

Paweł Kawalec

Przyczynowość stanów mentalnych w modelach naukowych

Próba alternatywnego uzasadnienia antynaturalizmu eksplanacyjnego w ujęciu Urszuli Żegleń

1. Wprowadzenie

Często słucham arii *Va pensiero* Giuseppe Verdiego w wykonaniu chóru Metropolitan Opera z 2001 roku, ponieważ podnosi mój nastrój. Podobne wyjaśnienia, w których odwołujemy się do naszych pragnień, zamierzeń, przekonań czy innych stanów mentalnych, są codziennością. I na serio traktujemy ich rolę przyczynową w wyjaśnieniu – fakt, że ktoś posiada pewien stan mentalny sprawia, że zachodzi określone zdarzenie, np. to, że lubię ten fragment opery Verdiego, sprawia, że często go słucham. Jeśli ktoś sądzi, że bardziej naukowe spojrzenie podważa takie wyjaśnienia, jest w błędzie. Pokazuję to w pozostałej części tekstu.

Nauka jest zaawansowaną formą ludzkiego poznania. Jej ustalenia można śledzić, by na nowo przemyśleć filozoficzne koncepcje umysłu bądź wybiórczo te ustalenia wkomponowywać w zamknięty już pogląd na naturę umysłu. Urszula Żegleń w *Filozofii umysłu* proponuje zrównoważone stanowisko:

[...] umysł może stać się samodzielnym przedmiotem badań filozofa. Mimo to jednak filozof umysłu nie może nie liczyć się z rezultatami badań naukowych. Jest zobowiązany do ich śledzenia i w ich świetle powinien rewidować swoje teorie i oceniać siłę argumentów stawianych w obronie głoszonych tez. (2003, s. 136–137)

Tego rodzaju postulat ma określić pewien wzór odniesień do badań naukowych nad umysłem dla filozofa pragnącego jednocześnie zachować autonomię badań filozoficznych, ale i ich zgodność z aktualnymi ustaleniami nauki. Taki wzór wymaga pewnego dookreślenia.

2. Postulat metodologiczny

Zasadniczym kryterium *przekonania* naukowców o prawdziwości „rezultatów badań” jest ich zgodność z danymi doświadczenia w obrębie tego, co jest obserwowalne, czyli ich adekwatność empiryczna (van Fraassen 1980, 12).

W odniesieniu do takich rezultatów badań naukowych, które mogą się różnić między sobą, będąc w równym stopniu adekwatne empirycznie, należałoby mówić o ich *akceptacji*, która dokonuje się tylko z uwagi na inne względy niż prawdziwość.

Postulat zgodności z rezultatami badań naukowych rozumieć należy zatem jako postulat zgodności z nauką w zakresie tego, co podlega ocenie adekwatności empirycznej, a nie innych, np. motywacyjnych czy ideologicznych względów przesądzających o akceptacji jednej z kilku adekwatnych empirycznie teorii czy hipotez.

Podobnie jak ma to miejsce w nauce, filozofa umysłu, który nie prowadzi badań empirycznych, może obowiązywać wyłącznie zgodność z rezultatami badań naukowych w tym zakresie, w jakim podlegają one ocenie prawdziwości w nauce. W tym natomiast zakresie, w jakim o ich przyjęciu (akceptacji) decydują inne kryteria, filozof może akceptować dowolną z filozoficznych teorii umysłu, pod warunkiem zachowania zgodności z rezultatami badań naukowych adekwatnych empirycznie. W szczególności zaś filozof umysłu nie może rościć *przekonania* w stosunku do tych rezultatów naukowych, które w nauce podlegają wyłącznie akceptacji na podstawie kryteriów pozaprawdziwościowych.

Podsumowując powyższe dopowiedzania, otrzymujemy tezę, do której w dalszej części tekstu będę się odnosił, określając ją mianem postulatu metodologicznego (w skrócie PM):

(PM) Filozoficzne teorie umysłu muszą być zgodne z rezultatami badań naukowych w zakresie danych doświadczenia.¹

3. Antynaturalizm eksplanacyjny

W dalszej części tekstu postaram się wykazać, że ten dość oczywisty postulat stanowić może ważką przesłankę w uzasadnianiu głównej tezy *Filozofii umysłu*. Tezy głównej, którą Żegleń formułuje następująco:

[...] zdarzeń mentalnych nie da się zredukować do zdarzeń fizycznych (procesów neuronalnych) ani wyjaśnić jedynie w oparciu o te procesy. (Żegleń 2003, s. 21)

W dalszej części tekstu odnosić się będę wyłącznie do drugiej części tezy głównej, którą skrótowo można określić jako tezę antynaturalizmu eksplanacyjnego:

¹ Chodzi nie tylko o zgodność z danymi doświadczenia, które dostępne są teraz, ale także z takimi, które zostaną uzyskane w przyszłości. To jest istotne dopowiedzenie, gdyż w przeciwnym wypadku wystarczyłyby nam tylko modele danych, a nie również modele teoretyczne.

(AE') zdarzeń mentalnych nie da się wyjaśnić jedynie w oparciu o procesy neuronalne.

Jaegwon Kim wszystkie wyjaśnienia podawane na gruncie filozofii umysłu traktuje jako wyjaśnienia przyczynowe: „aby dowolne zjawisko spełniało funkcję wyjaśniającą, jego obecność lub nieobecność musi powodować różnicę – *różnicę przyczynową*” (1998/2002, s. 41).

Ponieważ „powodować różnicę przyczynową” wydawać się może prostym błędem w języku polskim, więc zamiast dokonywać egzegezy tego sformułowania, podam, jak będę rozumiał wyrażoną w nim krótko ideę Kima dotyczącą wyjaśniania przyczynowego.²

(WP) Jeśli dla badanego przez nas wycinka rzeczywistości określamy pewien zbiór zjawisk *V* jako istotnych dla określonej perspektywy badawczej, to zjawisko *Z* z tego zbioru *V* pełni funkcję wyjaśniającą względem innego zjawiska *E* ze zbioru *V*, o ile *E* jest bezpośrednim lub pośrednim potomkiem³ *Z* w modelu przyczynowym *M* dla zjawisk *V* oraz o ile z *M* nie można usunąć zjawiska *Z* bez wpłynięcia na adekwatność empiryczną *M*.

Z tezą Kima przypisującą wyjaśnianiu przyczynowemu w sensie (WP) najważniejszą rolę w badaniach nad umysłem można się ogólnie zgodzić, a nawet – jak pokazuję bardziej szczegółowo w (Kawalec 2006) – potraktować ją jako szczególnie przypadek obowiązujący w badaniach nieeksperymentalnych.

I pozorna byłaby tu niezgodność z zarzutem podniesionym przez Stanisława Judyckiego:

[...] błędem różnych współczesnych wersji fizykalizmu jest nieuświadomione założenie, iż umysł można ‘wyjaśnić’, posługując się kategorią przyczynowości fizycznej. Inaczej mówiąc, fizykalizmy nie trafiają w prawdziwą strukturę kategorialną umysłu i stąd czerpią swoją siłę przekonywającą. (Judycki 1995, s. 21)

W przeciwieństwie do pojęcia wyjaśniania przyczynowego (WP) kategoria „przyczynowości fizycznej” pozostaje bowiem niezgodna z postulatami metodologicznym (PM). Przyczynowość fizyczna, o której tu mowa, jest w stosunku do (WP) poszerzona o elementy z tego obszaru teorii fizykalnych, który jest przedmiotem powyżej krótko scharakteryzowanej akceptacji, niezależnie od ich związku z adekwatnością empiryczną.

Kategorialne rozminięcie się z przedmiotem fizykalnych teorii umysłu, co podkreśla Judycki, jest, jak sądzę, innym jeszcze sposobem pokazania trafności

² Ta eksplikacja kontynuuje tę zaproponowaną przez Żegleń (2003, 187).

³ Krótkie wprowadzenie tych terminów podaję w punkcie 5 poniżej.

pojęcia wyjaśniania przyczynowego (WP) w stosunku do postulatu metodologicznego.

W sensie naiwnym – który jest kategorialnie nietrafny – podanie wyjaśnienia przyczynowego można by rozumieć jako wymóg przywoływania w eksplanansie wyłącznie takich zależności przyczynowych, których istnienie stwierdzono w badaniach naukowych lub nieprzywoływania w eksplanansie zależności, których niezachodzenie stwierdzono w badaniach naukowych. Nietrafność zachodzi także i wtedy, gdy to przywołanie – w obu przypadkach – dokonuje się tylko z dokładnością do adekwatności empirycznej.

Podany na początku artykułu przykład wyjaśnienia w postaci pewnego stwierdzenia z zakresu retrospekcji podmiotu nie jest jednak wyjaśnieniem przyczynowym ani w sensie naiwnym, ani w postaci wyeksplikowanej jako (WP). Podanie wyjaśnienia spełniającego wymogi wyjaśnienia przyczynowego w sensie (WP) jest jednak *zasadniczo* możliwe, pod warunkiem że określony zostałby zbiór relewantnych zjawisk, dysponowalibyśmy odpowiednią liczbą danych dla tych zjawisk oraz zastosowalibyśmy techniki analizy tych danych, pozwalające wskazać zachodzące wśród nich zależności przyczynowe.

Uwzględniając powyższe uwagi na temat wyjaśniania w odniesieniu do filozofii umysłu, można zaproponować następującą eksplikację antynaturalizmu eksplanacyjnego (AE'):

(AE) Modeli intencjonalnych⁴ nie da się zastąpić modelami neurofizjologicznymi bez uchylenia adekwatności empirycznej.

Podkreśliśmy, że w swojej argumentacji Żegleń, kontynuując strategię Judyciego (1995), doprecyzowuje (AE') w inny sposób. Wybiera ona rozwiązanie, zgodnie z którym stany mentalne nie podlegają zależnościom przyczynowym, lecz innego rodzaju uwarunkowaniom. Zasadnicza idea tej strategii polegałaby więc na stwierdzeniu kategorialnej rozbieżności w przypadku prób stosowania wyjaśnień przyczynowych w odniesieniu do stanów mentalnych. Taka strategia nie stoi w opozycji do zaproponowanej poniżej argumentacji, pod warunkiem wszakże, że – jak tu proponuję – koncepcja wyjaśniania przyczynowego jest zgodna z postulatem metodologicznym.

Zasadniczym celem pozostałej części tego artykułu jest wykazanie, że postulat metodologiczny (PM) uzasadnia tę część głównej tezy *Filozofii umysłu*, która wyżej została wyodrębniona jako antynaturalizm eksplanacyjny (AE).

⁴ Modele intencjonalne są takiego rodzaju modelami przyczynowymi, w których występują zmienne, najczęściej ukryte, reprezentujące stany mentalne.

4. Struktura argumentu wykluczającego

Wydaje się, że najważniejszym z argumentów przeciw (AE) jest argument wykluczający skuteczność przyczynową stanów mentalnych⁵, a tym samym – zgodnie z postulowaną przez Kima dla filozofii umysłu koncepcją wyjaśniania – wykluczający także rolę eksplanacyjną stanów mentalnych.

Przez argument wykluczający rozumiem zestaw następujących tez:

- (1) Stany mentalne superwenują na fizycznych stanach w świecie.
- (2) Każde zdarzenie fizyczne ma kompletną fizyczną historię przyczynową.
- (3) Jeśli stany mentalne mają skutki fizyczne, to dzięki własnościom mentalnym, które egzemplifikują.
- (4) Własności mentalne nie są identyczne z fizycznymi.

(5) Przyczynowość mentalna i przyczynowość fizyczna mają w podstawach ten sam charakter (homogeniczność przyczyn).

(6) Poza wyjątkowymi przypadkami naddeterminacji żadne zdarzenie nie ma więcej niż jednej kompletnej historii przyczynowej (teza o jedyności historii przyczynowej).

W odniesieniu do tak skonstruowanego argumentu wykluczającego postaram się wykazać, że (PM) jako eksplikacja sformułowanego w *Filozofii umysłu* postulatu metodologicznego prowadzi do zakwestionowania ostatniej tezy.

5. Dookreślenie tezy o homogeniczności przyczyn

Na początku doprecyzuję jednak tezę o homogeniczności przyczyn. Implikuje ona, że przyczynowość mentalna jest tego samego rodzaju przyczynowością, co fizyczna. W obu przypadkach mamy do czynienia z tego samego rodzaju zależnością, która może być odnoszona do stanów fizycznych, jak i do stanów mentalnych. Innymi słowy, określenia „mentalne” i „fizyczne” modyfikują człony relacji, ale nie samą relację między nimi zachodzącą. A zatem samo stwierdzenie zachodzenia zależności przyczynowej między zjawiskami *X* i *Y* nie przesądza jeszcze ich ontologicznego garnituru. Przyjrzyjmy się nieco uważniej temu stwierdzeniu.

Zgodnie z tym, co stwierdziliśmy wcześniej w punkcie 3, zjawisko *X* pełni funkcję eksplanacyjną względem zjawiska *Y*, o ile model przyczynowy, w którym *Y* jest bezpośrednim lub pośrednim potomkiem przyczynowym *X*, po usunięciu z niego *X* stałby się nieadekwatny empirycznie. Tezę o homogeniczności przyczyn rozumiem więc tak, że włączenie zależności między dwoma zjawi-

⁵ Anomalizm mentalny D. Davidsona i przyjęta przez niego koncepcja wyjaśniania (1963/2001; 1967/2001; 1970/2001; 1976/2001), oprócz rozlicznych trudności, ma również tę, że nie daje się uzgodnić z przyjętą tu koncepcją wyjaśniania przyczynowego ani tym bardziej przy podważeniu we współczesnych koncepcjach wyjaśniania przyczynowego roli eksplanacyjnej praw przyrody; por. (Woodward 2003; Kawalec 2006).

skami ze zbioru zmiennych V do modelu przyczynowego dokonuje się na podstawie zastosowania tych samych reguł, nawet gdyby były to zmienne różnego typu.

W graficznych modelach zależności przyczynowych, jakie stosuje się w badaniach empirycznych⁶, sposób rozstrzygania o tym, czy stwierdzoną między X a Y zależność włączyć do modelu przyczynowego dla zbioru zmiennych V jest – oględnie mówiąc – następujący. Jeśli z założenia, że Y jest bezpośrednim potomkiem przyczynowym X wynika, że Y nie ma innych bezpośrednich potomków w zbiorze V , to zależność między X a Y można włączyć do modelu przyczynowego dla zbioru V .

W przypadku natomiast, gdy między Y a X w modelu M istnieje szereg zmiennych ze zbioru V , takich, że każda jest bezpośrednim potomkiem poprzedzającej, to można powiedzieć, że Y jest pośrednim potomkiem przyczynowym X .⁷

Z powyższych dopowiedzeń wynika pewien ważny morał, który często pomijany jest w rozważaniach filozofów umysłu. Jeśli poważnie traktuje się zgodność tych rozważań z rezultatami badań naukowych w sensie określonym przez postulat metodologiczny (PM), to tezę o homogeniczności przyczyn należy czytać po pierwsze przez pryzmat danych, na podstawie których dokonuje się ustalenia, że zależność między X a Y jest przyczynowa, a po drugie uwzględnić, że jest to zależność przyczynowa *względem* modelu przyczynowego dla zjawisk ze zbioru V , do którego należą X i Y . Przyczyny są homogeniczne, ale nie samotne.

Zaproponowaną tu precyzację tezy o homogeniczności można zrealizować prawdopodobnie także bez przyjmowania postulatu metodologicznego (PM), choć wówczas inna byłaby konstrukcja modeli, do których odnoszona będzie zależność przyczynowa. Zwięźle taką propozycję wyraża Robert Piłat:

Ściśle biorąc, relacja przyczynowa zachodzi pomiędzy modelem umysłowym a czynnością, jednak w konstytucję modeli wchodzi pewne stany przekonaniowe [...]. (Piłat 1999, s. 113)

Rozważania poświęcone homogeniczności zakończmy, zwracając się ku próbom jej zakwestionowania. Szeroko znana strategia Franka Jacksona i Philipa Pettita (1990) polega na wprowadzeniu odrębnej kategorii przyczynowości w odniesieniu do tego co mentalne, w postaci relewancji przyczynowej, która jest odmiennym typem relacji od zależności przyczynowej. Stany mentalne, mimo że nie wchodzi w relacje przyczynowe, mogą pełnić funkcję eksplanacyjną, gdyż są przyczynowo relewantne. I tak, w wyjaśnieniu jakiegoś skutku fizycznego można odwoływać się do stanów mentalnych, które są dla niego relewantne przyczynowo. Nie wchodzi one w zależność przyczynową z nim,

⁶ Przegląd takich zastosowań podaję w (Kawalec 2005), a ich systematyczną dyskusję przeprowadzam w (Kawalec 2006).

lecz co najwyżej gwarantują, że na niższym poziomie ontologicznym istnieje stan fizyczny, który realizuje dany stan mentalny i który wchodzi w odpowiednią zależność przyczynową. To różni relewantne stany mentalne od pozostałych.

Przypomnijmy kilka podstawowych zarzutów (Menzies 2003), jakie podważają negację tezy o homogeniczności, zaproponowaną przez Jacksona i Pettita. Fakt, że podane przez Jacksona i Pettita warunki relewancji przyczynowej stanów mentalnych są nietrywialne i oddają intuicyjnie stwierdzaną różnicę między stanem mentalnym, który w ogóle nie ma związku z danym stanem fizycznym, np. ruchem ciała, a takim, który jest relewantny przyczynowo, nie zmienia jednak epifenomenalnego charakteru stanów mentalnych.

Ponadto, nawet jeśli byśmy uznali, że pojęcie relewancji przyczynowej jest wystarczającą podstawą do przypisania stanom mentalnym funkcji ekplanacyjnej, to jednak relewancja przyczynowa przysługująca stanom mentalnym ma niewiele wspólnego z tym, co uważamy za przyczynowość i jak potocznie określamy rolę przyczynową tych stanów.

Przypomnijmy też przywołaną w punkcie 3 strategię Judyckiego (1995; 2000), którą także stosuje w swojej książce Żegleń (2003, s. 185–201). Polega ona na wprowadzeniu kategorii uwarunkowania, która ma dotyczyć determinacji stanów mentalnych w sposób swoisty i nieprzyczynowy. Ta strategia, jak wyjaśniałem to wcześniej, nie wydaje mi się niezgodna z proponowaną tu, gdyż w obu przypadkach występuje odmienne pojęcie przyczynowości. Zauważmy jednak, że nawet gdybym był w błędzie i omawiana strategia podważałaby prawdziwość tezy o homogeniczności przyczyn, to i tak osiągnięty zostałby ten sam zasadniczy cel, jakim jest obalenie argumentu wykluczającego.

6. Podważenie tezy o jedyności historii przyczynowej

Przejdźmy obecnie do szczegółowej dyskusji dwóch powodów, które podważają jedną z tez argumentu wykluczającego – tezę (6) o jedyności historii przyczynowej. Kim wyraża zwięzłe motywujące ją obawy:

[...] obecność dwóch opisów przyczynowych, z których każdy oferuje kompletne wyjaśnienie przyczynowe danego zdarzenia, stwarza sytuację niepewną, wymagającą od nas wytłumaczenia, w jakiej relacji pozostają do siebie dwie wchodzące tutaj w grę przyczyny. (Kim 1998/2002, s. 77)

Sądzę, że obawy Kima mają źródło nie w możliwości podania dwóch kompletnych wyjaśnień przyczynowych, ale w tym, że przyczynowość jest przez niego kategoryzowana (Kim 1998/2002, s. 76) wedle metod J. S. Milla.⁸ „Przy-

⁷ Precyzyjną terminologię podają (Lauritzen 1996; Pearl 2000; Spirtes, Glymour, Scheines 2000).

⁸ Nie tyle istotne jest, że dziś mają one za sobą już długą historię, ale że od zarania nie

czyna konieczna" i „przyczyna wystarczająca" są z pewnością pożyteczne przy eksplikacji metody zgodności i różnicy. I tak np. w przypadku metody zgodności (bezpośredniej i odwróconej) na podstawie schematycznych zależności „Skutek=1 $\rightarrow$ Przyczyna=?" oraz „Przyczyna=? $\rightarrow$ Skutek=0" ustalamy odpowiednio przyczynę (towarzyszącą skutkowi) w sensie warunku koniecznego i (nieobecna przy nieobecnym skutku) w sensie warunku wystarczającego. Oczywiście, jest pewna analogia między podwójną metodą zgodności (bezpośredniej i odwróconej), a stosowanymi współcześnie metodami wykrywania zależności przyczynowych w badaniach obserwacyjnych (Rosenbaum 2002), ale eksplorowanie tej analogii może co najwyżej przynieść pożytek dydaktyczny i słuszne poczucie niepewności.

Sens „kompletnego wyjaśnienia przyczynowego" trudno jest inaczej wyeksplikować niż tak, że chodzi w tym przypadku o podanie modelu przyczynowego, który jest adekwatny empirycznie względem wszystkich posiadanych przez nas danych doświadczenia dla relewantnego zbioru zjawisk V . Nie dostrzegam jednak powodu, dla którego miałoby być tak, że jeśli zmienna Y jest elementem dwóch różnych zbiorów zjawisk V oraz W i dla każdego z nich podajemy model przyczynowy, to relacje przyczynowe, w jakich występuje Y , w obu modelach musiałyby być identyczne lub redukowalne.

Przywołajmy prostą ilustrację. W badaniach wpływu spożycia soli kuchennej S na poziom ciśnienia krwi C wykorzystuje się zarówno dane pochodzące z obserwacji ludzi, jak i dane z eksperymentów na zwierzętach.⁹ Niech V będzie zbiorem zjawisk, które badacze uznają za istotne w przypadku obserwacji ludzi, a W – zbiorem zjawisk istotnych np. w przypadku makaków. W oparciu o co mielibyśmy teraz zgłosić żądanie, aby zależność łącząca S i C po pierwsze występowała (lub nie) w obu modelach M_V oraz M_W i, po drugie, aby w przypadku gdy występuje, miała dokładnie taką samą postać, np. $C = f_S(S, \epsilon_S)$?

Wróćmy do przykładu podanego na początku. Jeśli to rzeczywiście byłoby wyjaśnienie na tyle ważne, żeby poświęcać mu wiele uwagi w badaniach naukowych, to również można opisać co najmniej dwa różne sposoby uformowania modeli przyczynowych, z których każdy odnosiłby się do częstości słuchania *Va pensiero* w związku z podnoszeniem nastroju. Utworzenie jednego z takich modeli bazowałoby na śledzeniu aktywności różnych części mózgu i określaniu towarzyszącego temu przebiegu reakcji fizyko-chemicznych w poszczególnych częściach mózgu. Drugi z takich modeli bazowałby na pomiarze częstości słuchania *Va pensiero*, towarzyszących temu stanów mentalnych u osób

były to metody naukowe ustalania zależności przyczynowych w sensie, który dopuszczany jest przez (PM).

⁹ Pomińmy fakt, że w niektórych modelach przyczynowych opartych na obserwacji nawyków żywieniowych różnych społeczności stwierdza się występowanie zależności przyczynowej między S a C , a w innych nie (Freedman, Pettiti 2002).

o usposobieniu, które byłoby dopasowane do mojego¹⁰, oraz ewentualnym śledzeniu innych zjawisk będących wskaźnikami stanów mentalnych.

Oczywiście operacjonalizacja stanów mentalnych, jak podnoszenie nastroju, która jest niezbędna do przeprowadzenia badań, i podanie wskaźników nie zastępuje samego stanu mentalnego. Przyznać więc należy, że drugi ze wskazanych powyżej modeli zawierałby zmienną niemierzalną, która reprezentowałaby ów stan mentalny i która byłaby przyczynowo związana w tym modelu ze zmiennymi mierzalnymi, które są wskaźnikami odnośnego stanu mentalnego. Jedno z najważniejszych odkryć współczesnych badań nad modelowaniem zależności przyczynowych dotyczy charakterystyki tego rodzaju modeli. Otóż, jak wykazuje Judea Pearl (2000, s. 77–83), przy spełnieniu pewnych kryteriów modele zawierające zmienne niemierzalne pozwalają na identyfikację istotnych parametrów, które następnie pozwalają na ocenę adekwatności empirycznej tych modeli.

Zilustruję to prostym przykładem. Załóżmy, że mamy model przyczynowy, w którym występują zmienne mierzalne X , Y , Z oraz zmienna niemierzalna U . Między nimi zachodzą następujące zależności przyczynowe: $X \rightarrow Y \rightarrow Z$ oraz $X \leftarrow U \rightarrow Z$. Zgodnie z ustaleniem Pearl'a, można obliczyć efekt przyczynowy X w stosunku do Z w taki sposób, aby zależał on tylko od wielkości znanych z obserwacji zmiennych mierzalnych. Tym samym ten model podlega ocenie pod względem adekwatności empirycznej, mimo iż zjawiska X oraz Z pozostają pod wpływem przyczynowym zmiennej niemierzalnej U .

Podkreślmy, że każdy z obu wspomnianych wyżej rodzajów modeli przyczynowych, niezależnie od określonej w nich relacji przyczynowej między słuchaniem *Va pensiero* a podnoszeniem nastroju, byłby sformułowany niezależnie od drugiego, spełniając warunki adekwatności empirycznej w stosunku do innego zbioru danych pozyskanych dla różnych zbiorów zjawisk.¹¹

Podejrzewałbym raczej, że tym, co faktycznie „stwarza sytuację niepewną” jest niekiedy właśnie przeciwieństwo sytuacji opisywanej przez tezę o jedyności historii przyczynowej. Mianowicie, niemożność podania kilku – komplementarnych i zarazem – kompletnych modeli przyczynowych danego zdarzenia. Ilustracjami w tym sensie „niepewnych” przypadków mogą być np. osoba o pseudonimie Elliot (Damasio 1994/1999) czy kobieta poddana operacji na ogromnego tętniaka tętnicy podstawnej (Sabom 1998; Greyson 2000, 339–342).

¹⁰ Chodzi w tym przypadku o stosowaną w badaniach obserwacyjnych, ustalających zależności przyczynowe, technikę statystyczną zwaną dopasowywaniem (ang. *matching*); por. (Rosenbaum 2002).

¹¹ Są powody uzasadniające sceptycyzm w stosunku do możliwości zbliżenia modelu neurofizjologicznego i modelu intencjonalnego. Po pierwsze, adekwatność empiryczna jest w przypadku obu modeli niezależna. Po drugie, z uwagi na wieloraką realizowalność stanów mentalnych dane o procesach je realizujących mogą się znacznie różnić w poszczególnych realizacjach; por. też (Menzies 2003, s. 31).

Ponieważ pierwszy z tych przypadków jest szeroko znany, więc zatrzymamy się krótko przy drugim:

Ze względu na rozmiar i umiejscowienie tętniaka niemożliwe było jego bezpieczne usunięcie przy wykorzystaniu standardowych technik neurochirurgii. Podjęto decyzję o hipotermicznym zatrzymaniu akcji serca, podczas którego temperaturę ciała kobiety obniżono do 60°F [ok. 17°C], zatrzymano rytm serca i oddychanie, fale mózgowe stały się płaskie oraz wypompowano krew z jej głowy. W tym momencie spełniała ona wszystkie ogólnie przyjęte kryteria śmierci mózgu: jej EEG było całkowicie płaskie, wskazując całkowity zanik aktywności elektrycznej kory mózgowej, zanikł wzbudzany słuchowo potencjał, wskazując na zaprzestanie funkcji pnia mózgu, a krew była całkowicie wypompowana z mózgu, co doprowadziło do braku występowania jakiegokolwiek jego funkcji. [...] w tym przypadku posiadamy ciągły zapis ciśnienia krwi, pojemności minutowej serca, częstości i rytmu akcji serca, natlenienia krwi, temperatury w podstawowych punktach ciała i mózgu, aktywności kory mózgowej i potencjału wzbudzanego w pniu mózgu. (Greyson 2000, 339; tłum. Paweł Kawalec)

Dzięki obniżeniu temperatury tempo metabolizmu w mózgu spadło na tyle, że można było utrzymywać ten stan przez ok. 45 minut. Zdarzenia mające miejsce w tym czasie na sali operacyjnej były jednak szczegółowo relacjonowane przez pacjentkę poddaną operacji:

Opisała ona dokładnie dwudziestoosobowe grono lekarzy, pielęgniarek i personelu technicznego obecne na sali operacyjnej, mimo że większości z nich nie znała wcześniej, a także kilka charakterystycznych szczegółów unikalnej piły pneumatycznej, której użyto do rozcięcia jej czaszki, zaskoczenie chirurga, że tętnica udowa jest zbyt wąska, by można było ją wykorzystać, a także muzykę, którą odtwarzano w sali operacyjnej, podczas gdy była nadal nieprzytomna. (Greyson 2000, 340; tłum. Paweł Kawalec)

Jak podkreśla Greyson, z uwagi na kompletność dokumentacji stanu pacjentki stwierdzającej zachodzenie „wyjątkowo przejrzystych procesów mentalnych” podczas operacji wykluczone są jakiegokolwiek modele fizjologiczne czy fizyczne, które byłyby adekwatne empirycznie.¹² Gdyby więc miał powstać w tym przypadku model przyczynowy, byłby to model z pewnością niezawierający zjawisk neurofizjologicznych, lecz pewne stany mentalne.

Drugim powodem, dla którego powątpiewać należy w tezę o jedności historii przyczynowej, jest pewna własność zależności przyczynowych. Eksplicując tezę o homogeniczności zależności przyczynowej, stwierdziliśmy, że określenie, czy zależność między X a Y jest przyczynowa, dokonuje się w stosunku do zmiennych występujących w tym samym modelu dla zbioru zjawisk V oraz w stosunku do danych, względem których ocenia się adekwatność empiryczną tego modelu. Z uwagi na rodzaj zjawisk, z jakimi mamy do czynienia w danym przypadku, możliwe jest wystąpienie sytuacji, gdy w równym stopniu adekwat-

¹² Najbardziej prawdopodobne z modeli fizjologicznych w tym przypadku rozważa – i odrzuca – Bruce Greyson (2000, 340-41).

ne empirycznie – i to w stosunku do tego samego zbioru danych – są dwa modele przyczynowe, które różni to, że jeden z nich zawiera zmienne odpowiadające niemierzalnym stanom mentalnym, podczas gdy drugi zawiera wyłącznie zmienne fizyczne.

Niektóre zjawiska, np. występowanie zaburzeń w funkcjonowaniu mózgu człowieka z powodu usunięcia fragmentów płatów czołowych, jak w przypadku wspomnianego wyżej Elliota, z zasady nie dopuszczają możliwości badań eksperymentalnych. W takim przypadku dane, jakimi dysponujemy dla określenia adekwatności empirycznej modelu przyczynowego, mają wyłącznie charakter statystyczny (w postaci łącznego rozkładu prawdopodobieństwa dla tych zmiennych występujących w modelu, które są mierzalne).¹³ Stąd też pod względem adekwatności empirycznej zrównane będą modele, które w różny sposób przewidują przyczynowe efekty manipulacji eksperymentalnej, dla której nie ma podstaw w danych do ich estymacji.

Zilustruję to prostym przykładem. Załóżmy, że w opracowanym na podstawie wyników badań obserwacyjnych Richarda Dolla i Bradforda Hilla z lat 50. XX wieku¹⁴ modelu przyczynowym zachodzi zależność między paleniem papierosów (P) a rakiem płuc (R): $P \rightarrow R$. Względem danych, jakimi posługiwali się w formułowaniu tego modelu Doll i Hill, adekwatny empirycznie – lecz odmienny – model przyczynowy zaproponował Ronald Fisher. Mianowicie skłonność do palenia papierosów oraz do zachorowalności na raka płuc są od siebie przyczynowo niezależne, gdyż mają wspólną przyczynę, jaką jest posiadanie szczególnego genu (G): $P \leftarrow G \rightarrow R$.

Oba modele, tj. model Dolla–Hilla oraz model Fishera, są równie adekwatne empirycznie, przy czym zasadnicza między nimi różnica polega na tym, że model Fishera wprowadza zmienną niemierzalną, jaką w tym przypadku jest G .

Podany tu przykład jest tylko uproszczoną ilustracją przypadku, który jest charakterystyczny dla obszaru badań naukowych, jaki określa się mianem badań obserwacyjnych (Rosenbaum 2002). Mimo że przypadki występowania w badaniach obserwacyjnych modeli przyczynowych, które są adekwatne empirycznie, choć podają różne „historie przyczynowe” jakiegoś zdarzenia, są nieuchronne, to jednak w tym przypadku słuszny może wydawać się niepokój zasygnalizowany przez Kima. Jeśli to tylko możliwe należy dążyć do pozyskania tego rodzaju danych, które pozwolą wyeliminować alternatywne historie przyczynowe danego zdarzenia. Nie zmienia to jednak faktu, że w pewnego rodzaju badaniach naukowych występowanie alternatywnych historii przyczynowych jest nagminne i nieuchronne, co filozof umysłu – zgodnie z postulatem (PM) – może jedynie odnotować.

¹³ Szczegółowe wyjaśnienie podaje (Pearl 2000, 65-105)

¹⁴ Etapy tych badań bardziej szczegółowo omawiam w (Kawalec 2005), gdzie podana jest także literatura źródłowa.

Podważenie przesłanki (6) argumentu wykluczającego prowadzi do konsekwencji takiej, że to samo zdarzenie może mieć kilka historii przyczynowych. Różnica pomiędzy tymi historiami nie jest oczywiście różnicą wyłącznie tego rodzaju, jaka jest dopuszczalna w przypadku naddeterminacji. Różne przyczyny danego zdarzenia w przypadku naddeterminacji nie wymagają wprowadzenia kilku różnych modeli przyczynowych. W przypadku jednak przyczynowości mentalnej model intencjonalny spełnia (lub nie), jak starałem się pokazać na podstawie postulatu (PM), warunki adekwatności niezależnie od modelu neurofizjologicznego dla danego zdarzenia. Z przyjętego tu pojęcia wyjaśniania przyczynowego wynika z kolei, że wyjaśnienia przyczynowe podane w oparciu o model neurofizjologiczny w żaden sposób nie implikują eksplanacyjnej redundancji modelu intencjonalnego, jeśli taki istnieje w danym przypadku.¹⁵ Podkreślmy jednak, że twierdzenia dotyczące przyczynowości, podawane w sposób zgodny – w sensie (PM) – z rezultatami badań naukowych, są relatywne do modelu, w związku z tym różne będą także wyjaśnienia przyczynowe.

7. Zakończenie

Jeśli przedstawiona powyżej argumentacja za antynaturalizmem eksplanacyjnym (AE) w oparciu o postulat zachowania w filozofii umysłu zgodności z rezultatami naukowymi w sensie (PM) jest słuszna, to prowadziłoby to nas do interesującego spostrzeżenia w odniesieniu do tezy o superweniencji (1) wchodzącej w skład argumentu wykluczającego. W dyskusji nad przyczynowością stanów mentalnych nie odnosiłem się do tej tezy, co pokazywałoby, że jest to hipoteza filozoficzna, która nie wpływa na adekwatność empiryczną filozoficznych teorii umysłu w zakresie przypisywania skuteczności przyczynowej stanom mentalnym. W tym sensie więc jej akceptacja uzależniona byłaby od innych niż prawdziwość względów, przede wszystkim, jak się zdaje, od przyjęcia światopoglądu materialistycznego.¹⁶

¹⁵ Potwierdza to także praktyka stosowana np. w badaniu psychopatologii, gdzie wykorzystuje się fakt, że występowanie niektórych stanów mentalnych jest symptomem zaburzeń psychicznych. W związku z tym konstruuje się – często zalgorytmizowane – instrumenty (np. ankiety, kwestionariusze), które pozwalają zaklasyfikować dane zaburzenie do jednej ze standardowych kategorii (International Classification of Diseases 10 WHO 1992; Diagnostic and Statistical Manual IV APA 1994), dzięki zestandaryzowanemu pomiarowi symptomów i oznak. Występujące tu podstawowe zmienne: wygląd i zachowanie; zaburzenia mowy; nastroju; treść i forma myśli; nienormalne percepcje; zaburzenia poznania świadczą o łączeniu w modelach zaburzeń umysłowych zmiennych kategorizowanych jako mentalne i niementalne (fizyczne, behawioralne), jak choćby w modelu przyczyn anhedonii. W przypadku modeli przyczynowych zasadniczą cechą istotną z punktu widzenia (PM) jest adekwatność empiryczna modelu. W dość intuicyjny sposób przyjmuje się nieredukowalność modeli intencjonalnych do neurofizjologicznych, gdyż jedno i drugie generują różne obserwowalne rozkłady; por. (Farmer i in. 2002).

Bibliografia prac przywołanych w tekście

- Damasio A. (1994/1999), *Błąd Kartezjusza. Emocje, rozum i ludzki mózg*, tłum. M. Karpiński, Poznań: Rebis.
- Davidson D. (1963/2001), *Actions, reasons, and causes*, [w:] Davidson 2001, s. 3–19.
- Davidson D. (1967/2001), *Causal relations*, [w:] Davidson 2001, s. 149–162.
- Davidson D. (1970/2001), *Mental events*, [w:] Davidson 2001, s. 207–225.
- Davidson D. (1976/2001), *Hempel on explaining action*, [w:] Davidson 2001, s. 261–275.
- Davidson D. (2001), *Essays on actions and events*, Oxford: Clarendon.
- Farmer A., McGuffin P., Williams J. (2002), *Measuring psychopathology*, Oxford: Oxford University Press.
- Freedman D. A., Petitti D. B., 2002, *Salt, blood pressure, and public policy*, „International Journal of Epidemiology”, t. 31, s. 319–320.
- Greyson B. (2000), *Near-death experiences*, [w:] *Varieties of anomalous experience. Examining the scientific evidence*, red. E. Cardena, S. J. Lynn i S. Krippner, Washington, DC: American Psychological Association, s. 315–352.
- Judycki S. (1995), *Umysł i synteza*, Lublin: RW KUL.
- Judycki S. (2000), *Transkauzalność a determinizm*, „Kognitywistyka i Media w Edukacji”, t. 3, s. 73–86.
- Kawalec P. 2005, *Understanding science of the new millennium*, <http://philsci-archive.pitt.edu/archive/00002558/>.
- Kawalec P. (2006), *Przyczyna i wyjaśnianie. Studium z metodologii i filozofii nauki*, Lublin: Wyd. KUL.
- Kim J. (1998/2002), *Umysł w świecie fizycznym*, tłum. R. Poczobut, Warszawa: IFiS PAN.
- Lauritzen S. (1996), *Graphical models*, Oxford: Clarendon.
- Menzies P. (2003), *The causal efficacy of mental states*, [w:] *Physicalism and mental causation. The metaphysics of mind and action*, red. S. Walter and H. Heckmann, Imprint Academic, s. 195–223.
- Pearl J. (2000), *Causality. Models, reasoning, and inference*, Cambridge: Cambridge University Press.
- Piłat R. (1999), *Umysł jako model świata*, Warszawa: IFiS PAN.
- Rosenbaum P. (2002), *Observational studies*, Nowy Jork: Springer.
- Sabom M. (1998), *Life and death. One doctor's fascinating account of near-death experiences*, Grand Rapids: Zondervan.
- Spirtes P., Glymour C., Scheines R. (2000), *Causation, prediction, and search*, Cambridge, MA.: MIT.
- van Fraassen B. (1980), *The scientific image*, Oxford: Clarendon.
- van Fraassen B. (2002), *The empirical stance*, New Haven: Yale University Press.
- Woodward J. (2003), *Making things happen: a theory of causal explanation*, Nowy Jork: Oxford University Press.
- Żegleń U. (2003), *Filozofia umysłu*, Toruń: Wydawnictwo Adam Marszałek.

¹⁶ W stosunku do poprzedniej wersji tekst jest znacznie ulepszony dzięki uwagom prof. Stanisława Judyckiego, za które jestem bardzo wdzięczny.

Umysł

Natura i sposób istnienia **Trzy debaty**

Pod redakcją
Zbysława Muszyńskiego

WYDAWNICTWO UNIWERSYTETU MARII CURIE-SKŁODOWSKIEJ
LUBLIN 2010

Spis treści

Od redaktora tomu	7
-------------------------	---

Debata pierwsza Natura i sposób istnienia umysłu

JÓZEF BREMER – Być albo nie być świadomym. Neurofilozoficzna prototeoria świadomości	11
MARCIN MIŁKOWSKI – Przypisywanie intencjonalności, czyli jak to z termostatem było	27
BARTELOMIEJ ŚWIĄT CZAK – Reprezentacja mentalna jako nośnik informacji. Na marginesie książki Urszuli Żegleń	37
PAWEŁ KAWALEC – Przyczynowość stanów mentalnych w modelach naukowych. Próba alternatywnego uzasadnienia antynaturalizmu eksplanacyjnego w ujęciu Urszuli Żegleń	45

Debata druga Forum dyskusyjne „Diametros”: Czym jest i jak istnieje umysł?

MARCIN MIŁKOWSKI, ROBERT PO CZOBUT – Czym jest i jak istnieje umysł? Wprowadzenie do dyskusji	61
STANISŁAW JUDYCKI – Dwa argumenty przeciwko materializmowi	81
TADEUSZ SZUBKA – Więcej (scjentyistycznego) dymu niż (filozoficznego) ognia. W związku z tekstem Marcina Miłkowskiego i Roberta Poczobuta	95
URSZULA ŻEGLEŃ – Czym jest i jak istnieje umysł? Dyskusja	101
ROBERT PIŁAT – Dyskusja z tekstem Marcina Miłkowskiego i Roberta Poczobuta Czym jest i jak istnieje umysł?	105
JANUSZ JUSIAK – O poznawczych walorach spekulacji. Głos w dyskusji dotyczący głównych tez i filozoficznych założeń sformułowanych w artykule Marcina Miłkowskiego i Roberta Poczobuta <i>Czym jest i jak istnieje umysł?</i> ...	111
KRZYSZTOF ŻUK – Specyfika badań nad umysłem	119
PIOTR MARKIEWICZ – Kartografia filozoficzna i nowe lądy kognitywistyczne. W związku z tekstem Marcina Miłkowskiego i Roberta Poczobuta	123
JAN WOŁEŃSKI – Nie mam nic przeciwko kognitywistyce... ..	129
ROBERT PIŁAT – Jeszcze o pojęciu umysłu	131
KAROL POLCYN – W kwestii związku pomiędzy filozofią umysłu a naukami empirycznymi	135
MAREK SUWARA, JAN WERSZOWIEC-PLĄZOWSKI – Jak działa umysł, czyli „chirurgia” filozoficzna	139