

BARTOSZ MAĆKIEWICZ\*

## NIEDOSTATKI METOD KWESTIONARIUSZOWYCH W FILOZOFII EKSPERYMENTALNEJ\*\*

### Abstract

THE SHORTCOMINGS OF THE QUESTIONNAIRE METHOD IN EXPERIMENTAL PHILOSOPHY

Surveys are the most widely used research tool in experimental philosophy. In this paper, I analyze two types of criticism of the questionnaire method put forward in the literature. The first type is metaphilosophical. It asserts that the surveys used by experimental philosophers are based on a flawed and unproductive method of cases. The second type is methodological. It argues that currently used questionnaires are unfit to measure philosophically relevant phenomena. I show that these objections can be met by a) improving questionnaires used in research practice and b) expanding the methodological repertoire of experimental philosophy.

*Keywords:* experimental philosophy, questionnaire methodology, metaphilosophy

---

Filozofia eksperymentalna od początku istnienia korzysta z metod kwestionariuszowych i do dziś są to metody najczęściej wykorzystywane przez jej zwolenników. Celem artykułu jest wykazanie, że wskazywane w dyskusjach nad filozofią eksperymentalną niedostatki metod kwestionariuszowych nie są dla niej zasadniczym problemem. Można bowiem ulepszać je tak, by pewnych błędów lub ryzyk unikać, a w wypadkach, gdy nie jest to możliwe, można korzystać z innych metod zbierania danych empirycznych.

Wykorzystywane przez filozofów eksperymentalnych kwestionariusze wzorowane są w znacznej mierze na metodzie przypadków charakterystycznej dla części nurtów współczesnej filozofii. Metodę tę można najogólniej opisać jako sposób prowadzenia badań filozoficznych polegający na konstruowaniu hipotetycznych sytuacji i rozważaniu ich wybranych aspektów w celu odpo-

---

\* Uniwersytet Warszawski, Wydział Filozofii, Krakowskie Przedmieście 3, 00-927 Warszawa, e-mail: b.mackiewicz@uw.edu.pl, ORCID: <https://orcid.org/0000-0002-9460-5742>.

\*\* Praca powstała w wyniku realizacji projektu badawczego o nr 2018/29/N/HS1/02863 finansowanego ze środków Narodowego Centrum Nauki.



wiedzenia na pewne pytanie filozoficzne<sup>1</sup>. Kwestionariusze zawierające opisy hipotetycznych sytuacji (tzw. scenariusze), często zresztą bezpośrednio nawiązujących do eksperymentów myślowych rozważanych w filozofii gabinetowej, to dominujący w filozofii eksperymentalnej sposób konstruowania badań. Również pytania towarzyszące scenariuszom nie odbiegają charakterem od tych stosowanych w metodzie przypadków. W wielu badaniach są to (bardziej lub mniej zamaskowane) pytania o stosowalność pewnego pojęcia. Metodę kwestionariuszową należy więc uznać za odzwierciedlenie tradycyjnej w filozofii analitycznej metody tworzenia i oceniania teorii filozoficznych.

W artykule przyjmować będę jako punkt wyjścia, że celem badań kwestionariuszowych w filozofii eksperymentalnej jest – podobnie jak celem stosowania tradycyjnej metody przypadków – zbadanie filozoficznych intuicji lub ustalenie treści filozoficznie istotnych pojęć. Intuicje zaś służą do oceny stanowisk filozoficznych. Należy mieć jednak na uwadze, że założenie to jest przedmiotem dyskusji zarówno jeśli chodzi o tradycyjną metodę przypadków (por. Machery 2017: 11-44, Cappelen 2012), jak i o filozofię eksperymentalną (por. Rose, Danks 2013).

Zarzuty wysuwane pod adresem metod kwestionariuszowych mają charakter metafizyczny i metodologiczny. Zarzuty metafizyczne uderzają przede wszystkim w gabinetową odmianę charakterystyczną dla współczesnej filozofii analitycznej. Filozofia eksperymentalna, opierając się na metodzie przypadków, zarzuty te poniekąd dziedziczy. W części 1 artykułu pokażę dwa rozwiązania, które mogą przyjąć filozofowie eksperymentalni, by oddalić zarzuty metafizyczne. Część 2 poświęcam analizie argumentów wskazujących problemy metodologiczne badań kwestionariuszowych przeprowadzanych przez filozofów eksperymentalnych. Rozważania te kończę propozycją metodologicznego udoskonalenia badań kwestionariuszowych oraz opowiadam się za stosowaniem w filozofii eksperymentalnej strategii metodologicznej triangulacji.

---

<sup>1</sup> Nie wszystkie przykłady wykorzystania w filozofii hipotetycznych sytuacji można zakwalifikować jako zastosowanie metody przypadków. Niektóre eksperymenty myślowe odgrywają jedynie funkcję ilustracji, np. według Machery'ego (2011) taką funkcję pełni bagnolud Davidsona (1987).

## 1. ARGUMENTY METAFILOZOFICZNE

### 1.1. Fundamentalna nierzetelność sądów na temat przypadków filozoficznych

Część wyników badań przeprowadzonych przez filozofów eksperymentalnych wskazuje, że intuicje zarówno filozofów, jak i niefilozofów na temat przypadków filozoficznych są niestabilne i wrażliwe na nieistotne z filozoficznego punktu widzenia czynniki o charakterze demograficznym (płeć, status socjoekonomiczny, wiek) lub związane ze sposobem prezentacji scenariuszy (kolejność, sformułowanie pytania, kontekst). Według niektórych dowodzi to nierzetelności intuicji ujawnianych przez eksperymenty myślowe (por. Nadelhoffer, Nahmias 2007, Alexander, Weinberg 2014, Kuś, Maćkiewicz 2016)<sup>2</sup>.

Machery (2017) rozważa kilka argumentów mających pokazać, że nierzetelność sądów<sup>3</sup> na temat przypadków filozoficznych ma swoje źródło w sposobie konstrukcji hipotetycznych sytuacji używanych przez filozofów w ich wywodach<sup>4</sup>. Celem Machery'ego jest pokazanie, że to sposób skonstruowania przypadków filozoficznych czyni sądy na ich temat wrażliwymi na czynniki filozoficznie nieistotne.

---

<sup>2</sup> Z problemu tego można wybrnąć, stwierdzając, że wcale nie dowodzi to nierzetelności, a raczej tego, że nasze pojęcia, o których treści mówią nam intuicje, są wrażliwe na kontekst (efekt kolejności) lub zmienne kulturowo i demograficznie. Jest to jednak rozwiązanie kłopotliwe z punktu widzenia roli, jaką intuicje odgrywają w filozofii, i uniwersalnego statusu, do którego pretendują koncepcje filozoficzne.

<sup>3</sup> Warto zauważyć, że Machery mówi nie o intuicjach, lecz o sądach. Uważa, że pojęcie intuicji jest w rozważaniach metafizycznych bezużyteczne. Wszystkie podawane przez niego argumenty dotyczące sądów można jednak bez dużej straty siły przekonywania zastosować także do intuicji. Dla precyzji wyводу będą jednak konsekwentnie, zgodnie ze stanowiskiem przyjmowanym przez Machery'ego, mówił o sądach.

<sup>4</sup> Samo pojęcie nierzetelności, tak jak go używa Machery, należy rozumieć w duchu dość standardowej definicji zgodnej z klasycznymi koncepcjami Goldmana (1979, 1988) i Alstona (1995): psychologiczny proces T wytwarzający sądy jest rzetelny w środowisku E wtedy i tylko wtedy, gdy albo w środowisku E proces T ma dyspozycję do wytwarzania wysokiej proporcji prawdziwych sądów, albo jeśli T jest procesem inferencyjnym i jego przesłanki są prawdziwe, to T ma dyspozycję do wytwarzania wysokiej proporcji prawdziwych sądów (Machery 2017: 96). W tej definicji warto zwrócić uwagę na trzy elementy. Po pierwsze, rzetelność rozumiana jest jako własność dyspozycyjna. Po drugie, możemy mówić o rzetelności pojedynczego wystąpienia jakiegoś procesu. Po trzecie, pojęcie rzetelności zrelatywizowane jest do środowiska, w którym dany proces zachodzi. Machery dostrzega również klasyczny dla reliabilizmu problem generalizowalności (por. Paprzycka 2013): jak właściwie scharakteryzować typ sądów, które miałyby być nierzetelne? Czy są to sądy dotyczące filozoficznych przypadków, sądy dotyczące przypadków gettierowskich, czy może sądy dotyczące przypadków gettierowskich, w których występują amerykańskie samochody? Machery uważa, że w tym wypadku decydujące jest to, że nie mamy żadnych powodów, by sądzić, że możliwe do wyróżnienia podtypy różnią się pod względem rzetelności.

Sytuacje rozważane przez filozofów mają według Machery'ego niecodzienne (*unusual*) charakter, co znaczy, że napotykamy je rzadko lub rzadko o nich czytamy. Niecodziennność jest więc różna z jednej strony od typowości (typowa sytuacja może być w przyjętym przez niego sensie niecodzienna, jeśli występuje często, ale nie rozważamy jej na co dzień), a z drugiej strony od fikcyjności czy hipotetyczności (możemy mieć fikcyjną sytuację, która nie jest niecodzienna). Jeśli niecodziennność filozoficznego przypadku polega na pewnej egzotyczności jego treści dosłownej, to potencjalnie możemy go zmodyfikować tak, aby tej niecodzienności unikał. Na przykład klasyczne przypadki Gettierowskie możemy pozbawić dziwnych sposobów nabywania przekonań przez bohatera (liczenie monet w kieszeni kolegi, wykonywanie arbitralnych operacji logicznych) i zastąpić je bardziej realistycznymi. Kiedy niecodziennność dotyczy samej istoty filozoficznego przypadku, jest to bardzo trudne lub wręcz niemożliwe.

Zdaniem Machery'ego niecodziennność przypadków filozoficznych jest niepokojąca, ponieważ im bardziej niecodzienna sytuacja, tym mniej jasne, jakie fakty w niej zachodzą. Odbiorcy danego scenariusza filozoficznego (zarówno filozofowie podczas swoich gabinetowych dociekań, jak i uczestniczące w badaniach osoby bez wykształcenia filozoficznego) mogą różnić się co do tego, jak pojmują daną sytuację. Ze względu na to, że sytuacja jest niecodzienna w przyjętym tu sensie, w jej pojmowaniu przez odbiorców pojawiają się luki. Są one wypełniane za pomocą wyobraźni, na przykład na podstawie nieistotnych z filozoficznego punktu widzenia informacji, takich jak powierzchowne własności opisywanej sytuacji czy treść wcześniej rozważanych przypadków. Taki mechanizm mógłby wyjaśnić wielokrotnie zaobserwowany w badaniach eksperymentalnych wpływ kolejności, sposobu prezentacji czy opisu na ocenę przypadków filozoficznych.

Machery twierdzi, że tak samo można wyjaśnić wpływ czynników demograficznych. Kiedy rozważana sytuacja jest na tyle niecodzienna, że musimy wypełnić luki pojawiające się w jej rozumieniu, uzupełniamy je w zgodzie z naszym światopoglądem czy doświadczeniem życiowym. Oceniając sytuacje (nawet fikcyjne), które nie są niecodzienne, nie musimy przy wypełnianiu luk w tak dużym stopniu odwoływać się do swojego światopoglądu czy doświadczenia życiowego. Pewna część filozoficznych eksperymentów myślowych jest, jak przekonuje Machery, na tyle egzotyczna, że rozważającym je osobom trudno nawet pojąć, jaką sytuację mają sobie wyobrazić. W takim wypadku to, jaką sobie faktycznie sytuację wyobrażą, zależy może od filozoficznie nieistotnych lub filozoficznie istotnych, ale niekontrolowanych czynników<sup>5</sup>.

---

<sup>5</sup> Wśród filozoficznie nieistotnych czynników mogących wpływać na intuicje zazwyczaj wymienia się różnice indywidualne, takie jak cechy osobowości, przynależność kulturową lub etniczną czy tożsamość płciową. Motywację tutaj stanowi pewnego typu egalitaryzm.

Zdaniem Machery'ego niecodziennosc przypadków filozoficznych związana jest z podstawową cechą konstrukcyjną właściwą niemal wszystkim występującym w literaturze eksperymentom myślowym, czyli „rozdzielaniem tego, co zwykle występuje razem” (*pulling apart what usually goes together*). Sytuacje przedstawiane w przypadku filozoficznym są konstruowane w ten sposób, żeby pewne własności, które w codziennym życiu współwystępują, w eksperymencie myślowym jednocześnie nie występowały.

Stosowanie metody przypadków wymaga odpowiednio skonstruowanych eksperymentów myślowych. Przykładowo stanowiska epistemologicznego internalizmu i eksternalizmu nie będą dawały innych przewidywań w codziennych i zwyczajnych sytuacjach przypisywania wiedzy. Różnicę między nimi możemy jednak uwypuklić, konstruując sytuację, w której proces prowadzący do powstania przekonania jest rzetelny, chociaż brakuje mu subiektywnego uzasadnienia. Za przykład może tutaj posłużyć Truetemp<sup>6</sup> (Lehrer 1990). Sytuacje tego typu są jednak albo bardzo rzadkie, albo w ogóle nie występują w rzeczywistości. Bez wątplenia w normalnych sytuacjach, kiedy mówię, że coś wiem, mam subiektywne uzasadnienie dla swoich przekonań, a jednocześnie moje przekonanie jest wynikiem rzetelnego procesu poznawczego.

Machery argumentuje, że w sytuacjach, w których „rozdzielone jest to, co zwykle występuje razem”, nasze zdolności rozpoznawania w świecie zjawisk takich jak wiedza, przyczynowość czy moralna dopuszczalność są upośledzone. Umiejętności te działają dobrze tylko w normalnych wypadkach. Kiedy własności, które służą nam do identyfikowania tych zjawisk, zostają rozdzielone w filozoficznym eksperymencie myślowym, jego odbiorca musi zdecydować, które z nich są rzeczywiście istotne. Nie chodzi tu przy tym po prostu o pewną statystyczną prawidłowość współwystępowania określonych własności pewnych sytuacji i prostą nieumiejętność wyobrażenia sobie, że pewne dwie własności nie występują razem, lecz istotny jest konflikt pewnych heurystyk, za pomocą których rozpoznajemy w świecie wiedzę, przyczynowość czy moralną dopuszczalność.

---

Gdyby na przykład okazało się, że na intuicje filozoficzne wpływa płeć, to o ile chcielibyśmy dalej używać intuicji w filozofii w dotychczasowy sposób, to bylibyśmy zmuszeni do traktowania w uprzywilejowany sposób intuicji jednej z płci. To wydaje się jednak nieuzasadnione.

<sup>6</sup> Eksperyment myślowy mający pokazać, że eksternalistyczne teorie wiedzy są błędne. Przedstawiony w nim bohater nabył przekonanie o aktualnej temperaturze powietrza dzięki rzetelnemu, to znaczy z reguły prowadzącemu do prawdziwych przekonań, procesowi poznawczemu, ponieważ naukowcy w tajemnicy wszczepili mu do mózgu urządzenie, które wytwarza w nim prawdziwe przekonania na temat temperatury. Mimo że warunek rzetelności procesu prowadzącego do przekonania jest spełniony (i jego przekonanie jest prawdziwe), to nie powiedzielibyśmy, że bohater dysponuje wiedzą na temat aktualnej temperatury.

Jak widzieliśmy, w wypadku Truetempa posiadanie subiektywnego uzasadnienia przekonania nabytego na drodze procesu poznawczego uznawanego za rzetelny służy nam do identyfikowania w świecie przypadków wiedzy. W historii o Truetempie występuje tylko jedna z tych dwóch własności – nabycie przekonania dzięki rzetelnemu procesowi. Musimy więc zdecydować, czy rzeczywiście to wystarcza, byśmy mogli przypisać bohaterowi wiedzę. Wymuszanie tego typu decyzji jest właśnie celem konstruowania eksperymentów myślowych. Nie możemy się bowiem odwołać do naszych wcześniejszych doświadczeń, ponieważ własności internalistyczne i eksternalistyczne wiedzy w większości znanych nam wypadków dotychczas współwystępowały. Nie sposób również skorzystać z dostępnych heurystyk, ponieważ one także opierają się na typowych przypadkach i tylko w ich kontekście są rzetelne. Ocena przypadku filozoficznego może więc być szczególnie wrażliwa na to, jakie eksperymenty myślowe rozważaliśmy chwilę wcześniej, do jakiej grupy demograficznej należymy czy też sposób, w jaki sytuacja została opisana.

Podsumowując, według Machery'ego sądy powstające pod wpływem rozważania przypadków filozoficznych są nie tyle przygodnie nierzetelne, ile nierzetelne fundamentalnie. Gdyby chodziło o przygodną nierzetelność, to być może udałoby się uczynić sądy bardziej rzetelnymi, na przykład lepiej kontrolując warunki przeprowadzania eksperymentu, niezależnie od tego, czy myślimy akurat o eksperymencie myślowym rozważanym w fotelu, czy o badaniu empirycznym. Fundamentalna nierzetelność polega na tym, że jest albo zupełnie nieusuwalna, albo nieusuwalna bez bardzo głębokich modyfikacji metody przypadków. Jeśli faktycznie tak jest i metoda ta jest epistemicznie ułomna, to należy zrezygnować z wykorzystania metody przypadków w filozofii, zaprzestać wzorowania na niej badań eksperymentalnych oraz zrewidować bardzo dużą część XX-wiecznej filozofii analitycznej, która na intuicyjnej analizie przypadków się opierała.

## 1.2. Konsekwencje krytyki metody przypadków dla filozofii eksperymentalnej

Krytykę metody przypadków przedstawioną przez Machery'ego należy uznać co do zasady za słuszną. Można próbować ją odpierać przez wskazanie, że zgubny efekt niepokojących własności widoczny jest tylko wówczas, gdy osobami rozważającymi hipotetyczne przypadki są ludzie bez treningu filozoficznego. Profesjonalni filozofowie mają bowiem odpowiednie przygotowanie umożliwiające im formułowanie rzetelnych sądów o filozoficznych eksperymentach myślowych. Ta strategia argumentacyjna, zwana w literaturze „obrońcą ekspercką”, była jednak przedmiotem krytyki eksperymentalistów, których badania pokazywały, że profesjonalni filozofowie są nie mniej niż niefilozofo-



wie wrażliwi na zaburzające czynniki, takie jak efekt kolejności czy sposób prezentacji scenariuszy (por. podsumowanie aktualnego stanu badań w Machery 2017: 45-89).

Zarzuty Machery'ego mają poważne konsekwencje dla standardowej praktyki badawczej filozofii eksperymentalnej. Sam Machery stoi na stanowisku, że charakterystyka metody przypadków jako sposobu na ustalanie treści pojęć na podstawie naszej kompetencji pojęciowej lub na wytwarzanie filozoficznych intuicji jest błędna, ponieważ zarówno pojęcie kompetencji pojęciowej, jak i intuicji filozoficznej jest puste. Istotne jest jednak to, że niezależnie od tego, czy Machery ma rację, to i tak metody kwestionariuszowe ze względu na elementy wspólne z filozoficzną metodą przypadków nie mogą dać nam żadnej pozytywnej wiedzy na temat filozoficznych intuicji czy treści pojęć, którymi posługują się ludzie. Sposób konstrukcji scenariuszy używanych przez filozofów eksperymentalnych jest bowiem analogiczny do tego wykorzystywanego przez filozofów gabinetowych. Nawet jeśli filozoficznie istotne intuicje istnieją, to kwestionariusze są ułomną metodą ich badania. Posługując się badaniami kwestionariuszowymi, możemy jedynie pokazać, że sądy dotyczące filozoficznych przypadków są faktycznie wrażliwe na nieistotne filozoficznie czynniki i że wrażliwość ta wynika z ich niecodzienności. Uważam, że jeśli Machery ma rację, to korzystająca z metod kwestionariuszowych filozofia eksperymentalna jest projektem dość krótkoterminowym. Gdy tylko wykaże niestabilność sądów ewokowanych przez eksperymenty myślowe, nie pozostanie jej już nic do zrobienia.

Rozszerzanie warsztatu metodologicznego przez filozofów eksperymentalnych może jednak osłabić krytykę Machery'ego. Skuteczność tego zabiegu zależy od tego, co właściwie filozofowie eksperymentalni chcą osiągnąć. David Rose i David Danks (2013) wyszczególniają dwie koncepcje filozofii eksperymentalnej – szeroką oraz wąską. Koncepcja wąska głosi, że filozofia eksperymentalna polega na przeprowadzaniu przez filozofów badań empirycznych w celu dowiedzenia się czegoś na temat intuicji filozoficznych ludzi. Z kolei zgodnie z koncepcją szeroką filozofia eksperymentalna jest połączeniem filozoficznego naturalizmu i praktyk badawczych kognitywistyki.

Jeśli przyjmiemy koncepcję wąską i co za tym idzie, uznamy, że interesuje nas badanie filozoficznych intuicji ludzi, to zarzuty Machery'ego można jedynie osłabić, a nie unieważnić. W takim wypadku argumenty Machery'ego możemy potraktować jako pewien zestaw wskazówek metodologicznych i zastanowić się, w jaki sposób wprowadzenie metod innych niż kwestionariuszowe może pomóc uniknąć wskazanych problemów.

Jeżeli niepokoi nas niecodziennność filozoficznych przypadków, to w kontekście badań empirycznych przede wszystkim powinniśmy unikać wpro-

dzania do używanych w eksperymentach scenariuszy nierealistycznych czy egzotycznych elementów na poziomie treści powierzchniowej. Oczywiście nie pozwoli nam to całkowicie uniknąć „rozdzielenia tego, co zwykle występuje razem” w sensie nakreślonym przez Machery’ego, ponieważ jest to podstawowa zasada konstrukcyjna wielu eksperymentów myślowych. Dołożenie starań, aby scenariusze uczynić jak najbardziej realistycznymi, może jednak zmniejszyć wrażliwość ocen na filozoficznie nieistotne czynniki. Przykładowo, Sara Weaver i John Turri (2018) przeprowadzili badanie, w którym przedstawiali badanym pochodzący od Dereka Parfita eksperyment myślowy dotyczący tożsamości osobowej, w którym bohater historyjki teleportuje się jednocześnie w dwa miejsca. Oczywiście, na ocenę badanych wpływać mogą wcześniejsze lektury i obejrzone filmy science-fiction<sup>7</sup>. Aby wykluczyć wpływ tego czynnika na wyniki eksperymentu, badacze powtórzyli badanie, używając nieco bardziej realistycznych scenariuszy, w których „rozdwojona” teleportacja została zastąpiona przez przeszczep po połowie mózgu do dwóch ciał.

Z niecodziennością filozoficznych przypadków możemy również walczyć, wprowadzając do filozofii eksperymentalnej nowe metody. Możemy skierować uwagę ku korpusom językowym i związanym z nimi metodom badawczym (Bluhm 2013, 2016, Sytsma i in. 2019, Maćkiewicz 2021). Korpusy językowe są konstruowane tak, aby pojawiające się w nich teksty odzwierciedlały do pewnego stopnia codzienną praktykę komunikacyjną danego rodzaju. Jakościowa oraz ilościowa analiza materiału językowego zgromadzonego w korpusach może umożliwić ustalenie na podstawie pewnych wzorców użycia treści filozoficznie istotnych pojęć potocznych<sup>8</sup>. Bluhm (2016) w korpusach widzi

---

<sup>7</sup> Ktoś obeznany z literaturą i filmem science-fiction może być lepszy niż laik w wyobrażeniu sobie sytuacji, w których występuje np. teleportacja i w analizowaniu konsekwencji takich wydarzeń. Problemem jest jednak to, że taka osoba może uzupełniać detale przedstawionego jej scenariusza (niekoniecznie zgodnie z treścią rozważanego eksperymentu myślowego) szczegółami zaczerpniętymi z napotkanych wcześniej w literaturze czy filmie sytuacji. Posłużmy się omawianym przykładem teleportacji. Wyobraźmy sobie, że badany bardzo często obcował z utworami science-fiction, w których Parfitowski model teleportacji (przesyłanie informacji i rekonstrukcja ciała w innym miejscu w opozycji do przesyłania materii) był elementem świata przedstawionego i postaci, które się teleportowały w taki właśnie sposób, traktowane były przez innych bohaterów tak, jakby odbyły zwyczajną podróż. Istnieje wówczas spora szansa, że przyzwyczajony do takiej teleportacji badany będzie traktował ją jako coś zupełnie niekontrowersyjnego i zaakceptuje bez większej refleksji rozstrzygnięcie Parfitowskiego problemu tożsamości osobowej domyślnie przyjmowane w świecie przedstawionym. Różnica między uprzednią wiedzą o rzeczywistym świecie a niejawnymi przekonaniem nabywanymi przez obcowanie ze światami fikcyjnymi jest tutaj wyraźna.

<sup>8</sup> Szczegółowa analiza zagadnienia, które techniki zaczerpnięte z językoznawstwa korpusowego mogą przynieść filozoficznie interesujące wyniki, wykracza poza ramy tego artykułu. Korpusy mają swoje ograniczenia. Jeśli chodzi o wykorzystanie ich do testowania hipotez (dotyczących np. treści pojęć czy poprawności), to można to robić tylko pod pewnymi wa-



narzędzie, które może służyć filozofom do testowania hipotez dotyczących treści pojęć w podobny sposób, jak robi się to za pomocą metody przypadków. Problemem tego podejścia jest jednak to, że filozoficzna analiza pojęciowa niewiele może zyskać na analizie typowych użyc i kontekstów, które z reguły znajdziemy w korpusie (por. Maćkiewicz 2021). Znacznie więcej dowiemy się, analizując przypadki graniczne stosowania interesującego pojęcia. Tego typu konteksty, zazwyczaj nieobecne w korpusie, eksploruje właśnie klasyczna analiza pojęciowa z charakterystyczną dla niej metodą przypadków. Nie oznacza to, że jest to jedyny sposób wykorzystania w filozofii korpusów. Jako przykład owocnego wykorzystania tego rodzaju narzędzi w dociekaniach filozoficznych przywołać można prowadzone przez Justina Sytsmę i współautorów (2019) badania nad potocznym pojęciem przyczyny. Ustalenie statystycznych wzorców współwystępowania jednostek leksykalnych odnoszących się do przyczyny z negatywnie nacechowanymi terminami pozwoliło badaczom uprawdopodobnić tezę, że – wbrew filozoficznej ortodoksji – pojęcie przyczyny ma wymiar normatywny.

Część luk w przedstawianych badanym scenariuszach może wynikać z tekstowej formy prezentacji bodźców. Luki te są w procesie interpretacji wypełniane przez skonfrontowanych z historyjkami badanych na podstawie filozoficznie nieistotnych informacji. Jeśli chcemy tego uniknąć, przydatne mogą być metody prezentacji sytuacji, w których niedookreśloność jest mniejsza i które charakteryzują się mniejszą dowolnością interpretacyjną. Obiecujące są w tym względzie graficzne formy prezentacji takie jak filmy, gry komputerowe czy technologia wirtualnej rzeczywistości (Alsmith, Longo 2019, Blascovich i in. 2002, Loomis, Blascovich, Beall 1999). Istnieją świadectwa za tym, że manipulacja formą prezentacji potrafi diametralnie zmienić oceny filozoficznych eksperymentów myślowych. Francis i jej zespół (2016) pokazali, że badani skonfrontowani z klasycznym filozoficznym dylematem kładki w wirtualnej rzeczywistości podejmują diametralnie inne decyzje, niż kiedy dylemat ten przedstawiony jest im w formie tekstowej.

Z kolei jeśli chcemy w jakiś sposób uporać się ze wskazanym przez Machery'ego problemem „rozdzielenia tego, co zwykle występuje razem”, to pomocne mogą okazać się metody jakościowe. W sytuacji, w której ocena filozoficznego przypadku wymaga rozstrzygnięcia konfliktu między kilkoma heurystycznymi zasadami, metody jakościowe są w stanie dać głębszy wgląd

---

runkami (por. przywoływany wcześniej Bluhm 2016). Należy również pamiętać, że ze swojej istoty korpusy językowe są niekompletne, co utrudnia (choć nie zawsze uniemożliwia) wyciąganie z nieobecności pewnych fraz w korpusie wniosków o ich niepoprawności czy niegramatyczności (por. omówienie argumentów przeciwników językoznawstwa korpusowego w: Stefanowitsch 2020 oraz ich krytykę).

w strategii przyjmowane przez badanych konfrontowanych z filozoficznymi eksperymentami myślowymi (Andow 2016b). W wywiadach pogłębionych można zapytać badanych o racje, które brali pod uwagę, formułując sąd na temat przedstawionego im przypadku, oraz o zasady rozstrzygania między nimi. Można również w toku wywiadu modyfikować sytuację, o której rozmawiamy, w taki sposób, aby ustalić treść pojęć, którymi posługują się badani. Jeśli przeprowadzenie pogłębionych wywiadów jakościowych jest niemożliwe ze względów organizacyjnych lub finansowych, podobny rezultat można osiągnąć, umieszczając w kwestionariuszu pytania otwarte zaprojektowane tak, aby dawały pewien wgląd w te strategie. Metoda ta stosowana była w filozofii eksperymentalnej. Na przykład Shaun Nichols i Joseph Ulatowski (2007) oraz Joshua Knobe i Jesse Prinz (2008) pytali w otwartej formule o uzasadnienia odpowiedzi udzielanych w kwestionariuszu. Odpowiedzi takie mogą być nawet systematycznie analizowane statystycznie po wcześniejszym zaklasyfikowaniu ich do jednej z kilku kategorii przez koderów, przed którymi zatajono cel badania.

Swoje nadzieje eksperymentalisci powinni jednak wiązać przede wszystkim z przyjęciem szerokiej koncepcji filozofii eksperymentalnej. Stanowisko to pozwala skutecznie unieważnić destrukcyjne dla filozofii eksperymentalnej konsekwencje krytyki Machery'ego. Jeżeli metody kwestionariuszowe modelowane na metodzie przypadków nie są, jakby wynikało z krytyki Machery'ego, odpowiednie do badania filozoficznie istotnych faktów empirycznych, to należy je po prostu porzucić. Skorzystać należy z najlepszych dostępnych metod, a tymi są obecnie metody szeroko rozumianej kognitywistyki. Jeżeli, przykładowo, odpowiednie fakty empiryczne dotyczące istotnych z filozoficznego punktu widzenia procesów poznawczych można ustalić, wykorzystując techniki neuroobrazowania<sup>9</sup>, to z punktu widzenia szerokiej koncepcji filozofii eksperymentalnej należy tak zrobić. Osobną kwestią jest oczywiście pokazanie, że fakty te są filozoficznie istotne. To jednak, jak wskazują Rose i Danks (2013), jest zadaniem przyjmowanego w danym kontekście filozoficznego naturalizmu.

---

<sup>9</sup> W proponowanym przez Rose'a i Danksa modelu uprawiania filozofii eksperymentalnej mieszczą się więc na przykład badania Brunona G. Bary i współautorów (2011), którzy wykorzystując fMRI, usiłowali sprawdzić, czy na poziomie neuronalnym istnieją postulowane przez Johna Searle'a intencje uprzednie i różne od nich intencje-w-działaniu. Należy zwrócić uwagę, że chociaż badanym prezentowane były bodźce w postaci (graficznych) historyjek, to były one bardzo daleko od przypadków konstruowanych przez filozofów.

## 2. ARGUMENTY METODOLOGICZNE

### 2.1. Pomiar intuicji w badaniach kwestionariuszowych

Filozofowie eksperymentalni zakładają, że kiedy konfrontują badanego ze scenariuszem (krótką historyjką lub opisem jakiejś sytuacji), a następnie proszą go o udzielenie odpowiedzi na pytania zawarte w kwestionariuszu, to mamy do czynienia z sytuacją *pomiaru*. Pojawia się więc pytanie, co w takim razie mierzyć mają zadawane w kwestionariuszu pytania. W kontekście filozofii eksperymentalnej najbardziej naturalna odpowiedź głosi, że mierzyć mają *intuicje* badanych. Nie jest to jednak niekontrowersyjna teza. John Bengson (2013), krytykując pewne odmiany programu filozofii eksperymentalnej, formułuje to metodologiczne założenie eksperymentalistów w dwóch wersjach. W wersji silniejszej brzmi ono:

[AEI] Wywołane odpowiedzi badanych wyrażają ich intuicje,

a w wersji słabszej:

[AEI\*] Wywołane odpowiedzi badanych w *typowych* lub *zwykłych* (normalnych itp.) sytuacjach wyrażają ich intuicje.

Bengson twierdzi, że założenie to niezależnie od wersji jest trudne do obrony. Uważa, że odpowiedzi badanych w eksperymentach mogą wyrażać coś innego niż intuicje, których odpowiadająca osoba może po prostu nie mieć<sup>10</sup>. Według Bengsona, po pierwsze, jest dalece nieoczywiste, czy założenia te są prawdziwe, można bowiem sobie wyobrazić realistyczne sytuacje, w których są one fałszywe. Po drugie, filozofowie eksperymentalni nie dostarczyli dotychczas odpowiednio silnych argumentów za przyjęciem nawet słabszej wersji tego założenia.

Bengson twierdzi, że przedstawiona przez niego argumentacja nie opiera się na żadnej specyficznej koncepcji intuicji. Uważam jednak, że siła jego argumentu zależy od tego, jaką koncepcję intuicji przyjmujemy. Jednym ze sposobów klasyfikowania koncepcji intuicji jest podział na koncepcje grube (*thick*) i cienkie (*thin*, por. Alexander, Weinberg 2014, Andow 2016a). Koncepcje grube są bardziej wymagające – to, czy ma się pewną intuicję, a nie jakiś inny stan mentalny zależy od wielu czynników. Bengsona zaliczyć trzeba do zwolenników koncepcji fenomenologicznych, zgodnie z którymi intuicji towarzyszy pewna

<sup>10</sup> Mogą to być np. zgadywanie, ocena zajścia jakiegoś faktu na podstawie domniemanego prawdopodobieństwa jego zajścia lub inne formy przewidywania.

określona fenomenologia, zwykle charakteryzowana w kategoriach oczywistości. Sam Bengson mówi, że pewne rzeczy narzucają się nam (*strikes us*) w pewien sposób.

Im bardziej wymagająca jest przyjmowana przez nas koncepcja intuicji, tym trudniej mieć intuicje i tym rzadziej się to zdarza. Jeśli samo posiadanie intuicji jest czymś niełatwym do osiągnięcia, to rzeczywiście można wątpić, czy uczestnicy eksperymentów je mają. Możliwość, że ich odpowiedzi nie mają nic wspólnego z tak rozumianymi intuicjami, lecz opierają się na przykład na zgadywaniu lub innych tego typu stanach mentalnych, jest w takim wypadku więcej niż prawdopodobna. Jeśli zaś przyjmie się koncepcję cienką, zgodnie z którą intuicje to po prostu sądy albo pewne inklinacje do wydawania określonych sądów, to wydaje się, że jest oczywiste, że AEI\* jest prawdziwe (por. Andow 2016a). Z kolei jeśli koncepcja wymaga posiadania przez intuicje pewnej etiologii, to – w zależności od tego, o jakiej etiologii myślimy – AEI może być bardziej lub mniej prawdopodobne. Jeśli na przykład interesują nas intuicje filozoficzne jako stany przekonaniowe wynikające z kompetencji pojęciowej (np. Ludwig 2007), to kwestionariusze niekoniecznie są najlepszą metodą ich pomiaru.

Sytuacja przedstawia się jeszcze inaczej, jeśli do kwestii intuicji podchodzimy w sposób bardziej naturalistyczny. Współcześnie wielu psychologów postuluje, że pewne sądy mogą być wytworem dwóch różnych typów procesów poznawczych. Procesy pierwszego typu, zwanego systemem 1, są nieświadome, niekontrolowane i charakteryzują się automatyzmem. Do systemu 2 należą procesy świadome i kontrolowane. Naturalistycznie nastawieni filozofowie, którzy zgadzają się z koncepcją podwójnego przetwarzania, muszą stwierdzić, że intuicjami są sądy wytwarzane przez system 1 (por. Fischer, Engelhardt 2016, 2017). Wówczas AEI należy uznać za fałszywe lub bardzo wątpliwe. Odpowiedzi w kwestionariuszu będą raczej wynikiem funkcjonowania systemu 2, to znaczy będą rezultatem świadomego rozumowania przeprowadzanego przez osoby badane. Aby wiarygodnie zmierzyć odpowiedź automatycznego szybkiego systemu 1, należałoby posłużyć się odpowiedniejszymi do tego metodami. Możliwość badania automatycznego przetwarzania charakterystycznego dla działania systemu 1 mogą dać na przykład metody pomiaru czasu reakcji (por. Arico i in. 2011).

Argumenty metodologiczne przedstawiane przeze mnie w dalszej części artykułu nie zakładają żadnej konkretnej koncepcji intuicji. Dotyczą raczej pewnych ogólnych własności metody kwestionariuszowej. Należy jednak pamiętać, że w zależności od przyjmowanej koncepcji intuicji zarzuty metodologiczne mogą mieć różną siłę. Na przykład kwestia niebezpośrednich miar może być ważniejsza dla zainteresowanych wynikami funkcjonowania systemu 1 natu-

ralistów, podczas gdy kwestie związane z pragmatyką badania ankietowego mają większe znaczenie dla zwolenników poglądu, że intuicje mają swe źródło wyłącznie w kompetencji pojęciowej.

## 2.2. Oczekiwania społeczne i błąd eksperymentatora

Robert L. Woolfolk (2013) uważa za fałszywe przyjmowane przez filozofów eksperymentalnych założenie, zgodnie z którym odpowiedzi badanych w kwestionariuszu stanowią odzwierciedlenie ich intuicji. Uczestnicy eksperymentów mogą odpowiadać zgodnie nie z tym, co naprawdę myślą, lecz z tym, co sądzą na temat tego, co eksperymentator chce usłyszeć lub co wydaje im się społecznie oczekiwane w danej sytuacji. Powinniśmy więc zachować dużą dozę krytycyzmu, wyciągając wnioski na temat przekonań badanych z ich odpowiedzi w kwestionariuszach. Tego typu zjawiska obserwuje się regularnie w naukach społecznych od ponad pięćdziesięciu lat (np. Cronbach 1950, Rosenthal 1976; w kwestii metod minimalizowania tendencji badanych do udzielania społecznie oczekiwanych odpowiedzi por. Nederhof 1985).

Argument Woolfolka ma jednak ograniczony zasięg. Zakłada po pierwsze, że badany będzie próbował zgadnąć, jakich odpowiedzi oczekuje od niego eksperymentator. Po drugie zakłada, że w eksperymentach przeprowadzanych przez filozofów eksperymentalnych istnieją odpowiedzi, które mogą być przez badanych interpretowane jako społecznie oczekiwane. Oba te założenia stosują się tylko do części badań z zakresu filozofii eksperymentalnej i tylko w ograniczonym stopniu. Stopień ten jest zależny od tego, jakiej problematyki filozoficznej dotyczą rozważane badania.

Przyjrzyjmy się na początek kwestii oczekiwań społecznych. Duża liczba badań przeprowadzanych przez filozofów eksperymentalnych dotyczy problemów ze swojej natury społecznych albo mających pewien aspekt społeczny. Badania dotyczące moralności są tu oczywistym przykładem. W eksperymencie Knobe'a (2003) badani byli konfrontowani ze scenariuszami opisującymi dyrektora przedsiębiorstwa wprowadzającego nowy program, który w zależności od warunku eksperymentalnego miał skutek uboczny w postaci zaszkodzenia środowisku lub poprawienia jego stanu. Przeważająca większość badanych zgadza się z twierdzeniem, że dyrektor umyślnie (*intentionally*) zaszкодził środowisku, oraz nie zgadza się, że dyrektor umyślnie środowisku pomógł. W świetle argumentu Woolfolka możemy rozsądnie przypuszczać, że w scenariuszu szkodzenia ludzie oceniają działanie dyrektora firmy jako umyślnie nie dlatego, że rzeczywiście uważają, że zaszкодził środowisku umyślnie, ale dlatego, że odpowiedź przeciwna implikuje w ich przekonaniu społecznie nieakceptowalne konsekwencje. Być może przyjmują, że odmawiając działaniu

dyrektora intencjonalności, zaprzeczyliby jednocześnie temu, że jest odpowiedzialny za negatywne dla środowiska skutki swoich działań. Wobec tego właściwą z punktu widzenia oczekiwań społecznych odpowiedzią jest ocena jego działania jako umyślnego, mimo że w rzeczywistości badani nie uważają, żeby dyrektor zaszkodził środowisku intencjonalnie<sup>11</sup>. W przypadku badań eksperymentalnych z zakresu etyki czy filozofii działania istnieje więc ryzyko, że oczekiwania społeczne wpłyną na odpowiedzi badanych. W tych wypadkach rzeczywiście wpływ pewnych oczekiwań zagraża wiarygodności danych uzyskiwanych przez filozofów eksperymentalnych.

Filozofia eksperymentalna nie ogranicza się jednak do badania kwestii związanych z moralnością czy ludzkim działaniem. Wiele z badań przeprowadzanych przez filozofów eksperymentalnych dotyczy bardziej „niewinnych” zagadnień, takich jak na przykład problem odniesienia nazw własnych. W takich wypadkach oczekiwania społeczne również mogą wpływać na odpowiedzi badanych, mimo że na pierwszy rzut oka może wydawać się to mało prawdopodobne. Rozważmy Kripke’owski (1980) scenariusz z Gödlem i Schmidtem<sup>12</sup>.

Zarówno filozofowie, jak i osoby niebędące filozofami mają w tym wypadku co do zasady intuicje zgodne z koncepcją Kripkego, chociaż istnieją powody, aby podejrzewać, że intuicje te mogą zależeć od kultury – osoby o pochodzeniu azjatyckim udzielają częściej odpowiedzi zgodnych z deskrypcjonizmem (por. Dongen i in. 2020). Można jednak rozsądnie utrzymywać, że to nie odmienne intuicje dotyczące funkcjonowania nazw własnych wyjaśniają różnicę w odpowiedziach Azjatów i osób należących do szeroko rozumianej kultury euroatlantyckiej, lecz większa waga, którą Azjaci przykładają do społecznych implikacji swoich odpowiedzi. Możemy na przykład wyobrazić sobie, że udzielenie odpowiedzi zgodnych z teorią przyczynowo-historyczną byłoby w tym wypadku umniejszaniem znaczenia prawdziwego odkrywcy tego dowodu albo stanowiłoby pochwałę złodzieja. Zauważmy, że różnice kulturowe w sposobie opowiadania w badaniach ankietowych są dobrze udokumentowane. Susanne Haberstroh i jej zespół (2002) pokazali, że Chińczycy wypełniający ankiety są bardziej wrażliwi na pewne pragmatyczne aspekty badań ankietowych. Z kolei

---

<sup>11</sup> Faktycznie tego typu wyjaśnienia zostały zaproponowane np. przez Freda Adamsa i Annie Steadman (2004). Ich wyjaśnienie nie jest przy tym powszechnie przyjmowane w literaturze. Ilustruje jednak, że można rozsądnie krytykować eksperymentalistów za brak uwzględnienia kwestii oczekiwań społecznych w przeprowadzanych przez nich badaniach.

<sup>12</sup> W tym eksperymencie myślowym prawdziwym odkrywcą dowodu twierdzenia o niezupełności jest nie Gödel, lecz jego przyjaciel Schmidt, któremu Gödel dowód ukradł. Kripke zadaje pytanie, do kogo odnosimy się, używając nazwy „Gödel”, jeśli jedyną deskrypcją, którą z tą nazwą kojarzymy, jest „odkrywca dowodu twierdzenia o niezupełności arytmetyki”? Zgodnie z koncepcją przyczynowo-historyczną, za którą opowiada się Kripke, odnosimy się do Gödla. Deskrypcjonista będzie twierdził, że w tym wypadku odniesieniem nazwy jest Schmidt.



Timothy Johnson i współautorzy (2005) odnotowują, że osoby pochodzące z kultur z większym dystansem władzy (*power distance*, por. Hofstede 1984) mają skłonność do udzielania w badaniach ankietowych bardziej skrajnych odpowiedzi.

Celem tego przykładu jest jedynie wskazanie możliwości, że nawet w kontekstach, w których społeczne oczekiwania nie powinny odgrywać większej roli (badania dotyczące intuicji semantycznych), może się zdarzyć, że wpływają one na odpowiedzi badanych. Rozsądne wydaje się więc stanowisko, zgodnie z którym ryzyko wpływu społecznych oczekiwań jest zależne od tego, jakiego typu kwestiom poświęcone jest badanie. Zarzut Woolfolka dotyczy więc w większym stopniu badań z zakresu eksperymentalnej etyki czy filozofii działania, a w mniejszym filozofii języka czy eksperymentalnej metafizyki. Należy też dodać, że w przypadku problematyki nie dotyczącej ze swej istoty kwestii społecznych (lub takiej, w której da się od nich w pewnej mierze abstrahować<sup>13</sup>), możemy do pewnego stopnia kontrolować wpływ społecznych oczekiwań na odpowiedzi badanych. W wypadku eksperymentu myślowego Kripkego możemy manipulować spodziewaną oceną moralną działania Schmidta i dzięki temu sprawdzić, czy rzeczywiście zaobserwujemy zmiany w odpowiedzi badanych. Jest to jednak trudniejsze w przypadku badań dotyczących kwestii nieodłącznie związanych ze sferą społeczną.

Drugim wskazanym przez Woolfolka czynnikiem, który może wpływać na odpowiedzi badanych, są oczekiwania eksperymentatora. Badani mogą próbować zgadywać, jaka jest hipoteza badania lub modyfikować swoje zachowanie z uwagi na pewne sygnały przekazywane im nieintencjonalnie przez eksperymentatora (por. Shadish, Cook, Campbell 2002, Cook, Campbell 1979).

Moim zdaniem w filozofii eksperymentalnej oddziaływanie oczekiwań społecznych i oczekiwań eksperymentatora na odpowiedzi badanych jest bardzo ograniczone. Po pierwsze, w filozofii eksperymentalnej, tak jak się ją współcześnie najczęściej uprawia, rola badacza w przebiegu eksperymentu jest stosunkowo skromna. Bardzo często badania przeprowadzane są przez internet za pomocą platform crowdsourcingowych, takich jak Amazon mTurk czy Clickworker. Jedynym kanałem komunikacji badacza z badanym jest kwestionariusz, który wypełnia uczestnik. Eksperymentator nie może więc obserwować zachowania uczestnika eksperymentu i aktywnie wywierać na niego presji.

---

<sup>13</sup> Można argumentować, że w filozofii nie znajdziemy żadnej kwestii, która nie ma wymiaru społecznego. Kwestia odniesienia nazw własnych, której użyłem jako ilustracji, jest przecież typowym przykładem zagadnienia, w którym aspekt społeczny może mieć niebagatelne znaczenie, o czym przypominają filozofowie stojący na stanowisku eksternalizmu semantycznego. Nie wszystkich kwestii oczekiwania społeczne dotyczą jednak w równym stopniu.

Jedyną realistyczną w tym wypadku możliwością wpływu eksperymentatora na wyniki badania jest tendencyjne formułowanie instrukcji dla badanego.

Po drugie, domyślenie się celu badania przez uczestników eksperymentu wydaje się stosunkowo trudne, szczególnie wówczas, gdy uwzględni się specyficzny charakter hipotez badawczych w filozofii eksperymentalnej. Zasadnicza część hipotez wyrasta z dość egzotycznych dla większości ludzi stanowisk filozoficznych. Nie ułatwia to badanym bez przygotowania filozoficznego domyślenia się celu eksperymentu. Można jednak argumentować, że chociaż odgadnięcie rzeczywistego celu badania wykracza poza kompetencje większości uczestników, to badani wciąż mogą próbować się go domyślać. Nawet jeśli nie uda im się zrekonstruować hipotez badawczych i odkryć prawdziwego celu eksperymentu, to ich domniemania na jego temat mogą wpływać na udzielane odpowiedzi. Zauważmy również, że w eksperymentach przeprowadzanych w schemacie wewnątrzgrupowym, to znaczy takich, w których (w przypadku badania kwestionariuszowego) jedna osoba konfrontowana jest z kilkoma scenariuszami, uczestnik może usiłować rozpoznać manipulowaną zmienną i wypracować systematyczny sposób odpowiadania. Ryzyko wystąpienia takiego podejścia u badanych jest szczególnie wysokie w sytuacji, gdy zadanie, przed którym stoi badany, jest trudne bądź niezrozumiałe (por. Scholl 2007).

Zagrożenie związane z potencjalnym odgadnięciem celu badania przez jego uczestników nie jest jednak tak poważne, jak twierdzi Woolfolk. Dysponujemy bowiem wynikami badań z zakresu filozofii eksperymentalnej, w których brali udział profesjonalni filozofowie, czyli osoby bez wątplenia posiadające wystarczające kompetencje do domyślenia się celu badania. W większości wypadków ich odpowiedzi nie różnią się systematycznie od odpowiedzi osób, które tych kompetencji nie mają. W badaniach z ich udziałem zaobserwowano efekt kolejności i efekt ramy odniesienia (por. Schwitzgebel, Cushman 2012, 2015), co powinno być niemożliwe, a przynajmniej mocno utrudnione, jeśli domyśliliby się celu badania<sup>14</sup>. Pozostaje jednak kwestia zagrożenia związanego z samą próbą odgadywania przez respondentów celu badania. Nawet jeśli badani nie będą w stanie domyślić się rzeczywistego celu eksperymentu, to ich domniemania na ten temat nadal mogą zniekształcać wyniki badania w systematyczny

---

<sup>14</sup> Zwróćmy uwagę, że dość łatwo wyobrazić sobie badanie z zakresu filozofii eksperymentalnej, w którym badacz świadomie decyduje się ułatwić badanemu dostrzeżenie celu badania i manipulowanych zmiennych. Przykładem może być eksperyment Wysockiego (2017), którego uczestnicy byli *explicite* proszeni o „mentalne” manipulowanie pewną zmienną w scenariuszu oraz konfrontowani z argumentami za uwzględnieniem pewnych czynników w swojej ocenie sytuacji. Oczywiście Wysocki w swoim eksperymencie chciał zbadać nie intuicje, a wpływ argumentów na ocenę przypadków Gettierowskich. Ukazuje to jednak pewną interesującą możliwość – całkowita „ślepotą” badanych na cel eksperymentu niekoniecznie jest bezwarunkowo pożądana.

sposób. Nie jesteśmy w stanie w pełni temu zapobiec, chociaż istnieją pewne techniki konstruowania scenariuszy, które ryzyko to pozwalają zmniejszyć. Literatura metodologiczna bogata jest w stosowne zalecenia (Rosenthal, Rosnow 1991, Nederhof 1985). W przypadku badań kwestionariuszowych dość użyteczną techniką jest sformułowanie instrukcji w taki sposób, aby przekonać osoby badane, że w ankiecie nie ma poprawnych i niepoprawnych odpowiedzi oraz że udzielone odpowiedzi nie będą służyć do oceny ich umiejętności oraz wiedzy.

### 2.3. Problem pomiaru i miary pośrednie w filozofii eksperymentalnej

Woolfolk stoi na stanowisku, że w wielu wypadkach lepsze od bezpośrednich pytań dotyczących kwestii interesujących filozofów eksperymentalnych byłoby skonstruowanie odpowiednich niebezpośrednich miar, które pozwoliłyby wnioskować o badanych zjawiskach (por. Payne, Gawronski 2010). Istnieją pewne problemy badawcze, w których samoopisowe miary nie zdają egzaminu. Woolfolk przywołuje w tym kontekście badania dotyczące rasizmu (por. Payne, Gawronski 2010). Ze względu na to, jak wrażliwą kwestią są postawy i zachowania rasistowskie, mamy powody sądzić, że miary samoopisowe nie pozwolą uzyskiwać wiarygodnych danych. Badani, nawet jeśli nie udzielają po prostu odpowiedzi zgodnych z oczekiwaniami społecznymi, to mogą nie zdawać sobie sprawy z własnych uprzedzeń i wynikających z nich rasistowskich przekonań czy zachowań. Nie wszystkie badania z zakresu filozofii eksperymentalnej dotyczą jednak kwestii związanych z tak silną presją społeczną, jak postawy rasistowskie. Woolfolk przecenia więc zagrożenia dla trafności badań związane z klasycznym paradygmatem narzędzi samoopisowych (*self-report*).

Postulat Woolfolka, zgodnie z którym w filozofii eksperymentalnej powinno używać się miar pośrednich, został na przykład zrealizowany w badaniach dotyczących odniesienia nazw własnych, gdzie pytania były zadawane uczestnikom na dwa sposoby. Pierwszy (pytanie metasemantyczne) to zadanie pytania o odniesienie przedmiotowe wyrażenia. Innymi słowy, w kwestionariuszu *explicite* pytano, do czego odnosi się badane wyrażenie. Drugi ze sposobów (pytanie semantyczne) polegał na zadaniu pytania o wartość logiczną zdania, w którym dane wyrażenie się pojawiało.

Z pytaniem metasemantycznym jako sposobem pomiaru intuicji semantycznych wiążą się dwa problemy. Po pierwsze, nie wiemy, co właściwie mierzymy za pomocą tak skonstruowanego pytania. Woolfolk zauważa, że nie mamy gwarancji, że badani interpretują pytania zadawane przez eksperymentatora w sposób przez niego zamierzony. Problem ten jest równie wyraźny, gdy w pytaniu pojawia się techniczny termin o dobrze ustalonym w filozofii znaczeniu, taki jak „odniesienie” („Do której osoby odnosi się nazwa *Gödel*?”), jak i gdy

próbuję się w pytaniu używać jakiegoś należącego do języka potocznego odpowiednika, który ma uchwycić rdzeń filozoficznego użycia słowa „odniesienie” („O której osobie mówi *S* (*is talking about*), kiedy używa nazwy *Gödel*?”) <sup>15</sup>.

Drugi problem związany jest z pewnego typu przesunięciem kategoryalnym. Jak wskazuje Genoveva Martí (2009), odpowiedzi na bezpośrednio zadawane pytania można potraktować jako korespondujące z wydawanymi przez badanych metajęzykowymi sądami o odniesieniu przedmiotowym użytych wyrażen. Filozofom eksperymentalnym niekoniecznie jednak chodzi o takie sądy – są raczej zainteresowani intuicjami z poziomu przedmiotowego, a nie z metapoziumu. Jeśli odpowiedzi wyrażają sądy metajęzykowe, to niewiele nam mówią o tym, do czego ludzie się odnoszą, używając danego wyrażenia. Różnicę łatwo uświadomić sobie, gdy rozważymy rodzaj kompetencji zaangażowanych w wydawanie sądów metajęzykowych oraz w interpretację odniesienia wyrażen. Żeby wydać metajęzykowy sąd dotyczący odniesienia jakiejś nazwy własnej, trzeba dysponować dość rozbudowanym aparatem pojęciowym („wyrażenie”, „desygnat”, „relacja odnoszenia się”). Natomiast żeby po prostu zinterpretować jakieś wyrażenie i ustalić jego odniesienie, wystarczy być kompetentnym użytkownikiem języka, do którego dane wyrażenie należy. Aby dowiedzieć się, jak kompetentni użytkownicy języka interpretują wyrażenia referencyjne, należy więc analizować nie sądy metajęzykowe dotyczące ich własności semantycznych, lecz ich językowe użycie (Martí 2009).

Drugą metodą pomiaru intuicji semantycznych jest pytanie o wartość logiczną zdania, w którym pojawia się interesujące wyrażenie (pytanie semantyczne). Zazwyczaj tego typu zdania testowe skonstruowane są tak, że przy jednej interpretacji wyrażenia zdanie jest prawdziwe, a przy drugiej fałszywe. W historii Kripke’owskiej moglibyśmy spytać, czy zdanie „Gödel jest odkrywcą dowodu twierdzenia o niezupełności” jest prawdziwe, czy fałszywe. Z udzielonej odpowiedzi moglibyśmy wywnioskować, czy badany interpretuje nazwę „Gödel” zgodnie z teorią przyczynowo-historyczną (zdanie ocenione jako prawdziwe), czy z deskrypcjonizmem (zdanie ocenione jako fałszywe). Mierzylibyśmy więc przyjmowane przez badanego odniesienie użytej nazwy własnej w sposób mniej bezpośredni, wywnioskowując ją z przypisywanej pewnemu zdaniu wartości semantycznej.

Zaletą tego rozwiązania jest to, że unikamy zadawania pytania o odniesienie (choć nie wszyscy badacze się z tym zgadzają, por. Martí 2013, 2014 oraz Devitt, Porot 2018). W konstrukcji pytania używamy jednak wciąż obciążonego teoretycznie terminu „prawdziwy”. Badani niekoniecznie muszą

---

<sup>15</sup> Ten drugi sposób użyty był na przykład w badaniu Machery’ego, Mallona, Nicholasa i Sticha (2004), które rozpoczęło międzykulturowy nurt badań w eksperymentalnej filozofii języka.

interpretować ten predykat tak, jak chciałby badacz, to znaczy jako pytanie o semantyczną prawdziwość. Mogą bowiem interpretować „prawdziwy” na przykład jako „pragmatycznie adekwatny w danej sytuacji”. Nie wiadomo również, jak interpretować udzielaną przez badanego odpowiedź „fałsz”. Nie można wykluczyć, że odpowiedź ta wybierana jest nie tylko wtedy, gdy osoba badana uważa, że zdanie jest fałszywe, lecz również wtedy, gdy badany nie chce zdaniu przypisać wartości logicznej prawdy z innego powodu niż jego fałszywość. Jeżeli zdanie jest pozbawione wartości logicznej albo źle zbudowane, to odpowiedź, że zdanie jest fałszywe, może być domyślnym wyborem (por. Abrisán, Szendroi 2013).

Interesującym przykładem wartym szczegółowego przeanalizowania jest badanie Michaela Devitta i Nicolasa Porota (2018) dotyczące odniesienia nazw własnych. Badanym przedstawione zostały scenariusze wzorowane na Kripke’owskich historyjkach o Gödlu i Schmidcie oraz o Jonaszu, użyte zostało zarówno pytanie metasemantyczne, jak i semantyczne. W jednym z warunków eksperymentalnych autorzy zdecydowali się dodatkowo na użycie zadania eksperymentalnego, które polegało albo na napisaniu przeznaczonej do podręcznika krótkiej notki o postaci opisanej w scenariuszu, albo na wyrażeniu o niej swojej opinii<sup>16</sup>.

Zaletą takiego podejścia jest to, że ani nie bada ono intuicji metajęzykowych, ani nie zakłada, że badani przypisują predykaty „prawdziwy” i „fałszywy”

<sup>16</sup> Przykładowo, w scenariuszu opisującym fikcyjnego chińskiego astronoma Tsu Chu’ung Chiha, który – analogicznie do Gödla w historyjce Kripkego – przypisał sobie odkrycia innego astronoma i był znany szerszej publiczności jedynie z tych odkryć, badacze zadawali uczestnikom eksperymentu pytanie: „Po przeczytaniu historyjki i przyjęciu założenia, że jest ona prawdziwa, jaka jest Twoja opinia o Tsu Chu’ng Chihu?”. Druga historyjka wzorowana była na Kripke’owskim przykładzie dotyczącym nazwy własnej „Jonasz”. W przykładzie tym zakładamy, że Jonasz faktycznie był postacią historyczną, ale żaden opis zawarty w Biblii (połknięty przez wieloryba itd.) nie jest prawdziwy o jakiegokolwiek rzeczywistej osobie. W badaniu Devitta i Porota scenariusz dotyczył nie postaci biblijnej, lecz fikcyjnego, legendarnego wodza Ambioryksa, który – jak to zostało odkryte przez historyków – wbrew obiegowej opinii nie walczył z Rzymianami. Zadawane badanym pytanie brzmiało: „W świetle tego odkrycia [dotyczącego „prawdziwego” Ambioryksa – BM], powiedz proszę, co belgijskie podręczniki historii powinny mówić o Ambioryksie, jeśli w ogóle powinny coś mówić. Zwiążcie uzasadnij”. W pierwszym wypadku z tego, jaką opinię o Tsu Ch’ung Chihu wyraził badany, wywnioskować można, do czego jego zdaniem odnosi się nazwa własna „Tsu Ch’ung Chih”. Jeśli bowiem opinia ta jest pozytywna (np. ktoś uważa go za świetnego naukowca), to domniemywać możemy, że badany interpretuje tę nazwę jako odnoszącą się do rzeczywistego odkrywcy opisanych w scenariuszu zjawisk astronomicznych. Jeśli opinia jest negatywna (np. ktoś uważa go za oszusta), to wnioskujemy, że interpretacją semantyczną „Tsu Ch’ung Chih” jest osoba, która ukradła wyniki badawcze i przypisała sobie ich odkrycie. Podobnie jest w przypadku Ambioryksa: jeśli ktoś opisze go jako legendarną, nieistniejącą postać, to zgodne jest to z teoriami deskrypcjonistycznymi, opisanie go jako rzeczywistego wodza, o którym opowieści obrosły legendami, świadczyć będzie o posługiwaniu się teorią przyczynowo-historyczną.

wypowiedziom w sposób, który odpowiada technicznemu użyciu tych pojęć. Jest to więc bardzo dobra niebezpośrednia miara zjawiska, które badacze chcieli zbadać. Wykorzystana metoda prowokowania wypowiedzi (*elicited production*) pozwala uchwycić fakty dotyczące rzeczywistego *użycia* pewnych środków językowych przez kompetentnych użytkowników języka. Interesującą obserwacją jest to, że w badaniu Devitta i Porota zaobserwowano dość duży stopień zgodności między tymi różnymi sposobami zadawania pytania. Na przykład w warunku eksperymentalnym, w którym badani mieli wyrazić swoją opinię o bohaterze (prowokowanie wypowiedzi), 95% z nich udzieliło odpowiedzi zgodnych z teorią przyczynowo-historyczną. Z kolei w warunkach eksperymentalnych, w których użyto pytania semantycznego, 95% i 81% badanych udzieliło takiej odpowiedzi. Kiedy użyto pytania metasemantycznego, było to 80% badanych. Pokazuje to, że chociaż badanym nie można przypisywać uprzywilejowanego dostępu do faktów dotyczących odniesienia przedmiotowego w ich języku, to ich sądy na ten temat są jednak dość wiarygodne.

Metodę prowokowania wypowiedzi Devitta i Porota (2018) można potraktować jako pośredni sposób mierzenia interesującego filozofów eksperymentalnych zjawiska (odniesienia przedmiotowego wyrażen referencyjnych, intuicji językowych). Po pierwsze, przykład ten pokazuje, że konstruowanie pośrednich miar jest do pewnego stopnia możliwe w filozofii eksperymentalnej nawet w standardowym paradygmacie badań ankietowych. Po drugie, postawić można hipotezę, że wskazane przez Woolfolka potencjalne zagrożenia metodologiczne wynikające z używania przez filozofów eksperymentalnych miar bezpośrednich nie są w praktyce tak groźne, jakby się wydawało. Świadczy o tym podobieństwo odpowiedzi uzyskanych za pomocą różnych „instrumentów” pomiarowych, to znaczy różnych sposobów zadawania pytań.

Pamiętać przy tym trzeba, że pośredni sposób pomiaru, który w opinii Woolfolka powinien znaleźć szersze zastosowanie w filozofii eksperymentalnej, również nie jest pozbawiony wad. Po pierwsze, skuteczność naszych operacjonalizacji zależy od szeregu dodatkowych założeń teoretycznych. Chociaż często da się je dobrze uzasadnić, to zawsze istnieje możliwość, że wykorzystana w eksperymencie operacjonalizacja jest niewłaściwa. Jak wskazałem, w przypadku pytania o prawdziwość zdania testowego operacjonalizacja mogłaby okazać się niewłaściwa, gdyby badani traktowali odpowiedzi „prawda” oraz „fałsz” inaczej niż jako odnoszące się do semantycznej wartości zdania.

Warto w tym miejscu zauważyć, że badania takie jak badanie Devitta i Porota (2018) odchodzą na tyle daleko od paradygmatu bezpośredniego pomiaru intuicji, że moglibyśmy sobie je wyobrazić w innej niż kwestionariuszowa formie. Zamiast pytania o prawdziwość pewnego stwierdzenia albo prowokowania wypowiedzi badanych, moglibyśmy prosić uczestników ekspe-



rymentu o wykonanie jakiegoś działania na wirtualnej reprezentacji bohaterów historyjki, które świadczyłyby o przyjętej przez badanego interpretacji nazwy własnej (np. „Teraz kliknij na nos Tsu Chu’ung Chiha”). Możliwa jest całkowita rezygnacja z proszenia badanych o wykonywanie zadań. Zamiast tego mierzyć można pewne nieświadome zachowania, które wskazywać mogą na przyjmowaną przez nich interpretację nazwy własnej. Na przykład w badaniu okulograficznym można przetestować, na którym z dwóch obrazków przedstawiających bohaterów historyjki badani koncentrują uwagę, słysząc odpowiednią nazwę własną. Wszystkie te metody pośredniego pomiaru przyjmowanego przez badanych odniesienia nazw są wykorzystywane w praktyce badawczej (por. np. Karabanov, Bosch, König 2007), chociaż w filozofii eksperymentalnej nie są jeszcze szeroko rozpowszechnione.

#### 2.4. Pragmatyka badania ankietowego

Badania ankietowe są sytuacją komunikacyjną, którą rządzą pewne pragmatyczne reguły konwersacji, nawet jeśli jest to konwersacja dość jednostronna. Znaczący to, że wypowiedzi badanych (odpowiedzi w kwestionariuszach filozofów eksperymentalnych) wyznaczane są zarówno przez treść semantyczną badania (scenariusze oraz pytania do nich), jak i przez czynniki pragmatyczne (por. Kauppinen 2007).

Dla filozofów eksperymentalnych zasadniczą kwestią powinno być odseparowanie czynników pragmatycznych od semantycznych, to znaczy ograniczenie wpływu nieistotnych czynników pragmatycznych na odpowiedzi respondentów. Antti Kauppinen (2007) uważa, że takiej operacji prawdopodobnie nie da się skutecznie przeprowadzić<sup>17</sup>. Zdaniem Kauppinena w przypadku zachowania komunikacyjnego w postaci zaznaczenia odpowiedzi w ankiecie nie możemy zrekonstruować istotnych aspektów kontekstu konwersacyjnego. Nie dysponujemy bowiem żadnymi informacjami o przekonaniach osoby czytającej ankietę, jej założeniach o przekonaniach drugiej strony konwersacji (badacza) itd. Nie jesteśmy w stanie wobec tego zrekonstruować całej pragmatyki konwersacji. Skoro zaś nie jest dla nas dostępna pragmatyka konwersacji, to nie możemy kontrolować czynników pragmatycznych w badaniu ankietowym.

<sup>17</sup> Jako przykład Kauppinen podaje omawiane wcześniej wyjaśnienie efektu Knobe’a zaproponowane przez Adamsa i Steadman (2004). Badacze ci uważają, że asymetrię w przypisywaniu intencjonalności zaobserwowaną przez Knobe’a (2003) można wyjaśnić przez odwołanie do czynników pragmatycznych: badani w ankiecie zaznaczają pozytywną odpowiedź na pytanie o umyślność działania polegającego na wywołaniu skutku ubocznego w warunkach szkodzenia, ponieważ zaznaczenie przeciwnej odpowiedzi pragmatycznie implikowałoby, że nie przypisują winy podmiotowi działającemu w scenariuszu.

Kauppinen (2007: 101) proponuje porzucenie w badaniach modelu ankiety (*Survey Model*) i przyjęcie modelu dialogicznego (*Dialogue Model*), w którym filozof eksperymentator zamiast przedstawiać badanemu filozoficzny przypadek, prowadziłby z nim rozmowę. W idealnej sytuacji ta metoda umożliwiłaby adekwatne poznanie pojęć używanych przez ludzi, dzięki stworzeniu warunków pozwalających na uzyskanie głębokich i stabilnych (*robust*) intuicji, czyli warunków takich, że na odpowiedź badanego wpływałyby wyłącznie względy semantyczne.

Zaproponowany przez Kauppinena model dialogiczny unikać ma ograniczeń badań kwestionariuszowych, w których jego zdaniem nie da się uzyskać idealnych warunków do rozważania przypadków filozoficznych. Ustalenie wyczerpującej listy takich warunków jest trudne, na pewno jednak zalicza się do nich umożliwienie pewnego typu ostrożnej czy głębszej refleksji nad rozważanymi przypadkami. W badaniach kwestionariuszowych, które często przeprowadzane są w salach wykładowych lub przez internet, stworzenie takich warunków jest prawdopodobnie niemożliwe. Postulat Kauppinena może zostać jednak w pewnym stopniu zrealizowany dzięki zastosowaniu metod jakościowych, w szczególności wywiadów pogłębionych (por. Andow 2016b).

W podobnym duchu co Kauppinen wypowiada się Simon Cullen (2010), który przekonuje, że w badaniach ankietowych przeprowadzanych przez filozofów eksperymentalnych niedostateczną uwagę poświęca się pragmatycznym aspektom ankiet. Cullen stoi jednak na stanowisku, że pragmatyczne efekty da się skutecznie zminimalizować. Aby to zrobić, należy uwzględnić dokonane poza filozofią eksperymentalną ustalenia z zakresu metodologii badań ankietowych. Sytuację badania ankietowego należy przy tym rozumieć jako rodzaj konwersacji, do której stosują się zwykle konwencje i normy komunikacji.

Przywoływane przez Cullena badania Norberta Schwarza (1995, 1996) pokazują, jak na odpowiedzi badanych wpływać może użyta w eksperymencie skala. Pragmatyczny efekt opisywany przez Schwarza polega na tym, że badani pytani „How successful would you say you have been in life?” inaczej odpowiadają w zależności od tego, czy skala rozciągała się od 1 do 10, czy od –5 do 5. Zdecydowanie więcej odpowiedzi znajdujących się po lewej stronie skali pojawia się w pierwszym wypadku niż w drugim. Zgodnie z zaproponowanym przez Schwarza wyjaśnieniem podpisy na krańcach skali („Not at all successful” oraz „Extremely successful”) są dwuznaczne i badani traktują wartość liczbową (która jest do pewnego stopnia arbitralna) jako wskazówkę interpretacyjną. „Not at all successful” interpretowane jest jako brak sukcesów, gdy towarzyszy mu liczba 1, a –5 badani interpretują jako wskazanie dużej liczby porażek. Odkrycie Schwarza jest według Cullena przykładem tego, jak zwykle pragmatyczne reguły konwersacji stosują się w sytuacji badania ankietowego.

Cullen przedstawia dwa dodatkowe studia przypadków z zakresu filozofii eksperymentalnej mające ilustrować wpływ pragmatyki ankietowej na odpowiedzi badanych. Pierwsze dotyczy eksperymentu zespołu Jonathana Weinberga (Weinberg, Nichols, Stich 2001), podczas którego badani byli konfrontowani ze scenariuszem modelowanym na filozoficznym przypadku Truetempa (Lehrer 1990) i mieli za zadanie określić, czy w przedstawionej im sytuacji mają do czynienia z wiedzą<sup>18</sup>. Weinberg i jego zespół użyli w swoim badaniu bardzo podobnej historyjki<sup>19</sup>. Historyjka kończyła się pytaniem, czy bohater miał wiedzę na dany temat. Badani wybierali jedną z dwóch odpowiedzi, albo „rzeczywiście widział” (*really knows*), albo „był tylko przekonany” (*only believes*)<sup>20</sup>.

Badanie to ma kilka niepokojących z punktu widzenia pragmatyki ankietowej elementów. Po pierwsze, jak zauważa Cullen, fraza *special feeling* za każdym razem ujęta jest w cudzysłów. Może to być jego zdaniem zrozumiane przez osoby badane jako wyraz nieprzychylniej opinii autora badania o szczególnym przeczuciu bohatera historyjki. To może natomiast sugerować im oczekiwaną przez autorów odpowiedź. Po drugie, również przedstawione badaniem możliwości odpowiedzi są sformułowane w sposób tendencyjny. Użycie przysłowka *really* sugeruje wysoki wymagany przez autorów standard wiedzy. Z kolei przysłówek *only* przy drugiej możliwości sugeruje niski standard przekonania. Cullen nie poprzestaje na teoretycznej krytyce metodologicznej. Przedstawia autorski eksperyment, w którym zmodyfikował możliwe odpowiedzi. Przedstawiał badanym scenariusz identyczny z tym wykorzystanym przez Weinberga i jego zespół, zamiast tendencyjnych *really knows* i *only believes* użył jednak pragmatycznie bardziej neutralnych *knows* i *doesn't know*. Zgodnie z oczekiwaniami badani byli w tej wersji eksperymentu bardziej skłonni przypisywać bohaterowi wiedzę niż w oryginalnym eksperymencie.

Efekty pragmatyczne stanowią więc istotne zagrożenie dla wiarygodności wyników uzyskiwanych przez filozofów eksperymentalnych. Szczegółowa analiza wykorzystywanych w kwestionariuszach materiałów badawczych może pozwolić zidentyfikować problematyczne z punktu widzenia pragmatyki

<sup>18</sup> Por. przypis 6.

<sup>19</sup> Pełna wersja użytej przez badaczy historyjki była następująca: Dave lubi grać w grę polegającą na rzucaniu monetą. Czasami, kiedy to robi, ma „szczególne przeczucie” (*special feeling*), że wynikiem następnego rzutu będzie orzeł. Jego „szczególne przeczucie” okazuje się mniej więcej w połowie przypadków trafne i mniej więcej w połowie przypadków błędne. Chwilę przed kolejnym rzutem Dave ma to „szczególne przeczucie”, w rezultacie czego jest przekonany, że wypadnie orzeł. Rzuca monetą i wyrzuca orła.

<sup>20</sup> W oryginale użyte zostało sformułowanie *only believed*, któremu w języku polskim lepiej odpowiada „tylko tak sądził”. Moim zdaniem jednak „tylko tak sądził” silniej niż *only believed* sugeruje, wbrew intencjom badaczy, że przekonanie podmiotu jest fałszywe.

sformułowania. Nie jest to więc nierozwiązywalny problem, a raczej wyzwanie metodologiczne dla eksperymentalistów.

## 2.5. Konsekwencje zarzutów metodologicznych dla filozofii eksperymentalnej

Wykorzystywana przez filozofów eksperymentalnych metoda kwestionariuszowa ma bardzo poważne ograniczenia. Odpowiedzi udzielane w badaniach ankietowych są narażone na zakłócający wpływ czynników pragmatycznych, oczekiwań społecznych i do pewnego stopnia również oczekiwań eksperymentatora.

Nie znaczy to, że kwestionariusze należy całkowicie porzucić. Dalej mogą być użyteczne w praktyce badawczej. Duża część omówionej krytyki metodologicznej nie ma wyłącznie charakteru destruktywnego, lecz może pomóc ulepszyć wykorzystywane przez filozofów eksperymentalnych narzędzia. W Tabeli 1 zebrana jest lista zaleceń sformułowana na podstawie omówionych zarzutów metodologicznych.

Problem	Zalecenia
Oczekiwania społeczne i oczekiwania eksperymentatora	<ul style="list-style-type: none"> <li>– Należy konstruować kwestionariusze w taki sposób, aby żadna z możliwych odpowiedzi nie była silnie preferowana tylko ze względu na oczekiwania społeczne.</li> <li>– Materiały badawcze używane do konstrukcji badań kwestionariuszowych (scenariusze itp.) powinny być w osobnych badaniach normowane ze względu na oczekiwania społeczne.</li> <li>– Osoby badane powinny otrzymywać jasne instrukcje zachęcające je do udzielania odpowiedzi oczekiwanych przez twórców eksperymentu (np. „W badaniu nie ma dobrych i złych odpowiedzi” itp.).</li> </ul>
Miary samoopisowe	<ul style="list-style-type: none"> <li>– Badania kwestionariuszowe należy replikować, wykorzystując inne miary niż samoopisowe (por. Devitta i Porota (2018) zadania prowokowania wypowiedzi).</li> </ul>
Pragmatyka badania ankietowego	<ul style="list-style-type: none"> <li>– Konstruowane badania kwestionariuszowe należy poddawać analizie ze względu na występowanie w nich niepożądanych czynników pragmatycznych.</li> <li>– Badania kwestionariuszowe należy replikować z innymi sposobami zadawania pytań i różnymi scenariuszami w celu upewnienia się, że odkryte efekty nie są wynikiem działania czynników pragmatycznych.</li> </ul>

Tabela 1. Lista problemów metodologicznych metody kwestionariuszowej wraz z zaleceniami pozwalającymi zminimalizować ich negatywny wpływ

Samo ulepszenie wykorzystywanych przez filozofów eksperymentalnych kwestionariuszy jednak nie wystarczy. Właściwym rozwiązaniem jest rozszerzenie zbioru metod wykorzystywanych przez filozofów eksperymentalnych. Nawiązać tutaj należy do pojęcia „metodologicznej triangulacji” (*methodological triangulation*, por. Campbell, Fiske 1959, Webb i in. 1966, Denzin 2007, Heesen, Bright, Zucker 2019). Metodologiczna triangulacja jest wywodzącą się z nauk społecznych strategią zakładającą, że w sytuacji, kiedy wyniki dostarczane przez pojedynczą metodę mogą być zakwestionowane lub kiedy nie możemy być ich pewni, powinniśmy skorzystać z wielu metod. Jeśli dostarczane przez kilka metod wyniki są zbieżne, wówczas możemy pokładać w rezultatach większe zaufanie niż w przypadku skorzystania z pojedynczej metody<sup>21</sup>.

W filozofii eksperymentalnej idea metodologicznej triangulacji również była dyskutowana<sup>22</sup>. W jej duchu w empirycznie zorientowanej filozofii matematyki postulowano połączenie metod kwestionariuszowych z metodami jakościowymi (Löwe, Kerkhove 2019). W moim przekonaniu metodologiczna triangulacja powinna obejmować nie tylko włączenie do filozofii eksperymentalnej metod jakościowych, ale też jak najszerszego repertuaru metod ilościowych współczesnej kognitywistyki.

## PODSUMOWANIE

Przedmiotem moich rozważań były niedostatki metod kwestionariuszowych. Główny zarzut metafizyczny głosi, że intuicje ewokowane przez filozoficzne eksperymenty myślowe są fundamentalnie nierzetelne. Ze względu na to, że kwestionariusze wykorzystywane w badaniach empirycznych są wzorowane na używanych w tradycyjnych rozważaniach filozoficznych eksperymentach myślowych, zarzut ten uderza również w filozofię eksperymentalną. W tradycyjnej filozofii z krytyki metafizycznej wynika postulat rewizji metod uprawiania filozofii. Dla filozofii eksperymentalnej krytyka ta powinna być sygnałem

<sup>21</sup> Należy wspomnieć, że idea metodologicznej triangulacji nie jest niekontrowersyjna i była w literaturze szeroko krytykowana. Przegląd zarzutów pod jej adresem można znaleźć w pracy (Heesen, Bright, Zucker 2019). Metodologiczni puryści, stojący w opozycji do postulujących triangulację pluralistów, uważają, że lepsza jest jedna metoda odpowiednia ze względu na badane zjawisko niż kilka niewłaściwych. W przypadku filozofii eksperymentalnej nie do końca wiadomo, która metoda miałaby być najwłaściwsza (a na pewno nie ma co do tego zgody wśród samych filozofów eksperymentalnych), dlatego uważam, że stanowisko puryzmu jest raczej mało obiecujące.

<sup>22</sup> Samo środowisko filozofów eksperymentalnych jest żywo zainteresowane rozszerzeniem repertuaru metodologicznego dyscypliny. Dowodem na to może być zorganizowana w 2017 roku przez Experimental Philosophy Group UK konferencja zatytułowana *Alternative Methods in Experimental Philosophy*.

do rozszerzenia katalogu wykorzystywanych w niej empirycznych metod badawczych i ulepszenia dotychczas używanych technik.

Oprócz krytyki metafizycznej w stronę filozofii eksperymentalnej wysuwane były również zarzuty metodologiczne. Najważniejsze z nich to: niepożądany wpływ społecznych oczekiwań i oczekiwań eksperymentatora na wyniki eksperymentu, problem bezpośrednich i niebezpośrednich miar oraz problem pragmatyki badania ankietowego. Wiele zagrożeń, na które wskazywali metodologiczni krytycy filozofii eksperymentalnej można zminimalizować, odpowiednio projektując kwestionariusze. Samo doskonalenie wykorzystywanych kwestionariuszy jest jednak niewystarczające. Ograniczenia metod kwestionariuszowych można przezwyciężyć jedynie uzupełniając repertuar metodologiczny filozofii eksperymentalnej o inne metody badawcze.

#### BIBLIOGRAFIA

- Abrusán M., Szendroi K. (2013), *Experimenting with the King of France: Topics, Verifiability and Definite Descriptions*, „Semantics and Pragmatics” 6(10), 1-43. <https://doi.org/10.3765/sp.6.10>
- Adams F., Steadman A. (2004), *Intentional Action in Ordinary Language: Core Concept or Pragmatic Understanding?*, „Analysis” 64(2), 173-181. <https://doi.org/10.1111/j.1467-8284.2004.00480.x>
- Alexander J., Weinberg J. M. (2014), *The Challenge of Sticking with Intuitions Through Thick and Thin* [w:] *Intuitions*, A. R. Booth, D. P. Rowbottom (eds.), Oxford: Oxford University Press, 187-212. <https://doi.org/10.1093/acprof:oso/9780199609192.003.0011>
- Alsmith A. J., Longo M. R. (2019), *Using VR Technologies to Investigate the Flexibility of Human Self-conception* [w:] *Methodological Advances in Experimental Philosophy*, E. Fischer, M. Curtis (eds.), London: Bloomsbury, 153-174. <https://doi.org/10.5040/9781350069022.ch-006>
- Alston W. (1995), *How to Think about Reliability*, „Philosophical Topics” 23(1), 1-29. <https://doi.org/10.5840/philtopics199523122>
- Andow J. (2016a), *Thin, Fine and with Sensitivity: A Metamethodology of Intuitions*, „Review of Philosophy and Psychology” 7(1), 105-125. <https://doi.org/10.1007/s13164-015-0247-2>
- Andow J. (2016b), *Qualitative Tools and Experimental Philosophy*, „Philosophical Psychology” 29(8), 1128-1141. <https://doi.org/10.1080/09515089.2016.1224826>
- Arico A., Fiala B., Goldberg R. F., Nichols S. (2011), *The Folk Psychology of Consciousness*, „Mind and Language” 26(3), 327-352. <https://doi.org/10.1111/j.1468-0017.2011.01420.x>
- Bara B. G., Ciaramidaro A., Walter H., Adenzato M. (2011), *Intentional Minds: A Philosophical Analysis of Intention Tested through fMRI Experiments Involving People with*



- Schizophrenia, People with Autism, and Healthy Individuals*, „Frontiers in Human Neuroscience” 5, 7. <https://doi.org/10.3389/fnhum.2011.00007>
- Bengson J. (2013), *Experimental Attacks on Intuitions and Answers*, „Philosophy and Phenomenological Research” 86(3), 495-532. <https://doi.org/10.1111/j.1933-1592.2012.00578.x>
- Blascovich J., Loomis J., Beall A. C., Swinth K. R., Hoyt C. L., Bailenson J. N. (2002), *Immersive Virtual Environment Technology as a Methodological Tool for Social Psychology*, „Psychological Inquiry” 13(2), 103-124. [https://doi.org/10.1207/s15327965pli1302\\_01](https://doi.org/10.1207/s15327965pli1302_01)
- Bluhm R. (2013), *Don't Ask, Look! Linguistic Corpora as a Tool for Conceptual Analysis* [w:] *Was dürfen wir glauben? Was sollen wir tun? Sektionsbeiträge des achten internationalen Kongresses der Gesellschaft für Analytische Philosophie*, M. Hoeltje, T. Spitzley, W. Spohn (Hrsg.), Duisburg–Essen: DuEPublico, 7-15.
- Bluhm R. (2016), *Corpus Analysis in Philosophy* [w:] *Evidence, Experiment and Argument in Linguistics and the Philosophy of Language*, M. Hinton (ed.), Frankfurt am Main: Peter Lang, 91-109.
- Campbell D. T., Fiske D. W. (1959), *Convergent and Discriminant Validation by the Multitrait-Multimethod Matrix*, „Psychological Bulletin” 56(2), 81-105. <https://doi.org/10.1037/h0046016>
- Cappelen H. (2012), *Philosophy without Intuitions*, Oxford: Oxford University Press. <https://doi.org/10.1093/acprof:oso/9780199644865.001.0001>
- Cook T. D., Campbell D. T. (1979), *Quasi-Experimentation: Design & Analysis Issues for Field Settings*, Boston: Houghton Mifflin.
- Cronbach L. J. (1950), *Further Evidence on Response Sets and Test Design*, „Educational and Psychological Measurement” 10(1): 3-31. <https://doi.org/10.1177/001316445001000101>
- Cullen S. (2010), *Survey-Driven Romanticism*, „Review of Philosophy and Psychology” 1(2), 275-296. <https://doi.org/10.1007/s13164-009-0016-1>
- Davidson D. (1987), *Knowing One's Own Mind*, „Proceedings and Addresses of the American Philosophical Association” 60(3), 441-458. <https://doi.org/10.2307/3131782>
- Denzin N. K. (2007), *Triangulation* [w:] *The Blackwell Encyclopedia of Sociology*, G. Ritzer (ed.), John Wiley & Sons. <https://doi.org/10.1002/9781405165518.wbeost050>
- Devitt M., Porot N. (2018), *The Reference of Proper Names: Testing Usage and Intuitions*, „Cognitive Science” 42(5), 1552-1585. <https://doi.org/10.1111/cogs.12609>
- Dongen, N. van, Colombo M., Romero F., Sprenger J. (2020), *Intuitions about the Reference of Proper Names: A Meta-Analysis*, „Review of Philosophy and Psychology”, 1-30. <https://doi.org/10.31234/osf.io/ez96q>
- Fischer E., Engelhardt P. E. (2016), *Intuitions' Linguistic Sources: Stereotypes, Intuitions and Illusions*, „Mind and Language” 31(1), 67-103. <https://doi.org/10.1111/mila.12095>
- Fischer E., Engelhardt P. E. (2017), *Diagnostic Experimental Philosophy*, „Teorema: Revista Internacional de Filosofía” 36(3), 117-137.

- Francis K. B., Howard C., Howard I. S., Gummerum M., Ganis G., Anderson G., Terbeck S. (2016), *Virtual Morality: Transitioning from Moral Judgment to Moral Action?*, „PloS one” 11(10), e0164374. <https://doi.org/10.1371/journal.pone.0164374>
- Goldman A. I. (1979), *What Is Justified Belief?* [w:] *Justification and Knowledge*, G. S. Papapas (ed.), 1-23, Dordrecht: Springer. [https://doi.org/10.1007/978-94-009-9493-5\\_1](https://doi.org/10.1007/978-94-009-9493-5_1)
- Goldman A. I. (1988), *Strong and Weak Justification*, „Philosophical Perspectives” 2, 51-69. <https://doi.org/10.2307/2214068>
- Haberstroh S., Oyserman D., Schwarz N., Kühnen U., Ji L. J. (2002), *Is the Interdependent Self More Sensitive to Question Context Than the Independent Self? Self-Construal and the Observation of Conversational Norms*, „Journal of Experimental Social Psychology” 38(3), 323-329. <https://doi.org/10.1006/jesp.2001.1513>
- Heesen R., Bright L. K., Zucker A. (2019), *Vindicating Methodological Triangulation*, „Synthese” 196(8), 3067-3081. <https://doi.org/10.1007/s11229-016-1294-7>
- Hofstede G. (1984), *Culture's Consequences: International Differences in Work-Related Values*, Newbury Park–London–New Delhi: Sage.
- Johnson T., Kulesa P., Cho Y. I., Shavitt S. (2005), *The Relation Between Culture and Response Styles: Evidence from 19 Countries*, „Journal of Cross-Cultural Psychology” 36(2), 264-277. <https://doi.org/10.1177/0022022104272905>
- Karabanov A., Bosch P., König P. (2007), „Eye Tracking as a Tool to Investigate the Comprehension of Referential Expressions” [w:] *Linguistics in Search of its Evidential Base*, S. Featherston, W. Sternefeld (eds.), Berlin–New York: De Gruyter Mouton, 207-226. <https://doi.org/10.1515/9783110198621.207>
- Kauppinen A. (2007), *The Rise and Fall of Experimental Philosophy*, „Philosophical Explorations” 10(2), 95-118. <https://doi.org/10.1093/acprof:osobl/9780199927418.003.0001>
- Knobe J. (2003), *Intentional Action and Side Effects in Ordinary Language*, „Analysis” 63(3), 190-194. <https://doi.org/10.1093/analys/63.3.190>
- Knobe J., Prinz J. (2008), *Intuitions about Consciousness: Experimental Studies*, „Phenomenology and the Cognitive Sciences” 7(1), 67-83. <https://doi.org/10.1007/s11097-007-9066-y>
- Kripke S. A. (1980), *Naming and Necessity*, Cambridge, MA: Harvard University Press.
- Kuś K., Maćkiewicz B. (2016), *Z rozmysłem, ale nie specjalnie. O językowej wrażliwości filozofii eksperymentalnej*, „Filozofia Nauki” 24(3) [95], 73-101.
- Lehrer K. (1990), *Theory of Knowledge*, Boulder: Westview Press.
- Loomis J. M., Blascovich J. J., Beall A. C. (1999), *Immersive Virtual Environment Technology as a Basic Research Tool in Psychology*, „Behavior Research Methods, Instruments, & Computers” 31(4), 557-564. <https://doi.org/10.3758/bfo3200735>
- Löwe B., Van Kerkhove B. (2019), *Methodological Triangulation in Empirical Philosophy* [w:] *Advances in Experimental Philosophy of Logic and Mathematics*, A. Aberdein, M. Inglis (eds.), London: Bloomsbury Methuen Drama, 15-38. <https://doi.org/10.5040/9781350039049.0005>

- Ludwig K. (2007), *The Epistemology of Thought Experiments: First Person Versus Third Person Approaches*, „Midwest Studies in Philosophy” 31(1), 128-159. <https://doi.org/10.2139/ssrn.3402517>
- Machery E. (2011), *Thought Experiments and Philosophical Knowledge*, „Metaphilosophy” 42(3), 191-214. <https://doi.org/10.1111/j.1467-9973.2011.01700.x>
- Machery E. (2017), *Philosophy within Its Proper Bounds*, Oxford: Oxford University Press. <https://doi.org/10.1093/oso/9780198807520.001.0001>
- Machery E., Mallon R., Nichols S., Stich S. P. (2004), *Semantics, Cross-Cultural Style*, „Cognition” 92(3), B1-12. <https://doi.org/10.1016/j.cognition.2003.10.003>
- Maćkiewicz B. (2021), „Korpusy w filozofii eksperymentalnej” [w:] *Efekt Knobe’a w świetle rozważań językoznawczych i metodologicznych. Studium teoretyczne i eksperymentalne efektu Knobe’a i problemu Butlera*, K. Kuś, B. Maćkiewicz, K. Paprzycka-Hausman (red.), Warszawa: Wydawnictwo Naukowe Semper, 123-136.
- Martí G. (2009), *Against Semantic Multi-Culturalism*, „Analysis” 69(1), 42-48. <https://doi.org/10.1093/analys/ann007>
- Martí G. (2013), *Empirical Data and the Theory of Reference* [w:] *Reference and Refering: Topics in Contemporary Philosophy*, W. P. Kabasenche, M. O’Rourke, M. H. Slater (eds.), Cambridge, Mass.: MIT Press, 63-82.
- Martí G. (2014), *Reference and Experimental Semantics* [w:] *Current Controversies in Experimental Philosophy*, E. Machery, E. O’Neill (eds.), New York: Routledge, 17-26. <https://doi.org/10.4324/9780203122884-2>
- Nederhof A. J. (1985), *Methods of Coping with Social Desirability Bias: A Review*, „European Journal of Social Psychology” 15(3), 263-280. <https://doi.org/10.1002/ejsp.2420150303>
- Nadelhoffer T., Nahmias E. (2007), *The Past and Future of Experimental Philosophy*, „Philosophical Explorations” 10(2), 123-149. <https://doi.org/10.1080/13869790701305921>
- Nichols S., Ulatowski J. (2007), *Intuitions and Individual Differences: The Knobe Effect Revisited*, „Mind and Language” 22(4), 346-365. <https://doi.org/10.1111/j.1468-0017.2007.00312.x>
- Paprzycka K. (2013), *Pojęcie wiedzy* [w:] *Przewodnik po epistemologii*, R. Ziemińska (red.), Kraków: Wydawnictwo WAM, 117-150.
- Payne B. K., Gawronski B. (2010), *A History of Implicit Social Cognition: Where Is It Coming from? Where Is It Now? Where Is It Going* [w:] *Handbook of Implicit Social Cognition: Measurement, Theory, and Applications*, B. Gawronski, B. K. Payne (eds.), New York–London: The Guilford Press, 1-15.
- Rose D., Danks D. (2013), *In Defense of a Broad Conception of Experimental Philosophy*, „Metaphilosophy” 44(4), 512-532. <https://doi.org/10.1111/meta.12045>
- Rosenthal R. (1976), *Experimenter Effects in Behavioral Research*, New York: Irvington.
- Rosenthal R., Rosnow R. L. (1991), *Essentials of Behavioral Research: Methods and Data Analysis*, McGraw-Hill Series in Psychology, New York: McGraw-Hill.

- Scholl B. J. (2007), *Object Persistence in Philosophy and Psychology*, „Mind and Language” 22(5), 563-591. <https://doi.org/10.1111/j.1468-0017.2007.00321.x>
- Schwarz N. (1995), *What Respondents Learn from Questionnaires: The Survey Interview and the Logic of Conversation*, „International Statistical Review” 63(2), 153-168. <https://doi.org/10.2307/1403610>
- Schwarz N. (1996), *Cognition and Communication: Judgmental Biases, Research Methods, and the Logic of Conversation*, Mahwah: Erlbaum.
- Schwitzgebel E., Cushman F. (2012), *Expertise in Moral Reasoning? Order Effects on Moral Judgment in Professional Philosophers and Non-philosophers*, „Mind and Language” 27(2), 135-153. <https://doi.org/10.1093/acprof:osobl/9780199927418.003.0014>
- Schwitzgebel E., Cushman F. (2015), *Philosophers' Biased Judgments Persist despite Training, Expertise and Reflection*, „Cognition” 141, 127-137. <https://doi.org/10.1016/j.cognition.2015.04.015>
- Shadish W. R., Cook T. D., Campbell D. T. (2002), *Experimental and Quasi-Experimental Designs for Generalized Causal Inference*, Boston: Houghton Mifflin.
- Stefanowitsch A. (2020), *Corpus Linguistics: A Guide to the Methodology*, Berlin: Language Science Press. <https://doi.org/10.5281/zenodo.3735822>
- Sytsma J., Bluhm R., Willemsen P., Reuter K. (2019), *Causal Attributions and Corpus Analysis [w:] Methodological Advances in Experimental Philosophy*, E. Fischer, M. Curtis (eds.), London: Bloomsbury Academic, 209-238. <https://doi.org/10.5040/9781350069022.ch-008>
- Weaver S., Turri J. (2018), *Personal Identity and Persisting as Many [w:] Oxford Studies in Experimental Philosophy*, vol. 2, T. Lombrozo, J. Knobe, S. Nichols (eds.), Oxford: Oxford University Press, 213-242. <https://doi.org/10.1093/oso/9780198815259.003.0010>
- Webb E. J., Campbell D. T., Schwartz R. D., Sechrest L. (1966), *Unobtrusive Measures: Non-reactive Research in the Social Sciences*, Chicago: Rand McNally.
- Weinberg J. M., Nichols S., Stich S. (2001), *Normativity and Epistemic Intuitions*, „Philosophical Topics” 29(1/2), 429-460. <https://doi.org/10.5840/philtopics2001291/217>
- Woolfolk R. L. (2013), *Experimental Philosophy: A Methodological Critique*, „Metaphilosophy” 44(1-2), 79-87. <https://doi.org/10.1111/meta.12016>
- Wysocki T. (2017), *Arguments over Intuitions?*, „Review of Philosophy and Psychology” 8(2), 477-499. <https://doi.org/10.1007/s13164-016-0301-8>