy, aby przekonanie było uzasadnione i prawdziwe, żeby było wiedzą.²⁹ W tym eksperymencie myślowym starający się o pewne stanowisko ma prawdziwe i uzasadnione przekonanie, że osoba z dziesięcioma monetami w kieszeni zostanie zatrudniona – czyli przekonanie to spełnia warunki z definicji – a mimo to nie wie, że właściciel dziesięciu monet otrzyma to stanowisko.

Schemat tych argumentów jest następujący:

- (1) eksperyment myślowy, opisujący sytuację podpadającą (lub nie) pod kryteria określone przez badaną definicję pojęcia;
- (2) ocena, czy sytuacja rzeczywiście podpada pod to pojęcie;
- (3) odrzucenie lub zaakceptowanie definicji w zależności od rezultatu (2).

Przykłady Platona i Gettier'a spełniały, odpowiednio, badaną definicję sprawiedliwości i definicję wiedzy, mimo że pierwszy z nich nie opisywał

²⁹ Według Gettier'a definicja wiedzy jako uzasadnionego prawdziwego przekonania pojawia się u Platona w *Teajecie*; E. Gettier, *Is Justified True Belief Knowledge?*

sprawiedliwego zachowania, a drugi przekonania, które byłoby wiedzą. Przykład Kripkego służył jednocześnie obaleniu deskryptywnej teorii znaczenia i ugruntowaniu jego własnej: teoria deskryptywna wskazywała na Schmidta jako referenta nazwy *Gödel*, teoria kauzalna wskazywała na Gödla, poprawny wynik dał ten drugi model.

I tu rodzi się podstawowe pytanie: skąd wiadomo, że zwrócenie broni szaleńcowi nie jest sprawiedliwe? Że w opisanym przez Kripkego przypadku referentem nazwy Gödel jest Kurt Gödel? Że starający się o pracę Smith nie wie, kto otrzyma stanowisko? Krótko mówiąc: co pozwala na wykonanie kroku (2) w opisanym przez mnie schemacie?

Stosując pojęcie do przedmiotu, a więc stwierdzając prawdziwości zdań postaci *przedmiot S podpada pod pojęcie P*, korzystamy z *intuicji filozoficznej*. Wiemy (stwierdzamy), że nie jest sprawiedliwe oddać nóż szaleńcowi, bo tak podpowiada intuicja filozoficzna, niezależnie od tego, czy dysponujemy zwerbalizowaną definicją sprawiedliwości, lub nawet jeżeli nie zgadza się to z uznaną przez nas definicją. Intuicja zgodna z rozstrzygnięciem implikowanym przez teorię/definicję potwierdza ją, sprzeczna stanowi przesłankę do zrewidowania tej teorii/definicji.

Definicja intuicji jest raczej negatywna – łatwiej jest powiedzieć, czym intuicja nie jest, niż czym jest:

[...] doświadczenia umysłowe [*intellectual happenings*], w których coś przedstawia się nam jako przypadek [podpadający pod zadane pojęcie], ale nie w wyniku wnioskowania, zmysłowego postrzegania lub przypomnienia sobie.³⁰

[...] niezmysłowa i nieintrospektywna świadomość stwierdzenia lub pojęcia. [...] osoba ma intuicję, że P, wtedy i tylko wtedy gdy P jawi się jako prawdziwe, możliwie prawdziwe lub koniecznie prawdziwe, przy czym to jawienie się jest [...] oparte na rozumieniu, a nie na postrzeganiu zmysłowym lub introspekcji.³¹

³⁰ J. Weinberg, *How to Challenge Intuitions Empirically Without Risking Skepticism*, s. 319.

³¹ B. Russel, *Intuition, addendum 2*, [w:] *Encyclopedia of Philosophy*, 2nd edition.

Zatem intuicja nie sprowadza się do percepcji zmysłowej lub introspekcji, nie jest formą dedukcji ani odwołaniem do pamięci.³² Jedyne wymienione pozytywne warunki to rozumienie – w obliczu pytania, czy pojęcie aplikuje się do danego przedmiotu, osoba może oprzeć się o swoją intuicję pod warunkiem, że to pojęcie rozumie.

Metoda argumentacji z intuicji filozoficznej, ze względu na punkt (1) schematu nazywana także metodą eksperymentów myślowych, doskonale pasuje do Strawsona analogii filozofii i gramatyki. Jeżeli użytkownik języka polskiego usłyszy zdanie *niestety, autobus nie przyjechało*, to oceni je jako niepoprawne, chociaż nie musi umieć sformułować reguły: *podmiot i orzeczenie zgadzają się co do rodzaju*. Ta ocena pochodzi z *intuicji gramatycznej*, jeżeli miałbym tym razem przeprowadzić analogię w przeciwnym kierunku i tak nazwać niezmysłowy, nieintrospektywny i niededukcyjny akt oceny gramatyczności wyrażenia językowego. Opisując gramatykę wybranego języka, gramatyk weryfikuje proponowane przez siebie reguły w oparciu o intuicję gramatyczną użytkowników tego języka. Odrzuca te reguły, które dopuszczają zdania ocenione intuicyjnie jako niegramatyczne, i te, które zakazują zdań ocenionych intuicyjnie jako gramatyczne.

Clarke zauważa, że nie wystarczy obserwować, jak ludzie używają słów – a więc na przykład badać jedynie korpusy językowe – aby zbudować trafną teorię tego, co w danym języku jest poprawne. Przytacza przykład mieszkańców stanu Georgia, którzy mówią z charakterystycznym akcentem. Język Amerykanów spoza Georgii, którzy odwiedzą ten stan, nie zostanie odebrany jako niepoprawny tylko dlatego, że brakuje im lokalnego akcentu.³³ Podobnie z pojęciami: fakt, że obserwacja zachowania użytkowników języka nie dostarczy żadnych przypadków, w których uznali przedmiot S za podpadający pod pojęcie P, nie oznacza, że nie uznaliby S za podpadający pod P, gdyby zostali o to zapytani.

Ta uwaga ma bezpośredni związek z przedstawionym przeze mnie schematem argumentacji z intuicji. Zgodnie z punktem (1) filozof

³² W kontekście przytoczonych definicji introspekcję należy rozumieć jako obserwację, której przedmiotem są przeżycia wewnętrzne.

³³ D. Clarke, *Philosophy's Second Revolution*, s. 101.

przedstawia pewien eksperyment myślowy, czyli opis przedmiotu S. Następnie pyta, czy przedmiot S spełnia analizowane pojęcie P. Zrozumienie opisu S pozwala na intuicyjne stwierdzenie, czy S podpada pod P. Ten sposób pozwala na wyjście poza przykłady dostarczane przez codzienne użycia języka i zbadanie intuicji filozoficznych w dowolnych kontekstach.

Oczywiście nie jest tak, że wystarczy jedna intuicja sprzeczna z definicją, aby ją obalić, podobnie jak w naukach szczegółowych jedna obserwacja sprzeczna z teorią nie wystarczy, żeby ją odrzucić. Intuicje nie muszą być, i przeważnie nie są, spójne. W przypadku niespójności można je ważyć – porównywać ze sobą i oceniać, na zachowaniu których bardziej nam zależy, a które możemy poświęcić, *poddać kwarantannie*.³⁴ Można też uznać, że na niektórych intuicjach nie należy polegać:

Na przykład możemy porzucić te sądy, co do których wahamy się lub których nie jesteśmy pewni. Albo te, które wysuwamy, gdy jesteśmy zdenerwowani lub przerażeni [...].³⁵

Procedurę uzgadniania intuicji Rawls nazwał *równowagą refleksyjną* i zastosował w *Teorii sprawiedliwości* przy uzasadnianiu swojej koncepcji z filozofii politycznej, ale metoda była wykorzystywana wcześniej i nie ogranicza się do żadnego wybranego obszaru filozofii.³⁶

Intuicja filozoficzna pozostaje jednym z podstawowych narzędzi filozofów analitycznych, chociaż budzi poważne zastrzeżenia. W tym miejscu chciałbym przytoczyć dwa spośród nich, które przyczyniły się do powstania filozofii eksperymentalnej. Część pozostałych problemów związanych z intuicją filozoficzną omówię w ostatnim rozdziale.

Autorzy argumentacji filozoficznej korzystają, siłą rzeczy, z własnej intuicji filozoficznej, jednak język tych argumentów zdradza coś zupełnie innego. Sokrates nie mówi Kefalosowi, że zwrócenia szaleńcowi broni nie nazwałby postępowaniem sprawiedliwym, mówi znacznie więcej: *przecież każdy przyzna*, że nie jest sprawiedliwie oddać broń takiej osobie. Gettier

³⁴ J. Horvath, *How (Not) to React to Experimental Philosophy*.

³⁵ J. Rawls, *A Theory of Justice, Revised Edition*, s. 42.

³⁶ N. Daniels, *Reflective Equilibrium* [w:] *The Stanford Encyclopedia of Philosophy (Fall 2008 Edition)*.

nie stwierdza, że według niego Smith nie wie, że osoba z dziesięcioma monetami w kieszeni dostanie pracę, ale że *jasne jest, że nie wie*. Kripke nie zdaje sprawy, do kogo on sam się odnosi, ale używa pierwszej osoby liczby mnogiej – *wydaje mi się jednak, się nie odnosimy [do Schmidta]. Po prostu nie odnosimy się*. Jego teoria ma opisać, do kogo/czego *zwykły człowiek w zamysle się odnosi*. Ostatecznie, kauzalna teoria referencji ma opisać *nasze* zachowanie językowe, *naszą* relację referencji, gdzie *my* to po prostu ludzie w ogóle.³⁷

Tak więc intuicje wykorzystywane przez filozofów w argumentacji mają pretensję do uniwersalności. Stwierdzając, że S jest instancją P, filozof nie tylko zdaje sprawę ze swojego osądu, ale zakłada, że przeważająca większość użytkowników języka zgodzi się z nim. Oczywiście jest możliwe, że dana intuicja jest powszechna, ale takie założenie jest z jednej strony bardzo silne, a z drugiej empirycznie sprawdzalne. Dlatego wypadaloby je zweryfikować.

Tak naprawdę zarzut nie jest taki prosty. Jeżeli filozof opublikuje tekst, w którym opiera argumentację na swoich intuicjach, częściową weryfikację wymusi sposób funkcjonowania wspólnoty akademickiej. Jeżeli tylko opublikowany tekst zawiera wystarczająco istotne wnioski – na chwilę przyjmę, że w świecie akademickim liczą się tylko względy merytoryczne – będzie czytany i komentowany przez innych filozofów, którzy skonfrontują swoje intuicje z intuicjami autora, a tym samym dostarczą (lub nie) potwierdzenia tych intuicji. Czy to wystarczy?

Grupa, w której taki tekst jest dyskutowany, to wąski zbiór ludzi o bardzo specyficznym wykształceniu i często podobnym profilu społecznym i kulturowym. Stąd bierze się podejrzenie, że intuicje tej grupy wcale nie są reprezentatywne dla ludzi w ogóle. Zamiast argumentu *w tym przypadku każdy zwykły człowiek przyzna, że...* powinno raczej pojawić się zdanie *w tym przypadku każdy zwykły filozof analityczny przyzna, że...* Oczywiście nie to chcą powiedzieć filozofowie, formułując swoje uzasadnienia.³⁸

³⁷ Por. podkreślenia w cytowanych trzech argumentach.

³⁸ Czy rzeczywiście filozofowie mają na myśli *każdego zwykłego człowieka*, czy może węższą grupę – te kwestie omawiam w ostatnim rozdziale, poświęconym zarzutom wobec filozofii eksperymentalnej.

Drugi zarzut wobec intuicji jako narzędzia filozoficznego, istotny z punktu widzenia filozofii eksperymentalnej, to brak uzasadnienia wiarygodności intuicji. W artykule *How to Challenge Intuitions Empirically Without Risking Skepticism* Jonathan Weinberg proponuje nazywać *obietującymi* [*hopeful*] te narzędzia poznania, które w mniejszym lub większym stopniu spełniają cztery warunki: są potwierdzone zewnętrznymi, wewnętrznymi spójne, znane są granice ich stosowania (lub można takie granice wyznaczyć) i znana jest zasada ich działania.³⁹ Narzędzie jest potwierdzone zewnętrznymi, kiedy dostarczone przez nie wyniki są potwierdzone przez inne środki poznania. Jest wewnętrznymi spójne, kiedy jego użycie przez inne (kompetentne) podmioty poznające i/lub wielokrotne użycie przez ten sam podmiot daje podobne rezultaty. Ma wyznaczone granice stosowania, kiedy wiadomo, w jakich okolicznościach można na nim polegać, a w jakich nie. Znana jest zasada działania, kiedy istnieje (zadowalający) model tłumaczący, jak to narzędzie działa, kiedy działa, i dlaczego zawodzi, kiedy zawodzi. Narzędzia poznania, które nie są obietujące, a więc które nie spełniają w wystarczającym stopniu podanych warunków, są *beznadziejnymi* [*hopeless*]. W oparciu o ten podział Weinberg formułuje i szeroko uzasadnia postulat, że nie należy korzystać z beznadziejnych narzędzi poznania.

Słyszę dźwięk rozbijanego szkła. Wniosuję, że rozbilo się coś szklanego. Wniosek wyciągnięty na podstawie wrażenia słuchowego zostaje potwierdzony w oparciu o zmysł wzroku, gdy podchodzę do skorup szklanki, i w oparciu o dotyk, gdy je zbieram. Inni też potwierdzają wnioski, wyciągnięte przeze mnie na podstawie wrażeń dostarczonych mi przez zmysł słuchu, pytając *też to słyszales? Chyba coś się rozbilo*. Z drugiej strony wiem, że kiedy jestem niewyspany, miewam halucynacje dźwiękowe, ale też umiem rozpoznać warunki, w których mogę na moim słuchu polegać. Co więcej, działanie moich uszu jest lepiej lub gorzej wytłumaczone przez medycynę. Stąd zmysł słuchu – jako narzędzie poznania – jest obietującym.

³⁹ J. Weinberg, *How to Challenge Intuitions Empirically Without Risking Skepticism*, s. 330. Angielskie określenie tych warunków: *external corroboration, internal coherence, detectability of margins i theoretical illumination*.

Weinberg podaje ciekawsze przykłady z historii: mikroskop elektronowy. Obrazy uzyskane dzięki takiemu mikroskopowi w danym laboratorium są spójne z wynikami uzyskanymi dzięki innym rodzajom mikroskopów i z wynikami uzyskanymi w mikroskopach elektronowych w innych laboratoriach. Wiadomo, w których przypadkach można się posłużyć mikroskopem elektronowym, a kiedy wyniki są niepoprawne. I, oczywiście, działanie mikroskopu elektronowego jest wytłumaczone przez odpowiednie teorie fizyczne – te same, które w ogóle pozwoliły inżynierom na skonstruowanie tego narzędzia. Z drugiej strony introspekcja, której psychologowie eksperymentalni po Wundcie starali się używać do badania bardziej złożonych fenomenów, jest przykładem beznadziejnego narzędzia, które, właśnie ze względu na to, że jest beznadziejne, zostało zarzucone.

Pytanie, które nasuwa się w kontekście metody eksperymentów myślowych, brzmi: czy intuicje są narzędziem obiecującym, czy beznadziejnym? W istocie celem artykułu Weinberga było zarysowanie odpowiedzi. Kwestia ma wielką wagę filozoficzną – rozumowania oparte o beznadziejne intuicje należałoby odrzucić, a z ich wniosków zrezygnować lub postarać się je dowieść obiecującymi narzędziami. Nie należy oceniać wszystkich intuicji za jednym razem, ponieważ niektóre rodzaje – na przykład intuicje używane w estetyce – mogą okazać się beznadziejne, a inne – na przykład te używane w logice – obiecujące. W projekcie odróżnienia intuicji obiecujących od beznadziejnych podstawową rolę pełnią badania empiryczne.

Przypuśćmy, że badamy intuicje etyczne dotyczące czynów określonego rodzaju. Kilkudziesięciu osobom prezentujemy odpowiednie eksperymenty myślowe i okazuje się, że ocena czynu zależy od kolejności prezentacji tych eksperymentów – a więc od czynnika etycznie nieistotnego. Przypuśćmy dalej, że nie ma dowodów, aby kolejność prezentacji materiału wywoływała podobne efekty w tych dziedzinach ludzkiego poznania, dla których istnieją jasne kryteria poprawności (odwrotnie niż, na przykład, upojenie alkoholowe, które wpływa negatywnie na cały wachlarz ludzkich zdolności: od rozwiązywania zadań z logiki po kierowanie samochodem). Skoro tak, to nie ma argumentów

na rzecz tego, aby ufać raczej intuicjom towarzyszącym takiej, a nie innej kolejności pytań. Podatność na założenia nieistotne czynniki oznacza, że ten rodzaj intuicji nie jest wewnętrznie spójny. Wewnętrzna niespójność razem z brakiem zewnętrznego potwierdzenia – o które trudno w przypadku ocen etycznych – czynią w tym przykładzie ten rodzaj intuicji beznadziejnym narzędziem.

Ten przykład pokazuje, jak możliwe są empiryczne badania o bezpośrednich filozoficznych konsekwencjach. Co więcej, metodologia tego hipotetycznego badania wydaje się bardzo prosta, ponieważ do zebrania materiału empirycznego wystarczy poprawnie skonstruowany kwestionariusz i liczba ankietowanych wystarczająca do wyprowadzenia statystycznie istotnych wniosków.

filozofia eksperymentalna: cele i metody

Dwa wymienione powody – niepotwierdzone pretensje filozofów do powszechności ich intuicji i niewystarczające uprawomocnienie intuicji jako metody uzasadniania – legły u podstaw filozofii eksperymentalnej. Filozofowie eksperymentalni zauważyli, że nie muszą czekać na wyniki nauk szczegółowych, tak jak to ma miejsce w przypadku filozofii empirycznej. Psychologowie, socjologowie i ekonomiści eksperymentalni – bo to ich badania najbardziej przypominają eksperymenty filozoficzne – mają własne cele, niekoniecznie zbieżne z celami filozofów. Dlatego filozoficznie istotne rezultaty trafiają się w tych badaniach przeważnie przez przypadek. Jednocześnie pewna część metod wykorzystywanych w tych dziedzinach jest łatwa do zastosowania (w odróżnieniu od, na przykład, narzędzi biologii lub fizyki) i nie wymaga specjalistycznej wiedzy. Stąd, jeżeli przed filozofią stanęły empiryczne pytania, dlaczego nie zaprojektować odpowiednich eksperymentów samodzielnie?

Niezależnie od silnego zakorzenienia w metodologii analizy pojęciowej, filozofia eksperymentalna czuje się spadkobierczynią tradycji filozoficznej, od której filozofia analityczna zdystansowała się w swoich początkach.

Kiedyś zgadzano się co do tego, że filozofia stawia pytania o kondycję ludzką. Filozofowie zajmowali się człowiekiem i tym, jak funkcjonuje umysł. Interesowali się rozumem i emocjami, kulturą i wrodzonymi ideami, pochodzeniem przekonań

moralnych i religijnych. W tym tradycyjnym poglądzie nie było szczególnie ważne, żeby filozofia była wyraźnie oddzielona od psychologii, historii lub nauk politycznych. Filozofowie zastanawiali się, w sposób bardzo ogólny, jak to wszystko łączy się ze sobą.

Nowy ruch filozofii eksperymentalnej szuka powrotu do tej tradycyjnej wizji. Tak jak filozofowie sprzed wieków, jesteśmy zainteresowani pytaniami o to, jakie naprawdę są istoty ludzkie. Zdajemy sobie sprawę, że to zadanie wymaga badania zjawisk nieuporządkowanych, uwikłanych i zmiennych w czasie i przestrzeni, ale nie widzimy, dlaczego nasz namysł miałby być przez to mniej filozoficzny. Wręcz przeciwnie, uważamy, że na wiele najważniejszych pytań filozoficznych można odpowiedzieć jedynie zgłębiając nieuporządkowane, uwikłane i zmienne prawdy o tym, jacy ludzie są rzeczywiście.⁴⁰

W *Manifestie filozofii eksperymentalnej* Joshua Knobe i Shaun Nichols ilustrują tę zmianę dającym do myślenia przykładem. Na pierwszym roku, na kursie filozofii starożytnej, studenci słyszą, że Platon zaproponował podział duszy na trzy części, co pozwoliło mu wyjaśnić niektóre zjawiska psychiczne jako konflikt lub współpracę tych części. Studenci, którzy zaczynają zastanawiać się, czy pomysł podziału umysłu na części jest prawdziwy i jaki materiał dowodowy należałoby zebrać, aby zweryfikować tę koncepcję, słyszą od wykładowców: *nie, nie, nie rozumiesz. Jeżeli chcesz uprawiać prawdziwą filozofię, musisz nauczyć się pozostawiać innym rozwiązanie takich problemów.*⁴¹ Takie podejście jest sprzeczne z pierwotnym duchem filozofii i, ograniczając badaczy, przynosi szkodę dyscyplinie.

Spora część filozofów analitycznych nie podziela tego wniosku. Początek wystąpienia Timothy'ego Williamsona *Armchair philosophy, metaphysical modality and counterfactual thinking*, wygłoszonego podczas jednego z zebrań *Aristotelian Society*, brzmiał: *Jeżeli cokolwiek można uprawiać w fotelu, to z pewnością filozofię. Jej tradycyjną metodą jest myślenie, bez obserwacji*

⁴⁰ J. Knobe, S. Nichols, *An Experimental Philosophy Manifesto*, [w:] J. Knobe, S. Nichols, *Experimental Philosophy*, s. 3.

⁴¹ J. Knobe, S. Nichols, *An Experimental Philosophy Manifesto*, [w:] J. Knobe, S. Nichols, *Experimental Philosophy*, s. 14.

i eksperymentów.⁴² Filozofowie postępujący w tym duchu, nazywani *gabinetowymi*, starają się odpierać oskarżenia o nieuzasadnione pretensje intuicji do powszechności i o ich podejrzaną status epistemiczny, jednocześnie odmawiając metodzie eksperymentalnej adekwatności w stosunku do problemów filozoficznych. Część wysuwanych przez nich zarzutów omówię w ostatnim rozdziale.

Najbardziej rzucająca się w oczy różnica między filozofami eksperymentalnymi a gabinetowymi polega na tym, że ci pierwsi w swojej argumentacji odwołują się do statystycznego rozkładu intuicji:

[...] analityk pojęciowy napisze raczej *w tym wypadku z pewnością powiemy, że...*, kiedy filozof eksperymentalny stwierdzi *w tym wypadku 79% badanych powiedziało, że...*⁴³

Jednocześnie jest to różnica najbardziej powierzchowna – istotne cechy filozofii eksperymentalnej dotyczą bardziej zasadniczych założeń metodologii filozoficznej. Artykuł *The Past and Future of Experimental Philosophy* jest poświęcony przede wszystkim odparciu pewnych zarzutów wysuwanych wobec filozofii eksperymentalnej, ale przy okazji tej obrony autorzy, Eddy Nahmias i Thomas Nadelhoffer, w bardzo jasny sposób opisali podstawowe założenia dziedziny, której bronią, i podziały wewnątrz niej. Wskazują, że filozofia eksperymentalna nie jest homogenicznym nurtem, ale składa się z trzech odrębnych projektów o odmiennych przesłankach, celach i wnioskach. Te projekty to: analiza eksperymentalna [*experimental analysis*], deskryptywizm eksperymentalny [*experimental descriptivism*] i restrykcjonizm eksperymentalny [*experimental restrictionism*].⁴⁴ Ich wspólnym mianownikiem są dwa założenia: nieufność wobec pretensji filozofów do powszechności ich intuicji filozoficznej oraz przekonanie, że intuicje filozoficzne można badać przy pomocy metod eksperymentalnych.⁴⁵

⁴² T. Williamson, *Armchair philosophy, metaphysical modality and counterfactual thinking*, s.2.

⁴³ J. Knobe, S. Nichols, *An Experimental Philosophy Manifesto*, [w:] J. Knobe, S. Nichols, *Experimental Philosophy*, s. 4.

⁴⁴ E. Nahmias, T. Nadelhoffer, *The Past and Future of Experimental Philosophy*, s. 126-129.

⁴⁵ E. Nahmias, T. Nadelhoffer, *The Past and Future of Experimental Philosophy*, s. 125.

Celem analizy eksperymentalnej jest badanie intuicji filozoficznych zwykłych ludzi tam, gdzie wyniki takiego badania mają konsekwencje filozoficzne. Jest to możliwe przede wszystkim w przypadku tych teorii filozoficznych, które ich autorzy starają się uprawomocnić odwołaniem do powszechnej intuicji [*folk intuition*]. Analiza eksperymentalna może obalić takie uzasadnienie, a w dalszej perspektywie wskazać teorię, która najlepiej oddaje intuicję zwykłych ludzi.

Nie jest tak, że najbardziej intuicyjna propozycja musi być prawdziwa. Autorzy tych mniej intuicyjnych mogą starać się wyjaśnić, dlaczego intuicje przeważającej części ludzi nie zgadzają się z ich własną i dlaczego ci ludzie myślą się w swoim osądzie, a więc sformułować odpowiednią teorię błędu [*error theory*]. Mogą też argumentować za zrewidowaniem tych intuicji wobec, na przykład, wyników nauk szczegółowych. Oczywiście, aby w ogóle rozpatrywać którąś z tych możliwości, najpierw trzeba stwierdzić nieprzystawanie teorii do powszechnych intuicji, a więc przeprowadzić analizę eksperymentalną.⁴⁶

W *Manifestie* Knobe i Nichols deklarują, że filozofia nigdy nie była konkursem popularności i filozofia eksperymentalna nie zamierza tego zmieniać.⁴⁷ W przypadku analizy eksperymentalnej nie jest to do końca prawda. Jeżeli filozof na poparcie swojej teorii odwołuje się głównie do powszechnej intuicji, to uprawiana przez niego filozofia ma wiele wspólnego z konkursami popularności, a jego teoria jest podatna na empiryczne sprawdzenie roszczenia do powszechności. Oczywiście w przypadku negatywnego wyniku można starać się bronić pomysł przy pomocy teorii błędu, ale to zasadniczo zmienia pierwotną formę argumentacji.

W centrum uwagi drugiej grupy filozofów eksperymentalnych, deskryptywistów eksperymentalnych, są źródła intuicji filozoficznych. Podstawowe stawiane przez nich pytanie brzmi: jakie mechanizmy psychologiczne odpowiadają za powstanie intuicji filozoficznych dotyczących zadanego problemu i na których spośród tych mechanizmów

⁴⁶ E. Nahmias, T. Nadelhoffer, *The Past and Future of Experimental Philosophy*, s. 126.

⁴⁷ J. Knobe, S. Nichols, *An Experimental Philosophy Manifesto*, [w:] J. Knobe, S. Nichols, *Experimental Philosophy*, s. 6.

można polegać, a na których nie, którym intuicjom ufać, a którym nie? Odkrycie tych mechanizmów służy potem jako argument za przyjęciem pewnych teorii filozoficznych i odrzuceniem innych.⁴⁸ Odwrotnie niż w przypadku klasycznej analizy pojęciowej, *postęp tu jest mierzony nie precyzją w opisananiu wzorca charakteryzującego intuicje ludzi, ale stopniem wyjaśnienia tych intuicji.*⁴⁹

W przeciwieństwie do analizy eksperymentalnej, deskrytywizm eksperymentalny nie traktuje wszystkich intuicji jako równie istotnych, ale stara się podzielić je na godne i niegodne zaufania na podstawie mechanizmu psychologicznego, którego są wynikiem. W tym sensie deskrytywizm stara się sięgnąć głębiej niż analiza. Jednocześnie badanie deskrytywistyczne wymaga przeprowadzenia analizy empirycznej, ponieważ, aby opisać mechanizm psychologiczny rodzący intuicje, potrzeba najpierw dowiedzieć się, jakie one w ogóle są.

Bywa, że projekty z analizy eksperymentalnej przeradzają się w badania deskryptywne. Tak było w przypadku badań Knobe'a nad intencjonalnością. Początkowo dane pochodziły wyłącznie z ankietowania studentów, a wnioski filozoficzne zostały oparte jedynie o statystyczny rozkład odpowiedzi.⁵⁰ Ale program badawczy rozszerzano, zbierano nowe dane. Knobe zaproponował wyjaśnienie zaobserwowanego efektu przy pomocy pewnego mechanizmu psychologicznego.⁵¹ W celu wykluczenia niektórych hipotez ankietowano między innymi osoby z uszkodzeniami przedczołowej kory brzuszno-przyśrodkowej, odpowiedzialnej za emocje.⁵² Pojawiły się także interpretacje inne niż wysunięta przez Knobe'a. Ich autorzy uznali część odpowiedzi ankietowanych za mylne, proponując pewną teorię błędu, na poparcie której przeprowadzili kolejne eksperymenty.⁵³

⁴⁸ E. Nahmias, T. Nadelhoffer, *The Past and Future of Experimental Philosophy*, s. 127.

⁴⁹ J. Knobe, S. Nichols, *An Experimental Philosophy Manifesto*, [w:] *Experimental Philosophy*, s. 5.

⁵⁰ J. Knobe, *Intentional Action and Side Effects in Ordinary Language*.

⁵¹ J. Knobe, *Person as Scientist, Person as Moralist*.

⁵² L. Young, F. Cushman, R. Adolphs, D. Tranel, M. Hauser (2006) *Does emotion mediate the effect of an action's moral status on its intentional status? Neuropsychological evidence*.

⁵³ A. Mele, *Acting intentionally: Probing folk intuitions*. [w:] Malle B., Moses L., Baldwin D., *Intentions and intentionality: Foundations of social cognition*.

Restrykcyjizm eksperymentalny opiera się na obserwacji, że intuicje filozoficzne zależą od nieistotnych filozoficznie czynników: m.in. płci, statusu socjo-ekonomicznego lub przynależności kulturowej. Jeżeli nie istnieje przy tym odpowiednia teoria błędu, która tłumaczyłaby na przykład, dlaczego należy przedkładać epistemologiczne intuicje obywateli rozwiniętych krajów Zachodu ponad intuicje Chińczyków lub Hindusów, intuicje nie mogą być narzędziem konstruowania – w tym wypadku epistemologicznych – teorii. Restrykcyjizm eksperymentalny dotyczy zasadniczo dwóch ostatnich spośród trzech rodzin wymienianych przez Stephena Sticha: projektów neoplatońskich i projektów normatywnych. To, czy dane przekonanie jest – obiektywnie – wiedzą, nie zależy od kulturowej przynależności orzekającego o tym obserwatora albo od kolejności prezentacji przykładów. Jeżeli nie ma powodów, dla których tę a nie inną kulturę lub taką a nie inną kolejność należy wyróżnić, intuicja epistemologiczna nie nadaje się do uzasadniania teorii epistemologicznych.

Tak więc restrykcyjista eksperymentalny szuka problemów filozoficznych, przy rozwiązywaniu których filozofowie odwołują się do intuicji filozoficznych zależnych od czynników, które sami uważają za filozoficznie nieistotne. W takim wypadku ciężar uprawomocnienia odwołania do intuicji spada na tych filozofów. Jeżeli nie mogą oni podać przekonującej teorii błędu, intuicja filozoficzna zostaje uznana za nieodpowiednie narzędzie. Krótko mówiąc: głównym zadaniem restrykcyjistów jest tropienie nieuzasadnionego użycia intuicji.⁵⁴

Mimo wspólnych cech, te trzy projekty mają różny stosunek do intuicji: analiza eksperymentalna szuka potwierdzenia lub obalenia teorii w oparciu o powszechne intuicje, eksperymentalny deskryptywizm stara się rozróżnić między intuicjami godnymi a niegodnymi zaufania, a eksperymentalny restrykcyjizm zasadniczo nie uznaje intuicji za prawomocne narzędzie w wielu obszarach filozofii. Te programy filozoficzne przeważnie nie są wzajemnie sprzeczne, są jedynie rozbieżne w swoich celach. Pierwsze dwa najczęściej odnoszą się do rodziny projektów analizy pojęciowej, jak je nazywa Stich, lub, zgodnie z Jacksonem, skromnej analizy pojęciowej. Eksperymentalny restrykcyjizm uderza z reguły w projekty neoplatońskie

⁵⁴ E. Nahmias, T. Nadelhoffer, *The Past and Future of Experimental Philosophy*, s. 128.

i normatywne lub, inaczej, w nieskromną analizę pojęciową. Z reguły – na przykład w jednej z pierwszych prac eksperymentalnofilozoficznych, *Semantics, cross-cultural style*, Stich i jego współpracownicy krytykują kauzalno-historyczną teorię referencji, którą należałoby raczej uważać za analizę pojęć niż za projekt neoplatoński, bo teoria Kripkego ma opisywać, do czego *zwykły człowiek odnosi się w swoim zamysle*, a nie idealną relację, która łączy znak i desygnat bez względu na zachowanie językowe ludzi.⁵⁵

Podstawowym narzędziem badawczym filozofii eksperymentalnej są ankiety. W ankiecie badacze opisują uczestnikom pewną sytuację, a następnie proszą o ocenę, czy ta sytuacja podpada pod określone pojęcie. Przedstawione sytuacje są często inspirowane klasycznymi eksperymentami myślowymi – na przykład we wspomnianej pracy *Semantics, cross-cultural style* jej autorzy użyli Kripke'owskiej opowieści o Gödlu w niemalże niezmienionej formie. Po opisie następuje pytanie o intuicję: czy mówiąc o Gödlu Y odnosi się do prawdziwego odkrywcy dowodu o niezupełności arytmetyki. W ten sposób ankiety symulują eksperyment myślowy.

W odpowiedziach używane są różne skale, od binarnych (*tak, nie*) do skali Likerta i bardziej szczegółowych. Oprócz odpowiedzi zbierane są także dane demograficzne, przy pomocy dodatkowych pytań badacze kontrolują, czy ankietowany zrozumiał przedstawiony mu opis, a pytania otwarte pozwalają na analizę argumentów uczestników. W ankiecie można użyć większej liczby opisów. W takim przypadku można kontrolować ich kolejność, w tym badać wpływ na przykład torowania [*priming*]. Kolejne dwa rozdziały zawierają przykłady ankiet filozoficznych.

W wielu badaniach metodę ankiet rozszerza się o manipulację zewnętrznymi czynnikami. Schnall i jej współpracownicy badali wpływ czystości bezpośredniego otoczenia ankietowanych na ich intuicje moralne, zostawiając kawałki starej pizzy na stole, przy którym uczestnicy wypełniali ankiety.⁵⁶ Inny eksperyment polegał na wyświetleniu jednej

⁵⁵ E. Machery, R. Mallon, S. Nichols, S. Stich, *Semantics, cross-cultural style*.

⁵⁶ S. Schnall, J. Benton, S. Harvey, *With a clean conscience: Cleanliness reduces the severity of moral judgments*.

grupie badanych fragmentu programu *Saturday Night Live*, innej – krajoznawczego filmu, i zbadaniu różnic w intuicjach moralnych tak uitorowanych osób.⁵⁷ Ostatecznie wyniki ankiet analizowane są przy pomocy narzędzi statystycznych (z reguły wykorzystuje się proste testy na równość średnich lub równość proporcji).

Filozofowie eksperymentalni, szczególnie deskryptywiści eksperymentalni, zainteresowani mechanizmami psychologicznymi odpowiedzialnymi za intuicje filozoficzne, używają także zaawansowanych technicznie narzędzi. W swoim badaniu Greene i Haidt użyli obrazowania metodą rezonansu magnetycznego do identyfikacji tych procesów w mózgu, które uaktywniają się podczas analizowania przez uczestników klasycznych eksperymentów myślowych z dziedziny etyki.⁵⁸

Filozofowie eksperymentalni deklarują, że używają metod psychologii eksperymentalnej.⁵⁹ W pewnym sensie nie jest to prawda. Ankiety, głównie, choć nie jedyne narzędzie badawcze filozofii eksperymentalnej, nie są często wykorzystywane w psychologii eksperymentalnej; częściej w psychologii społecznej i, przede wszystkim, w badaniach marketingowych. Brian Scholl, psycholog z Yale, w wykładzie *Two Kinds of Experimental Philosophy (and their methodological dangers)* wygłoszonym podczas spotkania *Pre-SPP Workshop on Experimental Philosophy* zaproponował inną metodę badania intuicji i leżących u ich podstaw procesów. Według niego należy raczej rozpocząć badanie od wybrania któregoś ze znanych mechanizmów, potwierdzonych w psychologii przez narzędzia dużo silniejsze niż ankiety, i eksperymentalnie sprawdzić, czy ten mechanizm leży u podstaw danych intuicji. Odwołanie się do rozpoznanych przez psychologię procesów (a więc procesów o znanych własnościach, łatwiejszych do identyfikacji) pozwoli na silniejsze, pod względem naukowym, wyjaśnienie wyników eksperymentu filozoficznego. Według Scholla taką metodę obrał np. Greene we wspomnianym już badaniu.⁶⁰ Uwagi Scholla dotyczą przede

⁵⁷ P. Valdesolo, D. DeSteno, *Manipulations of Emotional Context Shape Moral Judgment*.

⁵⁸ J. Greene, J. Haidt, *How (and where) does moral judgment work?*

⁵⁹ E. Nahmias, T. Nadelhoffer, *The Past and Future of Experimental Philosophy*, s. 126.

⁶⁰ B. Scholl, *Two Kinds of Experimental Philosophy (and their methodological dangers)*, nieformalna notatka ze spotkania *Pre-SPP Workshop on Experimental Philosophy*, 25 VI 2008.

wszystkim deskrytywizmu eksperymentalnego – analiza eksperymentalna i restrykcjonizm eksperymentalny nie są tak bardzo zainteresowane mechanizmami psychologicznymi.

Czy powstanie filozofii eksperymentalnej jest rewolucją filozoficzną w rozumieniu Clarke'a? Chyba nie. Filozofia eksperymentalna nie obrała sobie nowego przedmiotu – w centrum zainteresowania znajdują się nadal pojęcia oraz niektóre klasyczne pytania filozofii analitycznej. Tym samym filozofii eksperymentalnej brakuje najważniejszej cechy rewolucji: zmiany przedmiotu filozofii. Nie nastąpiło także oderwanie od filozofii kolejnej nauki szczegółowej; raczej odwrotnie, filozofia częściowo weszła w kompetencje empirycznych nauk społecznych, w szczególności psychologii.

Knobe i Nichols wyraźnie deklarują w *Manifestie*, że filozofia eksperymentalna dodaje nowe narzędzie do filozoficznego przybornika, natomiast nie zamierza pozbywać się niczego, co już się w nim znajduje.⁶¹ Co prawda wyniki filozofii eksperymentalnej wpisują się w długą tradycję antykognitywizmu w etyce teoretycznej, epistemologicznego relatywizmu i leżącej u ich podstaw ontologicznej oszczędności, ale tak naprawdę nic nie stoi na przeszkodzie, aby metody analizy eksperymentalnej lub eksperymentalnego deskrytywizmu wykorzystać w uzasadnianiu głęboko platońskich propozycji filozoficznych.

Podejrzliwość wobec gabinetowego użycia intuicji filozoficznej, którą dzielają wszyscy filozofowie eksperymentalni, nie oznacza całkowitej rezygnacji z tego instrumentu. Ostatecznie, jeżeli autorzy publikacji eksperymentalnofilozoficznych chcą przekonać do swoich wniosków, wykorzystane przez nich przesłanki, argumenty i podjęte decyzje metodologiczne muszą trafić w intuicję czytelników.

⁶¹ J. Knobe, E. Nahmias, *An Experimental Philosophy Manifesto*, s. 10.

W tym rozdziale przedstawiam trzy eksperymenty filozoficzne: z semiotyki, z epistemologii i z filozofii działania. Dwa pierwsze to eksperymenty restrykcyjny, ostatni należy do analizy eksperymentalnej, a inspirowane nim dalsze badania do nurtu deskrytywizmu eksperymentalnego. Wybór padł na te eksperymenty z dwóch powodów. Po pierwsze, ponieważ dotyczą trzech różnych gałęzi filozofii (i nie dotyczą etyki – badania nad intuicjami moralnymi są jednymi z najłatwiejszych do przeprowadzenia). Po drugie dlatego, że należą do najwcześniejszych eksperymentów filozoficznych i jednocześnie do najszerzej dyskutowanych.

deskryptywna a kauzalna teoria referencji

W swoich badaniach nad międzykulturowymi różnicami w intuicjach filozoficznych Stephen Stich jako pierwszy wykorzystał metodę eksperymentalną i w tym sensie można uważać go za założyciela kierunku. Te pierwsze eksperymenty dotyczyły intuicji epistemologicznych i intuicji semiotycznych. Obecnie stanowią one wzorzec dla wielu późniejszych projektów restrykcyjnych, tak jak badania empiryczne Joshui Knobe'a nad pojęciem intencjonalności legły u podstaw analizy eksperymentalnej. W tym podrozdziale omówię wyniki dotyczące teorii referencji, w następnym – teorii epistemologicznych.

We wspomnianym już artykule *Semantics, cross-cultural style* Mallon, Machery, Nichols i Stich wykorzystali wyniki badań nad intuicjami semiotycznymi w krytyce teorii referencji, a w artykule *Against Arguments From Reference*

wyprowadzili z tych wyników konsekwencje między innymi dla filozofii nauki i filozofii umysłu.⁶²

Przed opublikowaniem książki Saula Kripkego *Nazywanie i konieczność*, będącej zapisem jego wykładów na Princeton University, wśród filozofów bardzo popularna była deskryptywna teoria referencji. W odniesieniu do nazw własnych teoria deskryptywna głosi, że każda nazwa własna jest deskrypcją, to znaczy, że z każdą nazwą własną związany jest zbiór cech. Zgodnie z tą teorią referentem nazwy własnej jest jedyny przedmiot, który posiada te cechy całkowicie lub chociaż wystarczająco, przy czym każdy inny przedmiot charakteryzuje się tymi cechami w mniejszym stopniu. Jeżeli nie istnieje dokładnie jeden taki przedmiot, nie jest spełnione założenie, że nazwa jest nazwą własną.

Teorii deskryptywnej w odniesieniu do nazw własnych Kripke przeciwstawił teorię kauzalno-historyczną. Zgodnie z nią nazwa zostaje wprowadzona do wspólnoty językowej, aby umożliwić odnoszenie się do pojedynczego przedmiotu.⁶³ Każdy, kto posługuje się nazwą, odnosi się do tego samego obiektu co osoba, od której tę nazwę usłyszał. Taki łańcuch przekazywania sobie nazwy wraz z referentem prowadzi aż do tego, kto w akcie chrztu językowego pierwszy użył jej do oznaczenia referenta. I chociaż użytkownicy mogą kojarzyć z nazwą pewne opisy, nie ma to na relację oznaczania żadnego wpływu – przedmiot może zupełnie nie spełniać kojarzonego z nazwą zbioru cech i nadal być jej referentem.

Swoją teorię Kripke uzasadnił jedynie przez odwołanie się do intuicji słuchaczy swoich wykładów – nie poparł jej w żaden inny sposób. Aby udowodnić, że jego pomysł lepiej ujmuje powszechne intuicje niż model deskryptywny, podał takie użycie nazw, w którym teoria kauzalno-historyczna i teoria deskryptywna wyznaczają różnych referentów, a intuicja potwierdza werdykt tej pierwszej. Kripke przedstawił kilka takich przykładów, między innymi cytowany w poprzednim rozdziale przypadek Gödla.

⁶² E. Machery, R. Mallon, S. Nichols, S. Stich, *Semantics, cross-cultural style* oraz *Against Arguments From Reference*.

⁶³ Dalej będę używał słowa *nazwa* w znaczeniu *nazwa własna*, o ile tylko nie będzie to prowadziło do nieporozumień.

Mallon, Machery, Nichols i Stich postanowili zbadać, czy rzeczywiście intuicje Kripkego są tak powszechne, jak się przeważnie przyjmuje. Do sprawdzenia intuicji zwykłych ludzi, do których odwołuje się Kripke, użyli ankiet opartych o przypadek Gödla oraz przypadek Jonasza, który także pojawia się w *Nazywaniu i konieczności*. W obu eksperymentach myślowych przedmiot nie spełnia cech kojarzonych z nazwą, z którą został związany przez chrzest językowy. Według Kripkego szeroko podzielana intuicja stwierdza, że nazwa wskazuje na pierwotnego referenta, kiedy teoria deskrypcji wskazuje w przypadku Gödla na inny przedmiot (który lepiej spełnia skojarzone cechy), a w przypadku Jonasza orzeknie brak referenta.

Ankieta zawierała cztery historie, dwie analogiczne do przypadku Gödla i dwie analogiczne do przypadku Jonasza. Jedna z nich była nieznacznie tylko zmienioną opowieścią Kripkego:

Przypuśćmy, że w collegu Jan dowiedział się, iż Gödel to człowiek, który dowiódł ważnego twierdzenia matematycznego, nazywanego twierdzeniem o niezupełności arytmetyki. Jan jest całkiem dobry z matematyki i potrafi podać dowód tego twierdzenia. Ale to, że Gödel jest jego autorem, jest jedyną rzeczą, którą słyszał o Gödlu.

A teraz przypuśćmy, że Gödel wcale nie był autorem tego twierdzenia. Człowiek o nazwisku Schmidt, którego ciało odnaleziono wiele lat temu w Wiedniu w tajemniczych okolicznościach, w rzeczywistości dowiódł twierdzenia o niezupełności. Jego przyjaciel Gödel w jakiś sposób zdobył rękopis z dowodem i ogłosił go jako własny. Stąd znany jest jako człowiek, który dowiódł niezupełności arytmetyki. Większość osób, które słyszały nazwisko *Gödel*, jest jak Jan – to, że Gödel odkrył twierdzenie o niezupełności, jest jedyną rzeczą, którą kiedykolwiek słyszeli o Gödlu. Kiedy Jan używa nazwiska *Gödel*, mówi:

(A) o osobie, która naprawdę odkryła niezupełność arytmetyki?

(B) o osobie, która weszła w posiadanie rękopisu i przypisała sobie autorstwo twierdzenia⁶⁴

Ankiety wypełniło 31 studentów z Rutgers University (Stany Zjednoczone) i 41 z Uniwersytetu w Hongkongu (Chiny).⁶⁵ Wybór uczelni nie był przypadkowy. Rezultaty badań psychologicznych, w szczególności wyniki zespołu Richarda Nisbetta, wskazują na istnienie różnic w podstawowych procesach poznawczych między ludźmi wychowanymi w kulturze chińskiej a wychowanymi w kulturze zachodniej.⁶⁶ Różnice te układają się w ogólny wzorzec. Mieszkańcy Azji Wschodniej mają tendencję do rozpatrywania przedmiotów w zależności od ich powiązań z kontekstem, w którym są osadzone, i wyjaśniania ich zachowania w świetle tych powiązań. Dla mieszkańców Europy i Ameryki liczą się przede wszystkim samodzielne cechy przedmiotów, służące potem do przypisywania przedmiotom kategorii i wyjaśniania ich zachowania przez reguły powiązane z przypisaną kategorią.⁶⁷ W szczególności istotną rolę pełnią w tej kulturze wyjaśnienia przyczynowo-skutkowe. Pierwszy tryb myślenia Nisbett określa mianem holistycznego, drugi – analitycznego.

Zespół Sticha spodziewał się odnaleźć odbicie tych trybów myślenia w intuicjach semiotycznych. Postrzeganie holistyczne miałoby być powiązane z intuicjami deskryptywistycznymi ze względu na rolę kontekstu w tym sposobie poznania, a postrzeganie analityczne z intuicjami kauzalno-historycznymi ze względu na rolę powiązań przyczynowo-skutkowych. Tak więc wyprowadzona z modelu Nisbetta hipoteza brzmiała: studenci z New Jersey będą mieli silniejsze intuicje kauzalno-historyczne niż studenci z Hongkongu.

Każde z czterech pytań ankiety było punktowane: 0 za odpowiedź zgodną z teorią deskryptywną (w przypadku cytowanej wyżej ankiety

⁶⁴ E. Machery, R. Mallon, S. Nichols, S. Stich, *Semantics, cross-cultural style*, s. B6.

⁶⁵ W istocie ankietowanych było nieco więcej, ale niektóre ankiety zostały odrzucone ze względu na przynależność kulturową ankietowanych niespełniającą warunków eksperymentu.

⁶⁶ R. Nisbett, K. Peng, I. Choi, A. Norenzayan, *Culture and systems of thought: Holistic vs. analytic cognition*.

⁶⁷ R. Nisbett, K. Peng, I. Choi, A. Norenzayan, *Culture and systems of thought: Holistic vs. analytic cognition*, s. 291.

to odpowiedź A) i 1 za odpowiedź zgodną z teorią kauzalno-historyczną (B). Studenci odpowiadali następująco:

Tabela 1. Wyniki ankiet w eksperymencie dotyczącym intuicji semiotycznych pytanie:

przypadek Gödla:	
uczestnicy ze Stanów Zjednoczonych	1,13 (0,88)
uczestnicy z Chin	0,63 (0,84)
przypadek Jonasza:	
uczestnicy ze Stanów Zjednoczonych	1,23 (0,96)
uczestnicy z Chin	1,32 (0,76)

Wyjaśnienia: μ oznacza średnią z odpowiedzi, σ odchylenie standardowe. Maksymalna możliwa wartość μ to 2 (dwie odpowiedzi kauzalno-historyczne), a minimalna to 0 (dwie odpowiedzi deskrytywistyczne).

Źródło: E. Machery, R. Mallon, S. Nichols, S. Stich, *Semantics, cross-cultural style*.

Testy statystyczne wykazały istotną różnicę w odpowiedziach Chińczyków i Amerykanów dla przypadku Gödla i brak takiej różnicy dla przypadku Jonasza.⁶⁸ Dla tego drugiego wyniku badacze proponują dwa wyjaśnienia. Albo problem stanowiły zbyt długie opisy (przypadek Jonasza wymaga bardziej skomplikowanej historyjki), albo – co byłoby ciekawsze – badani mają opory z uznaniem, że nazwa może nie mieć referenta.

Niezależnie od braku różnicy w odpowiedziach dla przypadku Jonasza, filozoficznie istotne są dwa fakty. Po pierwsze, w eksperymencie myślowym z Gödlem i Schmidtem intuicje Chińczyków bardzo różnią się od intuicji Amerykanów, co autorzy tłumaczą właśnie modelem zaproponowanym przez Nisbetta. Studenci z Hongkongu wykazali silne intuicje deskrytywne, potwierdzając wysuniętą hipotezę. Po drugie, nawet intuicje Amerykanów nie potwierdzają teorii Kripkego, ponieważ średnia punktów plasuje się w połowie skali, a odchylenie standardowe wskazuje na zróżnicowanie odpowiedzi. Konsekwencje dla teorii referencji są jasne. Teoria kazualno-historyczna nie ujmuje zachowania językowego Chińczyków, a tym samym nie jest uniwersalna. Argument z intuicji –

⁶⁸ Wykorzystano test t-Studenta o hipotezie alternatywnej stwierdzającej nierówność średnich punktów. Dla przypadku Gödla *p-wartość* < 0,05.

jedyny, którego Kripke używa na poparcie swojego pomysłu – upada, przynajmniej w swojej dotychczasowej formie. Budowanie teorii referencji, w której kultura wspólnoty językowej nie gra roli, nie jest możliwe na bazie intuicji semiotycznych członków tej wspólnoty, skoro różne wspólnoty mają różne intuicje. Jeżeli teoria Kripkego (jak każda teoria referencji, której uargumentowanie sprowadza się do odwołania do gabinetowych intuicji) miałaby się utrzymać, jej obrońcy musieliby uzasadnić, dlaczego ich intuicje semiotyczne są trafniejsze niż intuicje, na przykład, mieszkańców Chin.

Jeżeli takie uzasadnienie nie jest możliwe, wyjściem wydaje się program badawczy analogiczny do lingwistycznego projektu Noama Chomskiego. Jego celem byłoby opisanie *ukrytej teorii* [*implicit theory*] odpowiedzialnej za używanie nazw przez użytkowników języka. Oczywiście, różni użytkownicy języka mogą mieć różne teorie, w szczególności przynależność do różnych kultur będzie miała tu znaczenie. Dlatego metodologia tego programu musiałaby być empiryczna i wykorzystywać badanie intuicji językowych w tych wspólnotach komunikacyjnych, co do których istnieje uzasadnione podejrzenie, że nazwy mogą funkcjonować w nich inaczej niż we wspólnocie akademickiej filozofów analitycznych.

Jednak nawet ten program jest wątpliwy. W artykule *Against Arguments From Reference* ci sami autorzy wysuwają roboczy wniosek, częściowo potwierdzony przez wyniki eksperymentu: nikt nie wie, jaka teoria referencji opisuje jego zachowanie językowe. Przede wszystkim, intuicje semiotyczne różnią się nawet wewnątrz tak homogenicznych grup jak studenci jednego uniwersytetu, o podobnym statusie socjo-ekonomicznym i zapleczu kulturowym. Dlatego nawet przynależność kulturowa nie jest wystarczającą przesłanką za określeniem, w jaki sposób dana osoba używa nazw własnych, a więc jaką ukrytą teorią dysponuje. Każdego należałoby rozpatrywać osobno. Co więcej, dane pokazują, że ta sama jednostka może mieć w jednym przypadku intuicje deskryptywistyczne, a w innym kauzalno-historyczne, co opisująca jej zachowanie językowe teoria referencji powinna uwzględnić. To bardzo komplikuje program inspirowany lingwistyką Chomskiego. Dalej, nawet otwarcie sformułowanym przez filozofów twierdzeniom o preferowanej przez nich

samych teorii referencji niekoniecznie należy wierzyć. Do czasu powstania *Nazywania i konieczności* filozofowie przyjmowali przeważnie taką lub inną wersję teorii deskryptywnej, a po lekturze tej książki w większości przyznali, że w istocie ich zachowanie językowe lepiej opisuje teoria kauzalno-historyczna.

Przyjęcie wniosku, że nikt nie wie, jaka teoria referencji opisuje jego zachowanie językowe, ma konsekwencje wszędzie tam, gdzie używane są argumenty z oznaczania [*arguments from reference*]. Argumenty z oznaczania mają wspólną strukturę. Najpierw w odniesieniu do pewnego terminu lub klasy terminów filozof przyjmuje określoną teorię referencji. Następnie stwierdza, że referent terminu (terminów) ma, zyskuje lub traci pewne cechy, przez co relacja oznaczania zachodzi, przestaje zachodzić lub termin zaczyna wskazywać na inny przedmiot. Na tej podstawie wyciąga wnioski na temat wiedzy o referencie lub nawet na temat jego istnienia.

W filozofii umysłu taką strategię stosują na przykład eliminatywiści [*eliminativists*]. Eliminatywizm stwierdza, że postawy [*propositional attitudes*] takie jak przekonania lub pragnienia, których ludzie używają do wyjaśniania zachowania, nie istnieją. Uzasadnienie jest następujące. Terminy z języka codziennego, opisujące stany umysłowe, są zdefiniowane przez swoją rolę w potocznej psychologii. Jeżeli ich referenty istnieją, muszą spełniać te definicje. Skoro w świetle obecnej psychologii i kognitywistyki potoczna psychologia jest fałszywa, nic nie spełnia tych definicji, a tym samym referenty terminów potocznej psychologii nie istnieją.

W filozofii nauki Feyerabend i Kuhn używali argumentu z oznaczania, aby wykazać, że nie ma (absolutnego) postępu naukowego. Podobnie jak w poprzednim wypadku, znaczenie terminów – w tym wypadku naukowych – jest wyznaczone przez ich rolę w teoriach naukowych. Podczas rewolucji znaczenie terminów zmienia się tak bardzo, że zaczynają wskazywać na inne referenty. Nauka przed i po rewolucji opisuje coś zupełnie innego – porównanie starej i nowej teorii nie jest możliwe, a tym samym nie można stwierdzić, czy zaszedł jakiś postęp.

W obu przykładach argumentacja zakłada deskryptywną teorię znaczenia. Jeżeli przyjąć teorię kauzalno-historyczną w odniesieniu do terminów

psychologii potocznej, ich referenty nie muszą spełniać konotowanej definicji. Tym samym nie można wyprowadzić wniosku, że nie istnieją, nawet jeżeli nic nie podpada pod skojarzone z terminami definicje. Analogicznie, kauzalno-historyczna teoria znaczenia zapewnia stałość relacji oznaczenia mimo rewolucji naukowych. Rewolucje zmieniają znaczenie terminów, ale ich referenty pozostają te same, co daje grunt dla porównywania ze sobą teorii naukowych.

Którą z teorii referencji – deskryptywną czy kauzalno-historyczną – należy przyjąć w odniesieniu do tych dwóch problemów filozoficznych? Stich, Mallon, Machery i Nichols twierdzą, że nie można odpowiedzieć na to pytanie, bo jedyne narzędzie, używane dotąd w rozstrzyganiu tego typu kwestii – intuicja semiotyczna towarzysząca eksperymentom myślowym – okazało się zawodne. Wobec tego filozofowie powinni w ogóle zrezygnować z argumentów z oznaczania i postarać się o inne uzasadnienia dla swoich propozycji.

badania nad pojęciem wiedzy

W artykule *Normativity and Epistemic Intuitions* Weinberg, Nichols i Stich opisują wyniki eksperymentu dotyczącego intuicji epistemologicznych.⁶⁹ Argumentacja ma strukturę podobną do tej z artykułu dotyczącego teorii referencji; także niektóre hipotezy empiryczne autorzy wysunęli na podstawie modelu Nisbetta. Badanie Weinberga, Nicholasa i Sticha jest pierwszym eksperymentem filozoficznym i to od niego można liczyć początek filozofii eksperymentalnej.

Autorzy poddają krytyce grupę strategii, nazwaną przez nich *romantyzmem opartym o intuicje* [*intuition-driven romanticism*] (w skrócie: ROI), których celem jest dostarczenie norm epistemicznych – reguł postępowania mówiących, jak należy formułować i weryfikować sądy. Nazwa wzięła się stąd, że według ROI właściwe normy epistemiczne są ukryte gdzieś głęboko w umyśle i epistemolodzy mogą je wydobyć, analizując swoje intuicje,

⁶⁹ J. Weinberg, S. Nichols, S. Stich, *Normativity and Epistemic Intuitions*. *Epistemic intuition* tłumaczę jako *intuicja epistemologiczna*, ponieważ termin *intuicja epistemiczna* znaczyłby raczej *intuicja poznawcza* (a więc umożliwiająca poznanie), kiedy autorzy mają na myśli intuicje służącą do kwalifikacji przekonania jako wiedzy.

podobnie jak dziewiętnastowieczni romantycy szukali wiedzy o człowieku i świecie, zagłębiając się we własne przeżycia psychiczne.

Strategie ROI mają wspólny schemat. Na wejściu ROI przyjmuje intuicje epistemologiczne, być może także obok danych innego rodzaju (na przykład pewnych wyników nauk empirycznych), a na wyjściu ROI podaje twierdzenia (otwarcie lub skrycie) normatywne. Wyjście zależy od podanych na wejściu intuicji epistemologicznych, a więc dostarczenie istotnie różnych intuicji zwróci istotnie różne normy. Przeważnie wyjście ROI jest wyprowadzone z wejścia w procesie osiągania równowagi refleksyjnej. Przykłady strategii podam przy omawianiu poszczególnych ankiet.

Weinberg, Nichols i Stich wysuwają dwie hipotezy:

1. intuicje epistemologiczne są różne w różnych kulturach,
2. intuicje epistemologiczne są różne w różnych grupach socjo-ekonomicznych.⁷⁰

Autorzy twierdzą, że jeżeli te hipotezy są prawdziwe, bardzo trudno będzie uzasadnić obrońcom strategii ROI obowiązywalność norm przez nie generowanych. W każdym razie na pewno ciężar dowodu spocznie na filozofach, którzy uważają ROI za poprawną metodę pozyskiwania reguł epistemicznych.

Pierwszy eksperyment dotyczy toczonej w epistemologii debaty eksternalizm-internalizm. Internalizm w odniesieniu do wiedzy głosi, że to, czy przekonania są wiedzą, zależy wyłącznie od czynników introspektywnie dostępnych posiadaczowi tych przekonań. Eksternalizm to pogląd przeciwny. Przykładowo, jeżeli prawdziwe przekonanie zostało uzyskanie dzięki władzy poznawczej godnej zaufania, internalista uzna to przekonanie za wiedzę pod warunkiem, że podmiot zdaje sobie sprawę, iż może polegać na tej władzy poznawczej. Eksternaliście do uznania

⁷⁰ W istocie w artykule autorzy proponują jeszcze dwie hipotezy – *intuicje epistemiczne zależą od kolejności prezentowania przypadków poddanych intuicyjnej ocenie oraz intuicje epistemiczne danej osoby zależą od liczby kursów filozoficznych, na które uczęszczała* – ale nie weryfikują ich w eksperymentach, które przytaczają.

przekonania za wiedzę wystarczy w tym wypadku niezawodność władzy poznawczej, nawet jeżeli podmiot nie jest tej niezawodności świadomy.

Badacze zaprojektowali następujące ankiety:

Pewnego dnia spadający kamień uderza Karola, powodując zmiany w jego mózgu – od tego czasu Karol zawsze potrafi trafnie oszacować temperaturę otoczenia. Karol nie jest świadomy zmiany, jaka zaszła w jego mózgu. Kilka tygodni później zmiana w mózgu Karola powoduje, że Karol uważa, iż w jego pokoju temperatura wynosi 22°C. Oprócz tego nie ma żadnych innych powodów uważać, że temperatura wynosi 22°C. I faktycznie, temperatura w jego pokoju wynosi 22°C. Czy Karol wie, że temperatura wynosi 22°C, czy tylko w to wierzy?

naprawdę wie

tylko wierzy⁷¹

Karol uzyskuje swoje przekonanie dzięki niezawodnemu mechanizmowi, ale nie jest świadomy tej niezawodności. Zatem odpowiedź *naprawdę wie* jest eksternalistyczna, a odpowiedź *tylko wierzy* internalistyczna.

Druga ankieta wykorzystywała podobną historię. Tym razem jednak cała społeczność zyskała niezawodny mechanizm, którego istnienia jednak nikt z jej członków nie jest świadomy.

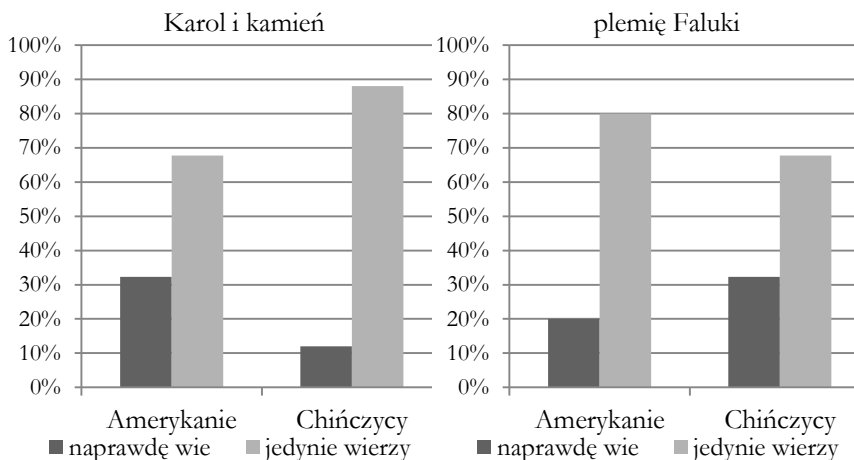
Faluki są dużą, ale zżytą wspólnotą zamieszkującą odległą wyspę. Pewnego dnia w wyspę uderzył radioaktywny meteoryt, zmieniając strukturę chemiczną mózgow członków plemienia – od tego czasu zawsze potrafią trafnie oszacować temperaturę otoczenia. Faluki nie są świadomi, że w ich mózgach zaszła taka zmiana. Kal należy do plemienia Faluki. Kilka tygodni po uderzeniu meteorytu Kal przechadza się plażą. Zmiany w jego mózgu powodują, że uważa, iż temperatura otoczenia to 22°C. Oprócz tego nie ma żadnych innych powodów uważać, że temperatura wynosi 22°C. I faktycznie, temperatura otoczenia wynosi w tym czasie 22°C. Czy Kal wie, że temperatura wynosi 22°C, czy tylko w to wierzy?

naprawdę wie

tylko wierzy⁷²

⁷¹ J. Weinberg, S. Nichols, S. Stich, *Normativity and Epistemic Intuitions*, s. 443.

Ponownie, odpowiedź *naprawdę wie* jest zgodna z eksternalizmem, a *tylko wierzy* z internalizmem.



Rysunek 1. Karol i kamień, plemię Faluki.

Źródło: Weinberg J., Nichols S., Stich S., *Normativity and Epistemic Intuitions*.

Ankiety wypełniali studenci Rutgers University. Próbę dobrano w podobny sposób jak później w badaniu intuicji semiotycznych, a więc wybrano uczestników z zachodniego i wschodnioazjatyckiego (chińskiego) kręgu kulturowego.

W obu ankietach obie grupy wykazywały przeważnie intuicje internalistyczne. Jednak w przypadku Karola i kamienia różnica w intuicjach między grupami była statystycznie istotna: studenci o kulturze zachodniej chętniej przyznawali przekonaniom Karola status wiedzy niż studenci z kultury wschodniej.⁷³ Drugą ankietę, z historią o plemieniu, autorzy skonstruowali kierując się sformułowaną przez Nisbetta opozycją myślenie holistyczne-analityczne, w szczególności opozycją społeczność-jednostka. Spodziewali się znieść różnicę w odpowiedziach dzięki przyznaniu zdolności precyzyjnego określania temperatury nie pojedynczej osobie, ale całej społeczności. I istotnie, tym razem między odpowiedziami obu grup nie wystąpiła istotna statystycznie różnica.⁷⁴

⁷² J. Weinberg, S. Nichols, S. Stich, *Normativity and Epistemic Intuitions*, s. 445.

⁷³ Wykorzystano test Fishera, *p-wartość* = 0,0201.

⁷⁴ test Fishera, *p-wartość* = 0,2527.

W drugim eksperymencie autorzy postanowili przebadac, czy różnice kulturowe także znajduj odzwierciedlenie w intuicjach dotyczcych przypadkw Gettiera. Przypadki Gettiera to sytuacje, w ktrych osoba ma prawdziwe i dobrze uzasadnione przekonanie, przy czym samo uzasadnienie jest fałszywe lub jedynie przypadkiem prawdziwe (w pierwszym rozdziale przytoczyłem oryginalnie podany przez Gettiera przykad). W ankiecie autorzy przedstawili nastpujc historyjk:

Bob ma przyjacilk Jill, ktra jeździła Buickiem wiele lat. Dlatego Bob uważa, że Jill jeździ amerykańskim samochodem. Nie zdaje sobie sprawy, że jej Buick został ostatnio skradziony i że Jill kupiła Pontiac, ktry także jest amerykańskim samochodem. Czy Bob naprawd wie, czy tylko wierzy, że Jill jeździ amerykańskim samochodem?

naprawd wie

tylko wierzy⁷⁵

Hipotez autorzy ponownie wyprowadzili z modelu Nisbetta, według ktrego ludzie z zachodniej kultury czściej odwołuj si do zwizkw przyczynowo-skutkowych w opisywaniu świata, kiedy wychowani w kulturze wschodnioazjatyckiej czściej odwołuj si do podobieństwa. W przytoczonym przypadku Gettiera przekonanie uzasadniajce (prawdziwy) sd Boba, iż Jill jeździ amerykańskim samochodem, okazuje si fałszywe. Dlatego autorzy spodziewali si, że studenci wychowani w kulturze zachodniej czściej niź studenci ze wschodnioazjatyckiego krgu kulturowego bd twierdzić, że przekonania Boba nie s wiedz.

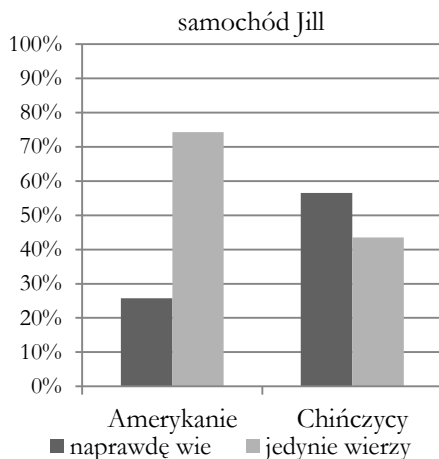
Badania potwierdziły hipotez: 26% ankietowanych z krgu kultury zachodniej uznało przekonania Boba za wiedz, kiedy w przypadku kultury wschodnioazjatyckiej tak intuicj przejawiało 56% badanych.⁷⁶

Rżnice w intuicjach osb o rżnym statusie socio-ekonomicznym autorzy znaleźli dla przypadkw Gettiera, w ktrych uzasadnienie sdu jest prawdziwe, ale rwnie dobrze mogłoby być fałszywe. To znaczy, przyznanie statusu wiedzy przekonaniom osb opisywanych w ankiecie

⁷⁵ J. Weinberg, S. Nichols, S. Stich, *Normativity and Epistemic Intuitions*, s. 449.

⁷⁶ Test Fishera, *p-wartość* = 0,006.

zależy od oceny alternatywnych – możliwych, choć niezachodzących – sytuacji.



Rysunek 2. Amerykański samochód Jill.

Źródło: Weinberg J., Nichols S., Stich S., *Normativity and Epistemic Intuitions*.

Pierwsza ankieta opisywała Pata, który wybrał się na wycieczkę do zoo:

Pat poszedł ze swoim synem do zoo. Kiedy podchodzą do wybiegu zebry, Pat wskazuje na zwierzę i mówi: *To zebra*. Ma rację — to jest zebra. Jednak z odległości, z jakiej oglądają zwierzę, Pat nie mógłby odróżnić prawdziwej zebry od muła ucharakteryzowanego na zebkę. I gdyby oglądane przez nich zwierzę było sprytnie przebranym za zebkę mułem, Pat dalej uważałby, że patrzy na zebkę. Czy Pat naprawdę wie, czy tylko wierzy, że oglądane przez niego zwierzę to zebra?

naprawdę wie

tylko wierzy⁷⁷

W drugiej ankiecie badani mieli ocenić, czy Jim wie, że przyjmowanie nikotyny nie zwiększa prawdopodobieństwa zachorowania, nawet jeżeli dowody na to mogły zostać sfabrykowane.

To jasne, że palenie papierosów zwiększa ryzyko zachorowania na raka. Jednak istnieją badania dowodzące, że samo zażywanie nikotyny bez palenia (np. w postaci tabletek) nie zwiększa ryzyka

⁷⁷ J. Weinberg, S. Nichols, S. Stich, *Normativity and Epistemic Intuitions*, s. 451.

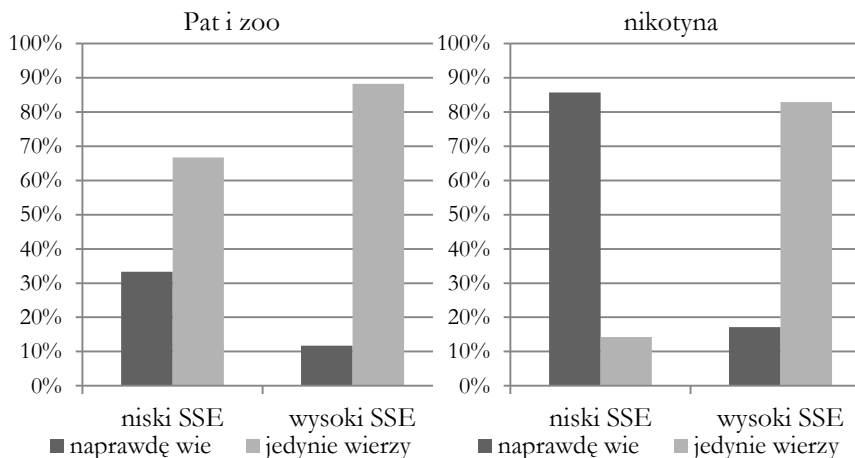
zachorowania na raka. Jim wie o tych badaniach, przez co wierzy, że zażywanie nikotyny nie zwiększa ryzyka zachorowania na raka. Jest możliwe, że właściciele koncernów tytoniowych nieuczciwie zmyślili i opublikowały dowody na nieszkodliwość nikotyny i że te dowody są w istocie fałszywe i mylące.

W istocie właściciele koncernów tytoniowych nie zmyślili tych dowodów, ale Jim nie jest tego świadomy. Czy Jim naprawdę wie, czy tylko wierzy, że zażywanie nikotyny nie zwiększa ryzyka zachorowania na raka?

naprawdę wie

tylko wierzy⁷⁸

Zgodnie z uznaną praktyką w badaniach z psychologii społecznej autorzy przyjęli, że ankietowany ma wysoki status socjo-ekonomiczny, jeżeli uczęszczał do college'u przez co najmniej rok; w przeciwnym przypadku ma niski status. O wypełnienie ankiet poproszono przechodniów z głównych ulic miasta New Brunswick, oferując im jako wynagrodzenie bony do restauracji McDonald's.



Rysunek 3. Pat i zoo, nikotyna.

Źródło: Weinberg J., Nichols S., Stich S., *Normativity and Epistemic Intuitions*.

W obu ankietach badani o wysokim statusie socjo-ekonomicznym istotnie mniej chętnie określali przekonania mianem wiedzy, przy czym

⁷⁸ J. Weinberg, S. Nichols, S. Stich, *Normativity and Epistemic Intuitions*, s. 452.

w przypadku ankiety dotyczącej firm tytoniowych różnica była ogromna: stosunek odpowiedzi został odwrócony.

Autorzy przyznają, że nie dysponują żadnym modelem psychologicznym lub socjologicznym, który wyjaśniałby otrzymane wyniki. Proponują dwie hipotezy. Pierwsza opiera się na obserwacji, że w przytoczonych historiach istotną rolę grają niezachodzące, ale możliwe sytuacje. W swoich ocenach osoby o wysokim statusie socjo-ekonomicznym miałyby brać pod uwagę kontryfaktualne sytuacje chętniej niż osoby o niskim statusie. Druga hipoteza mówi, że osoby o niskim statusie socjo-ekonomicznym częściej przyznają przekonaniom status wiedzy, ponieważ ich minimalne wymagania względem wiedzy są niższe.

Tak naprawdę w obu badaniach wyjaśnienia różnic w intuicjach epistemologicznych pełnią jedynie pomocniczą rolę. Filozoficzne konsekwencje ma sam fakt występowania istotnych różnic w intuicjach między kulturami lub grupami o różnym statusie socjo-ekonomicznym. Dostarczenie wyjaśnień służy co najwyżej uwiarygodnieniu stwierdzenia, że przyczyną zaobserwowanych rozbieżności w intuicjach są właśnie różnice w kulturze lub statusie socjo-ekonomicznym. Różne intuicje przekładają się na różne normy lub reguły epistemiczne otrzymane przez użycie ROI. Istotność wyników polega na tym, że nie są to grupy, dla których łatwo sformułować teorię błędu. Jak z oczywistych przyczyn epistemolog nie będzie martwił się, że jego teorii przeczą intuicje epistemologiczne, powiedzmy, pięcioletnich dzieci, tak trudno mu będzie w ten sam sposób postąpić z intuicjami Chińczyków, którzy nie tylko stanowią prawie jedną piątą populacji świata, ale przede wszystkim są członkami dojrzałej i zaawansowanej technicznie kultury. Jednym słowem, dopóki nie pojawi się przekonująca teoria błędu, strategie ROI nie powinny służyć za narzędzie formułowania norm epistemicznych.

działanie intencjonalne

Eksperyment Joshuy Knobe'a z roku 2003 zapoczątkował w filozofii eksperymentalnej dyskusję nad pojęciem działania intencjonalnego

[*intentional*].⁷⁹ Knobe wysunął tezę, że na to, czy działanie jest postrzegane jako intencjonalne, ma wpływ ocena moralna: ludzie traktują pewne akty moralnie złe jako intencjonalne, zaś analogiczne akty moralnie dobre jako nieintencjonalne. Co ważne, wpływ czynników normatywnych na przypisanie aktowi deskryptywnej cechy (bo intencjonalność jest cechą deskryptywną) należy, według Knobe'a, do kompetencji leżącej u podstaw tego przypisania. Oznacza to, że źródłem tego efektu nie są względy pragmatyczne ani błąd wykonawczy [*performance error*].

W kolejnych publikacjach Knobe uogólnił swoją hipotezę. W artykule *Person as Scientist, Person as Moralist* skrytykował metaforę umysłu jako uniwersytetu, zgodnie z którą umysł funkcjonuje podobnie jak uniwersytet podzielony na odrębne wydziały:

Istnieją procesy psychologiczne poświęcone kwestiom religijnym (wydział teologiczny umysłu), estetyce (wydział sztuk pięknych), moralności (wydział filozoficzny)... i są też procesy poświęcone wyłącznie pytaniom o charakterze naukowym. Te ostatnie działają odmiennie od tych wykorzystywanych w myśleniu nad pytaniami dotyczącymi, powiedzmy, moralności lub estetyki.⁸⁰

Według Knobe'a ta metafora jest nietrafna. Na podstawie wyników własnych eksperymentów oraz zainspirowanych nimi cudzych badań wysunął tezę, że nie ma wyraźnej granicy między procesami odpowiedzialnymi za moralność a procesami, które można by określić mianem naukowych. Jednym słowem, kompetencje umysłowe odpowiedzialne za dziedziny deskryptywne w istocie uwzględniają czynniki normatywne.

Pierwszy, klasyczny już eksperyment opierał się na dwóch ankietach:

⁷⁹ J. Knobe, *Intentional Action and Side Effects in Ordinary Language*. Tłumaczę termin *intentional* jako *intencjonalny/intencjonalnie*. Są dobre odpowiedniki tego terminu w języku polskim, m.in. *umysłny, zamierzony*. Jednak zdecydowałem się na tłumaczenie najbliższe oryginalnemu brzmieniu z jednego względu: wyniki eksperymentów były wrażliwe na zmianę słów, nawet jeżeli wykorzystano terminy uważane za synonimy (np. *intended* – zamierzony). Fenomenologiczny termin *intencjonalność* nie ma prawie nic wspólnego z używanym tutaj terminem *intencjonalność* z analitycznej filozofii działania i nie należy ich ze sobą wiązać.

⁸⁰ J. Knobe, *Person as Scientist, Person as Moralist*, s. 315.

Wiceprezes pewnej firmy spotkał się z prezesem i zapytał: *Mamy możliwość rozpoczęcia nowego projektu. Projekt zwiększy zyski, a także pomoże środowisku naturalnemu. Czy mamy wdrożyć projekt?* Prezes odpowiedział: *Nie obchodzi mnie, czy projekt pomoże środowisku. Chcę tylko zwiększyć zyski. Rozpocznijcie wdrażanie projektu.*

Projekt został wdrożony i, zgodnie z przewidywaniami, pomógł środowisku.⁸¹

W drugiej ankiecie projekt zamiast szkodzić, pomagał środowisku:

Wiceprezes pewnej firmy spotkał się z prezesem i zapytał: *Mamy możliwość rozpoczęcia nowego projektu. Projekt zwiększy zyski, a także zaszkodzi środowisku naturalnemu. Czy mamy wdrożyć projekt?* Prezes odpowiedział: *Nie obchodzi mnie, czy projekt zaszkodzi środowisku. Chcę tylko zwiększyć zyski. Rozpocznijcie wdrażanie projektu.*

Projekt został wdrożony i, zgodnie z przewidywaniami, zaszkodził środowisku.⁸²

Ankietowani mieli odpowiedzieć, czy prezes intencjonalnie zaszkodził/pomógł środowisku. W obu historyjkach wpływ na środowisko nie jest celem prezesa, ale efektem ubocznym podjętego przez niego działania. Co ważne, prezes jest świadomy, że opisany efekt wystąpi. Jedyną różnicą w ankietach jest ocena moralna działania – Knobe przyjął, że ankietowani oceniają szkodenie środowisku jako moralnie złe, a pomaganie jako moralnie dobre.

W oryginalnym badaniu ankietę wypełniło 78 spacerowiczów z jednego z nowojorskich parków, którzy otrzymali losowo jedną z dwóch wersji ankiety. 82% wypełniających ankietę o prezesie szkodzącym środowisku uznało, że zaszkodził środowisku intencjonalnie. 33% wypełniających ankietę o prezesie pomagającym środowisku odpowiedziało, że pomógł środowisku intencjonalnie. Stosunek odpowiedzi w dwóch wersjach ankiety został niemal odwrócony; oczywiście różnica była statystycznie istotna. Wyniki pozwoliły Knobe'owi sformułować wniosek, że użycie pojęcia intencjonalności działania zależy, między innymi, od oceny moralnej tego działania.

⁸¹ J. Knobe, *Intentional Action and Side Effects in Ordinary Language*, s. 192.

⁸² J. Knobe, *Intentional Action and Side Effects in Ordinary Language*, s. 192.

Potocznej psychologii przypisuje się rolę wyjaśniania i przewidywania zachowania, a więc analogiczną do psychologii naukowej. Można rozpoznać tę funkcję w takich zdaniach jak *wyszredł z kina, bo nie lubi filmów akcji* albo *zaczęła studia, bo chce więcej zarabiać*. Konsekwencją wniosku Knobe'a jest zaprzeczenie temu przekonaniu – pojęcie intencjonalności, z pozoru czysto deskryptywne, okazuje się wrażliwe na czynniki normatywne. Knobe nie twierdzi, że potoczna psychologia nie może pełnić funkcji predykcyjnych lub eksplanacyjnych, ale że odpowiadające za nią kompetencje nie ograniczają się wyłącznie do tych funkcji.

Pojawiły się inne wyjaśnienia efektu Knobe'a – bo taką nazwę zyskał w literaturze wpływ czynników normatywnych na stosowanie pojęcia intencjonalności – w których kompetencja leżąca u podstaw pojęcia działania intencjonalnego jest niewrażliwa na czynniki normatywne, natomiast efekt ma źródło w procesach psychologicznych zewnętrznych wobec tej kompetencji. Przede wszystkim chodzi tu o względy pragmatyczne: odpowiadając na pytanie o intencjonalność nieekologicznego działania prezesa, ankietowany chce w istocie powiedzieć, że prezes zasługuje na karę. Odpowiedź *tak, intencjonalnie zaszkodził środowisku* jest traktowana jako sposób orzeczenia winy.

Ten zarzut został podjęty w kolejnych badaniach, ciekawych z metodologicznego punktu widzenia. W jednej z wersji ankiety badani zostali utworzeni przed przedstawieniem właściwego eksperymentu myślowego. Najpierw przeczytali historijkę o pijanym kierowcy, który potracił pieszego, i odpowiedzieli na pytania, czy kierowca (a) intencjonalnie potracił przechodnia (przeważająca odpowiedź – nie), (b) ponosi winę za to zdarzenie (przeważająca odpowiedź – tak). Dzięki tym pytaniom ankietowani uświadomili sobie, że działanie nie musi być intencjonalne, aby można było być za nie odpowiedzialnym. Następnie przedstawiono im historijkę o prezesie zarządu; efekt Knobe'a wystąpił – tak jak w eksperymencie bez torowania.

Eksperyment powtórzono też na czteroletnich dzieciach (z odpowiednio zmodyfikowaną ankietą). Kompetencje językowe dzieci dopiero się rozwijają, dlatego występowanie efektu Knobe'a sugerowałoby, że odpowiedzialny za niego mechanizm psychologiczny jest dla umysłu

podstawowy. Przeprowadzono też badanie nad osobami z zespołem Aspergera. Dotknięci tym zaburzeniem rozumieją słowa w ich najbardziej dosłownym sensie i mają trudności w posługiwaniu się metaforami. Zaobserwowanie efektu Knobe'a w odpowiedziach osób z zespołem Aspergera oznaczałoby, że zależność od czynników normatywnych rzeczywiście należy do kompetencji odpowiedzialnej za użycie pojęcia działania intencjonalnego. Eksperyment powtórzono także dla różnych grup językowych. We wszystkich przypadkach odpowiedzi ułożyły się w ten sam wzorzec, co odpowiedzi spacerowiczów z nowojorskiego parku.

Mimo to dyskusja nadal trwa. Alfred Mele i Fiery Cushman zwracają uwagę, że nadal pozostaje silna mniejszość odpowiedzi, w których badani albo oceniają, że obu przypadkach (proekologicznym i nieekologicznym) mamy do czynienia z działaniem intencjonalnym, albo w obu przypadkach z działaniem nieintencjonalnym.⁸³ Formułują pewne reguły, którym według nich podlegają odpowiedzi ankietowanych, i na tej podstawie wnioskuje, że należy mówić o dwóch różnych pojęciach działania intencjonalnego. Thomas Nadelhoffer wskazuje na praktyczne konsekwencje efektu Knobe'a.⁸⁴ Prawodawcy rozróżniają przeważnie między przestępstwami dokonanymi intencjonalnie i nieintencjonalnie, w pierwszym przypadku przewidując wyższą karę. Jeżeli na uznanie działania za intencjonalne wpływa jego ocena moralna, a z kolei intencjonalność działania jest potem jednym z czynników branych pod uwagę przy ocenie moralnej tego działania, powstaje błędne koło. *Z pewnością nie chcemy, aby samo oskarżenie kogoś o czyn naganny moralnie powodowało, że sędziowie chętniej uznają tę osobę za winną [ze względu na uznanie działania za intencjonalne]. Tak więc wygląda na to, że mimo wszystko mamy do czynienia z błędem wykonawczym.*⁸⁵

⁸³ A. Mele i F. Cushman, *Intentional Action: Two-and-a-Half Folk Concepts?* [w:] J. Knobe, S. Nichols, *Experimental Philosophy*, s. 171-187.

⁸⁴ T. Nadelhoffer, *Bad Acts, Blameworthy Agents, and Intentional Actions: Some Problems for Jury Impartiality*.

⁸⁵ T. Nadelhoffer, *Bad Acts, Blameworthy Agents, and Intentional Actions: Some Problems for Jury Impartiality*, s. 206.

We wrocławskim Kole Filozofii Eksperymentalnej zaprojektowaliśmy i przeprowadziliśmy kilka eksperymentów filozoficznych. W tym rozdziale przedstawiam dwa z nich: eksperyment z filozofii politycznej i z filozofii religii.

sprawiedliwość i zasada maksimum u Rawlsa

Pierwszy eksperyment, który zaprojektowaliśmy, był inspirowany pomysłem Johna Rawlsa. W książce *Teoria Sprawiedliwości* Rawls zaproponował zasady, którym powinien czynić zadość urząd społeczny, aby był sprawiedliwy. Jako jeden ze sposobów uzasadnienia swojej wizji Rawls podaje *argument z zasłony niewiedzy*.

zasłona niewiedzy i zasada dyferencji

Argument z zasłony niewiedzy brzmi następująco. Zgromadźmy grupę osób, której zadaniem będzie zaprojektowanie ustroju. Projektowanie odbędzie się poprzez dyskusję, w której będą ścierały się pomysły dyskutantów, aż do osiągnięcia konsensusu. Dopiero po zaprojektowaniu ustroju każda z osób zajmie w nim losowe miejsce, tak więc podczas projektowania nie będzie znała swojej przyszłej przynależności do klasy socjo-ekonomicznej, swoich talentów i upośledzeń, płci, rasy, orientacji seksualnej, a nawet epoki historycznej. Co więcej, nie będzie znała także swoich przyszłych celów życiowych, awersji do ryzyka, skłonności do optymizmu lub pesymizmu.⁸⁶ Osoby znajdujące się za taką zasłoną niewiedzy (inaczej: osoby *w sytuacji pierwotnej*) będą bezstronne, a zatem

⁸⁶ J. Rawls, *A Theory of Justice, Revised Edition*, s. 118.

ustrój przez nie skonstruowany będzie sprawiedliwy. Pozostaje pytanie, jaki w takim razie ustrój zaprojektują.

Odpowiedź Rawlsa brzmi: byłby to ustrój (a) respektujący podstawowe wolności, w którym (b) nierówności społeczne sprzyjałyby najgorzej uposażonym. Warunek (a) nosi nazwę *pierwszej zasady sprawiedliwości*. Przez podstawowe wolności, do których odwołuje się ta zasada, Rawls rozumie prawa wyborcze, wolność słowa i sumienia, prawa własności i inne, niekontrowersyjne w tradycji liberalizmu, prawa i wolności.

Warunek (b) nosi nazwę *drugiej zasady sprawiedliwości*, albo *zasady dyferencji*. Precyzyjne sformułowanie zasady dyferencji w kategoriach użyteczności mówi, że ze zbioru możliwych ustrojów najlepszy jest ten, który gwarantuje najwyższą użyteczność najsłabszej jednostki (klasy społecznej), czyli tej o najmniejszej użyteczności spośród wszystkich jednostek (klas) w tym społeczeństwie. Zbiór możliwych ustrojów ogranicza się do tych respektujących pierwszą zasadę i w tym sensie pierwsza zasada ma prymat nad drugą. Zasadę dyferencji można wyrazić w języku teorii gier. Jeżeli ustrój zdefiniować jako wektor wypłat (użyteczności) każdej klasy społecznej, zasada dyferencji będzie wynikiem gracza stosującego kryterium maksimum. Zasłona niewiedzy oznacza, że gracz nie zna prawdopodobieństw wypłat (a więc znalezienia się w określonej klasie w danym ustroju).

Ze względu na nieznaną własnych preferencji przez osoby w sytuacji pierwotnej, przy ocenie konkurencyjnych ustrojów Rawls odwołuje się do pojęcia *dóbr pierwotnych*, które są tym, *czego racjonalny człowiek pragnie, czegokolwiek poza nimi jeszcze by nie pragnął*.⁸⁷ Zalicza do nich, między innymi, dochód, majątek, dostępność do stanowisk, szacunek do samego siebie. Zatem dobra pierwotne są środkami, które pozwalają jednostkom na realizację ich planów życiowych

Za zasadą dyferencji stoi także następująca etyczna argumentacja. Nierówności będące konsekwencją talentów, upośledzeń, miejsca i czasu urodzenia i innych, niezależnych od jednostki czynników, są arbitralne. Niemożliwe jest ich wyrównanie, ponieważ nie da się redystrybuować

⁸⁷ J. Rawls, *A Theory of Justice*, s. 78.

wrodzonych talentów i upośledzeń. Ale można dokonać niedoskonałej symulacji takiej redystrybucji – zaprzęgnąć talenty do pracy na rzecz wszystkich, nie tylko ich posiadaczy (lub dzierżawców, jak raczej należałoby nazwać takie osoby w świetle teorii Rawlsa). Zakłada się bowiem, że obdarzeni talentami będą osiągnęli lepsze wyniki niż ludzie ich pozbawieni. Nie można obarczyć osób najlepiej uposażonych za dużymi obciążeniami, ponieważ będzie to demotywujące i, w wyniku redystrybucji, najgorzej uposażeni dostaną mniej, niż by mogli przy mniejszym opodatkowaniu najbogatszych. Nie można zmuszać utalentowanych osób do pracy, ponieważ klóciłoby się to z pierwszą zasadą sprawiedliwości.

metoda. Związki z Teorią Sprawiedliwości

W naszym eksperymencie postanowiliśmy jak najlepiej odtworzyć zasłone niewiedzy, aby sprawdzić, czy uczestnicy rzeczywiście wybiorą to, czego oczekiwali Rawls. Eksperyment przebiega następująco.

Zbieramy grupę osób, którym obiecujemy wypłacić pewną sumę pieniędzy. Suma ta będzie każdemu uczestnikowi przydzielana losowo; uczestnicy nie znają rozkładu losów ani dokładnych kwot możliwych do wylosowania. Przedstawiamy im także przebieg całego eksperymentu.

Przed losowaniem badani mają, poprzez wypełnienie ankiety, wybrać plan redystrybucji tych sum. Do wyboru jest pięć planów, od najbardziej libertariańskiego (1) do najbardziej egalitarnego (5):

Ankieta. Czy chcesz, aby:

- (1) każdy zatrzymał 100% swojej wygranej;
- (2) każdy zatrzymał 75% swojej wygranej,
a 25% całości wygranych zostało rozdzielone po równo;
- (3) każdy zatrzymał 50% swojej wygranej,
a 50% całości wygranych zostało rozdzielone po równo;
- (4) każdy zatrzymał 25% swojej wygranej,
a 75% całości wygranych zostało rozdzielone po równo;
- (5) całość wygranych została rozdzielona po równo między uczestników?

Kiedy uczestnicy wypełnią ankiety, zbieramy je, a oni sami rozpoczynają dyskusję na temat najlepszej procedury podziału. W czasie dyskusji badani

mają za zadanie notować zarówno własne jak i cudze argumenty, które uważają za przekonujące (niezależnie od tego, czy rzeczywiście spowodowały zmianę decyzji). Po dyskusji uczestnicy wypełniają taką samą ankietę jeszcze raz.

Po zebraniu ankiet losujemy jedną, która wyznaczy sposób dystrybucji wygranych, a uczestnicy losują liczby odpowiadające wysokości wygranych. Zdecydowaliśmy się na losowanie ankiety, ponieważ dzięki temu badani odpowiadają zgodnie z własną oceną. Gdyby na przykład redystrybucja była wyznaczona przez dominantę, niektórzy mogliby udzielać odpowiedzi najbliższej im spośród tych, które uważają za najbardziej popularne (ale nie tej, którą uważają za najlepszą, jeżeli postrzegają ją jako niepopularną). Natomiast przekonanie innych uczestników do własnej odpowiedzi powoduje, że prawdopodobieństwo jej wylosowania rośnie. Po wylosowaniu sposobu dystrybucji rozdajemy wygrane w wysokości zależnej od wylosowanej liczby.

Uważam, że w eksperymencie udało się nam dobrze oddać niektóre cechy zasłony niewiedzy. Po pierwsze, dzięki zastosowaniu wypłat pieniężnych, w eksperymencie nie liczą się indywidualne preferencje uczestników co do wykorzystania tych środków, podobnie jak przy podejmowaniu decyzji za zasłoną niewiedzy. Po drugie, nie przekazujemy żadnych informacji na temat rozkładu losów ani wysokości wypłat, przez co uczestnicy nie mogą podejmować decyzji na podstawie prawdopodobieństwa. Faza dyskusji między pierwszym a drugim wypełnieniem ankiety odpowiada dyskusji decydentów za zasłoną niewiedzy. Natomiast do zakończenia eksperymentu nie wymagamy uzgodnienia przez wszystkich uczestników wspólnej odpowiedzi.

Badanie nie spełnia jednego, bardzo ważnego warunku eksperymentu myślowego Rawlsa. W *Teorii Sprawiedliwości* osoby za zasłoną niewiedzy były racjonalne, co oznaczało, między innymi, podejmowanie decyzji ze względu na dążenie do dóbr pierwotnych, a bez względu na szczegółową wizję dobrego życia (a więc na przykład tę implikowaną przez poglądy religijne albo szczegółowe zasady moralne, zaszczerpione w procesie socjalizacji). Oczywiście tego nie dało się zapewnić w naszym eksperymencie i, co więcej, spełnienie tego warunku wydaje się bardzo

trudne, jeżeli nie niemożliwe. Chcąc interpretować argument z zasłony niewiedzy jako hipotezę empiryczną – tak, jak to miało miejsce badanie nad kazualno-historyczną teorią referencji – trzeba przyjąć, że dotyczy ona doskonale racjonalnych jednostek. Ewentualny wynik eksperymentu niezgodny z przewidywaniami *Teorii Sprawiedliwości* nie przeczyłby wnioskowi, że racjonalne (na sposób Rawlowski) jednostki w pozycji pierwotnej będą się kierowały zasadą dyferencji, ponieważ zawsze można powiedzieć, że taki a nie inny wybór uczestników badania był skutkiem ich nieracjonalności.

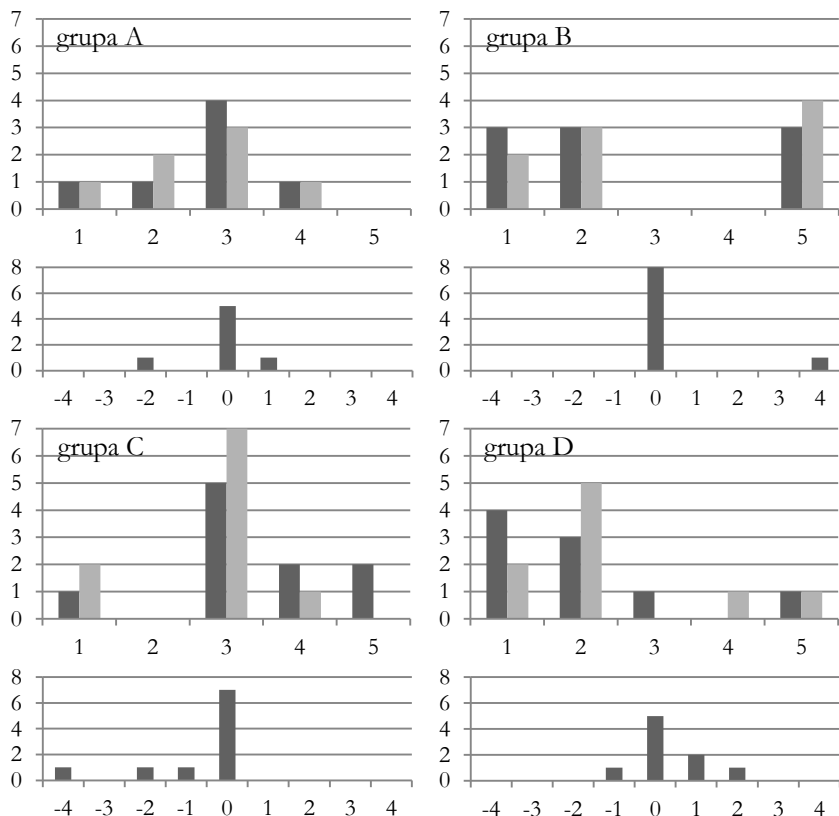
Po co w takim razie w ogóle przeprowadzać taki eksperyment? Po pierwsze, jego wynik jest *jakaś* wskazówką odnośnie rozstrzygnięcia *Teorii Sprawiedliwości* – jeżeli rezultat dyskusji byłby rażąco różny od tego rozstrzygnięcia, należałoby zapytać, dlaczego tak się stało. Podczas dyskusji uczestnicy notują argumenty, co daje podstawę dla odpowiedzi na to pytanie i, ewentualnie, dla prób podzielenia tych argumentów na motywowane wyłącznie istotnymi (racjonalnymi) czynnikami i na takie, które odwołują się do czynników uznanych przez Rawlsa za etycznie nieistotne. Po drugie, nawet jeżeli przyjąć, że osoby w pozycji pierwotnej zaprojektowałyby ustrój zgodnie z zasadą maksimumu, pozostaje pytanie, dlaczego ich wybór miałby obowiązywać rzeczywistych ludzi. Jednym z warunków koniecznych, żeby argument miał taką siłę normatywną, jest założenie, że cała konstrukcja dobrze ujmuje to, co jest ważne w ludziach jako istotach moralnych. I tak na przykład Sen, krytykując *Teorię Sprawiedliwości*, zgłasza wątpliwości właśnie do tego warunku. Według niego analiza nie powinna opierać się o pojęcie dóbr pierwotnych, ale o pojęcie zdolności realizacji celów życiowych.⁸⁸ Jeżeli w eksperymencie, w fazie dyskusji, pojawiłby się jakiś powtarzający się argument nieuwzględniony w *Teorii Sprawiedliwości*, mogłoby to sugerować – chociaż z pewnością nie dowodzić – że stoi za nim pewna, pominięta przez Rawlsa, cecha ludzi jako istot moralnych.

wyniki i interpretacja

Badanie przeprowadziliśmy na czterech grupach, liczących od 7 do 12 osób. W trzech wypadkach były to grupy studentów, jedna grupa była

⁸⁸ A. Sen, *Nierówności. Dalsze rozważania*. tłum. I. Topińska, s. 101.

złożona z ludzi w różnym wieku, przeważnie pracujących zawodowo (grupa B). Wyniki przedstawiłem na rysunku 4.



Rysunek 4. Rozkłady odpowiedzi przed i po fazie dyskusji.

Wyjaśnienia: na wykresach z dwukolorowymi kolumnami przedstawiłem rozkład odpowiedzi przed (ciemny kolor) i po (jasny) fazie dyskusji; odpowiedź 1 jest najbardziej libertariańska, odpowiedź 5 najbardziej egalitarystyczna. Wykresy z jedną kolumną przedstawiają rozkład zmian po fazie dyskusji: wysokość kolumny dla argumentu k odpowiada liczbie uczestników, która w wyniku dyskusji zmieniła numer odpowiedzi o k .

W przypadku ogólnym zasada dyferencji nie implikuje pełnej równości dóbr pierwotnych, ale eksperyment został tak zaprojektowany, że z kryterium maksyminowym zgodna jest najbardziej egalitarystyczna odpowiedź (5).

Najbardziej oczywisty wniosek z eksperymentu brzmi, że badani ani przed dyskusją, ani po niej nie wybierają jednogłośnie odpowiedzi zgodnej z teorią Rawlsa. Co więcej, nie jest to nawet najpopularniejsza odpowiedź; w jednej grupie (A) nie została zaznaczona przez nikogo, w innej grupie (C) po fazie dyskusji dwie osoby zrezygnowały z tej odpowiedzi. Rozkłady odpowiedzi są różne w różnych grupach i trudno doszukać się między nimi podobieństwa.

Drugi ważny wniosek jest taki, że dyskusja zmienia odpowiedzi, ale u niewielu uczestników. Przeważająca liczba osób pozostaje przy swoim wyborze i, co więcej, trudno zauważyć wzorec zmian wspólny wszystkim grupom. W dwóch grupach (A, D) znajdowały się jednocześnie osoby, które w drugiej ankiecie zmieniły odpowiedź na bardziej egalitarystyczną i osoby, które po dyskusji wybrały odpowiedź bardziej libertariańską. Sparowany test t-Studenta dla próby złożonej ze wszystkich grup nie potwierdził istnienia różnicy w odpowiedziach przed i po dyskusji (hipoteza alternatywna jednostronna, p -wartość = 0,1015). P-wartości dla grup A-D badanych osobno, otrzymane w testach Wilcozona dla hipotez jednostronnych, to A: 0,5, B: 0,5, C: 0,0907, D: 0,1724 – czyli tylko w jednej grupie (C) dyskusja istotnie wpłynęła na odpowiedzi.

Natomiast bardzo silna regularność występuje w argumentacji uczestników. We wszystkich czterech grupach już na początku dyskusji pojawiał się argument z arbitralności: *to, co wylosuje, nie zależy ode mnie, nie jest wynikiem moich starań, a tym samym nie zasługuje na większe pieniądze bardziej niż ktokolwiek inny*. Uczestnicy starający się przekonać grupę do wyboru (mniej lub bardziej) egalitarystycznego uważali to uzasadnienie za bardzo silne – na argumenty za bardziej libertariańską dystrybucją odpowiadali *ale przecież to jest losowe, nie ma tu żadnej zasługi*, po prostu odwołując się ponownie do tego samego uzasadnienia. Dla nich argument z arbitralności był ostateczny w tym sensie, że kontrargumenty pozostałych uczestników nie podważały jego zasadności.

Z drugiej strony pojawiał się argument libertariański: *to, co wylosuje, jest moje, więc nikt inny nie ma do tego prawa; nikt nie powinien mnie zmuszać do dzielenia się, jeżeli będę chciał(a), to podzielę się dobrowolnie*. Jest to argument ciekawy o tyle, że implikuje zaskakujące pojęcie własności. Uczestnicy mieli wybrać plan

dystribucji przed losowaniem, a więc zanim znali swoje wypłaty, jednak używający tego argumentu uważali te przyszłe, potencjalne, nieznanne wypłaty za swoją własność, a redystrybucję uważali za naruszenie tej własności. Być może był to częściowo efekt błędu metodologicznego: w ankiecie używaliśmy zdań takich jak każdy zatrzymuje 100% *swojej* wygranej, sugerując, że wygrana najpierw staje się własnością uczestnika, a potem jest redystrybuowana.

Inny argument brzmiał: *gdyby wszystko było pewne, życie byłoby nudne albo ryzyko jest częścią życia*. Ten argument także często pojawiał się w dyskusji i był używany do uzasadnienia jednego z nieskrajnych wyborów (odpowiedzi o numerach 2, 3 lub 4), w którym część wygranej uczestnik zatrzymuje, a częścią się dzieli. Temu argumentowi towarzyszyło dalsze uzasadnienie: *dlatego niech każdy otrzyma jakąś wypłatę, aby nie odszedł z pustymi rękami, ale nie dzielimy się wszystkim, żeby pozostał czynnik ryzyka*.

Czy te wyniki mają jakieś filozoficzne konsekwencje? Po pierwsze, zwolennik argumentu z zasłony niewiedzy powinien podać wytłumaczenie faktu, że dominujący wybór badanych nie był zgodny z zasadą dyferencji. W *Teorii Sprawiedliwości* takie wytłumaczenie jest już obecne: żadne rzeczywiste osoby nie mogą znaleźć się w pozycji pierwotnej, ponieważ nie mogą abstrahować od swoich szczegółowych preferencji i koncepcji dobra, ani nie mogą spełnić wymogów Rawlowskiej racjonalności. Jednak dlaczego wyniki różnią się aż tak bardzo od zasady dyferencji? Czy całą winą można obarczyć etycznie nieistotne czynniki, które wpływają na wybór ankietowanych? Analiza argumentacji pomaga wskazać te argumenty, które wydają się niezależne od takich czynników.

Jednym z nich byłoby odwołanie się do innej koncepcji człowieka – u Rawlsa, ze względu na arbitralność właśnie, człowiek jest dzierżawcą swoich talentów; argumenty przeciwników opcji egalitarystycznej zdradzają, że w ich rozumieniu arbitralność i własność nie wykluczają się. Można otrzymać coś w arbitralny sposób i być tego wyłącznym posiadaczem.

Z drugiej strony na korzyść koncepcji Rawlsa świadczy uniwersalność i siła argumentu z arbitralności. Ponieważ w każdej grupie uczestnicy

przywoływali go przez całą dyskusję, chociaż nikt z nich nie znalazł teorii Rawlsa, wydaje się, że ujmuje on bardzo silne intuicje etyczne.

eksperyment w filozofii religii

W tym podrozdziale przedstawiam wyniki eksperymentu inspirowanego pewną popularną próbą odparcia argumentu ze zła. Argument ze zła głosi, że nie można pogodzić istnienia Boga z istnieniem zła, a w każdym razie z istnieniem tak ogromnej ilości zła, jakiej świat doświadczał i doświadcza. Przeciwko temu zarzutowi teiści wysuwają kontrargument: być może zło jest konieczne, a tym samym Bóg nie mógł go wyeliminować. Celem badania było pokazanie, że teista w tym kontrargumencie wykorzystuje schemat oceniania niezgodny z praktyką językową, a tym samym używa terminów niezgodnie z ich właściwym znaczeniem (tj. znaczeniem, w którym sami użytkownicy chcą tego pojęcia używać). Badanie przeprowadziliśmy na przykładzie pojęcia sprawiedliwości.

Badanie należy do analizy eksperymentalnej – empirycznie sprawdzamy (niekoniecznie uświadomione) warunki, pod którymi ludzie używają pojęcia sprawiedliwości, a następnie staramy się pokazać, że te warunki nie są spełnione w kontrargumencie teisty.

teodycea – uwagi wstępne

Jednym z wielu zarzutów podnoszonych przeciwko wierze w istnienie Boga jest *argument ze zła*. Istnienie Boga – definiowanego jako wszechmogąca i moralnie doskonała osoba, która stworzyła świat – jest nie do pogodzenia z istnieniem zła.⁸⁹ Moralnie doskonała istota dąży do usunięcia wszelkiego zła, które można usunąć. Przyjmuje się, że ograniczenia płynące z logiki nie przeczą boskiej wszechmocy, to znaczy, że Bóg jest wszechmocny, dopóki potrafi zrealizować każdy stan rzeczy, który jest logicznie możliwy. Nie byłby wszechmocny, gdyby istniał jakiś możliwy stan rzeczy, którego Bóg nie mógłby zrealizować.

Próba obrony polega na wskazaniu, że zło należy do sfery konieczności. Tak na przykład postępuje Leibniz w swojej *Teodycei*, twierdząc, że zło jest konieczne dla istnienia wolności, czyli że żaden możliwy świat, w którym

⁸⁹ Z tej definicji korzysta np. J. L. Mackie, *Evil and Omnipotence*.

żyją wolne jednostki, nie jest pozbawiony zła, natomiast świat bez wolnych jednostek jest zawsze gorszy niż świat, który je zawiera. Współcześnie argumentację opartą o podobny schemat przeprowadza Alvin Plantinga w książce *Bóg, wolność, zło*.⁹⁰ Najbardziej ogólnie usprawiedliwienie Boga Plantinga formułuje jako zdanie: *Bóg ma poważne powody, aby istnienie zła dopuścić*.⁹¹ Pojawia się pytanie, jakie są to powody, albo chociaż czy można w ogóle wskazać cokolwiek, co mogłoby stanowić taki powód.

Niektórzy wierzący filozofowie podejmują się formułowania takich powodów. I tak Richard Swinburne stara się pokazać dobre konsekwencje zła moralnego (na przykładzie detonacji bomby atomowej) i zła naturalnego (na przykładzie cierpienia zwierząt).

[...] likwidacja pewnego możliwego lub istniejącego zła oznacza likwidację pewnego dobra. Każde niewielkie zwiększenie liczby istniejących lub możliwych złych stanów powoduje niewielkie zwiększenie liczby istniejących lub możliwych dobrych stanów. Przypuśćmy, że jedna osoba mniej została spalona przez bombę atomową zrzuconą na Hiroszimę. Wtedy byłoby mniej okazji dla wykazania się odwagą lub współczuciem; mniej informacji o efektach napromieniowania, mniej osób (krewnych tej osoby) miałyby silną motywację do prowadzenia kampanii na rzecz rozbrojenia nuklearnego i przeciwko imperializmowi. I tak dalej.⁹²

Ból, na który w swoim życiu narażone są zwierzęta, Swinburne uzasadnia w podobnym stylu:

Bóg mógł stworzyć świat, w którym zwierzęta przeżywają w swoim życiu najwyżej dreszcze [a nie prawdziwe cierpienie], ale ich życie jest bogatsze dzięki złożoności i trudności zadań, z którymi się spotykają, i problemów, na które muszą odpowiednio zareagować.⁹³

⁹⁰ A. Plantinga, *Bóg, wolność, zło*, tłum. K. Gruba.

⁹¹ A. Plantinga, *Bóg, wolność, zło*, s. 62.

⁹² R. Swinburne, *The Existence of God, 2nd ed.*, s. 264.

⁹³ R. Swinburne, *The Existence of God, 2nd ed.*, s. 245.

Uważam, że taki argument z bilansu – równoważenie wybuchu bomby atomowej wzrostem liczby członków ruchu na rzecz rozbrojenia, a potworne męczarnie zwierząt ginących w pożarach lasów złożonością zadań, z którymi zwierzęta spotykają się w swoim życiu – jest w najzwyklejszy sposób niemoralny. Przedstawione przez Swinburne przypadki dobra, które są konsekwencją przypadków zła, w żadnym razie nie rekompensują tego zła. Ale mój kontrargument może przekonać tylko tych, którzy podzielają intuicję moralną, na której się on opiera – być może fakt, że liczba pacyfistów wzrosła po zrzuceniu bomby na Hiroszimę, jest dla kogoś wystarczającym usprawiedliwieniem boskiej nieingerencji w lot bombowca lub w działanie zapalnika.

Istnieje kontrargument, który nie odwołuje się do intuicji moralnej. Przyjmijmy, że przypadkom zła w ostatecznym rozrachunku towarzyszą rekompensujące je z nawiązką przypadki dobra. W takim razie być może świat byłby lepszy, gdyby zła było więcej? Jeżeli trzęsienia ziemi i epidemie skutkują aktami wielkiego poświęcenia, dlaczego nie występują jeszcze częściej? Do obrony tezy, że świat nie mógłby być lepszy, nie wystarczy wskazać, jak to robi Swinburne, że mniejsza ilość zła skutkowałaby mniejszą ilością dobra i, w ostatecznym rozrachunku, gorszym od aktualnego światem. Trzeba jeszcze dodać, że większa ilość zła w świecie nie przełożyłaby się na dodatkową ilość dobra, przez którą taki świat byłby lepszy od aktualnego. Jednym słowem, argument z bilansu musi stwierdzić, że zła jest w świecie aktualnym dokładnie tyle, że bilans przyjmuje wartość maksymalną spośród wszystkich możliwych światów.

To dużo silniejsza (a więc jeszcze mniej prawdopodobna) hipoteza niż ta, że usunięcie zła spowodowałoby usunięcie jakiegoś dobra i zmniejszenie bilansu. Broniąc jej, Swinburne nie powinien poprzestać na wskazaniu faktu, że po zrzuceniu bomby atomowej rośnie liczba zwolenników rozbrojenia, a trzęsienia ziemi są okazją do wykazania się odwagą. Powinien jeszcze pokazać, że gdyby trzecia bomba została zrzucona na Kioto, liczba osób, które dołączyłyby do ruchu na rzecz rozbrojenia nuklearnego, byłaby zbyt mała, aby zrekompensować liczbę ofiar tej bomby, a trzęsienie ziemi w Polsce we wrześniu roku 2010 (nie było takiego trzęsienia) nie byłoby okazją do wykazania się odwagą przez

wystarczającą liczbę ludzi, w odróżnieniu od trzęsienia ziemi na Haiti w styczniu roku 2010 (było takie trzęsienie). Ciężar dowodu, że pożary, które nie wybuchły, i bitwy, które się nie rozegrały, nie byłyby źródłem wystarczającej ilości dobra w odróżnieniu do faktycznych pożarów i bitew, spoczywa na Swinburnie. Być może tak rzeczywiście jest, być może liczba pożarów i bitew jest optymalna, ale – co ważne dla dalszej argumentacji – nie ma na to żadnych dowodów. Nie istnieje argument, który bez odwołania do optymalności rozkładu katastrof uzasadniałby, że dobre skutki trzęsień ziemi, które się nie wydarzyły, różniłyby się istotnie od dobrych skutków faktycznych trzęsień ziemi.

Przedmiotem tego badania są intuicje moralne dotyczące sprawiedliwości (zgodnie z przyjętą definicją Bóg jest sprawiedliwy, przy założeniu, że sprawiedliwość jest konsekwencją moralnej doskonałości). Próba uzasadnienia, iż Bóg jest sprawiedliwy mimo objawiającej się w świecie niesprawiedliwości, opiera się na tym samym schemacie, co argumentacja Swinburne'a: być może w dalszej perspektywie pozornie niesprawiedliwe boskie działania okażą się sprawiedliwe. Takiej perspektywy ma dostarczyć *hipoteza testu*.

Bóg, według wielu religii, przekazał ludziom pewne prawa i nakazał ich przestrzegać. W judaizmie odczytuje się je z Tory, w chrześcijaństwie z Biblii, w islamie z Koranu. Przykładem takich praw jest Dekalog, w wielu religiach obok prawa zapisanego mówi się także o prawie naturalnym lub o zgodności postępowania z sumieniem. Jednak przestrzeganie tego prawa nie wiąże się z nagrodami, a łamanie go – z karami. Można przestrzegać prawa i mieć poczucie zmarnowanego życia, można go nie przestrzegać i uważać się za szczęśliwego. Czy taki Bóg może być sprawiedliwy?

Teista usprawiedliwiający brak boskiej ingerencji przeważnie stwierdza, że nagroda lub kara czekają człowieka dopiero po śmierci. To wyjaśnienie nazwę hipotezą testu, ponieważ w tej perspektywie życie pełni rolę egzaminu, od wyniku którego zależy pośmiertna nagroda lub kara. Z tym wytłumaczeniem wiążą się pewne problemy:

(a) nikt nie jest poddawany takiemu samemu testowi. Niektóre dzieci nie dożywają wieku, w którym można by je obarczyć odpowiedzialnością

za ich działania. Niektórzy rodzą się w spokojnym miejscu w spokojnych czasach i nie są poddawani tym samym próbom co ludzie, którzy doświadczyli jakiegoś kataklizmu. W spokojnym życiu zarówno dużo trudniej złamać jakiś moralnie istotny zakaz, jak i wykazać się niektórymi ważnymi cnotami. Co więcej, wybory życiowe są przynajmniej częściowo, zdeterminowane warunkami zewnętrznymi – wychowaniem, kulturą – których jednostka nie wybierała. W ten sposób ludzie wychowani niezgodnie z boskim prawem z większym prawdopodobieństwem są skazani na łamanie go. Nie wiadomo, jak należałoby potraktować ludzi niezdolnych do rozróżniania dobra od zła z przyczyn czysto zewnętrznych (np. jakichś neurologicznych zaburzeń). Czy mieliby zyskać tę zdolność po śmierci? Jeżeli tak, to dlaczego nie mogli mieć jej od początku?

(b) nawet jeżeli Bóg przekazał ludziom prawo, to nie przekazał wiarygodnych narzędzi interpretacyjnych. Albo można brać prawo dosłownie – wtedy na przykład dekalog obowiązuje tylko mężczyzn, bo tylko oni są jego adresatami, oraz uprawomocnia niewolnictwo (*Nie będziesz pożądał żony bliźniego twego, ani jego niewolnika, ani jego niewolnicy, [...]*⁹⁴) – albo można prawo interpretować. W tym drugim przypadku trzeba się zastanowić, czy za złą interpretację, dokonaną w dobrej intencji, ponosi odpowiedzialność interpretator, który dołożył wszelkich starań, aby tekst zrozumieć, czy raczej autor przekazu, który sformułował myśli niewystarczająco jasno. W przypadku niejasnych ustaw wini się raczej ustawodawcę; jego obarczyć winą za cierpienie zadane w imię boskich nakazów zinterpretowanych źle, choć w dobrej wierze?

(c) być może łatwiej można wyobrazić sobie potencjalne kary, ale na pewno nie jest prosto wyobrazić sobie sposób, w jaki Bóg miałby wynagrodzić cierpienia ludziom, którzy przestrzegają nadanego przez niego prawa. Jak można komuś wynagrodzić stratę rodziny w pożarze albo ciężką traumę z dzieciństwa, która wpłynęła istotnie na resztę życia? Jak mogą cieszyć się szczęściem po śmierci, jeżeli mam świadomość, że ktoś bardzo mi bliski doznaje kary za łamanie boskich praw?

⁹⁴ Księga Wyjścia 20:17, *Biblia Tysiąclecia*. Por. także Księga Powtórzonego Prawa, 5:21.

(d) nie wiadomo też, jak takie nagrody mogłyby wyglądać wobec faktu ogromnego zróżnicowania ludzi na przestrzeni epok i kultur, między sobą i w trakcie pojedynczego życia. Inne cele życiowe i kryteria szczęścia stawia sobie dziecko, inne stawiali sobie ludzie z epoki kamiennej, inne stawiamy sobie my – dorośli uczestnicy kultury zachodniej; inne kryteria stawiam sobie ja, a inne mój sąsiad. Z drugiej strony w takim stopniu, w jakim kryteria zależą od czynników zewnętrznych (kultury, wychowania), są arbitralne, bo otoczenie, w którym człowiekowi przychodzi dorastać, nie jest jego wyborem. Większość kryteriów, jeśli nie wszystkie, związana jest z życiem doczesnym. Dlatego wierzący może powiedzieć, że człowieka posłusznego Bogu czeka jakaś wspaniała nagroda, ale z pewnością nie może wskazać, co czyni ją taką wspaniałą.

Jest jeszcze jeden problem, ogólniejszy niż zarzuty do hipotezy testu. Większość wierzących filozofów, od Augustyna, przez Leibniza, po cytowanego już Plantingę, w swoich teodyceach stwierdza, że świat z wolnymi jednostkami jest zawsze lepszy od świata zaludnionego automatami, a więc że zło jest kosztem, który należy ponieść dla umożliwienia wolności. John Mackie dowodzi, że możliwe są wolne jednostki, które zawsze wybierają dobro.⁹⁵ Jest to element argumentacji przeciwko istnieniu Boga jako dobrej istoty: jeżeli rzeczywiście jest tak, jak twierdzi Mackie, to Bóg mógł stworzyć świat lepszy niż aktualny, a przecież dobra istota nie stwarza niekoniecznego zła. Plantinga odiera ten argument jedynie poprzez zaprzeczenie stwierdzeniu Mackiego, że możliwe są wolne jednostki nieczyniące zła (a więc że zło istniejące w świecie jest konieczne dla istnienia wolności).

Są co najmniej dwa zarzuty przeciwko temu kontrargumentowi. Po pierwsze, jeżeli zło jest konieczne dla istnienia wolności, to czy ludzie po śmierci będą wolni? Jeżeli nie, to miejsce, do którego trafią, będzie gorsze niż to, z którego przybyli – zgodnie ze stwierdzeniem, że lepszy świat ze złem i z wolnością, niż bez nich obu. Jeżeli tak, to albo będą dopuszczać się zła, albo nie. W pierwszym przypadku sytuacja człowieka niewiele będzie się różnić, niż za życia – zło, co najmniej to, które jest konsekwencją ludzkich wyborów, będzie obecne. Drugi przypadek stoi

⁹⁵ J. L. Mackie, *Evil and Omnipotence*.

w sprzeczności z tezą, że Bóg nie mógł stworzyć świata, w którym wolne jednostki wybierają tylko dobro – skoro może stworzyć stan, w którym ludzie są jednocześnie wolni i całkowicie dobrzy, to dlaczego aktualny świat taki nie jest? Jeżeli odpowiedzią ma być *po to, aby człowieka przetestować*, to znowu powraca pytanie, dlaczego niektórzy, na przykład wcześniej zmarłe dzieci, nie zostają żadnemu testowi poddani.

Drugi zarzut brzmi następująco. Podejmując działania etycznie istotne, ludzie wybierają dobro z częstością p i zło z częstością $1 - p$. Oczywiście p można mierzyć dla różnych hierarchii wartości, różnych grup ludzi, sytuacji, epok, zamiast jedną liczbą można posługiwać się wektorem prawdopodobieństw, odpowiadających różnym rodzajom czynów, lub różnym rodzajom czynów można przypisać różne wagi i w wyznaczaniu p używać średniej ważonej – to nie ma znaczenia dla tego argumentu. Z doświadczenia widzimy, że $p < 1$, powiedzmy, $p = 0,2$, jeżeli jesteśmy surowi, bądź $p = 0,8$, jeżeli pobłażliwi – znajomość dokładnej wartości nie jest potrzebna, wystarczy zauważyć, że $p < 1$. Teraz wyobraźmy sobie świat M , w którym $p_M > p$, przy zachowaniu warunku $p_M < 1$. W takim świecie zła byłoby mniej, a więc byłby lepszy, a jednocześnie ludzie nie byłoby automatami. Zwiększanie p przy spełnionym warunku $p < 1$ powodowałoby, że zła na świecie by ubywało, chociaż człowiek nadal byłby wolny.

Wobec tych uwag pozostaje poddać się i przyznać, że być może istnieje dobry i sprawiedliwy Bóg, ale człowiek nie potrafi przedstawić sobie sposobu, w jaki to dobro i sprawiedliwość mogłyby się zrealizować.

Swinburne nie zgodzi się z tą diagnozą – uważa, że potrafi wyjaśnić obecność zła, a w każdym razie wskazać, gdzie takiego wyjaśnienia można szukać. Natomiast Plantinga – któremu z pewnością ateistycznych intencji przypisać nie można – formułuje możliwe wyjaśnienia, ale zauważa, że są nieprzekonujące lub co najmniej niewystarczające. Jednak nie jest to dla niego argumentem przeciwko istnieniu moralnie doskonałego Boga – to, że nie potrafimy wyjaśnić istnienia zła, jeszcze nie oznacza, że takie wyjaśnienie nie istnieje:

Dlaczego przypuszczać, że jeśli Bóg rzeczywiście ma swoje powody, by dopuszczać zło, teści mieliby pierwsi o nich

wiedzieć? Być może Bóg ma ważny powód, ale jest on zbyt skomplikowany, by go zrozumieć. Albo może nie objawił nam go z jakichś innych przyczyn.⁹⁶

To, że teista nie wie, dlaczego Bóg dopuszcza zło, jest, być może, interesującym faktem na temat teisty, ale sam w sobie niewiele albo zgoła nic nie mówi na temat racjonalności wiary w Boga. Potrzeba o wiele więcej, by móc choćby rozpocząć argumentację antyteologiczną.⁹⁷

I w istocie, fakt, że teista nie wie, dlaczego Bóg dopuszcza zło (lub dlaczego nie manifestuje swojej sprawiedliwości), jest bardzo interesujący.

Przeprowadzone badanie dotyczy intuicji w odniesieniu do pojęcia sprawiedliwości. Wyniki rzucają nieco światła na zaproponowany przez Plantingę schemat *nie mam pewności, dlaczego to zrobiłeś, a dopóki nie mam pewności, wstrzymam się z osądem Twojego działania*. Główna hipoteza brzmi, że ludzie stosują inny schemat – *navet jeżeli nie mam pewności, ocenię Cię zgodnie z najbardziej prawdopodobnym wyjaśnieniem Twojego działania* – a tym samym teista, dla którego niewiedza na temat powodów istnienia zła nie stanowi wystarczającej motywacji przyznania, że Boga nie ma lub nie jest sprawiedliwy, nie stosuje pojęcia *sprawiedliwy* konsekwentnie.

metoda badania

Aby zbadać intuicje filozoficzne użytkowników języka, skonstruowaliśmy ankietę w dwóch wersjach: MB i BM. Ankieta MB zawierała dwa opisy i po dwa pytania dotyczące każdego opisu:

Opis 1. Przypuśćmy, że Pani Elżbieta ma czwórkę dzieci. Ustaliła pewne nakazy i zakazy: np. należy myć zęby po jedzeniu, nie wolno przeklinać ani bić się, należy odrabiać lekcje. Jednak Pani Elżbieta nagradza i karze dzieci zupełnie losowo: dziecko, które naruszyło zakaz lub nie spełniło nakazu może być równie dobrze ukarane, co nagrodzone. Dziecko, które wypełnia ustalenia, może być nagrodzone, ale także może być ukarane. Często też Pani Elżbieta nie podejmuje żadnej akcji w stosunku do dzieci.

⁹⁶ A. Plantinga, *Bóg, wolność, zło*, s. 60.

⁹⁷ A. Plantinga, *Bóg, wolność, zło*, s. 50.

Z punktu widzenia obecnych (oraz dawnych) koncepcji psychologicznych i pedagogicznych jej zachowanie nie da się wytłumaczyć żadnym sensownym celem wychowawczym (co nie znaczy, że Pani Elżbieta nie ma jakiegoś ukrytego planu, zgodnie z którym postępuje).

Pytanie 1: czy, według Ciebie (zgodnie z Twoim osądem moralnym lub definicją sprawiedliwości, którą uważasz za słuszną/trafną), Pani Elżbieta jest sprawiedliwa w stosunku do swoich dzieci?

- (a) Tak (b) Nie (c) Nie potrafię określić

Pytanie 2: w przypadku odpowiedzi (c). Nie potrafię określić, ponieważ:

Opis 2. Przypuśćmy, że Bóg istnieje, ma możliwość wpływu na świat i przekazał ludziom (np. w formie dekalogu, mitów lub świętych tekstów) prawa, których nakazał im przestrzegać. Jednak nie zachodzi związek między przestrzeganiem tych praw a życiowym dobrostanem. Życie ludzi, którzy przestrzegają boskiego prawa, bywa szczęśliwe i spełnione, a także nieszczęśliwe i pełne niepowodzeń. Wśród tych, którzy tego prawa nie przestrzegają, również znajdują się ludzie od bardzo szczęśliwych po bardzo nieszczęśliwych. Ludzie bywają nieszczęśliwi z różnych powodów: niektórzy w wyniku działań innych ludzi, niektórzy z winy własnych wyborów, ale bardzo wiele osób jest nieszczęśliwych z przyczyn, za które żaden człowiek nie ponosi odpowiedzialności, np. chorób, trzęsień ziemi, powodzi, susz. Tak samo jest z ludźmi szczęśliwymi: niektórzy osiągają szczęście dzięki własnej pracy, ale już np. to, że ktoś urodzi się w kraju bogatym, a nie biednym, jest zupełnie arbitralne.

Z punktu widzenia obecnego stanu wiedzy nie można wytłumaczyć, dlaczego tak się dzieje (co nie znaczy, że Bóg nie ma jakiegoś planu, który w ten sposób realizuje).

Pytanie 3: czy, według Ciebie (zgodnie z Twoim osądem moralnym lub definicją sprawiedliwości, którą uważasz za słuszną/trafną), Bóg z *opisu 2* jest sprawiedliwy w stosunku do ludzi?

- (a) Tak (b) Nie (c) Nie potrafię określić

Pytanie 4: w przypadku odpowiedzi (c). Nie potrafię określić, ponieważ:

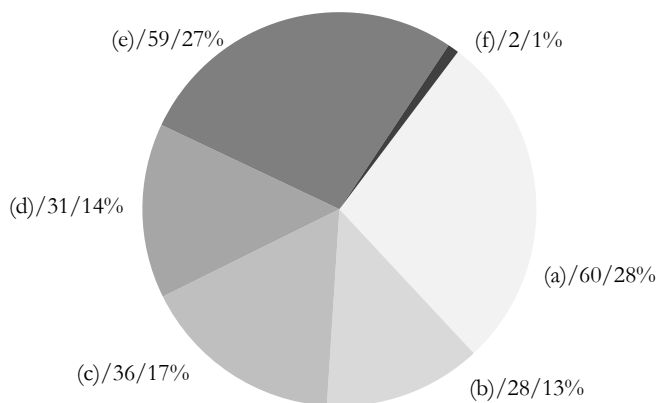
Ankieta BM różniła się od MB jedynie kolejnością opisów (stąd też nazwy ankiet: MB – matka/Bóg, BM – Bóg/matka). Obie ankiety kończyły się metryczką, która, oprócz standardowych pytań o wiek, płeć i wykształcenie, zawierała pytanie dotyczące wiary:

Uważam, że:

- (a) Bóg nie istnieje i jestem tego (zupełnie lub prawie) pewna/y,
- (b) Bóg nie istnieje, ale czasem poważnie zastanawiam się, że może jednak istnieje,
- (c) zastanawiałam/em się i znajduję tyle samo argumentów za istnieniem Boga, co przeciw (albo nie znajduję żadnych argumentów za i przeciw),
- (d) Bóg istnieje, ale czasem poważnie zastanawiam się, że może jednak nie istnieje,
- (e) Bóg istnieje i jestem tego (zupełnie lub prawie) pewna/y,
- (f) nigdy mnie to szczególnie nie interesowało i nie zastanawiałam/em się.

Zebrałiśmy 216 ankiet. Ankiet MB było 104 (czyli 48%), ankiet BM 112 (52%).

Do badania ankiet wykorzystaliśmy narzędzia statystyczne, przede wszystkim test t-Studenta dotyczący równości wartości średnich oraz test równości proporcji, oparty na statystyce χ^2 . Stosujemy oba testy, aby wyniki były bardziej wiarygodne. Obliczenia zostały przeprowadzone za pomocą skryptów języka R.



Rysunek 5. Struktura próby, na której przeprowadzono badanie.

Wyjaśnienia: (a) Bóg nie istnieje i jestem tego (zupełnie lub prawie) pewna/y; (b) Bóg nie istnieje, ale czasem poważnie zastanawiam się, że może jednak istnieć; (c) zastanawiałam/em się i znajduję tyle samo argumentów za istnieniem Boga, co przeciw (albo nie znajduję żadnych argumentów za i przeciw); (d) Bóg istnieje, ale czasem poważnie zastanawiam się, że może jednak nie istnieć; (e) Bóg istnieje i jestem tego (zupełnie lub prawie) pewna/y; (f) nigdy mnie to szczególnie nie interesowało i nie zastanawiałam/em się.

Odpowiedziom z pytań 1 i 3 przypisaliśmy liczby: tak – 2, nie – 0, trudno powiedzieć – 1. W teście t-Studenta między odpowiedziami nie zakłada się żadnego związku (skala nominalna). W teście χ^2 porządek na liczbach odzwierciedla porządek odpowiedzi (skala porządkowa). Odpowiedź *nie potrafię określić* jest problematyczna – z jednej strony jej miejsce w porządku wszystkich odpowiedzi zakłada, że ankietowany tak samo nie jest skłonny potępić, jak usprawiedliwić ocenianego postępowania – gdyby odpowiedzi interpretować jako glosy (*za, przeciw, wstrzymuje się*), zamiana odpowiedzi *nie potrafię określić* na parę *tak i nie* nie zmieniałaby wyniku głosowania. Z drugiej strony, przypisanej tej odpowiedzi jakiegokolwiek wartości porównywalnej z pozostałymi wydaje się arbitralne. Dlatego przeważnie analizę przedstawiamy w obu skalach (nominalnej i porządkowej), jak również raz z uwzględnieniem, a raz bez uwzględnienia odpowiedzi *nie potrafię określić*. W ten sposób staramy się ochronić badanie przed zarzutem, że wyniki zależą od dosyć problematycznej skali.

Za wierzące uznaliśmy osoby, które na pytanie 6 udzieliły odpowiedzi (d) lub (e). Wierzącymi-pewnymi nazywamy osoby, które dały odpowiedź (d), wierzącymi-wątpiącymi – odpowiedź (e). Ateistów identyfikujemy poprzez odpowiedzi (a) i (b). Niewierzącymi będziemy określać ankietowanych, którzy zaznaczyli odpowiedź (a), (b), (c) lub (f). Należy więc pamiętać, że zgodnie z tą konwencją ateści są podzbiorem niewierzących i słowa te nie są synonimami. Na rysunku 5 przedstawiliśmy udział każdej odpowiedzi.

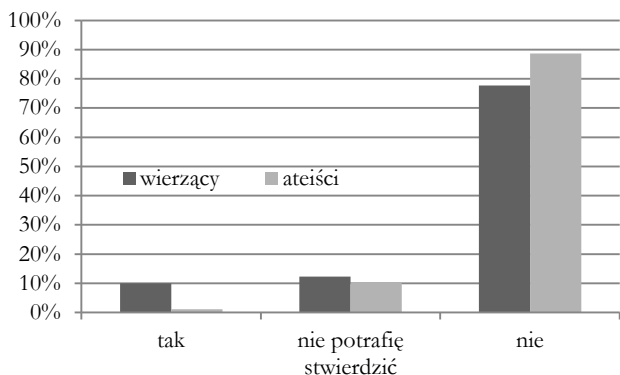
czy to te same pojęcia?

Zebrane dane pozwoliły na uzyskanie dwóch bardzo ważnych wyników. Po pierwsze, wierzący inaczej oceniają Elżbietę niż ateści i niż niewierzący.

Tabela 2. Różnica w odpowiedziach na pytanie *czy Elżbieta jest sprawiedliwa?*

porównywane grupy:	$\mu_1 (\sigma_1)$	$\mu_2 (\sigma_2)$	T	P	$n_1 + n_2$
wierzący / ateści	0,322 (0,650)	0,125 (0,365)	0,0068	0,0125	90 + 88
wierzący / ateści *	0,228 (0,639)	0,025 (0,225)	0,0046	0,0111	79 + 79
wierzący / niewierzący	0,322 (0,650)	0,135 (0,365)	0,0074	0,0022	90 + 126
wierzący / niewierzący *	0,228 (0,639)	0,018 (0,190)	0,0029	0,0021	79 + 111

Wyjaśnienia: $\mu_i (\sigma_i)$ oznacza średnią i odchylenie standardowe, indeks i – pierwszą lub drugą grupę z pary porównywanych. Składniki sumy $n_1 + n_2$ oznaczają liczbę ankiet w pierwszej i w drugiej grupie. Symbol * oznacza, że usunięto odpowiedzi *nie potrafię stwierdzić*. Kolumna T zawiera p-wartości wyznaczone dla testu t-Studenta, gdzie hipotezą alternatywną jest $H1: \mu_1 > \mu_2$. Kolumna P zawiera wartości p dla testu równości proporcji χ^2 o hipotezie alternatywnej: proporcja odpowiedzi *tak* jest większa w pierwszej grupie niż w drugiej.



Rysunek 6. Udział odpowiedzi na pytanie *czy Elżbieta jest sprawiedliwa?*

Różnice występują także wewnątrz grup. Widać różnicę między oceną Elżbiety przez wierzących-wątpiących a niewierzących (tabela 3), ale testy nie dają jednoznacznego rozstrzygnięcia (test χ^2 wskazuje na różnicę przy skali nominalnej, test t-Studenta nie; przy skali porządkowej p-wartości dla obu testów są bliskie standardowo przyjmowanemu poziomowi ufności 0,1, ale nie są od niego mniejsze). Między wierzącymi-pewnymi a wierzącymi-wątpiącymi testy także nie pozwalają jednoznacznie stwierdzić różnicy.

Tabela 3. Różnica w odpowiedziach na pytanie *czy Elżbieta jest sprawiedliwa?*
Podział wewnątrz grup.

porównywane grupy:	μ_1 (σ_1)	μ_2 (σ_2)	T	P	$n_1 + n_2$
wierzący-wątpiący / niewierzący	0,161 (0,522)	0,135 (0,365)	0,3961	0,0919	31 + 126
wierzący-wątpiący / niewierzący *	0,133 (0,507)	0,180 (0,190)	0,1155	0,1096	30 + 111
wierzący-pewni / wierzący-wątpiący	0,407 (0,698)	0,161 (0,522)	0,0320	0,3286	59 + 31
wierzący-pewni / wierzący-wątpiący *	0,286 (0,707)	0,133 (0,507)	0,1349	0,2516	49 + 30

Wyjaśnienia: znaczenie symboli jak w tabeli 2. W wierszu trzecim od dołu hipotezą alternatywną jest $H_1: \mu_1 < \mu_2$.

Wobec istotniej różnicy między wierzącymi a niewierzącymi w ocenie Elżbiety przyjmujemy, że za tę różnicę odpowiedzialne są przede wszystkim ankiety wierzących-pewnych.

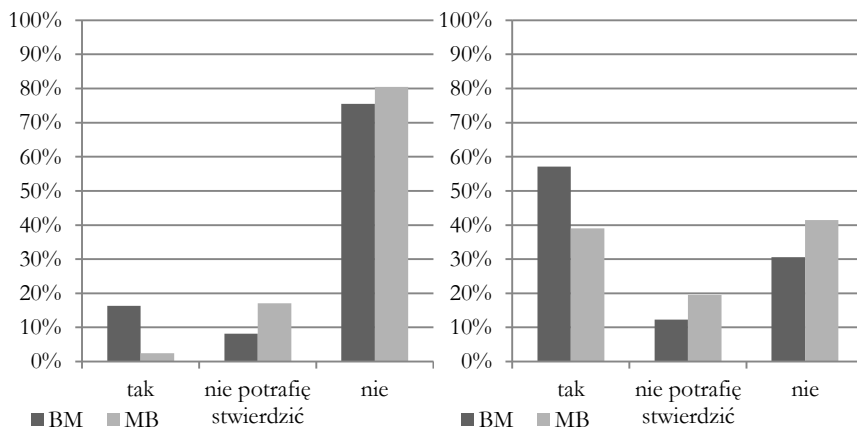
kolejność ma znaczenie

Zaobserwowaną różnicę można wytłumaczyć na gruncie samych danych. W ankietach osób wierzących odpowiedzi silnie zależą od kolejności pytań (tabela 3). Kiedy pytanie o Elżbietę pojawia się w ankiecie jako drugie, wierzący chętniej odpowiadają, że matka jest sprawiedliwa. Bóg także jest oceniany częściej przez wierzących jako sprawiedliwy, kiedy nie poprzedza go pytanie o Elżbietę. Natomiast zależności od kolejności nie da się zaobserwować wśród ateistów (tabela 4).

Tabela 4. Zależność odpowiedzi od kolejności pytań w ankietach osób wierzących.

porównywane grupy:	μ_1 (σ_1)	μ_2 (σ_2)	T	P	$n_1 + n_2$
BM / MB. Elżbieta	0,408 (0,761)	0,219 (0,475)	0,0779	0,0333	49 + 41
BM / MB. Elżbieta *	0,355 (0,773)	0,059 (0,342)	0,0126	0,0448	45 + 34
BM / MB. Bóg	1,265 (0,908)	0,975 (0,908)	0,0677	0,0667	49 + 41
BM / MB. Bóg *	1,302 (0,964)	0,970 (1,015)	0,0763	0,1110	43 + 33
BM / MB. Suma	1,673 (1,375)	1,195 (1,166)	0,0388	0,4074	49 + 41
BM / MB. Suma *	1,700 (1,471)	0,896 (1,145)	0,0065	0,3669	40 + 29

Wyjaśnienia: znaczenie symboli jak w tabeli 2.



Rysunek 7. Udział odpowiedzi na pytanie czy *Elżbieta jest sprawiedliwa?* oraz czy *Bóg jest sprawiedliwy?* w zależności od kolejności pytań; uwzględnione tylko ankiety osób wierzących.

Wyjaśnienia: w ankiecie BM opis przypadku Boga jest pierwszy, a Elżbiety drugi. W MB kolejność jest odwrócona.

Tabela 5. Zależność odpowiedzi od kolejności pytań w ankietach ateistów.

porównywane grupy:	$\mu_1 (\sigma_1)$	$\mu_2 (\sigma_2)$	T	P	$n_1 + n_2$
BM / MB. Elżbieta	0,095 (0,297)	0,152 (0,420)	0,2310	0,5	42 + 46
BM / MB. Elżbieta *	0 (0)	0,049 (0,312)	0,1617	0,5	38 + 41
BM / MB. Bóg	0,357 (0,617)	0,478 (0,722)	0,1996	0,2876	42 + 46
BM / MB. Bóg *	0,182 (0,584)	0,333 (0,756)	0,1764	0,2824	33 + 36
BM / MB. Suma	0,452 (0,771)	0,630 (0,951)	0,1679	0,4699	42 + 46
BM / MB. Suma *	0,182 (0,584)	0,400 (0,945)	0,1271	0,3867	33 + 35

Wyjaśnienia: znaczenie symboli jak w tabeli 2, ale hipotezą alternatywną jest $H_1: \mu_1 < \mu_2$.

Pozostaje jeszcze pytanie, czy kolejność jest jedynym powodem zaobserwowanej różnicy. Wyniki testów statystycznych dla samej ankiety MB nie pokazują już istotnej różnicy między ateistami a wierzącymi w pytaniu dotyczącym Elżbiety (tabela 6), chociaż nadal średnia odpowiedzi wierzących jest większa niż odpowiedzi ateistów. Być może należy wobec tego twierdzić, że są jeszcze jakieś inne powody różnicy między ocenami wierzących a ateistów.

Tabela 6. Różnica odpowiedzi na pytanie *czy Elżbieta jest sprawiedliwa?* w ankietach MB.

porównywane grupy:	$\mu_1 (\sigma_1)$	$\mu_2 (\sigma_2)$	T	P	$n_1 + n_2$
wierzący / ateści	0,219 (0,475)	0,152 (0,420)	0,2349	0,5	41 + 46
wierzący / ateści *	0,059 (0,343)	0,049 (0,312)	0,4479	0,5	34 + 41

Wyjaśnienia: znaczenie symboli jak w tabeli 2.

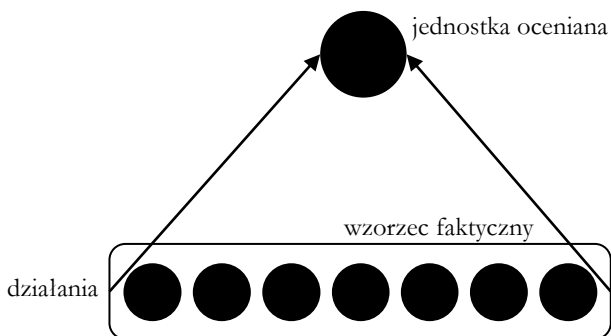
interpretacja wyników – dwa modele oceniania

Po pierwsze, wygląda na to, że wierzący i ateści posługują się innymi pojęciami sprawiedliwości. Dla ateisty fakt, że nie można podać prawdopodobnego uzasadnienia decyzji Elżbiety, jest wystarczającą przesłanką do określenia jej zachowania jako niesprawiedliwego. Natomiast wierzący częściej dopuszcza możliwość, że Elżbieta może być sprawiedliwa, pomimo że nie potrafi wskazać wzorca tłumaczącego jej zachowanie jako sprawiedliwe. Im bardziej ktoś jest pewny swojej wiary, tym chętniej usprawiedliwia Elżbietę, zgadzając się na taki argument. Podobny schemat argumentacyjny stosowany jest przez Plantingę – jest możliwe, aby Bóg był doskonały moralnie, ponieważ jest możliwe,

że istnienie zła jest konieczne, nawet jeżeli człowiek nie potrafi wytłumaczyć, na czym ta konieczność mogłaby polegać.

Okazuje się jednak, że w przypadku oceny Elżbiety intuicje niewierzących i wierzących nie różnią się, gdy opisu przypadku matki nie poprzedza opis przypadku Boga. Różnice pojawiają się dopiero, kiedy pytanie o Elżbietę pada po pytaniu o Boga. Uważam, że to zjawisko można wyjaśnić następująco.

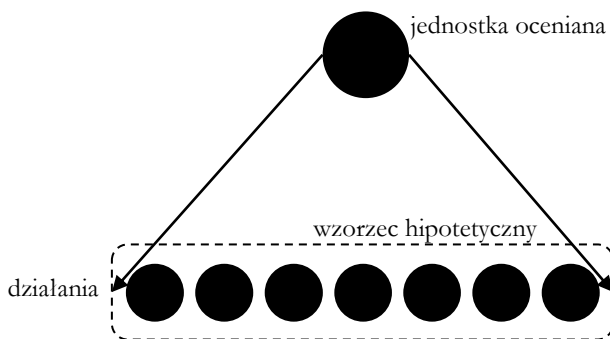
Rysunki 8 i 9 zawierają schematy dwóch modeli oceniania. Pierwszy model (rysunek 8) nazwaliśmy *w górę*. W tym modelu przedmiot oceny (w przykładzie Bóg lub Elżbieta) jest odpowiedzialny za pewne działania. Działania można starać się opisać pewnymi wzorcami. Spośród nich oceniający wybiera wzorec najlepiej dopasowany do zaobserwowanych działań, a następnie na podstawie tego wyboru dokonuje oceny działającego podmiotu. Tego modelu używali respondenci, oceniając Elżbietę. Najpierw uznali jej zachowanie, tj. arbitralne nagradzanie i karanie, za lepiej dopasowane do wzorca *niesprawiedliwe* niż *sprawiedliwe*, a następnie ocenili zgodnie z tym wzorcem samą Elżbietę jako niesprawiedliwą.



Rysunek 8. Model oceniania *w górę*.

W modelu *w dół* (rysunek 9) jednostka oceniana od początku ma przypisaną pewną cechę. Ta cecha przenosi się na działania ocenianego według reguły *jeżeli oceniany ma cechę A, to i jego działania są A*. Następnie

oceniający zauważa, że obserwowane działania nie układają się w żaden spójny wzorzec zgodny z A, ale stwierdza, że jest możliwe, aby taki wzorzec istniał, a jedynie nie można go rozpoznać ze względu na niewystarczającą wiedzę o ocenianym albo zawilość samego wzorca. Tak więc oceniający przyjmuje istnienie hipotetycznego wzorca zgodnego z A, który dobrze opisuje działania ocenianej jednostki.



Rysunek 9. Model oceniania *w dół*.

Tego modelu używała znaczna część wierzących, oceniając Boga. Ocenianie przebiegało w następujący sposób. Bóg jest z definicji sprawiedliwy, a więc i jego działania są sprawiedliwe. Co prawda najlepszy dostrzeżony wzorzec sugeruje, że są one niesprawiedliwe, ale jest możliwe, że istnieje inny wzorzec, którego nie można dostrzec ze względu na niewystarczającą informację (np. o pośmiertnych nagrodach i karach). Ten hipotetyczny wzorzec jest wystarczającą przesłanką do odrzucenia najlepszego dostrzeżonego wzorca, zgodnie z którym Boskie działania (w tym zaniechania) są niesprawiedliwe.

Zaobserwowana w ankietach osób wierzących wysoka wrażliwość odpowiedzi na kolejność – usprawiedliwianie Elżbiety i odmawianie sprawiedliwości Bogu, kiedy pytania o nich pojawiają się w ankiecie jako drugie – bierze się stąd, że zauważają oni fakt, że posługują się dwoma różnymi modelami. Dwa różne modele oceniania odpowiadają dwóm różnym pojęciom sprawiedliwości, ponieważ do znaczenia pojęcia należy reguła jego stosowania. Chcąc posługiwać się jednym pojęciem, wierzący

poświęcają swoją pierwotną intuicję na rzecz spójności. Zmiana jest silniejsza w przypadku oceny Boga (zmiana odpowiedzi *sprawiedliwy* na jedną z pozostałych) niż w przypadku Elżbiety (zmiana odpowiedzi *niesprawiedliwa*). Sugeruje to, że intuicje do przypadku Elżbiety są silniejsze niż do przypadku Boga. Co więcej, w dodatkowym pytaniu o powody udzielenia odpowiedzi *nie potrafię określić*, część ankietowanych pisała: jeżeli nagroda/kara czekają na ludzi dopiero po śmierci, to Bóg jest sprawiedliwy, jeżeli nie, to nie jest. Ankiety nie zawierały argumentów na rzecz tezy, że hipoteza testu jest bardzo problematyczna. Prawdopodobnie gdyby ankietowana zawierała te powody, wyniki badania byłyby jeszcze bardziej jednoznaczne.

Te wyniki stanowią pewien głos w debacie na temat istnienia doskonałego moralnie Boga. Ich najważniejsza konsekwencja brzmi, że Bóg nie może być sprawiedliwy, jeżeli sprawiedliwość bierzemy w zwykłym znaczeniu tego słowa. Stosowanie pojęcia *sprawiedliwość*, które można odczytać z codziennego użycia, wymaga przyjęcia modelu *w górę*, polegającego na znalezieniu wzorca najbardziej dopasowanego do działań ocenianej jednostki i jej ocenie zgodnie z tym wzorcem. Stwierdzenie Plantingi, iż *to, że teista nie wie, dlaczego Bóg dopuszcza zło, jest, być może, interesującym faktem na temat teisty, ale sam w sobie niewiele albo zgoła nic nie mówi na temat racjonalności wiary w Boga* opiera się na innym modelu, *w dół*, który stoi w sprzeczności z codzienną praktyką językową. Tak więc kiedy teista używa tego argumentu, zmienia znaczenia pojęć, a tym samym przestaje się odnosić do tego, przeciwko czemu argumentował.

uwagi i zarzuty

Dobro a sprawiedliwość. W oryginalnym sformułowaniu argument ze zła opiera się na pojęciach dobra/zła, a nie na pojęciach czynu sprawiedliwego/niesprawiedliwego. Jest możliwe, że granica między czynami (etycznie) dobrymi a złymi oraz granica między czynami sprawiedliwymi i niesprawiedliwymi nie pokrywają się.⁹⁸ Są dwie tradycje, w których można rozumieć pojęcie sprawiedliwości. Pierwsza to grecko-rzymskie rozumienie sprawiedliwego postępowania, sformułowane przez

⁹⁸ M. Magdziak, *Uwagi o zasadzie równej miary*, [w:] *Czy sprawiedliwość jest możliwa?* red. D. Probuca, s. 85.

Ulpiana jako *oddawanie każdemu, co mu się należy*. Druga to wczesnochrześcijańskie powiązanie *czynu sprawiedliwego i czynu usprawiedliwionego*. W tym wypadku sam czyn może być zły, ale sprawiedliwy, bo usprawiedliwiony (niektórzy za takie racje uważają na przykład niewystarczającą wiedzę lub dobre intencje).⁹⁹

Elżbieta może być niesprawiedliwa w sensie Ulpiana, ponieważ nie oddaje swoim dzieciom tego, co im się należy w świetle ustanowionych przez nią praw, ale sprawiedliwa w sensie wczesnochrześcijańskim. Na przykład jeżeli jej zachowanie jest dla kogoś wystarczającą przesłanką do rozpoznania zaburzenia psychicznego, może ją usprawiedliwiać, przypisując jej dobre intencje mimo nieumiejętności działania zgodnie z nimi. Bóg może być niesprawiedliwy w sensie Ulpiana w swoim postępowaniu wobec poszczególnych ludzi, ale usprawiedliwiony, na przykład, koniecznością dopuszczenia zła dla umożliwienia wolności. Oczywiście może także być sprawiedliwy w sensie Ulpiana – to założenie leży u podstaw hipotezy testu.

Zamiast przy pomocy modeli *w górę i w dół*, można próbować zinterpretować wyniki badania przy pomocy tych dwóch pojęć. Należałoby wyjaśnić dwie ważne prawidłowości: asymetrię oceny Boga i Elżbiety oraz wpływ kolejności na odpowiedzi. Ponownie, efekt kolejności sugeruje, że ankietowanie chcą być konsekwentni i używać jednego pojęcia zamiast dwóch różnych. Jednak zebrane dane nie wystarczają do sformułowania bardziej szczegółowych wyjaśnień. Dlatego bardziej naturalnym przedmiotem badania byłyby po prostu intuicje dotyczące pojęcia dobra (dobrego postępowania). Jeżeli eksperyment zostanie powtórzony, nowe ankiety będą uwzględniały tę uwagę.

Nieanalogiczne ankiety. Dwie wersje historii o prezesie firmy, które Knobe zaprojektował dla zbadania pojęcia działania intencjonalnego, różniły się tylko jednym terminem – słowo *zaszkodził* zostało zamienione na *pomógł*.¹⁰⁰ W porównaniu z nimi ankiety o Bogu i Elżbiecie są powiązane dużo luźniej. Jednak gdyby analogia była ściślej, podobna tej z eksperymentu Knobe’a, zaobserwowany efekt byłby jeszcze silniejszy. Co

⁹⁹ M. Magdziak, *Uwagi o zasadzie równej miary*, s. 86.

¹⁰⁰ Por. Rozdział II, podrozdział *działanie intencjonalne*.

natomiast jest ważne w tym badaniu, to że nigdzie w interpretacji wyników nie pojawia się założenie, że opowieści są analogiczne, a badani, którzy ich za analogiczne nie uznali, nie są kompetentnymi użytkownikami języka. To uczestnicy badania oceniali, na ile przedstawione historie są analogiczne, i zgodnie z tą oceną wypełniali ankietę.

Teistyczny najlepiej dopasowany wzorzec. W książce *The existence of God* Swinburne uzasadnia, że najlepszy wzorzec opisujący relację Boga do świata (przy założeniu, że Bóg istnieje) to *Bóg jest doskonały moralnie w swoim postępowaniu wobec świata*. Staralem się wskazać, że ten wzorzec wcale nie jest najlepszy, ponieważ jego przyjęcie wiąże się z takimi absurdalnymi, niepotwierdzonymi empirycznie konsekwencjami, jak stwierdzenie, że rozkład epidemii i katastrof jest optymalny. Jednak teiście, który przyjmuje Swibrune'owskie wyjaśnienie istnienia zła, nie obowiązują wnioski tego badania. Wyjaśnienie Swinburne'a jest, przynajmniej w teorii, zgodne z modelem *w górę*.

Boska a ludzka natura. Główny zarzut, z którym spotykało się to badanie, brzmi, że nie można porównywać Boga z Elżbietą, ponieważ w pierwszym wypadku mamy do czynienia z istotą z definicji doskonalszą od człowieka, a w drugim po prostu z człowiekiem. Tak więc dopuszczalne jest założenie, że boski plan istnieje, ale jest zbyt skomplikowany dla zrozumienia przez człowieka, ale niedopuszczalne założenie, że motyw Elżbiety są niemożliwe do zrozumienia przez nikogo innego.

Twierdzę, że jest to niebezpieczny argument. Dlaczego?

Wyobraźmy sobie następującą sytuację. W pewnym państwie po wojnie domowej doszedł do władzy dyktator. W czasie wojny wsławił się odwagą, ale też miłością dla wrogów, dlatego objęcie przez niego funkcji prezydenta zostało przyjęte z pewną ulgą. Teraz dyktator urzęduje w swoim pałacu w stolicy kraju, z rzadka ukazuje się swoim rodakom, natomiast z jego gabinetu co rusz wychodzą rozkazy: rozstrzelać tego, uniewinnić tamtego, przesiedlić tę wioskę, pomóc mieszkańcom tamtej, tym zarekwirować plony, tamtym oddać majątek. Najbardziej prawdopodobny wzorzec, tłumaczący zachowanie dyktatora, to *oszałał albo jest złym człowiekiem*. Jednak urzędnicy prezydenta bronią go, mówiąc,

że nikt nie zna planu, zgodnie z którym prezydent postępuje, i że w ostatecznym rozrachunku okaże się, że wszystko to doprowadzi do większego dobra obywateli. Wszelką krytykę urzędnicy ucinają: *dopóki nie udowodnisz, że prezydent nie ma ukrytego planu, nie będziemy cię słuchać, boiemy nie wierzymy w jego mądrość i dobre intencje.*

Co ważne dla tego argumentu: zgodnie z tym, co mówią urzędnicy, prezydent jest istotą doskonalszą od zwykłych ludzi – pokonał wrogów, wykazał się podczas wojny niebywałą odwagą, ma szlachetne urodzenie i niebiosy mu sprzyjają (czyż w dniu urodzin Kim Ir Sena nie rozkwitły kwiaty, a jaskółki nie przemówiły ludzkim głosem?). Dodanie do systemu przekonań urzędników założenia o wyższości prezydenta nad zwykłymi obywatelami, które to założenie pozwala bronić go przed krytyką, jest w pełni analogiczne przyjęciu, że Bóg z definicji jest doskonalszy, a więc nie należy dziwić się nieujawnianiu przez niego swoich planów, których zapewne ludzie, jako istoty ułomne, i tak by nie zrozumieli.

Teraz, jeżeli mechanizm usprawiedliwiania niepoznawalnym planem zostanie uznany za poprawny sposób argumentacji, to otwiera to drzwi do usprawiedliwiania postępowania zarówno Kim Ir Sena, jak i Elżbiety. Omawiany eksperyment pokazał, że nie jest to możliwość czysto teoretyczna, ale że wierzący uatorowani wcześniej przez historię o Bogu częściej usprawiedliwiali Elżbietę niż wierzący, którzy nie zastanawiali się, jak wyjaśnić boską niechęć do przeciwdziałania złu. Wśród niewierzących ten efekt nie występował.

Jeżeli chciałbym ten efekt zobrazować w przejawiony sposób, powiedziałbym, że wierzący, który zobaczy na ulicy matkę bijącą dziecko za odrobienie lekcji, zareaguje, chyba że chwilę wcześniej czytał *Teodyceę* Leibniza, Swinburne'a lub Plantingę. Ateista zareaguje natomiast niezależnie od tego, czy pod ręką ma akurat Leibniza, czy *Anię z Zielonego Wzgórza*.

Filozofia eksperymentalna od momentu swojego powstania spotykała się z zarzutami. Niektóre odnosiły się do wyników konkretnych badań, ale przeważnie były wysuwane przeciwko kierunkowi jako całości. Przede wszystkim zarzuty dotyczyły dominującego w filozofii eksperymentalnej narzędzia, ankiet, oraz samych podstaw nurtu – nieufności wobec intuicji gabinetowej oraz założenia, że intuicje można badać empirycznie.

Co ciekawe, ale i zrozumiale, filozofię eksperymentalną ciepło przyjęli przedstawiciele społecznych nauk empirycznych, tacy jak psychologowie, socjologowie lub ekonomiści (w szczególności ekonomiści eksperymentalni), natomiast u tradycyjnie zorientowanych filozofów wzbudziła silną negatywną reakcję. Przekonanie Wiliamsona, podzielane przez większość filozofów, że *jeżeli cokolwiek można uprawiać w fotelu, to z pewnością filozofię*, kłóci się z obrazem filozofa rozdającego ankiety i testującego hipotezy statystyczne.

Zarzuty przedstawicieli empirycznych nauk społecznych i filozofów gabinetowych są podobne, ale wnioski tych pierwszych są zupełnie inne niż tych drugich. Zgodnie z zarzutami obu grup ankiety często są w stanie ująć intuicji filozoficznych, ponieważ odpowiedzi ankietowanych są funkcją nie tylko intuicji, ale także przekonań co do oczekiwań badacza i wrażliwości na reguły konwersacyjne. Zarzuty w stosunku do metod ankietowych zawiera artykuł Simona Cullen'a *Survey Driven Romanticism* (tytułem nawiązujący do *Intuition Driven Romanticism* z publikacji Weinberga, Nicholisa i Sticha)¹⁰¹ oraz wspomniana przeze mnie

¹⁰¹ S. Cullen, *Survey-Driven Romanticism*.

w pierwszym rozdziale notatka Scholla, ale obaj autorzy uważają, że te problemy da się pokonać, stosując bardziej rzetelne metody badawcze, wypróbowane już w psychologii.

Z drugiej strony w artykule *The Rise and Fall of Experimental Philosophy* Antti Kauppinen na podstawie podobnych zarzutów dochodzi do przekonania, że nie jest możliwa metoda eksperymentalna, która dawałaby dostęp do intuicji filozoficznych.¹⁰² Artykuł Nahmiasa i Nadelhoffer'a, *The Past and Future of Experimental Philosophy*, jest odpowiedzią na tę pesymistyczną konkluzję. Autorzy przyznają, że część zarzutów jest trafna, ale formułują zgoła inny wniosek. Analogicznie do hasła Dewey'a, że *jedynym lekiem na braki demokracji jest więcej demokracji*, stwierdzają, że lekiem na braki metody eksperymentalnej jest po prostu więcej metodologicznie zaawansowanych eksperymentów.

Jednym słowem, zarzuty stawiane przez badaczy z empirycznych nauk społecznych i przez filozofów gabinetowych są podobne – metodologia ankiet często nie potrafi uchwycić tego, do uchwycenia czego została zaprojektowana – ale ci pierwsi traktują to jako problem techniczny, możliwy do przewyżczenia przez poprawę metodologii, a ci drudzy jako wrodzoną i nieuleczalną chorobę filozofii eksperymentalnej, która podważa sens istnienia kierunku. U podstaw zarzutów drugiego rodzaju leży bardzo silne i trudne do obrony założenie, że nie tylko obecne, ale w ogóle wszystkie przyszłe metody filozofii eksperymentalnej nie będą potrafiły dotrzeć do intuicji filozoficznie istotnych, a z drugiej strony, że gabinetowemu użyciu eksperymentów myślowych udaje się dotrzeć do tych intuicji.

W tym rozdziale przedstawię kilka zarzutów oraz postaram się pokazać, dlaczego metoda dyskusji, zastosowana przez nas w eksperymencie dotyczącym *Teorii Sprawiedliwości*, stanowi odpowiedź na dwa z nich.

filozofowie nie odwołują się do intuicji

Zarzut, który chcę omówić na początku, nie jest podnoszony przez krytyków filozofii eksperymentalnej, ale spotkałem się z nim kilka razy przy okazji wykładów, prezentacji i rozmów o filozofii eksperymentalnej,

¹⁰² A. Kauppinen, *The Rise and Fall of Experimental Philosophy*.

jakie miałem przyjemność przeprowadzić. Wynika on, być może, z niewystarczającego przemyślenia metodologii eksperymentów myślowych. A brzmi tak: *odwołanie do opinii zwykłego człowieka jest zwrotem retorycznym i w istocie filozof ma na myśli po prostu, że jego teoria jest prawdziwa lub w swoim przykładzie Kripke nie stwierdza, jaka dla zwykłego człowieka jest relacja nazw własnych do przedmiotów, ale jaka być powinna, albo Gettier stwierdza jedynie, jaka jest jego własna intuicja, i jest mu obojętne, jakie są cudze intuicje.*

Gdyby było tak, jak stwierdza ten zarzut, to odwołania do eksperymentów myślowych przestałyby pełnić funkcję argumentów. Przypuśćmy, że argument Gettier'a ma strukturę: (1) przywołanie definicji wiedzy jako uzasadnionego prawdziwego przekonania, (2) przedstawienie pewnej sytuacji, (3) stwierdzenie: *mimo że przekonanie z sytuacji (2) jest prawdziwe i uzasadnione, to ja, Gettier, nie uważam go za wiedzę.* Wtedy ten argument nie dowodziłby istnienia uzasadnionych i prawdziwych przekonań, które nie są wiedzą, ale jedynie, że dla Gettier'a (i możliwe, że dla nikogo poza nim) istnieją uzasadnione i prawdziwe przekonania, które nie są wiedzą.

Jeżeli Kripke, po wyłożeniu swojej teorii referencji nazw własnych, napisalby, że każdy powinien zgodnie z nią się zachowywać (a nie, że każdy zachowuje się zgodnie z nią), to byłby to postulat, a nie uzasadnienie. Stwierdzenie, że zwykły człowiek poprzez użycie nazw własnych tak a tak odnosi się do przedmiotów, jest argumentem za taką a taką teorią referencji, natomiast stwierdzenie, że zwykły człowiek tak a tak powinien odnosić się do przedmiotów, nie jest argumentem, ale samo domaga się uzasadnienia. Podobnie, jeżeli proponuję pewną definicję lub teorię, i na wezwanie *uzasadnij!* odpowiadam *bo moja definicja jest trafna* lub *bo moja teoria prawdziwa*, to niczego w istocie nie uzasadniłem.

zarzut z refleksji i zarzut z intuicji austińskich

Pod koniec artykułu *Normativity and Epistemic Intuitions* Weinberg, Nichols i Stich odpowiadają na kilka zarzutów wobec ich badania, między innymi na *zarzut z refleksji* [*the reflection objection*] i *zarzut z intuicji austińskich* [*Austrian intuition objection*]. Pierwszy zarzut mówi, że filozoficznie istotne są nie spontaniczne intuicje [*first-off intuitions*], ale takie, które pojawiają się po czasie namysłu. Ankiety nie ujmuje takich intuicji lub co najmniej nie

można przy ich pomocy rozróżnić między spontaniczną a przemyślaną intuicją [*minimally reflective intuition*].

Filozofowie eksperymentalni odpierają ten zarzut na dwa sposoby. Po pierwsze, w ankiecie można do pewnego stopnia kontrolować, czy intuicje są przemyślane; Weinberg, Nichols i Stich zrobili to, prosząc o krótkie pisemne wyjaśnienie swojej odpowiedzi. Po drugie, postawiona hipoteza empiryczna dotyczyła różnic kulturowych i nawet jeżeli część odpowiedzi wypływała ze spontanicznych intuicji, bardziej prawdopodobne wydaje się, że namysł nie zlikwidowałby różnic kulturowych. Co więcej, właściwie równie dobrze mógłby je pogłębić. Zatem to na obrońcy gabinetowego stosowania intuicji semiotycznych leży ciężar udowodnienia, że przemyślane intuicje są niewrażliwe na czynnik kulturowy.¹⁰³

Drugi zarzut mówi, że filozoficznie istotne są tylko *intuicje austyńskie*, czyli takie, które pojawiają się po czasie refleksji i dyskusji. Intuicje austyńskie mogą dać zupełnie inne wyniki, ponieważ – zgodnie założeniem towarzyszącym temu zarzutowi – *kiedy rozsądni ludzie namyślają się razem, są wszelkie powody, by przypuszczać, że ostatecznie dojdą do wspólnego zdania* [*reach a meeting of the minds*].¹⁰⁴

Weinberg, Nichols i Stich odpowiadają na ten zarzut podobnie jak na poprzedni. Przyznają, że rzeczywiście nie zostały przebadane intuicje austyńskie, ale wydaje się, że dyskusja niczego by nie zmieniła. Jeżeli argument odwołuje się do intuicji, to uczestnicy dyskusji mający pewien wspólny kierunek swoich, nawet spontanicznych, intuicji – wynikający na przykład z kultury – będą raczej w swoich rozstrzygnięciach wzajemnie się utwierdzać.

Uważam, że odparcie zarzutu z refleksji jest skuteczne – przy pomocy dodatkowych pytań można kontrolować, czy ankietowany wystarczająco przemyślał swoją odpowiedź. Utrzymanie tego zarzutu wymagałoby precyzyjnego określenia, kiedy namysł nad odpowiedzią jest wystarczająco głęboki lub długotrwały. Poprzeczka musiałaby być ustawiona

¹⁰³ J. Weinberg, S. Nichols, S. Stich, *Normativity and Epistemic Intuitions*, s. 20.

¹⁰⁴ J. Weinberg, S. Nichols, S. Stich, *Normativity and Epistemic Intuitions*, s. 21.

wystarczająco wysoko, aby wykluczyć wszystkie metody empiryczne, ale nie na tyle, aby rzucić podejrzenia na filozofów stosujących metodę gabinetową. Natomiast zarzut z intuicji austynskich jest trudniejszy w obronie i nie da się go oddalić stosując dodatkowe pytania w ankiecie. Pytanie brzmi, czy w ogóle da się dotrzeć do tego rodzaju intuicji przy pomocy metod eksperymentalnych. Pewne rozwiązanie przedstawię w dalszej części rozdziału. Niezależnie od tych wątpliwości pozostaje jeszcze kwestia, dlaczego filozoficznie istotne są intuicje austynskie, a nie na przykład jedynie intuicje przemyślanie dobrze, ale w samotności.¹⁰⁵

model dialogu i model refleksji

Koniecznym elementem każdego eksperymentu myślowego jest intuicyjna ocena przedstawionej w nim sytuacji. Odwołanie do intuicji przybiera formę stwierdzenia postaci *zwykły człowiek w sytuacji S stwierdzi, że X jest C* albo *widzimy, że X jest C*, albo *to oczywiste, że X jest C*, albo *każdy teraz powie, że X jest C*. Według zarzutu Kauppinena sformułowanego w *The Rise and Fall of Experimental Philosophy* filozofowie eksperymentalni źle interpretują te stwierdzenia. Ich interpretacja jest taka:

(E) *W sytuacji S powiemy, że X jest (nie jest) C* jest predykcją o tym, że większość niespecjalistów odpowie, iż X jest (nie jest) C, jeżeli zostanie im przedstawiona sytuacja S.¹⁰⁶

kiedy w istocie

(A) W sytuacji S powiemy, że X jest C jest hipotezą o tym, jak (1) kompetentny użytkownik zadanego pojęcia odpowiedzialby, gdyby (2) rozważył przypadek w wystarczająco dobrych [*sufficiently ideal*] warunkach i (3) na jego odpowiedź miałyby wpływ jedynie względy semantyczne [*semantic considerations*].¹⁰⁷

Kauppinen argumentuje, że, po pierwsze, ankiety nie testują poprawnej hipotezy (A), ale filozoficznie nieistotną hipotezę (E), i, po drugie – co musi być częścią argumentu, aby Kauppinen nie podważył jednocześnie tradycyjnej metodologii – eksperymenty myślowe przeprowadzane przez

¹⁰⁵ To pytanie zadaje np. M. Liao, *A defense of intuitions* i nie znajduje satysfakcjonującej odpowiedzi.

¹⁰⁶ A. Kauppinen, *The Rise and Fall of Experimental Philosophy*, s. 104.

¹⁰⁷ A. Kauppinen, *The Rise and Fall of Experimental Philosophy*, s. 108.

filozofów gabinetowych testują (A). Co w takim razie uprawnia filozofów do wypowiedzania się z gabinetu za innych?

Kauppinen uzasadnia to uprawnienie dwoma modelami w epistemologii pojęć, które nazywa *modelem dialogu* i *modelem refleksji*. Przypuśćmy, że filozof chce ustalić znaczenie pewnego pojęcia. W tym celu rozmawia z ludźmi, przedstawia im różne przykłady, wysłuchuje opinii, konfrontuje rozmówców z konsekwencjami ich stanowisk i sprawdza, czy nadal przy nich pozostają, prezentuje różne interpretacje, starając się odróżnić względy semantyczne od pragmatycznych i tak dalej. Dzięki temu potrafi wyłowić te intuicje, które spełniają założenia (A).¹⁰⁸

Kauppinen najwyraźniej zdaje sobie sprawę, że filozofowie, analizując dane pojęcie, nie zagadują nieznanym – deklaruje, że tradycyjna metodologia przemyślenia w gabinecie eksperymentów myślowych wystarczy, ale uzasadnia ją, wychodząc od modelu dialogu. Przez sam fakt rozmawiania z innymi w ciągu swojego życia, ucząc się posługiwania językiem i tym samym ucząc się posługiwania pojęciami, filozofowie posiadli wiedzę na temat pojęć i sięgają do tej wiedzy właśnie poprzez intuicję. Refleksja w gabinecie spełnia swoje zadanie, ponieważ proces zakładany przez model dialogu odbył się dzięki dotychczasowemu uczestniczeniu w życiu społeczności:

uczestnicząc w odpowiednich grach językowych, byliśmy [my, filozofowie] poprawiani i utwierdzaliśmy się w naszych osądach podczas tej nauki, dzięki czemu posiadliśmy wiedzę normatywną na temat kryteriów poprawnego zastosowania pojęć [...]. Wiemy, jakie ruchy są akceptowane w naszej wspólnocie językowej (tj. akceptowane przez tych, którzy podzielają nasze pojęcia), ponieważ zarówno otrzymaliśmy informację zwrotną [*feedback*] od innych, jak i im jej udzielaliśmy.¹⁰⁹

Główny problem krytyki Kauppinena polega na tym, że albo z góry zakłada, albo nie budzi jego zastrzeżeń to, że filozofowie z reguły pochodzą z określonych klas społeczno-ekonomicznych, i te *odpowiednie gry językowe*, w których uczestniczyli, nie muszą być reprezentatywne dla

¹⁰⁸ A. Kauppinen, *The Rise and Fall of Experimental Philosophy*, s. 124.

¹⁰⁹ A. Kauppinen, *The Rise and Fall of Experimental Philosophy*, s. 127.

całego społeczeństwa. Ich wiedza o pojęciach może być wiedzą o pojęciach angielskiej lub amerykańskiej wyższej klasy średniej z dużych miast, a nie o pojęciach ludzi w ogóle lub chociaż osób anglojęzycznych w ogóle. Badania restrykjonistów eksperymentalnych często dotyczą różnic w intuicjach osób z różnych grup kulturowych i społeczno-ekonomicznych. Bodźcem do ich przeprowadzenia była rezygnacja z założenia, że filozofowie uczestniczyli w wystarczającej liczbie zróżnicowanych gier językowych. Skoro wyniki badań pokazują, że (niekoniecznie przemyślane) intuicje (niekoniecznie kompetentnych) użytkowników języka są różne w różnych kulturach/grupach społecznych, to jeżeli nawet nie podważa to założenia, że intuicje filozofów są reprezentatywne dla ludzi w ogóle, to z pewnością przenosi na filozofów gabinetowych ciężar uzasadnienia tego założenia.

Opis nabywania wiedzy o pojęciach sugeruje też, że filozof podczas nauki używania pojęć jest przeźroczysty – że jego intuicje są elastyczne i jeżeli w wyniku doświadczenia rozpoznaje nowe reguły stosowania pojęcia, sprzeczne z dotychczas przez niego przyjętymi, to jego intuicje dostosowują się. Ale rozpoznana przez psychologię prawidłowość jest raczej taka, że osobie, która przyjęła pewien pogląd, dużo trudniej zmienić go na inny (sprzeczny z poprzednim), niż przyjąć ten inny pogląd komuś, kto jeszcze nie ma wyrobionej opinii. Jeżeli ten mechanizm funkcjonuje także podczas uczenia posługiwania się pojęciami, filozof, rozmawiając z osobami o innych intuicjach, stwierdzi raczej, że to one używają pojęcia niepoprawnie, niż że jego wiedza o pojęciu wymaga uaktualnienia. Podobnie prawidłowość, zgodnie z którą ludzie wolą przebywać z podobnymi do siebie, może powodować, że użytkownicy pojęć dużo częściej są utwierdzani przez otoczenie w swoich intuicjach, niż wynikałoby to z rzeczywistego rozkładu intuicji w społeczeństwie.

Co więcej, ze sformulowaniem (A) wiążą się pewne wątpliwości. Termin *wystarczająco dobre warunki* niesie ze sobą nierozwiązywalne jak na razie problemy. Kauppinen podaje przykłady niewystarczająco dobrych warunków:

[...] nie traktujemy jako miarodajnych sądów o kolorach, wydanych przy pewnych typach oświetlenia, albo sądów o odpowiedzialności, wydanych w stanie wzburzenia.¹¹⁰

To są przykłady dość oczywiste. Gorzej jest już z wynikami badań, które sugerują, że sądy moralne zależą od tego, czy wydający je oglądał wcześniej *Saturday Night Live*, czy film krajoznawczy.¹¹¹ Tu trudno jest uzasadnić, dlaczego intuicje moralne po obejrzeniu *Saturday Night Live* miałyby być bardziej lub mniej odpowiednie dla uzasadniania sądów etycznych niż intuicje widzów krajoznawczego filmu. Kauppinen stwierdza, że nie ma żadnej ogólnej procedury wyznaczania wystarczająco dobrych warunków dla badania danego pojęcia, ale

[...] odczytujemy te warunki z samych praktyk rozwiązywania rozbieżności w czasie lub między podmiotami – zauważając różnice w odpowiedziach i szukając dla tych różnic wyjaśnienia, ostatecznie odrzucamy sądy wydane w pewnych okolicznościach.¹¹²

Po pierwsze, w tym argumencie pojawia się bardzo silne założenie – że w ogóle istnieją jakieś ogólne praktyki rozwiązywania rozbieżności. Po drugie, sądzę, że argument popada w błędne koło. Odrzucamy te intuicje, które nie prowadzą do właściwego wyniku, ale to, który wynik jest właściwy, otrzymujemy przez odwołanie do intuicji. Jeżeli postanowilibyśmy odrzucić intuicje inne niż te, które faktycznie odrzuciliśmy, właściwy okazałby się inny wynik i ten wynik uzasadniałby następnie tę inną decyzję o odrzuceniu intuicji.

Druga wątpliwość dotycząca sformułowania (A) wiąże się z terminem *kompetentny użytkownik*, za którego użyciem kryje się konkretna, wcale nie oczywista wizja funkcjonowania pojęć.

Powiedzieć, że ktoś jest kompetentnym użytkownikiem, to powiedzieć, że jest w stanie stosować pojęcie poprawnie w wystarczającej liczbie przypadków [...] i dzięki temu zająć

¹¹⁰ A. Kauppinen, *The Rise and Fall of Experimental Philosophy*.

¹¹¹ Mam tu na myśli wspomniany już artykuł Valdesolo P, DeSteno D., *Manipulations of Emotional Context Shape Moral Judgment*.

¹¹² A. Kauppinen, *The Rise and Fall of Experimental Philosophy*, s 114.

stanowisko w sprawie tego, co jest poprawnym użyciem. [...] takie normatywne roszczenia do poprawnego użycia (a tym samym do kompetencji) nie mogą być wyprowadzone z faktycznego użycia lub (prostych) dyspozycji do stosowania pojęcia.¹¹³

Kompetentni użytkownicy to ci, którzy stosują pojęcia zasadniczo zgodnie z normami pojęciowymi panującymi we wspólnocie językowej. Aby wyłowić niekompetentnych użytkowników, trzeba zidentyfikować co najmniej najważniejsze normy rządzące pojęciem. Te normy nie mogą pochodzić z faktycznego użycia lub prostych dyspozycji, indywidualnych lub zbiorowych, ponieważ sama normatywna idea roszczenia do poprawności otwiera przepaść między tym, co ludzie są skłonni powiedzieć na temat konkretnego przypadku, a tym, co powinni, w świetle własnych pojęć, powiedzieć.¹¹⁴

Według Kauppinena pojęcia są publiczne, każda wspólnota ma swoje własne pojęcia, które mogą być używane poprawnie, podobnie (tu powrócę do analogii gramatyki i analizy pojęciowej) jak język polski może być używany poprawnie – czyli zgodnie z uchwałami Rady Języka Polskiego. Rozstrzyganie o poprawności użycia należy do kompetentnych użytkowników.

Kompetentny użytkownik to ten, który używa pojęć poprawnie, ale o poprawności użycia mogą decydować tylko kompetentni użytkownicy, to znaczy tylko oni mogą odkryć reguły rządzące poprawnym użyciem pojęć. Aby nie popaść w błędne koło, przy definiowaniu kompetencji trzeba odwołać się do innych niż poprawność kryteriów – na przykład stwierdzić, że kompetentny użytkownik pojęcia to taki, który wystarczająco długo i głęboko zastanawiał się nad tym pojęciem. A ta definicja odsyła do poprzednich zarzutów.

Kauppinen w jednym z przypisów podaje prosty przykład niekompetencji pojęciowej. Zdarzyło się, że został poproszony o podanie jednego z dwóch wiszących obok siebie ręczników – ręcznika koloru brzoskwiowego. Nie potrafił spełnić prośby, ponieważ ręczniki, spośród

¹¹³ A. Kauppinen, *The Rise and Fall of Experimental Philosophy*, s 110.

¹¹⁴ A. Kauppinen, *The Rise and Fall of Experimental Philosophy*, s 112.

których miał wybrać, były dla niego w równym stopniu brzoskwińowe (choćby gdyby drugi ręcznik miał kolor istotnie różny, na przykład granatowy, Kauppinen wiedziałby, który podać).¹¹⁵ Ma to być przykład na to, że niekompetencja pojęciowa jest możliwa. Problem polega na tym, że w tym przykładzie użytkownik pojęcia *kolor brzoskwińowy* sam stwierdza, że jest niekompetentny. Nie do końca wiadomo, jak miałoby wyglądać wykluczenie niekompetentnych użytkowników, którzy twierdzą, że są kompetentni. Sytuacja, w której wspólnota językowa dzieli się na grupy, których członkowie uważają siebie za kompetentnych użytkowników jakiegoś pojęcia, a członków pozostałych grup za niekompetentnych, jest nie tylko możliwa, ale w przypadku pojęć politycznych często spotykana. W każdej takiej grupie pojęcie rządziłoby się nieco innymi regułami i sposób używania tego pojęcia, poprawny w jednej grupie, byłby niepoprawny w świetle reguł innych grup.

Istnieje inny pogląd na kompetencję, nawiązujący do programu Chomsky'ego w lingwistyce i mający oparcie w modularnej teorii umysłu. Jak według Chomsky'ego istnieją tylko idiolekty (języki prywatne), a o języku polskim można mówić tylko dlatego, że niektóre z tych idiolektów wydają się wystarczająco podobne, tak istnieją tylko pojęcia prywatne, a pojęcia publiczne pojawiają się tam, gdzie pojęcia jednostek są wystarczająco do siebie zbliżone. W tym modelu to pojęcia prywatne mają pierwszeństwo i w żadnym wypadku nie można wyciągać normatywnych konsekwencji z reguł rządzących pojęciami publicznymi.

W takim razie gdzie tu jest miejsce na kompetencję? Zasadniczo – nie ma. Każdy jest kompetentnym użytkownikiem własnego idiolektu i tak samo każdy jest kompetentnym użytkownikiem swoich prywatnych pojęć. Można ratować pojęcie kompetencji w oparciu o pojęcie błędu wykonawczego. Według tezy o modularności, co najmniej część zdolności umysłowych przyjmuje postać modułów – wyspecjalizowanych mechanizmów umysłowych o określonych właściwościach.¹¹⁶ Mogę mówić niegrammatycznie we własnym idiolektcie, jeżeli wypowiedziane przeze mnie słowa nie zgadzają się z wyjściem modułu odpowiedzialnego

¹¹⁵ A. Kauppinen, *The Rise and Fall of Experimental Philosophy*, s. 109.

¹¹⁶ J. Fodor, *Modularity of Mind*, s. 37.

za gramatykę mojego idiolektu. Jeżeli jękająca się osoba powie *parę na na na na na niego!* prawdopodobnie mówi niegrammatycznie we własnym idiolektie – jękanie się powoduje błąd wykonawczy, czyli zakłócenia na linii modul-aparat mowy. Łatwo to zweryfikować, prosząc taką osobę o napisanie tego samego zdania. Można mieć także pewność, że ktoś wypowiedział się w swoim idiolektie niegrammatycznie, jeżeli po chwili sam swoją wypowiedź poprawił, a więc przyznał się do niegrammatycznego użycia (jest to przypadek analogiczny do ręczników Kauppinena). Taką niegrammatyczną (ze względu na idiolekt) wypowiedź można próbować określić mianem aktu niekompetencji, chociaż w dalszym ciągu trudno sobie wyobrazić permanentnie niekompetentnych użytkowników idiolektów, jeżeli wykluczyć osoby z poważnymi zaburzeniami neurologicznymi.

Podobnie jest z użyciem pojęć. Poprawnie używam (swojego) pojęcia, jeżeli robię to zgodnie z tym, jak działa leżący u jego podstaw modul. Używam pojęcia niepoprawnie, jeżeli na wejściu lub na wyjściu modulu ma miejsce błąd wykonawczy. W pierwszym wypadku jestem kompetentny w odniesieniu do tego pojęcia, w drugim nie – i jest to chyba jedyne pojęcie kompetencji, które da się utrzymać w świetle tej teorii. Konsekwencją tego podejścia, istotną wobec zarzutów Kauppinena, jest fakt, że użytkownicy pojęć są raczej kompetentni niż niekompetentni i do kontrolowania, czy ankietowany jest kompetentny w użyciu swojego pojęcia, wystarczą właściwie proste zabiegi w postaci prośby o uzasadnienie odpowiedzi. A te są stosowane w większości eksperymentów filozoficznych. We współczesnej filozofii analitycznej ta teoria odzwierciedla się w samej nomenklaturze, kiedy przy analizie pojęcia mówi się o kompetencji leżącej u jego podstaw [*underlying competence*].

Można starać się utrzymywać, że ludzie z reguły nie są kompetentni w użyciu własnych pojęć, skoro czasem wystarczy kilka dodatkowych pytań, aby zaczęli zmieniać swoje poglądy (takie sytuacje Platon przedstawił w swoich wczesnych dialogach, każąc Sokratesowi rozmawiać z przechodniami na ulicach Aten). Jeżeli takie kryterium wystarczy do uznania kogoś za niekompetentnego użytkownika pojęcia, należałby

za niekompetentnych uznać także filozofów, skoro książka Kripkego i artykuł Gettier'a (które były właściwie prezentacją kilku hipotetycznych przykładów) wystarczyły do diametralnej zmiany rozpoznania przez filozofów własnych pojęć referencji i wiedzy.

Najciekawszą konsekwencją publikacji Kauppinena jest propozycja metodologiczna, którą Nadelhoffer i Nahmias wysunęli w odpowiedzi na zarzut z modelu dyskusji. Według nich nic nie stoi na przeszkodzie, żeby to, co opisuje model dyskusji, zrealizować w sformalizowanej postaci eksperymentu, pozbywając się tym samym wad modelu nieformalnego.¹¹⁷ W modelu dialogu nie ma żadnych elementów, których nie dałoby się zasymulować w kontrolowanym środowisku laboratoryjnym. Przewagą *eksperymentalnego modelu dialogu i refleksji*, jak go nazywają autorzy, polega między innymi na tym, że dane mogą być zbierane z reprezentatywnej (a w każdym razie dużo bardziej reprezentatywnej) próby, szybciej, mniej przypadkowo i, przede wszystkim, zebrane dane mogą być przechowywane i przetworzone do postaci ilościowej, kiedy w modelu nieformalnym z przebiegu dyskusji pozostaje niewiele poza ogólnymi wrażeniami, zacierającymi się w czasie wspomnieniami rozmówców.

jeszcze jeden eksperyment

Pierwszy w ogóle eksperyment filozoficzny, jaki przeprowadziliśmy, dotyczył *Teorii Sprawiedliwości* Rawlsa i polegał na badaniu odpowiedzi uczestników badania przed dyskusją i po niej.¹¹⁸ W tym czasie nie znaleźliśmy jeszcze zarzutów Kauppinena ani propozycji metodologicznych Nadelhoffera i Nahmiasa, natomiast zdawaliśmy sobie sprawę, że dla filozofii eksperymentalnej nasza metoda jest nowa. Kiedy natrafiłem na artykuł *The Past and Future of Experimental Philosophy*, okazało się, że metoda dyskusji odpowiada temu, co autorzy nazywają eksperymentalnym modelem dialogu i refleksji.

Co więcej, metoda dyskusji odpowiada na dwa pozostałe zarzuty – zarzut z refleksji i zarzut z intuicji austynskich. Jeżeli do stwierdzenia, że ankietowany przemyślał swoją intuicję, nie wystarczy dodatkowa

¹¹⁷ E. Nahmias, T. Nadelhoffer, *The Past and Future of Experimental Philosophy*, s. 131.

¹¹⁸ Rozdział III, podrozdział *sprawiedliwość i zasada maksimumu u Rawlsa*.

prośba o wyjaśnienie odpowiedzi, to powinna wystarczyć konieczność wysłowienia i uargumentowania swojego stanowiska podczas dyskusji, konieczność odpowiedzi na pytania innych uczestników badania oraz możliwość sformułowania własnych pytań odnośnie cudzych intuicji. Jeżeli to nie jest wystarczającym warunkiem dla uznania, że intuicja została przemyślana, to trudno powiedzieć, co miałyby nim być.

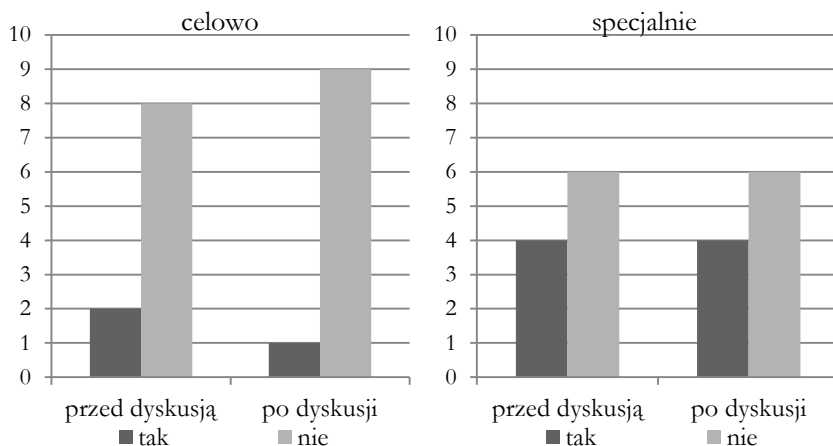
Intuicje austinskie zdefiniowane są jako intuicje pojawiające się po długim czasie refleksji i dyskusji. W kontekście metody dyskusji zasadnicze pytanie brzmi, czy czas eksperymentu jest wystarczający, aby wyłonione w jego trakcie intuicje nazwać austinskimi. Nasze eksperymenty trwały około pół godziny do godziny – trudno nazwać ten czas *długim* – ale kończyły się dopiero wtedy, kiedy uczestnicy stwierdzali, że wystarczająco omówili swoje argumenty i nie podadzą już nowych. Z tego względu przyjmuję, że odpowiedzi uzyskane w badaniach przeprowadzanych wraz z intuicjami austinskimi lub przynajmniej pokazują, w którym kierunku odpowiedzi by podążały, gdyby dyskusja była prowadzona na tyle długo (np. w postaci kilku lub kilkunastu osobnych spotkań), aby ten okres uznać za wystarczający w świetle definicji intuicji austinskich.

W zarzutach z intuicji przemyślanych i intuicji austinskich występuje założenie, że, odpowiednio, namysł i dyskusja będą miały istotny wpływ na odpowiedzi – gdyby powierzchowne i przemyślane/austinskie intuicje były takie same, nie byłoby powodu na przypisywanie pierwszym mniejszej wagi niż drugim. W badaniu dotyczącym zasłony niewiedzy nie znaleźliśmy statystycznie istotnej różnicy w intuicjach przed i po dyskusji. To było inspiracją dla jeszcze jednego eksperymentu.

W badaniu wykorzystaliśmy historie o prezesie pomagającym lub szkodzącym środowisku.¹¹⁹ Wybór padł na badanie Knobe'a tylko dlatego, że chcieliśmy użyć przypadku, co do którego spodziewaliśmy się dużej różnicy w intuicjach. Eksperyment został przeprowadzony na dwóch dziesięcioosobowych grupach studentów ekonomii, w wieku 21-23 lata. Ankiety zawierały dwie historie – jedną, w której prezes pomagał środowisku, i drugą, w której szkodził. W każdej grupie rozdaliśmy pięć

¹¹⁹ Rozdział II, podrozdział: *działanie intencjonalne*.

ankiet o kolejności przykładów pomógł-zaszкодził i pięć o kolejności zaszкодził-pomógł (aby kontrolować ewentualny wpływ kolejności). Ankieta rozdana pierwszej grupie (nazwę ją C) zawierała pytanie *czy prezes celowo zaszкодził (pomógł) środowisku?* z odpowiedziami *tak, prezes zaszкодził (pomógł) środowisku celowo* i *nie, prezes nie zaszкодził (pomógł) środowisku celowo*. Ankieta w drugiej grupie (S) zawierała pytanie *czy prezes specjalnie zaszкодził (pomógł) środowisku,* z analogicznymi odpowiedziami do wyboru.



Rysunek 10. Wpływ dyskusji na intuicje filozoficzne – rozkład odpowiedzi na pytania *czy prezes celowo/specjalnie zaszкодził środowisku?*

W grupie C przed dyskusją nikt nie stwierdził, że prezes pomógł środowisku celowo, a dwie osoby (na dziesięć) stwierdziły, że zaszкодził celowo. Po dyskusji nikt nie uważał, że prezes pomógł celowo i jedna osoba uważała, że zaszкодził celowo – zdanie zmieniła jedna osoba. W grupie S przed dyskusją nikt nie uważał, że prezes pomógł środowisku specjalnie i cztery osoby (na dziesięć) uważały, że zaszкодził specjalnie. Po dyskusji rozkład odpowiedzi nie zmienił się. Testy statystyczne nie wykazały istotnej różnicy w odpowiedziach przed i po dyskusji.¹²⁰ Natomiast, co ważne, uczestnicy rozumieli nawzajem swoje argumenty

¹²⁰ Dla grupy C w teście Fishera i w teście proporcji z hipotezą zerową *stosunek liczb odpowiedzi na pytanie o szkodzenie jest taki sam przed i po dyskusji* i hipotezą alternatywną *jest większy p-wartość wyniosła 0,5*. W tych testach dla grupy S (z hipotezą alternatywną *jest różny*) p-wartość wyniosła 1.

i przyznawali kontrargumentom pewną siłę przekonywania – jednak na tyle dużą, aby zmienić zdanie.

Zatem metodą dyskusji zostało przeprowadzonych sześć eksperymentów, cztery dotyczące intuicji sprawiedliwości rozdzielczej i dwa dotyczące terminów *celowo* i *specjalnie*. W tych pierwszych dyskusja zmieniała intuicje nieznacznie, w tych drugich nie zmieniała intuicji prawie w ogóle. Czy w takim razie można przyjąć, że dyskusja nie ma istotnego wpływu na intuicje filozoficzne? Oczywiście, jeszcze nie. Zostały zbadane intuicje dotyczące tylko dwóch wybranych pojęć, przez co wyników nie można uogólniać na wszystkie intuicje filozoficzne. Natomiast w tych dwóch zbadanych przypadkach można uznać, że dyskusja nie zmienia odpowiedzi na tyle, aby rozróżnić między intuicjami austińskimi a powierzchownymi.

Nikt, kto został przekonany lub miał okazję kogoś przekonać w dyskusji, w której ważną formą argumentacji były eksperymenty myślowe, nie zgodzi się, że sądy oparte na intuicjach nie poddają się zmianom. Co więcej, wyniki restrykcyjnych demonstrujących, jak intuicje zależą od tak różnych czynników jak kolejności pytań w ankiecie albo czystość biurka, przy którym badany ją wypełnia, z pewnością nie pozwalają przyjąć, że intuicje są stałe. W związku z tym można zadać ciekawe pytanie badawcze: co powoduje, że intuicje filozoficzne mogą być wrażliwe na czynniki tak nieistotne, jak czystość biurka, a niewrażliwe na tak istotne, jak wymiana argumentów? Lub które intuicje są niewrażliwe na nieistotne czynniki, które są niewrażliwe na argumenty, i jak te dwa podziały mają się do siebie? Oczywiście odpowiedzi nie da się udzielić z wnętrza gabinetu.

filozof jako ekspert

Zarzut z biegłości [*expertise*] stwierdza, że dla filozofii istotne są tylko intuicje filozofów, a tym samym wyniki eksperymentów przeprowadzanych na osobach bez wykształcenia filozoficznego nie mają żadnego znaczenia. Dlaczego? Najprostsza odpowiedź brzmi: *To byłoby absurdalne, gdyby fizyk lub biolog przeprowadzał ankiety dotyczące intuicji zwykłych ludzi dotyczących fizyki lub biologii. [...] Tak jak fizyk nie bierze pod uwagę laickich poglądów na fizykę [folk physics], tak filozof nie musi brać pod uwagę filozofii laików*

[folk philosophy].¹²¹ Ta uwaga jest trafna w odniesieniu do technicznych lub specjalistycznych – jeżeli można je tak nazwać – problemów filozoficznych, na przykład czy zadany system formalny, którego celem jest opisanie logiki indukcji, nie jest sprzeczny, albo czy umysł ma modularną budowę. Zresztą przy rozwiązywaniu takich problemów nie stosuje się z reguły eksperymentów myślowych, a filozofowie nie argumentują w takim wypadku, że *zwykli człowiek uznałby, że formalizacja jest sprzeczna, a umysł modularny*. Natomiast niektóre podstawowe dla filozofii zagadnienia, takie jak tożsamość w czasie przedmiotów lub osób, istnienie wolnej woli albo reguły sprawiedliwego ustroju, nie dałyby się nawet sformułować, gdyby nie intuicje laików. *Pomyśl, jak marginalizujące byłoby stwierdzić „my, filozofowie, dużo napisaliśmy na temat tego, co nazywamy «odpowiedzialnością moralną», ale nasze pojęcie jest zupełnie niezwiązane z tym, co zwykli ludzie rozumieją przez homonimiczny termin «odpowiedzialność moralna»”*.¹²² Jednocześnie nie jest też prawdą, że specjaliści z dziedzin niefilozoficznych nie biorą pod uwagę intuicji laików – na przykład pojęcie choroby w medycynie nie może za nadto odbiec od tego, co rozumieją pod nim zwykli ludzie.

Dlatego uzasadnienie zarzutu z biegłości trzeba sformułować inaczej: istotne są tylko intuicje filozofów, ponieważ dzięki swojemu wykształceniu i doświadczeniu w przeprowadzaniu eksperymentów myślowych analizują pojęcia po prostu lepiej i są (wystarczająco) odporni na filozoficznie nieistotne czynniki, na przykład na zmianę kolejności prezentacji eksperymentów myślowych. W szczególności, wyniki restrykcjonistów eksperymentalnych nie stanowią argumentu przeciwko zasadności użycia intuicji w projektach neoplatońskich i normatywnych.

W artykule *Are Philosophers Expert Intuiters?* Jonathan Weinberg, Chad Gonnerman, Cameron Buckner i Joshua Alexander, opierając się na psychologicznych studiach nad biegłością, argumentują, że trening filozoficzny nie jest wystarczający dla uznania filozofów (osób z wykształceniem filozoficznym, czytających, komentujących i publikujących artykuły w periodykach filozoficznych) za ekspertów.

¹²¹ E. Machery, R. Mallon, S. Nichols, S. Stich, *Semantics, cross-cultural style*, s. 8.

¹²² E. Machery, R. Mallon, S. Nichols, S. Stich, *Semantics, cross-cultural style*, s. 9.

Skoro intuicje osób bez wykształcenia filozoficznego są wrażliwe na nieistotne czynniki, na filozofach gabinetowych spoczywa ciężar dowodu, że ich intuicje zachowują się inaczej.¹²³ Z kolei według Williamsona obowiązuje tu domniemanie niewinności i to filozofowie eksperymentalni muszą udowodnić, że trening filozoficzny nie uodparnia filozofów na czynniki filozoficznie nieistotne.¹²⁴

Nawet jeżeli zarzut z biegłości poddaje w wątpliwość dotychczasowe wyniki, nie stanowi zagrożenia dla filozofii eksperymentalnej jako kierunku. Jak zauważa Joachim Horvath, wobec tego zarzutu restrykcjoniści eksperymentalni (bo ich przede wszystkim on dotyczy) powinni po prostu przeprowadzić eksperymenty na filozofach.¹²⁵ Pierwsze takie badania są już prowadzone. Vaesen i Peterson pokazali, że intuicje filozofów zależne są od ich ojczystego języka.¹²⁶ W badaniu Cushmana i Greene'a, wykorzystującym dylemat wagonika [*trolley problem*], odpowiedzi zarówno filozofów, jak i niefilozofów są wrażliwe na kolejność pytań. Od kolejności zależy też, czy filozofowie stwierdzają, że stosują zasadę podwójnego efektu.¹²⁷ Zatem jak na razie hipoteza o biegłości – intuicje filozofów są odporne na nieistotne czynniki – nie potwierdza się.

Williamson jako jeden z argumentów na rzecz hipotezy o biegłości wymienia fakt, że poziom niezgody co do wyników eksperymentów myślowych jest niższy wśród doświadczonych filozofów niż wśród studentów filozofii. Fakt ten miałby wynikać stąd, że trening filozoficzny powoduje, iż intuicje filozoficzne są coraz trafniejsze, a tym samym odpowiedzi, w miarę trwania treningu, zbliżają się do siebie. Williamson ma świadomość, że konwergencja intuicji może także wynikać stąd, że osoby z intuicjami innymi niż standardowe w końcu przejmują standardowe (tj. podzielane przez większość filozofów) intuicje w wyniku socjalizacji w środowisku akademickim, albo – jeżeli upierają się przy swojej pierwotnej intuicji – są wyłączeni ze obiegu naukowego: nie są

¹²³ J. Weinberg, Ch. Gonnerman, C. Buckner i J. Alexander, *Are Philosophers Expert Intuiters?*

¹²⁴ T. Williamson, *Philosophical Expertise and the Burden of Proof*, w recenzji.

¹²⁵ J. Horvath, *How (Not) to React to Experimental Philosophy*, s. 471-472.

¹²⁶ K. Vaesen, M. Peterson, *Knowledge and the Quality of Information*, w recenzji.

¹²⁷ F. Cushman, J. Greene, *The philosopher in the theater*, [w:] *Social psychology of morality: The origins of good and evil*, red. M. Mikulincer, P. Shaver.

zatrudniani na uniwersytetach, a ich publikacje nie są akceptowane. Jednak dopóki nie ma dowodów, że taki proces wykluczania istnieje i przybiera zauważalne rozmiary, jest to nic więcej niż teoria spiskowa.¹²⁸

Takie dowody zaczynają się pojawiać. W artykule *Gender and Philosophical Intuition* Stephen Stich i Wesley Buckwalter przywołują dane dwojakiemu rodzaju: dane z dotychczas przeprowadzonych eksperymentów filozoficznych oraz dane statystyczne z wydziałów filozofii Rutgers University (dane z 10 lat) i z University of Minnesota.¹²⁹ Okazało się, że na początku studiów na kursy filozoficzne równie chętnie co studenci zapisują się studentki, ale z czasem odsetek studentek drastycznie maleje (z 46% na początku studiów do 26% pod koniec studiów – dane z Rutgers University). Wśród zatrudnionych w wydziałach filozoficznych na pozycjach naukowych tylko 20% to kobiety, w tym 16% to profesorowie (dane z 20 najlepszych amerykańskich wydziałów filozoficznych). Jednocześnie z badań przeprowadzonych przez filozofów eksperymentalnych wynika, że kobiety mają inne niż mężczyźni intuicje w odniesieniu do najczęściej przedstawianych na zajęciach eksperymentów myślowych.¹³⁰ Jaki związek łączy te fakty?

Sugerujemy, że co najmniej część dysproporcji płci w filozofii można wytłumaczyć *efektem selekcji* [*selection effect*]. Studenci zaczynają kursy filozoficzne z różnymi intuicjami na temat wielu typowych eksperymentów myślowych i, jak pokazaliśmy, w wielu przypadkach istnieją statystycznie istotne różnice w intuicjach kobiet i mężczyzn. Jednak większość członków wydziału [filozofii], których pozycja [jako prowadzących zajęcia] pozwala im rozstrzygać, które intuicje są poprawne (i *oczywiste*), jest i zawsze była płci męskiej. Tak więc studentki częściej niż studenci dowiadują się, że ich intuicje dotyczące eksperymentów myślowych omawianych na zajęciach odstają od intuicji prowadzącego. Jeżeli rzeczywiście studenci (o boję płci) mniej chętnie zapisują się na kolejne kursy filozoficzne, gdy nie dzielają intuicji osób prowadzących zajęcia, to wszystkie

¹²⁸ T. Williamson, *Philosophical Expertise and the Burden of Proof*, s. 20.

¹²⁹ W. Buckwalter i S. Stich, *Gender and Philosophical Intuition*.

¹³⁰ Autorzy przebadali sześć eksperymentów, między innymi *skrzypka, mózg w naczyniu i chiński pokój*.

elementy potężnego i kumulatywnego efektu selekcji zgadzają się – efektu selekcji, który „odfiltrowuje” więcej kobiet niż mężczyzn.¹³¹

W kontekście zarzutu z biegłości istotne jest to, że efekt selekcji – jeżeli rzeczywiście zachodzi – wyjaśnia także konwergencję intuicji. Zgoda wśród filozofów co do oceny eksperymentów myślowych przestaje świadczyć na rzecz istnienia czegoś takiego jak biegłość w korzystaniu z intuicji filozoficznej, a tym samym intuicje osób o wykształceniu filozoficznym nie są wcale bardziej godne zaufania niż intuicje osób bez takiego wykształcenia.

Dla mnie zarzut z biegłości jest interesujący z jednego względu. Aby trening prowadził do biegłości w jakiejś dziedzinie, potrzeba między innymi, żeby jego elementem była powtarzana praktyka z szybką i trafną informacją zwrotną. Williamson uważa, że posługiwanie się eksperymentami myślowymi realizuje ten wymóg: podczas studiów filozoficznych student czyta artykuły i książki, w których eksperymenty myślowe stanowią kluczowy element argumentacji, rozmawia o nich w czasie zajęć, sam też formułuje eksperymenty myślowe, które następnie są oceniane przez prowadzącego lub grupę.¹³² Ale na jakiej podstawie otrzymuje informację zwrotną, czyli skąd wie, że przeprowadził eksperyment myślowy poprawnie? Lekarza-eksperta można poznać po tym, że przepisane przez niego leki lub przeprowadzone operacje pozwalają wyleczyć pacjenta. Strażak jest ekspertem w gaszeniu pożarów, jeżeli robi to skutecznie i szybko. Jak odróżnić filozofa-eksperta, biegłego w przeprowadzaniu eksperymentów myślowych, od amatora?

Próbując rozwiązać jakiś problem przy pomocy eksperymentu myślowego, student będzie zadowolony, jeżeli przekona do swojego rozwiązania pozostałe osoby z grupy ćwiczeniowej lub prowadzącego, czytającego zaliczeniowy esej. Informacja zwrotna, którą otrzymuje, to *tak, przekonałaś/eś mnie* lub *nie, nie przekonałaś/eś mnie*, a dokładniej *tak, Twój eksperyment wywołał we mnie intuicję na poparcie Twojej tezy* lub *nie, Twój eksperyment nie wywołał we mnie intuicji, na których Ci zależało*. Czasem student

¹³¹ W. Buckwalter i S. Stich, *Gender and Philosophical Intuition*, s. 30.

¹³² T. Williamson, *Philosophical Expertise and the Burden of Proof*, s. 14-15.

może usłyszeć, dlaczego jego eksperyment był nieskuteczny – na przykład, że w swojej ocenie nie zwrócił uwagi na pewne szczegóły, a gdyby zwrócił, to jego intuicje byłyby inne. Prowadząc lub studenci z grupy mogą mu w odpowiedzi dać do przemyślenia inny eksperyment myślowy, skonstruowany przez analogię, ale w założeniu wywołujący inne intuicje. *Jeżeli ze swojego eksperymentu wysuniesz takie a takie konsekwencje, to w takim razie musiałbyś się zgodzić także na podobną ocenę następującej sytuacji* (tu zostaje sformułowany nowy eksperyment myślowy), *ale Twoja intuicja dotycząca tej sytuacji zapewne będzie inna*. Wtedy student może próbować przeformułować swój eksperyment tak, żeby ocalić pierwotny wniosek, ale obronić się przed zarzutami. Jeżeli tak przebiega trening filozoficzny (w odniesieniu do eksperymentów myślowych), to miarą sukcesu w formułowaniu eksperymentów myślowych jest odsetek osób, w których eksperyment wywołał takie intuicje, na jakich zależało jego autorowi. Wyształcenie filozoficzne może prowadzić do biegłości – w projektowaniu eksperymentów myślowych dla uzyskania spodziewanych intuicji i w zwracaniu uwagi na te elementy eksperymentu myślowego, których nieuwzględnienie może prowadzić do innej intuicji filozoficznej niż w przypadku pełnego zrozumienia.

W tym miejscu pomocna może być pewna analogia. Mogę twierdzić, że pewien żart jest śmieszny, nawet jeżeli nikt inny tak nie uważa – ponieważ ludzie mogą tego żartu po prostu nie rozumieć. Jeżeli jednak dostarczę słuchaczom informacji potrzebnych do zrozumienia żartu, oni go rozumieją, ale nadal w przeważającej większości nie uznają za śmieszny, to dalej mogę twierdzić, że ten żart jest śmieszny dla mnie, ale nie, że moi słuchacze mają gorsze ode mnie poczucie humoru.

Przypuśćmy, że dwie osoby, Mikołaj i Natalia, dyskutują na temat dopuszczalności aborcji, przy czym M uważa aborcję za niedopuszczalną moralnie. Jak N mogłaby argumentować przeciwko stanowisku M? Mogłaby, między innymi:

(a) przedstawić eksperyment myślowy, którego celem byłaby zmiana intuicji M. *Budzisz się podłączony przy pomocy aparatury medycznej do nieprzytomnego, bardzo sławnego skrzypka. Jest chory i tylko pozostawienie go podłączonego do Ciebie w ten sposób przez dziewięć miesięcy pozwoli*

*mu wyzdrowieć; jeżeli zostanie odłączony wcześniej, umrze. Po dziewięciu miesiącach będzie mógł zostać odłączony i żyć dalej bez przeszkód. Czy można odmówić Ci prawa do odłączenia się od skrzypka przed upływem tego czasu?*¹³³ N zakłada przy tym, że, po pierwsze, M intuicyjnie oceni, iż nie można odmówić osobie z eksperymentu prawa do odłączenia się i, po drugie, M uzna problem dopuszczalności aborcji i dopuszczalności odłączenia się od skrzypka za wystarczająco analogiczne, aby odpowiedź na eksperyment była jednocześnie odpowiedzią w sprawie aborcji.

(b) N mogłaby także opisać naprawdę okropne warunki, w których wychowują się dzieci w domach dziecka – na przykład gwałty dokonywane na młodszych dzieciach przez starszych wychowanków – i przedstawić wyniki badań, zgodnie z którymi dopuszczenie aborcji prowadzi do zmniejszenia przestępczości.¹³⁴ W takim wypadku N spodziewa się, że M zachowa swoje intuicje co do samego zabiegu, ale w procesie równowagi refleksyjnej uzna, że silniejsze są intuicje za dopuszczalnością aborcji, wywołane przez opisane przez N fakty.

(c) N mogłaby próbować pokazać, że intuicje moralne M mają źródło w etycznie nieistotnych czynnikach. Może poprosić M zastanowienie się, czy miałby podobne intuicje, gdyby rozwój płodu u człowieka przebiegał nieco inaczej. Na przykład, że – dopóki nie zacznie odczuwać – płód nie przypomina wyglądem noworodka, ale na przykład zieloną galaretowatą masę. Jeżeli M stwierdziłby, że w tym wypadku nie miałby tak silnych intuicji przeciwko dopuszczalności aborcji, oznaczałoby to, że są one wywołane przez fakt, że płód przypomina kształtem dziecko. Jeżeli dla M wygląd jest czynnikiem etycznie nieistotnym, nie powinien brać pod uwagę tych intuicji, które są wywołane tylko przez zewnętrzne podobieństwo, a więc nawet jeżeli je nadal żywi, powinien je odizolować.

Te trzy argumenty są przykładami strategii argumentowania przeciwko stanowisku uzasadnionemu intuicjami. Można próbować (a) zmienić intuicje przez przywołanie wystarczająco analogicznego eksperymentu myślowego o innym wyniku. Można próbować (b) zmienić ostateczną ocenę przez dołączenie dodatkowych intuicji. Można (c) podważyć

¹³³ Na podstawie: J. Thomson, *A Defense of Abortion*.

¹³⁴ Por. S. Levitt, S. Dubner, *Freakonomia. Świat od podszewki*, s. 163.

zaufanie do danej intuicji. W przypadku (a) przekonywany nie musi uznać analogii ani nie musi ocenić tego analogicznego eksperymentu tak, jak rozmówca. Jeżeli ktoś uznaje, że osoba podłączona do skrzypka nie ma prawa się odłączyć, nie można po prostu powiedzieć, że jego intuicje są niepoprawne. W przypadku (b) wynik refleksyjnej równowagi może być inny, niż u przekonującego; jeżeli dla kogoś fakt, że oddane do domów dziecka dzieci z niechcianych ciąż często przeżywają tam piekło, nie przechyla szali na rzecz dopuszczenia aborcji, nie można zarzucić mu błędu. W przypadku (c) powody, dla których jeden podważa trafność intuicji, dla innego mogą być niewystarczające – może nie zgadzać się na kryteria, które mają odróżniać poprawne intuicje od niepoprawnych.

Właściwie czyjąś intuicyjną ocenę można otwarcie uznać za błędną jedynie w przypadku niespójności logicznej lub wyprowadzeniu jej z fałszywych przekonań faktycznych. Na przykład, jeżeli M przyznaje, że jego intuicje wywołane są zewnętrznym podobieństwem płodu do dziecka, uznaje, że wygląd jest czynnikiem nieistotnym etycznie, i jednocześnie nadal ocenia aborcję na podstawie tych intuicji, jest logicznie niespójny. Jeżeli M sprzeciwia się aborcji, ponieważ jest przekonany, że płód cierpi podczas zabiegu, jego intuicje można odrzucić, ponieważ opierają się na założeniach niezgodnych z wynikami badań medycznych. Poza tymi przypadkami, jeżeli N przedstawi M swoje eksperymenty myślowe, ale nie wywoła intuicji, których się spodziewała, nie da się nic więcej zrobić – M ma nieprzystające do N, ale spójne intuicje, i nie ma (prostych) kryteriów, według których można odróżnić między intuicjami M a N. Natalia nie może po prostu powiedzieć: *Mikołaj, jesteś spójny w tym, co mówisz, ale Twoje intuicje są gorsze.*

W ten sposób powraca platoński obraz filozofa zadającego pytania, ale nienarzucającego swojego zdania – akuszeru asystującego przy porodzie. Student filozofii trenuje się w zadawaniu pytań i może dojść w tym do biegłości, ale nie może być ekspertem w odpowiadaniu na nie. W klasycznej metodologii eksperymentów myślowych filozof nie tylko zadaje pytania, ale sam na nie odpowiada i, zgodnie z zarzutem z biegłości, jego odpowiedzi są lepsze niż odpowiedzi laików. W filozofii eksperymentalnej filozof zadaje pytania i formułuje hipotezy na temat

tego, jak badani odpowiedzą. Jeżeli wynik eksperymentu obali hipotezę, filozof musi się z tym pogodzić niezależnie od tego, jak sam odpowiedzialby na pytania z własnej ankiety.

Większość dotychczasowych eksperymentów wykorzystywała proste ankiety do badania intuicji, ale można zaimplementować postaci eksperymentu, na przykład, hipotetyczną rozmowę Mikołaja z Natalią. W zależności od odpowiedzi ankietowanego można prezentować mu kolejne eksperymenty myślowe lub zwracać uwagę na te szczegóły, których mógł nie uwzględnić. Filozof eksperymentalny wykorzystuje swoją biegłość przy projektowaniu eksperymentów myślowych i zadawaniu dodatkowych pytań, ale intuicyjną ocenę pozostawia już ankietowanym. Mechanizm doboru pytań w zależności od dotychczasowych odpowiedzi badanego może służyć ocenie nie tylko pojedynczych eksperymentów, ale całej dyskusji filozoficznej nad wybranym problemem: po prezentacji wybranego eksperymentu myślowego można przedstawić do oceny ankietowanym kontrargumenty, które pojawiły się w rzeczywistej dyskusji na łamach pism filozoficznych, kontrargumenty do tych kontrargumentów i tak dalej.

Jeżeli odpowiedzi z takiej ankiety służą potem do wnioskania o tym, co składa się na badane pojęcie, to jest to program analizy eksperymentalnej. Jeżeli celem badania będzie rozróżnienie intuicji poprawnych od niepoprawnych przez odkrycie mechanizmu psychologicznego, będzie to deskryptywizm eksperymentalny. Takie badanie łamie regułę, że nie można powiedzieć Mikołajowi, że jego intuicje są gorsze – dlatego właśnie napisałem, że nie ma *prostych* kryteriów, według których można odróżnić między intuicjami M a N. Przypuśćmy, że w oparciu o argumenty N powstanie ankieta i że okaże się, iż ankietowani, którzy znają blisko osoby wychowane w domach dziecka lub same zostały w nich wychowane, istotnie częściej uważają argument N za przeważający (lub co najmniej silny), niż ankietowani, którzy takich osób nie znają. Można stwierdzić, że intuicje M są nietrafne (nawet jeżeli M jest przekonany, że opowieści byłych wychowanków sierocińców nie wpłynęłyby na jego odpowiedź), tylko przyjmując, że ankietowani znający takie osoby lepiej potrafią ocenić siłę argumentu N. Zatem wniosek będzie obowiązywał

tylko tych, którzy zgadzają się na takie kryterium rozróżniania intuicji trafnych od nietrafnych. W badaniach restrykcjonistów sytuacja jest prostsza: wyniki z reguły wskazują, że w odniesieniu do pewnych problemów nie ma lepszych i gorszych intuicji, bo wszystkie są zle.

Jeżeli filozofia jako dyscyplina nauki – a dokładniej ta jej część, w której istotną rolę gra intuicja filozoficzna – funkcjonuje w ten sposób, że publikowane i dyskutowane są eksperymenty myślowe najlepiej ujmujące intuicje czytelników, to minimalny spójny model filozofii mówiłby, że filozofowie zajmują się zasadniczo wzajemnym trafianiem w swoje intuicje. Teorie filozoficzne byłyby jak utwory muzyczne: niektóre wpadają w ucho, inne nie, ale te pierwsze nie mają żadnej wartości ponad to właśnie, że wpadają w ucho. Model metafizycznie ciężki mówi, że intuicja jest narzędziem poznania świata. Wtedy teorie filozoficzne oparte o intuicje pełnią podobną rolę, co teorie nauk empirycznych – opisują obiektywny świat. Pomiędzy tymi modelami znajduje się obraz filozofii jako analizy pojęciowej – mimo, że pojęcia są ludzkim wytworem (i niezależnie od tego, na ile są intersubiektywne, a na ile tylko subiektywne), rozjaśnianie ich znaczenia jest istotne chociażby dlatego, że są używane przy formułowaniu życiowych potrzeb i celów. Analiza eksperymentalna i deskrytywizm eksperymentalny mogą być uprawiane zgodnie z każdym z tych modeli. Wyniki restrykcjonistów eksperymentalnych sugerują, że w odniesieniu do wielu problemów model metafizycznie ciężki, a czasem nawet analiza pojęciowa, są niepoprawne, bo intuicja nie spełnia wymaganych od niej warunków. Albo trzeba znaleźć inne narzędzie do rozstrzygania tych problemów, albo trzeba te problemy zarzucić, albo trzeba przyjąć wizję filozofii jako dość oryginalnej formy niezobowiązującej rozrywki.

Rzadko można mieć szczęście obserwować narodziny nowego ruchu filozoficznego. Taką przyjemność mieli Ayer i Quine, uczęszczający na spotkania Koła Wiedeńskiego, lwowscy uczniowie Tadeusza Kotarbińskiego i słuchacze Moore'a i Russella po ich wyzwoleniu się z wpływu heglizmu. Być może w przypadku filozofii eksperymentalnej mamy do czynienia z podobną sytuacją.

Czy filozofia eksperymentalna jest nowym kierunkiem filozoficznym, tak jak kiedyś filozofia analityczna lub fenomenologia? Filozofowie eksperymentalni są w pełni świadomi, że stosowane przez nich metody są dla współczesnej filozofii nowe. Z drugiej strony podkreślają, że eksperymenty są dodatkowym narzędziem filozofa i nie mają na celu wyprzeć dotychczasowych środków badawczych, i deklarują powrót do zapomnianej przez gabinetową filozofię analityczną tradycji zajmowania się człowiekiem takim, jaki jest naprawdę. Filozofia eksperymentalna, niezależnie od tego, czy traktować ją jako nowy kierunek, czy nurt wewnątrz filozofii analitycznej, zdecydowanie stanowi nową jakość we współczesnej filozofii.

To nie oznacza, że filozofia eksperymentalna ma przed sobą świetlaną przyszłość. Tradycyjnym filozofom wydaje się nie do pomyślenia, żeby praca filozoficzna obejmowała przeprowadzanie ankiet i analizę statystyczną. Stosunek filozoficznej wspólnoty akademickiej do filozofii eksperymentalnej ewoluował od ignorowania jej do otwartej wrogości. Ale kierunek ma też istotne atuty. Projektowanie i przeprowadzanie eksperymentów jest fascynującą czynnością i żadna inna metodologia filozoficzna nie dostarcza dreszczyku emocji związanego

z weryfikowaniem samodzielnie wysuniętych hipotez. Co więcej, ze względu na metodologię inspirowaną empirycznymi naukami społecznymi – w szczególności psychologią i ekonomią eksperymentalną – publikacje eksperymentalnofilozoficzne akceptowane są też przez redakcje czasopism z tych dziedzin. W świecie akademickim, w którym *publikować* znaczy *być*, daje to istotną przewagę nad uprawiającymi filozofię z gabinetu.

W książce starałem się przedstawić teoretyczne podstawy kierunku oraz, przedstawiając wybrane eksperymenty filozoficzne, pokazać, że przeprowadzanie własnych badań w tej dziedzinie wcale nie musi być trudne. Mam nadzieję, że ten cel został osiągnięty.

- Appiah K., *Experiments in Ethics*, Cambridge: Harvard University Press, 2008.
- Arystoteles, *Metafizyka*, tłum. K. Leśniak, Warszawa 2003,
- Buckwalter W. i Stich S., *Gender and Philosophical Intuition*, „Social Science Research Network, Working Papers Series” (2010),
url = <<http://ssrn.com/abstract=1683066>>,
- Clarke D., *Philosophy's Second Revolution: Early and Recent Analytic Philosophy*, Open Court, 1997,
- Cullen S., *Survey-Driven Romanticism*, „Review of Philosophy and Psychology” (2010) 1 (2):275-296,
- Cushman F., Greene J., *The philosopher in the theater*, [w:] *Social psychology of morality: The origins of good and evil*, red. Mikulincer M., Shaver P., APA Press,
- Daniels N., *Reflective Equilibrium* [w:] *The Stanford Encyclopedia of Philosophy (Fall 2008 Edition)*,
url = <<http://plato.stanford.edu/archives/fall2008/entries/reflective-equilibrium/>>,
- Dummett M., *Origins of Analytical Philosophy*, Duckworth 1993,
- Fodor J., *Modularity of Mind*, MA: MIT Press, Cambridge 1983,
- Gettier E., *Czy uzasadnione i prawdziwe przekonanie jest wiedzą?* „Analysis” 23 (1963),
tłum. J. Hartman, J. Rabus,
- Greene J., Haidt J. *How (and where) does moral judgment work?* „Trends in Cognitive Science” (2002) 6 (12),
- Hergenhahn B., *An Introduction to the History of Psychology*, Wadsworth Publishing Co., 2000,
- Horvath J., *How (Not) to React to Experimental Philosophy*, „Philosophical Psychology”, Vol. 23 (3), 2010,
- Jackson F., *From Metaphysics to Ethics: a defense of conceptual analysis*, Clarendon Press 1998,
- Kauppinen A., *The Rise and Fall of Experimental Philosophy*, „Philosophical Explorations” 10 (2):95 – 118,

- Knobe J., *Intentional Action and Side Effects in Ordinary Language*, „Analysis” 63 (2003), , 190-193,
- Knobe J., *Person as Scientist, Person as Moralist*, „Brain and Behavioral Sciences” 33: 315-329 (2010),
- Knobe J., Nichols S., *Experimental Philosophy*, Oxford University Press 2008,
- Kripke S., *Nazywanie a konieczność*, tłum. B. Chwedeńczuk, Warszawa 1988,
- Levitt S., Dubner S., *Freakonomia. Świat od podszewki*, tłum. A. Sobolewska, Warszawa 2006,
- Liao M., *A defense of intuitions*, „Philosophical Studies” 140 (2) 2008: 247-262,
- Machery E., Mallon R., Nichols S., Stich S., *Semantics, cross-cultural style*, „Cognition” 92 (3), (2004),
- Mackie J. L., *Evil and Omnipotence*, „Mind, New Series”, 1955/254,
- Magdziak M., *Uwagi o zasadzie równej miary*, [w:] *Czy sprawiedliwość jest możliwa?* red. Probučka D.
- Martinich A., Sosa D., *A Companion to Analytic Philosophy*, Blackwell Ltd. 2001,
- Mele A., *Acting intentionally: Probing folk intuitions*. [w:] Malle B., Moses L., Baldwin D., *Intentions and intentionality: Foundations of social cognition*, Cambridge, Mass.: MIT Press. 2001,
- Nahmias E., Nadelhoffer T., *The Past and Future of Experimental Philosophy*, „Philosophical Explorations”, Vol. 10, No. 2, lipiec 2007,
- Nisbett R., Peng K., Choi I., Norenzayan A., *Culture and systems of thought: Holistic vs. analytic cognition*, „Psychological Review” 108 (2001),
- Papineau D., *What is X-Phi Good For?*, „The Philosophers’ Magazine” (2010),
- Platinga A., *Bóg, wolność, zło*, Kraków 1995,
- Platon, *Państwo*, Kęty 2002, tłum. W. Witwicki,
- Prinz J., *Empirical Philosophy and Experimental Philosophy* [w:] Knobe J., Nichols S., *Experimental Philosophy*, Oxford University Press 2008,
- Rawls J., *A Theory of Justice*, Revised Edition, Harvard University Press 1999
- Russel B., *Intuition, addendum 2*, [w:] *Encyclopedia of Philosophy, 2nd edition*, MacMillan 2009,
- Schnall S., Benton J., Harvey S., *With a clean conscience: Cleanliness reduces the severity of moral judgments*, *Psychological Science*, (2008) 19,
- Scholl B., *Two Kinds of Experimental Philosophy (and their methodological dangers)*, notatka ze spotkania *Pre-SPP Workshop on Experimental Philosophy*, 25 VI 2008. Notatka jest nieformalna i nie ma charakteru publikacji; źródło: http://experimentalphilosophy.typepad.com/experimental_philosophy/files/scholl_xphi_notes.pdf
- Sen A., *Nierówności. Dalsze rozważania*. tłum. I. Topińska, Kraków 2000,
- Strawson P., *Analiza i metafizyka*, tłum. A. Grobler, Kraków 1994,
- Swinburne R., *The Existence of God, 2nd ed.*, Oxford University Press 2004,

- Thomson J., *A Defense of Abortion*, „Philosophy and Public Affairs” 1:1 (jesień 1971),
- Vaesen K., Peterson M., *Knowledge and the Quality of Information*, 2010, w recenzji,
- Valdesolo P, DeSteno D., *Manipulations of Emotional Context Shape Moral Judgment*, „Psychological Science” 17, (2006),
- Weinberg J., Gonnerman Ch., Buckner C. i Alexander J., *Are Philosophers Expert Intuiters?*, „Philosophical Psychology” 23(3) 2010,
- Weinberg J., *How to Challenge Intuitions Empirically Without Risking Skepticism*, „Midwest Studies in Philosophy” 31 (1):318–343,
- Weinberg J., Nichols S., Stich S., *Normativity and Epistemic Intuitions*, „Philosophical Topics” (2001) 29 (1&2),
- Williamson T., *Armchair philosophy, metaphysical modality and counterfactual thinking*, „Proceedings of the Aristotelian Society” 105 (2005),
- Williamson T., *Philosophical Expertise and the Burden of Proof*, 2010, w recenzji,
- Young L., Cushman F., Adolphs R., Tranel D., Hauser M. (2006) *Does emotion mediate the effect of an action’s moral status on its intentional status? Neuropsychological evidence*, „Journal of Cognition and Culture” 6:291-304.