

JAN ROSTEK*

KOMPATYBILISTYCZNA ODPOWIEDŹ NA PROBLEM WYKLUCZENIA PRZYCZYNOWEGO

Abstract

THE COMPATIBILIST REPLY TO THE CAUSAL EXCLUSION PROBLEM

The causal exclusion problem, as presented by Jaegwon Kim, is aimed to show that non-reductive physicalism is self-contradictory, as it is impossible for effects already having a sufficient physical cause to have a distinct mental cause. One possible reply to Kim's argument is compatibilism – a view within non-reductive physicalism that states that psychophysical supervenience is able to sustain the coexistence of efficacious mental and physical causes. The aim of this paper is to describe the currently most popular compatibilist strategy, as well as to present and evaluate some objections to it. First, I sketch the causal exclusion problem and show how compatibilists could undermine the causal exclusion principle using counterfactual analysis. Then I point out why this is a troublesome and insufficient tool and what challenges await compatibilism in its current form.

Keywords: non-reductive physicalism, compatibilism, causal exclusion, counterfactuals, overdetermination

Problem wykluczenia przyczynowego, pierwotnie przedstawiony przez Jaegwona Kima (2002: 48-49), miał w zamierzeniu działać jako argument przeciw nieredukcyjnemu fizykalizmowi. Z tego względu ma on niebagatelną wagę dla metafizyki umysłu: jeśli przesłanki i rozumowanie Kima są poprawne, to nieredukcyjny fizykalizm, stanowisko dość popularne w ostatnich dziesięcioleciach, może podzielić los dualizmu substancjalnego i znaleźć się na marginesie debaty. Warto więc uważnie przyglądać się działalności kompatybilistów – nieredukcyjnych fizykalistów, którzy próbują utrzymać swoje stanowisko dzięki odrzuceniu zasady wykluczenia przyczynowego. W artykule przedstawię argumenty kompatybilistów, skupiając się na najczęściej używa-

* Instytut Filozofii, Uniwersytet Jagielloński, ul. Grodzka 52, 31-044 Kraków, e-mail: jank.rostek@gmail.com.

nym przez nich narzędziu — analizie nierzeczywistych okresów warunkowych — a także opiszę dalsze problemy, które napotykają. W pierwszej części przedstawię pokrótce formy nieredukcyjnego fizykalizmu, opiszę problem wykluczenia przyczynowego i sposoby, na które można na niego odpowiedzieć. Następnie w części drugiej dokładniej przedstawię strategię kompatybilistów, a w części trzeciej przytoczę wybrane zarzuty im stawiane i możliwe próby odpowiedzi na nie.

1. PROBLEM Z NIEREDUKCYJNYM FIZYKALIZMEM

Problem wykluczenia przyczynowego powstaje wskutek zestawienia kilku twierdzeń, które — choć samodzielnie dość intuicyjne — razem stanowią zestaw sprzeczny i wymagają odrzucenia przynajmniej jednego z nich. Jak zobaczymy, są to zasady: przyczynowania mentalnego, nieredukowalności, domknięcia przyczynowego świata fizycznego oraz wykluczenia przyczynowego. Zdaniem Kima nieredukcyjni fizykaliści, w zależności od tego, które stwierdzenie odrzucają, muszą albo przyjąć redukcjonizm lub eliminatywizm, albo całkiem odrzucić fizykalizm i pójść w kierunku dualizmu. Kompatybiści twierdzą jednak, że mają podstawy do odrzucenia zasady wykluczenia przyczynowego, nie muszą więc rezygnować z nieredukcyjnego fizykalizmu. Zanim zobaczymy, na czym polegają poszczególne rozwiązania problemu i co je motywuje, ustalmy, czym właściwie jest nieredukcyjny fizykalizm.

1.1. NIEREDUKCYJNY FIZYKALIZM

Zgodnie z definicją Kima (2008: 77-78) wszyscy nieredukcyjni fizykaliści przyjmują ontologiczny fizykalizm, czyli twierdzenie, zgodnie z którym cały świat ma zasadniczo charakter fizyczny. Jednocześnie wyróżniają w nim zjawiska mentalne, które mają naturę fundamentalnie odmienną od zjawisk fizycznych i nie są do nich redukowalne. Kim (2008: 78-91) wyróżnia trzy główne teorie materialistyczne¹ odrzucające redukcjonizm. Pierwszą z nich był — obecnie już raczej niemający zwolenników — monizm anomalny Donalda Davidsona. Davidson zakładał, że zdarzenia mentalne mogą wchodzić w relacje przyczynowe z innymi zdarzeniami mentalnymi lub fizycznymi. Jednocześnie utrzymywał, że relacje przyczynowe muszą być oparte na prawach, a nie istnieją żadne prawa, które dotyczyłyby sfery mentalnej — ani

¹ Będę tu używał słów *fizykalizm* i *materializm*, jak również ich pochodnych, wymiennie.

o charakterze psychologicznym, ani psychofizycznym. Davidson rozwiązywał ten problem, twierdząc, że przyczynowanie mentalne odbywa się dzięki działaniu praw fizycznych, ponieważ każde zdarzenie mentalne można przedstawić za pomocą opisu fizycznego, a zatem jest ono zdarzeniem fizycznym (Davidson 2001). Między zdarzeniami mentalnymi i fizycznymi zachodzi identyczność egzemplarzy, tzn. każdemu wystąpieniu zdarzenia mentalnego możemy przyporządkować jedno wystąpienie zdarzenia fizycznego. Natomiast brak praw dotyczących sfery psychicznej nie pozwala na identyczność typów zdarzeń mentalnych i fizycznych.

Nieredukcyjny fizykalizm może też wyrastać ze ściśle powiązanej z funkcjonalizmem tezy o wielorakiej realizowalności (Bechtel, Mundale 1999: 176). Teza ta, szeroko przyjęta szczególnie w naukach kognitywnych, głosi, że zdarzenia mentalne są realizowane przez zdarzenia fizyczne. Jeden typ zdarzenia mentalnego może być jednak realizowany przez więcej niż jeden typ fizyczny: np. ból pleców może występować zarówno u ludzi, jak i u kotów, ale za ludzki ból będą odpowiadać inne wyładowania neuronalne. Również u samych ludzi ból pleców może być powodowany przez różne zdarzenia neuronalne. Uogólniając, wieloraka realizacja zachodzi w sytuacjach, w których (przynajmniej) dwa systemy są identyczne pod względem psychicznym, natomiast różne neurofizjologicznie (Polger, Shapiro 2016). Nie istnieje zatem jednoznaczny przekład predykatów mentalnych na fizyczne, co nie pozwala na tworzenie ogólnych praw psychofizycznych, potrzebnych do redukcji poziomu mentalnego do poziomu fizycznego. Znowu mamy więc do czynienia z identycznością egzemplarzy bez identyczności typów.

Trzeci rodzaj nieredukcyjnych teorii materialistycznych opiera się na superwencji psychofizycznej. Relacja superwencji łączy elementy dwóch zbiorów własności w taki sposób, że element z jednego zbioru (własność superwientna) jest zależny od elementu z drugiego zbioru (własności bazowej) i z nim współzmienny, ale jednocześnie nie jest do niego redukowalny (Kim 2001: 205). W przypadku superwencji psychofizycznej własności mentalne superwenują na własnościach fizycznych, czyli własności mentalne zdarzenia zawsze zależą od jego własności fizycznych i nie mogą ulec zmianie bez zmiany tych drugich.

1.2. PROBLEM WYKLUCZENIA PRZYCZYNOWEGO

Kim przedstawił zastrzeżenia do wszystkich wskazanych tu teorii, jednak najszerszym echem odbił się jego ogólny argument przeciw wszystkim formom nieredukcyjnego fizykalizmu. Opiera się on na problemie wykluczenia przyczynowego, który był opisywany na różne sposoby. W ujęciu Dwayne'a

Moore'a (2017: 21), w nieco zmienionej wersji, problem polega na niekompatybilności następujących czterech zasad:

- (I) Zasada przyczynowania mentalnego: niektóre zdarzenia fizyczne mają przyczyny mentalne².
- (II) Zasada nieredukowalności: przyczyny mentalne różnią się od przyczyn fizycznych.
- (III) Zasada domknięcia przyczynowego świata fizycznego: każdy fizyczny skutek ma swoją wystarczającą przyczynę fizyczną.
- (IV) Zasada wykluczenia przyczynowego: jeśli zdarzenie ma swoją wystarczającą przyczynę, to nie może mieć żadnej innej przyczyny w tym samym momencie, jeśli nie jest przedeterminowane.

Zdarzenie jest przedeterminowane, jeśli ma przynajmniej dwie przyczyny, które same w sobie są wystarczające do zajścia tego zdarzenia i nie potrzebują wzajemnie swojej pomocy (Carey 2011: 252)³. Paradygmatycznym przykładem przedeterminowania jest sytuacja, w której dwóch członków plutonu egzekucyjnego strzela do skazańca dokładnie w tym samym momencie: oba strzały są przyczyną śmierci, natomiast gdyby jeden z nich nie miał miejsca, to drugi wciąż byłby skuteczny. Przedeterminowanie występuje niezwykle rzadko, a więc jeśli przyczynowanie mentalne ma być powszechne, to jego opis nie powinien zakładać przedeterminowania.

Jeśli zdarzenia fizyczne muszą mieć wystarczające przyczyny fizyczne, a jednocześnie wykluczamy posiadanie więcej niż jednej przyczyny wystarczającej, to nie ma miejsca na przyczyny mentalne, które byłyby czymś innym niż przyczyny fizyczne. Zasady I-IV są zatem nie do pogodzenia i wydaje się, że przynajmniej jedną z nich należy odrzucić. Problem nieredukcyjnych fizykalistów polega na tym, że nie mogą wykreślić żadnej z nich bez zmiany stanowiska lub popadnięcia w (pozorny) absurd. Przyczynowanie mentalne jest szeroko uznawane za przedteoretyczną intuicję, która stanowi podstawę naszego myślenia o świecie, nieredukowalność jest oczywistą podstawą nieredukcyjnego fizykalizmu, a domknięcie przyczynowe trudno odrzucić, przyjmując materializm; np. Jesper Kallestrup (2006: 471) pisze, że o ile nie jest ono konieczne dla minimalnego fizykalizmu, o tyle trzeba je przyjąć, jeśli chcemy mocnego stanowiska materialistycznego. Natomiast zasada wyklu-

² Jako przyczyny i skutki podaje się w literaturze zarówno zdarzenia, jak i własności. W przedstawianej tu debacie te same argumenty mają zastosowanie w obu przypadkach; będę pisał głównie o zdarzeniach fizycznych i mentalnych, ale wszystko to może stosować się również do własności.

³ Dokładniejsze definicje przedeterminowania podam w kolejnej części artykułu.

czenia przyczynowego jest motywowana na kilka różnych sposobów, z których Moore (2017: 22-23) wyszczególnia pięć. Są to:

1. Intuicja — zasada wykluczenia przyczynowego jest przez wielu uznawana za intuicyjnie oczywistą prawdę o świecie.
2. Oszczędność ontologiczna — nauka powinna postulować istnienie tak niewielu bytów, jak to możliwe. Jeśli skutek ma już jedną przyczynę wystarczającą, to wprowadzanie każdej kolejnej może być uznane jedynie za bezproduktywne poszerzanie ontologii.
3. Sporadyczność przedeterminowania — przypadki wydarzenia się dwóch wystarczających przyczyn są w świecie fizycznym bardzo rzadkie i odosobnione. Co prawda jest możliwe, żeby jabłko spadło z drzewa jednocześnie przez podmuch wiatru i ruch ręki człowieka, ale prawdopodobieństwo wystąpienia takich dwóch zdarzeń w dokładnie tym samym czasie jest niezwykle niskie. Przyczynowanie mentalne jest zaś powszechne, a więc gdyby wiązało się z przedeterminowaniem, to i ono byłoby powszechne, co przeczy naszym obserwacjom i intuicjom.
4. Addytywność energii i pędu — energia i pęd są wielkościami fizycznymi, których wartość dla układu jest sumą wartości dla części składowych tego układu. Zatem jeśli przyczynowanie wymaga przekazania energii lub pędu, a jedna przyczyna wystarczająca przekazuje całą energię lub pęd potrzebne do wywołania skutku, to każda kolejna przyczyna albo dostarcza nadmiarowej energii lub pędu, co powinno zwiększyć efekt (a przecież sam fakt, że nasz ruch ręką jest wolicjonalny, nie daje mu dodatkowej siły), albo też nie jest w stanie przekazać energii lub pędu (czyli jest niejako „bezwładna”).
5. Argument z konieczności — jeśli zdarzenie fizyczne ma wystarczającą przyczynę fizyczną, to ta przyczyna jest wszystkim, czego potrzeba, by to zdarzenie zaszło, a zatem do jego zajścia nie jest konieczna żadna przyczyna mentalna. Przyczyny mentalne są bezcelowe i zbędne, tak więc należy je wykluczyć. Zdarzenie mentalne może towarzyszyć pewnej przyczynie fizycznej, ale nie będzie stanowiło przyczyny, tak jak huk wystrzału zawsze towarzyszy strzałowi z pistoletu, ale na pewno nie jest przyczyną morderstwa dokonanego za pomocą tego strzału. W dodatku argument z konieczności sugeruje, że jeśli przyjmujemy nieredukowalność, to

musimy odrzucić przyczynowanie mentalne lub domknięcie przyczynowe (wątek ten rozwinę w dalszej części artykułu).

1.3. SPOSOBY RADZENIA SOBIE Z PROBLEMEM WYKLUCZENIA PRZYCZYNOWEGO

Podstawowym sposobem na rozwiązanie problemu wykluczenia przyczynowego jest odrzucenie jednej z zasad I-IV. Odrzucenie zasady przyczynowania mentalnego prowadzi do epifenomenalizmu – stanowiska, zgodnie z którym nieredukowalna sfera mentalna istnieje, ale nie ma żadnych mocy przyczynowych; zdarzenia mentalne wyłącznie towarzyszą niektórym zdarzeniom fizycznym (Horowitz 1999: 421). Epifenomenalizm jest jednak niepopularną teorią, która napotyka wiele problemów i uznawana jest za sprzeczną z naszymi podstawowymi intuicjami (Kim 2002: 41-42, Hyslop 1998: 65-66). Co więcej, według Kima nieuchronnie prowadzi on do eliminatywizmu, czyli poglądu, zgodnie z którym własności i zdarzenia mentalne w ogóle nie istnieją. Zasady nieredukowalności nie uznają fizykaliści redukcjoni, tacy jak choćby właśnie Kim. Dualizm substancjalny odrzuca zaś zasadę domknięcia przyczynowego świata fizycznego (Moore 2017: 21), ale liczba dualistów drastycznie się skurczyła od czasów Kartezjusza, między innymi wskutek rozwoju nauki. Można wreszcie podważyć zasadę wykluczenia przyczynowego. Czynią to niektórzy nieredukcyjni fizykaliści, zwani kompatybilistami. Nie akceptują oni stwierdzenia Kima, że jedynymi możliwościami są eliminatywizm, redukcjonizm oraz dualizm, i próbują dowieść możliwości współistnienia przyczyn fizycznych i mentalnych w obrębie świata fizycznego. Od powodzenia ich projektu zależy zasadność zarzutów Kima oraz ogólny układ sił na polu zmagania między zwolennikami konkurencyjnych stanowisk dotyczących metafizyki umysłu.

2. KOMPATYBILIZM

Kompatybilistyczne podawanie w wątpliwość zasady wykluczenia przyczynowego najczęściej zasadza się na utrzymywaniu silnej relacji, np. superwencji, między zdarzeniami fizycznymi i mentalnymi, która to relacja nie występuje między konkurującymi przyczynami w przypadku przedeterminowania. Na różnice między przyczynowaniem mentalnym a paradygmatycznymi przypadkami przedeterminowania wskazują między innymi Karen Bennett (2003), Thomas Kroedel i Moritz Schulz (2016), Brandon Carey (2011) i Jesper Kallestrup (2006). Najpopularniejszą metodą wykazywania problemów z zasadą wykluczenia przyczynowego jest zapisanie jej w postaci nierze-

czywistych okresów warunkowych i ich analiza. Nie jest to jedyna taktyka kompatybilistów, inne proponują np. Stephen Yablo (1992), Kroedel i Schulz (2016) czy Derk Pereboom (2002). Skupię się tu jednak na okresach warunkowych jako narzędziu najpowszechniejszym i prowadzącym do najciekawszych wniosków. Warto też przyrzeć się innym definicjom przedeterminowania i sprawdzić, czy rzeczywiście zagrażają nieredukcyjnym fizykalistom. Zaczniemy od okresów warunkowych.

2.1. OMINIĘCIE ZASADY WYKLUCZENIA DZIĘKI NIERZECZYWISTYM OKRESOM WARUNKOWYM

Nierzeczywiste okresy warunkowe to zdania opisujące stan rzeczy, który miałby miejsce w przypadku zajścia innego, czysto hipotetycznego stanu rzeczy. Mają one ogólną budowę „Gdyby było tak, że ϕ , to byłoby tak, że ψ ” (Sendlak 2014: 92), co można zapisać jako $\phi \Box \rightarrow \psi$. Często używamy ich w języku potocznym, na przykład do opisywania potencjalnych konsekwencji zdarzeń, które nie miały miejsca („Gdybym urodził się w Toronto, to byłbym Kanadyjczykiem”), ale mogą też posłużyć do opisanego modelu przyczynowości. Najbardziej znanym modelem opartym na nierzeczywistych okresach warunkowych jest pierwsza teoria Davida Lewisa, zgodnie z którą zdarzenie \checkmark jest zależne przyczynowo od zdarzenia ϕ , jeśli prawdziwe są dwa nierzeczywiste okresy warunkowe: „Gdyby zaszło ϕ , to zaszłoby ψ ” ($\phi \Box \rightarrow \psi$) oraz „Gdyby nie zaszło ϕ , to nie zaszłoby ψ ” ($\neg \phi \Box \rightarrow \neg \psi$)⁴ (Lewis 1973: 563)⁵.

Prawdziwość nierzeczywistych okresów warunkowych ocenia się według Lewisa, korzystając z pojęcia światów możliwych. Każda potencjalna kombinacja stanów rzeczy (niekoniecznie zachodzących w świecie rzeczywistym) zachodzi w pewnym świecie możliwym, który znajduje się w pewnej „odległości” od innych światów. Światy są tym bliżej siebie, im są do siebie podobniejsze, a każdy świat jest najbliższy samemu sobie. Nierzeczywiste okresy warunkowe są prawdziwe wtedy, gdy ich poprzednik jest fałszywy w każdym świecie możliwym (czyli jest metafizycznie niemożliwy) lub gdy najbliższy świat, w którym zarówno poprzednik, jak i następnik są prawdziwe, jest bliżej naszego świata rzeczywistego niż dowolny świat, w którym poprzednik jest prawdziwy,

⁴ Jeśli chcemy być dokładni, to powinniśmy takie zdania zapisywać w formie $Z(\phi) \Box \rightarrow Z(\psi)$ i $\neg Z(\phi) \Box \rightarrow \neg Z(\psi)$, gdzie $Z(x)$ oznacza „zachodzi x ”. Dla uproszczenia zapisu często pomija się ten predykat; również i ja będę go pomijał, jednak należy pamiętać, że korzystamy z ukrytego skrótu.

⁵ Należy zauważyć, że ta teoria, jak i cała opisywana tu dyskusja, zakłada deterministyczną wizję przyczynowości. Przyjęcie paradygmatu indeterministycznego zmusiłoby nas m.in. do przyjęcia zmienionej wersji zasady domknięcia przyczynowego, która nie prowadziłaby do problemu wykluczenia przyczynowego (Tse 2015).

a następnik fałszywy (Lewis 1973: 560). Ten pierwszy warunek opiera się na prawie logiki, zgodnie z którym z fałszu wynika wszystko. Okresy warunkowe z niemożliwym poprzednikiem są w takim ujęciu trywialnie prawdziwe i nie mówią nic o świecie. Dlatego będą tu określane mianem „jałowych” (*vacuous*), w odróżnieniu od okresów warunkowych prawdziwych na mocy drugiego warunku, które będą po prostu nazywał prawdziwymi.

Okresów warunkowych można użyć do sformułowania argumentu z konieczności na rzecz zasady wykluczenia przyczynowego. Otóż jeśli pewne zdarzenie fizyczne f^* ma dwie odrębne wystarczające przyczyny – zdarzenie fizyczne f i superwenujące na nim zdarzenie mentalne m – które mają prowadzić do jego przedeterminowania, to prawdziwe są następujące okresy warunkowe (Moore 2017: 29):

(1) Gdyby zaszło f bez m , to zaszłoby f^* : $(f \wedge \neg m) \square \rightarrow f^*$.

(2) Gdyby zaszło m bez f , to zaszłoby f^* : $(m \wedge \neg f) \square \rightarrow f^*$.

Jeśli (1) i (2) są prawdziwe, to f^* jest przedeterminowane. Nie chcąc do tego dopuścić, kompatybiści powinni zaprzeczyć prawdziwości przynajmniej jednego z tych okresów warunkowych. Jak zauważa Moore (2017: 30-31), może to jednak rodzić niepożądane konsekwencje. Przyjmijmy, że postulujemy fałszywość (1), czyli twierdzimy, że samo f bez m nie może powodować f^* . Jednak jeśli f potrzebuje „pomocy” m , to nie jest przyczyną wystarczającą f^* . W tej sytuacji fałszywa jest zasada domknięcia przyczynowego świata fizycznego, ponieważ wymaga ona, żeby f było przyczyną wystarczającą f^* . Mogłoby się wydawać, że w takim razie kompatybiści powinni zaakceptować (1), ale tu też napotykają problem. Jeśli f^* może zajść bez m , to m nie jest przyczyną konieczną, a jeśli dowolne zdarzenie fizyczne może zajść bez udziału przyczyny mentalnej, to fałszywe jest przyczynowanie mentalne. Okres warunkowy (2) też na pierwszy rzut oka nie może zostać uznany za prawdziwy, ponieważ jeśli f^* może zajść bez f , to zdarzenia fizyczne nie mają przyczyn fizycznych i zasada domknięcia przyczynowego jest fałszywa. Część kompatybilistów może też nie chcieć stwierdzać fałszywości (2), ponieważ gdy m potrzebuje f do wywołania f^* , to nie jest przyczyną wystarczającą, a według niektórych interpretacji przyczynowanie mentalne zakłada, że przyczyna mentalna powinna być wystarczająca (np. Bennett 2003: 481).

Wydaje się, że kompatybiści są w potrzasku. Nie dość, że nie mogą przyjąć prawdziwości (1) i (2) jednocześnie, ponieważ prowadziłyby to do powszechnego przedeterminowania, to jeszcze przyjęcie nawet jednego z nich prowadzi do sprzeczności z założeniami nieredukcyjnego fizykalizmu, podobnie jak odrzucenie (1) i, być może, odrzucenie (2). Sytuacja nie jest jednak

beznadziejna: kompatybiliści mogą uporać się z tymi problemami lub je omiąć. W obróceniu analizy okresów warunkowych na swoją korzyść może pomóc im utrzymywanie silnej relacji zależności między przyczynami fizycznymi i mentalnymi. Taką relacją jest np. superweniencja psychofizyczna, zgodnie z którą zdarzenia mentalne są zależne od realizujących je zdarzeń fizycznych w ten sposób, że zdarzenie mentalne m koniecznie potrzebuje zajścia jakiegoś realizującego je zdarzenia fizycznego f , które z kolei koniecznie musi zachodzić razem z m (Kallestrup 2006: 459-460). Te dwie konieczności możemy zapisać jako $f \rightarrow m$ i $m \rightarrow \cup f$, gdzie $\cup f$ jest alternatywą wszystkich takich zdarzeń fizycznych f , które realizują m (żeby $\cup f$ było prawdziwe, musi zajść przynajmniej jedno z tych f -ów). Nie ma jednak zgody co do natury tych zależności. Raczej niekontrowersyjne jest uznanie ich za konieczności nomologiczne (McLaughlin, Bennett 2014), czyli wynikające z praw istniejących w świecie rzeczywistym, które są jednak przygodne i nie zachodzą w pewnych światach możliwych. Niektórzy materialiści idą o krok dalej i uznają $f \rightarrow m$ i/lub $m \rightarrow \cup f$ za metafizycznie konieczne, czyli prawdziwe w każdym świecie możliwym (Wilson 2005: 436)⁶.

Wyżej pokazałem problemy wynikające z prawdziwości lub fałszywości (1) i (2). Jednak nierzeczywiste okresy warunkowe mogą też być jałowe. Jałowe okresy warunkowe, jako że nie mówią nic o świecie, traktujemy inaczej niż okresy niejako prawdziwe i nie musimy przyjmować konsekwencji, które pociągają za sobą ich prawdziwość. Widać to dobrze na przykładzie analizy (1) i (2), której mogliby dokonać redukcjoniści. Według fizykalistów redukcyjnych zdarzenia mentalne można utożsamiać z fizycznymi, a zatem poprzedniki okresów warunkowych (1) i (2) brzmią dla nich: $f \wedge \neg f$. Oczywiście sprzeczność logiczna pociągą niemożliwość metafizyczną prawdziwości takiego poprzednika, a zatem (1) i (2) są dla redukcjonistów jałowe. Jednocześnie jest oczywiste, że nie mamy tu do czynienia z przedeterminowaniem, ponieważ pojedyncza przyczyna nie może sama przedeterminować swojego skutku. Widać zatem, że jałowość nierzeczywistych okresów warunkowych nie pociągają za sobą tych samych konsekwencji co ich prawdziwość.

Mając to na uwadze, można zająć się analizą (1) i (2). Zaczniemy od (1). Jeśli pewne zdarzenie fizyczne f realizuje pewne zdarzenie mentalne m , to zajście f bez m jest wykluczone przez łączącą je relację superweniencji. Jeżeli uznajemy $f \rightarrow m$ za konieczność metafizyczną, to w żadnym świecie możliwym f nie zachodzi bez m , a zatem poprzednik w (1) zawsze będzie fałszywy. Okres warunkowy (1) jest zatem jałowy. Problem natomiast pojawia się, gdy uznamy

⁶ Szersze omówienie kwestii charakteru konieczności w relacji superweniencji wykracza poza ramy tego artykułu.

$f \rightarrow m$ za konieczność zaledwie nomologiczną. Wtedy występowanie f bez m musimy uznać za metafizycznie możliwe, a najbliższy świat, w którym prawdziwe jest $f \wedge \neg m$, jest światem, w którym zachodzi f^* (ponieważ f jest przyczyną wystarczającą f^*). W tej sytuacji okres (1) jest prawdziwy, co stoi w sprzeczności z przyczynowaniem mentalnym. Kompatybiściści powinni więc twierdzić, że zajście f metafizycznie wymusza zajście m , przez co (1) jest jałowy.

Analizując (2), można przyjąć kilka strategii. Jeśli komuś nie zależy na tym, żeby m było przyczyną wystarczającą f^* , to zadowoli się wykazaniem fałszywości (2). Potrzebne do tego będzie uznanie relacji przyczynowej między m a f^* za „wrażliwą na realizację”, czyli za taką, w której jakakolwiek zmiana w zdarzeniu realizującym m prowadzi do zerwania relacji przyczynowej (Zhong 2017: 635). Przyjmijmy, że najbliższy świat, w którym prawdziwe jest $m \wedge \neg f$, jest światem, w którym zachodzi jakieś f' , nieznacznie różniące się od f i również realizujące m . Jednak f' może być przyczyną dla pewnego f^{**} , które nieznacznie różni się od f^* . Przykładowo, pewne wyładowanie neuronalne może powodować skurcz mięśni prowadzący do ruchu ręką. Jeśli w tym wyładowaniu weźmie udział nieznacznie inna wiązka neuronów, to może ono prowadzić do bardzo podobnego skurczu, ale jednak nieco innego, np. wolniejszego. W takim przypadku nie możemy powiedzieć, że oba wyładowania mają ten sam skutek. Najbliższy świat możliwy, w którym zachodzi m bez f , jest zatem światem bez f^* , co czyni (2) fałszywym. Jak już jednak wskazywałem wcześniej, jest to problem dla tych, którzy twierdzą, że przyczyny mentalne są wystarczające dla skutków fizycznych, dlatego to rozwiązanie nie odpowiada wszystkim kompatybilistom.

W podobny sposób można stwierdzić prawdziwość (2). Wymaga to uznania relacji przyczynowej między m a f^* za niewrażliwą na realizację. W takim wypadku również przyjmujemy, że najbliższym światem możliwym, w którym zachodzi m bez f , będzie świat z f' . Tym razem jednak nic nie przeszkadza m w byciu przyczyną f^* , a zatem zarówno poprzednik, jak i następnik (2) są prawdziwe, co pozwala na uznanie jego prawdziwości. Jak widzieliśmy, akceptacja (2) miała wiązać się z odrzuceniem zasady domknięcia przyczynowego świata fizycznego. By uniknąć tej konsekwencji, wystarczy założyć, że $m \rightarrow \cup f$ jest koniecznością metafizyczną. Dzięki temu w każdym możliwym świecie f^* ma swoją przyczynę fizyczną, którą jest jedno ze zdarzeń zdolnych do realizacji m , a tym samym domknięcie przyczynowe jest zachowane.

Zdaniem niektórych obie opisane strategie błędnie wyznaczają najbliższy świat możliwy. Lewis (2000: 190) po latach stwierdził, że jeśli szukamy najbliższego świata, w którym nie występuje pewne zdarzenie, to zastąpienie go innym, bardzo podobnym zdarzeniem może doprowadzić do wniosków sprzecznych z intuicjami. Zamiast tego powinniśmy całkowicie wyciąć to zda-

rzenie i nie zastępować go niczym innym. Rozważmy teraz (2): najbliższe światy możliwe, w których prawdziwe jest $m \wedge \neg f$, są, uogólniając, światami, w których prawdziwe jest $m \wedge \neg \cup f$, ponieważ nie występuje w nich żadne zdarzenie w zamian za f . Jednak jeśli znów założymy, że $m \rightarrow \cup f$ jest koniecznością metafizyczną, to nie istnieją żadne światy, w których prawdziwe byłoby $m \wedge \neg \cup f$. Zatem i w szczególności $m \wedge \neg f$ jest zawsze fałszywe — czyli (2) jest jałowy i nie pociąga za sobą żadnych zobowiązań.

Podsumowując: kompatybiści, którzy chcą ominąć problem wykluczenia przyczynowego za pomocą analizy w kategoriach nierzeczywistych okresów warunkowych, mogą (i muszą) utrzymywać, że okres warunkowy (1) „Gdyby zaszło f bez m , to zaszłoby f^* ” jest jałowy. Muszą przy tym przyjmować, że relacja superweniencji psychofizycznej zakłada metafizyczną konieczność zachodzenia m , jeśli zachodzi f . Więcej możliwości kompatybiści mają w ocenie okresu warunkowego (2) „Gdyby zaszło m bez f , to zaszłoby f^* ”. Jeśli przyjmują metafizyczną konieczność zachodzenia któregoś z f -ów realizujących m , gdy zachodzi m , to muszą uznać (2) za jałowy lub prawdziwy, w zależności od wybranej metody wyznaczania najbliższych światów możliwych. Prawdziwość (2) wymaga jednak przyjęcia dodatkowego założenia, że relacja przyczynowa między m a f^* jest niewrażliwa na realizację. Wrażliwość tej relacji na realizację jest zaś konieczna do uznania fałszywości (2) — wtedy jednak nie można utrzymywać, że przyczyny mentalne mogą być wystarczające dla skutków fizycznych. Dotychczasowe rozważania są jednak zasadne tylko przy przyjętej definicji przedeterminowania. Warto sprawdzić, czy inne jej sformułowania będą tak samo życzliwe dla kompatybilistów.

2.2. PRZEDETERMINOWANIE

Opisany sposób na odrzucenie zasady wykluczenia przyczynowego polega na przedstawieniu przedeterminowania (do którego miałyby prowadzić ta zasada w połączeniu z założeniami nieredukcyjnego fizykalizmu) w postaci twierdzenia o prawdziwości dwóch nierzeczywistych okresów warunkowych. Nie jest to jednak jedyna możliwa definicja. Carey (2011) przedstawia kilka innych opisów przedeterminowania, które również pozwalają na odrzucenie zasady wykluczenia przyczynowego, ewentualnie na przyjęcie powszechnego przedeterminowania w przyczynowaniu mentalnym. Najprostsza z definicji wygląda następująco:

- (P1) Skutek e jest przedeterminowany, gdy ma wystarczającą przyczynę p i pewną inną, nieidentyczną przyczynę q .

W takim ujęciu przedeterminowanie okazuje się jednak całkowicie niegroźne. Weźmy pod uwagę łańcuch przyczynowy, w którym następują kolejno po sobie takie zdarzenia p , q i e , że p jest wystarczającą przyczyną q , a q wystarczającą przyczyną e . Relacja przyczynowości jest przechodnia, a zatem p również jest wystarczającą przyczyną e , które w tej sytuacji jest skutkiem przedeterminowanym. Takie łańcuchy przyczynowe są jednak powszechne w świecie fizycznym. Przyjęcie (P1) oznacza, że wszystkie skutki fizyczne są przedeterminowane, nie tylko te mające przyczyny mentalne. Kompatybilści mogą więc spokojnie przyjąć powszechne przedeterminowanie w przyczynowaniu mentalnym.

Definicja (P1) z oczywistych względów nie powinna być satysfakcjonująca dla nikogo. Weźmy więc pod uwagę inną:

- (P2) Skutek e jest przedeterminowany, gdy ma dwie wystarczające i jednoczesne przyczyny p i q , które są od siebie metafizycznie, nomologicznie i przyczynowo niezależne.

(P2) brzmi już dużo wiarygodniej, ale nie pozwala na przyjęcie zasady wykluczenia przyczynowego. Można bowiem znaleźć przykłady takich sytuacji, w których zdarzenie fizyczne ma więcej niż jedną przyczynę wystarczającą w danym momencie, a jednak nie są to przyczyny spełniające warunek przedeterminowania. Tak jest chociażby w przykładach relacji determinacji, które opisuje Yablo (1992). Przytoczmy jego przykład dotyczący własności zdarzeń. Budynki w pewnym mieście wytrzymują pomniejszych trzęsienia ziemi. Ich konstrukcja sprawia jednak, że zawałają się w przypadku nasilonego wstrząsu (*violent*), czyli powyżej 5 stopni w skali Richtera. Jeśli nastąpi taki wstrząs i budynki się zawałają, to oczywistą (i wystarczającą) tego przyczyną będzie fakt, że wstrząs był nasilony. Przypuśćmy jednak, że wstrząs ten był ledwo nasilony (*barely violent*), czyli miał powyżej 5, ale poniżej 6 stopni w skali Richtera („umiarkowany”, ale nie „silny”). Bycie wstrząsem ledwo nasilonym też stanowi wystarczającą przyczynę zawalenia się budynków, a zatem jedno zdarzenie ma już dwie przyczyny wystarczające (Yablo 1992: 257-258). Zatem zgodnie z zasadą wykluczenia przyczynowego powinno być ono przedeterminowane. Jednak własności bycia wstrząsem nasilonym i bycia wstrząsem ledwo nasilonym nie są od siebie niezależne: zachodzi między nimi relacja determinowania, w której bycie wstrząsem ledwo nasilonym determinuje (wymusza) bycie wstrząsem nasilonym (ale nie w drugą stronę). Tak więc, z jednej strony, zawalenie się budynku ma dwie jednoczesne i wystarczające przyczyny, czyli zgodnie z zasadą wykluczenia przyczynowego powinno być przedeterminowane. Z drugiej strony, przyczyny te nie są od siebie niezależne, czyli zgodnie z (P2) nie prowadzą do przedeterminowania. Zachodzi sprzeczność, zatem jeśli przyjmujemy (P2), to zasada wykluczenia przyczynowego musi być fałszywa.

Trzecia definicja, którą podaje Carey, odwołuje się do nierzeczywistych okresów warunkowych. Została już ona omówiona w poprzednim punkcie, przejdźmy więc od razu do definicji:

- (P3) Skutek *e* jest przedeterminowany, gdy relacja między jego przyczynami *p* i *q* uprawomocnia przynajmniej jeden zarzut przeciwko przedeterminowaniu.

Oczywiście taka definicja jest niejasna i trudna do przyjęcia, ale może odpowiadać zwolennikom argumentu z wykluczenia przyczynowego, a więc warto się nad nią pochylić. Zarzuty przeciw przedeterminowaniu zostały podzielone przez Theodore'a Sidera (2003) na trzy kategorie: zarzut metafizyczny (przedeterminowanie jest w jakiś sposób niespójne metafizycznie z naturą przyczynowania), epistemiczny (nie mamy powodów, żeby wierzyć w przedeterminowanie) i ze zbiegu okoliczności (powszechne przedeterminowanie musiałoby być nieprawdopodobnym zbiegiem okoliczności). Przy ich analizie będę brał pod uwagę przyczyny, między którymi zachodzi relacja superwencji — tylko wtedy będą to przypadki relewantne, ponieważ taka jest według kompatybilistów relacja między przyczynami fizycznymi i mentalnymi.

Rozważmy najpierw zarzut metafizyczny. Jeśli ma się stosować, to natura przyczynowości musi wykluczać powszechne przedeterminowanie przez przyczyny pozostające w relacji superwencji. Weźmy pod uwagę sytuację, w której uderzenie piłki rozbija szybę. Dzieje się tak dlatego, że piłka jest twarda — można uznać to za warunek wystarczający rozbicia szyby. Własność bycia twardą superwenuje jednak na pewnej „twardej” strukturze atomowej: możliwych struktur odpowiedzialnych za twardość jest wiele, tak samo jak wiele zdarzeń fizycznych może być odpowiedzialnych za jakieś zdarzenie mentalne. Struktura atomowa też jest więc wystarczającą przyczyną rozbicia szyby. Zgodnie z zarzutem metafizycznym jedna z tych przyczyn (zapewne ta bazowa, czyli struktura atomowa) powinna wykonywać całą „pracę przyczynową”, co pozostawiałoby drugą z nich (twardość piłki) bez żadnej roli przyczynowej. Brzmi to dziwacznie — w takim wypadku sama piłka nie może być przyczyną rozbicia szyby⁷. Zarzut metafizyczny odmawia więc obiektom makroskopowym mocy przyczynowych (lub istnienia), na co mamy prawo się nie zgadzać. Zarzut można więc odrzucić.

Pozostałe zarzuty brzmią już mniej groźnie. Zarzut epistemiczny zakłada, że powinniśmy wątpić w jednoczesne występowanie zdarzeń fizycznych i men-

⁷ Wygląda to jednak dziwacznie tylko, gdy zakładamy, że piłki i szyby istnieją i mają moce przyczynowe. Założenie to odrzucają eliminatywiści w kwestii nieożywionych obiektów makroskopowych, np. Trenton Merricks (2001). Załóżmy jednak, że poruszamy się w kręgu bardziej standardowych ontologii i jesteśmy realistami w odniesieniu do piłek.

talnych powiązanych relacją superwencji. Taka wątpliwość nie jest jednak częsta: wiara w istnienie zdarzeń mentalnych jest raczej rozpowszechniona i niezależna od tego, czy towarzyszą im zdarzenia fizyczne. Można też powątpiewać w słuszność odrzucania jakichkolwiek teorii dotyczących metafizyki umysłu na podstawie samych intuicji (por. Schwitzgebel 2014)⁸. Także zarzut ze zbiegu okoliczności nie może zostać postawiony, skoro relacja superwencji psychofizycznej jasno zakłada, że współwystępowanie przyczyn fizycznej i mentalnej nie jest dziełem przypadku, lecz powszechnego prawa. Żadnego z tych zarzutów nie da się wysunąć, zatem zgodnie z (P3) superwencja psychofizyczna nie prowadzi do przedeterminowania w przyczynowaniu mentalnym i powinniśmy przyjąć fałszywość zasady wykluczenia przyczynowego. Wydaje się więc, że stanowisku kompatybilistycznemu nie zagraża żadna potencjalna definicja przedeterminowania.

3. ZARZUTY WOBEC KOMPATYBILIZMU

Analiza nierzeczywistych okresów warunkowych, chociaż początkowo wydaje się dobrym narzędziem, może okazać się bronią obosieczną i podważyć nie tylko zasadę wykluczenia przyczynowego, lecz także te zasady, które stoją u podstaw kompatybilizmu. Najpierw wyjaśnię, dlaczego okresy warunkowe powodują kłopoty z zachowaniem domknięcia przyczynowego świata fizycznego, a następnie przeanalizuję sposób, w jaki podają w wątpliwość zasadę przyczynowania mentalnego.

3.1. ZACHOWANIE DOMKNIĘCIA PRZYCZYNOWEGO

To, w jaki sposób kompatybiliści oceniają okresy warunkowe (1) i (2), budzi poważne wątpliwości, na co zwraca uwagę Moore (2017: 35-36). Jeśli zgadzamy się z Lewisem i jego uaktualnioną metodą oceny prawdziwości nierzeczywistych okresów warunkowych, to zdanie (2) jest jałowe. Wymaga to uznania konieczności metafizycznej warunku $m \rightarrow \cup f$. Moore widzi w tym problem, ponieważ nawet niektórzy fizykaliści mogą nie chcieć odrzucać metafizycznej możliwości istnienia bezcielesnych duchów. Stanowiska materialistyczne, które dopuszczały, że w pewnych światach możliwych zachodzą zdarzenia mentalne bez fizycznych podstaw, zajmowali m.in. Lewis (1983) i Jackson (1994). Z drugiej strony, wielu autorów twierdzi, że fizykalizm powinien przyjmować metafi-

⁸ Oczywiście, w takim wypadku powinniśmy też znaleźć inne niż intuicyjne motywacje dla utrzymania przyczynowania mentalnego (i odrzucenia epifenomenalizmu).

zyczną konieczność superwencji psychofizycznej (czyli metafizyczną konieczność obu warunków: $f \rightarrow m$ i $m \rightarrow \cup f$). Na przykład, Chalmers (1996) wpisuje tę konieczność w definicję fizykalizmu, a Francescotti (2014) przekonuje, że fizykaliści muszą ją przyjmować, nawet jeśli nie traktują samego fizykalizmu jako prawdziwego we wszystkich światach możliwych. Możemy więc zdefiniować dwa rodzaje stanowisk fizykalistycznych: pierwszy to fizykalizm słaby, w którym superwencja psychofizyczna (a przynajmniej jej warunek $m \rightarrow \cup f$) zachodzi z koniecznością zaledwie nomologiczną⁹, a drugi to fizykalizm mocny, który zakłada metafizyczną konieczność superwencji¹⁰. Otóż jałowość okresu warunkowego (2) jest problematyczna wyłącznie dla zwolenników fizykalizmu słabego, a zatem kompatybiliści będący zarazem fizykalistami mocnymi mogą lepiej radzić sobie z problemem wykluczenia przyczynowego.

Z zarzutów opisanych przez Moore'a dużo poważniejszy jest drugi, który pokazuje, że kompatybiliści mogą nie być w stanie utrzymać domknięcia przyczynowego świata fizycznego. Ustaliliśmy już, że tylko uznanie okresu warunkowego (1) za jałowy nie jest sprzeczne z założeniami nieredukcyjnego fizykalizmu. Należy jednak zauważyć, że zasada domknięcia przyczynowego głosi, iż przyczyny fizyczne muszą być wystarczające dla skutków fizycznych. W rezultacie przyczyna fizyczna f powinna być w stanie samodzielnie spowodować skutek f^* , czyli zdanie (1) powinno być prawdziwe. Jego jałowość nie mówi zaś nic o tym, czy f rzeczywiście jest przyczyną wystarczającą, ponieważ jałowe okresy warunkowe nie dają nam żadnych informacji o świecie. Kompatybiliści nie mogą za pomocą analizy nierzeczywistych okresów warunkowych wykazać, że każda przyczyna fizyczna jest przyczyną wystarczającą, a zatem nie są w stanie udowodnić prawdziwości domknięcia przyczynowego świata fizycznego. To już brzmi problematycznie, ponieważ powinno się unikać włączania do teorii tez, których nigdy nie będziemy w stanie udowodnić w ramach tej teorii (nawet gdybyśmy posiadli całą wiedzę o świecie).

Niemożność wysunięcia jednoznacznych argumentów za prawdziwością lub fałszywością tezy nie czyni jednak stanowiska, które tę tezę przyjmuje, fałszywym. Kompatybilizm nie musi być zatem odrzucony przez swoich zwolenników. Niemniej, powinni oni znaleźć sposób na udowodnienie zasady domknięcia przyczynowego za pomocą innego narzędzia niż okresy warunkowe.

Opisany zarzut można jednak pogłębić. Jeśli f nie może zająć bez m , to nigdy nie będzie prawdą, że f^* zostało spowodowane wyłącznie przez f . Jed-

⁹ Niektórzy fizykaliści, jak chociażby Jackson (1994), uznają $f \rightarrow m$ za konieczność metafizyczną, a $m \rightarrow \cup f$ za konieczność nomologiczną, natomiast stanowisk odwrotnych raczej nie znajdziemy w literaturze.

¹⁰ Należy przy tym zauważyć, że ani Chalmers, ani Francescotti nie uznaliby takiego podziału: to, co określiłem mianem fizykalizmu słabego, dla nich w ogóle nie jest fizykalizmem.

nakże, jeśli f ma być przyczyną wystarczającą f^* , to f musi być w stanie powodować f^* bez obecności m . Tak się nigdy nie dzieje, ponieważ sytuacja ta jest metafizycznie niemożliwa. Zatem nie możemy uznać f za przyczynę wystarczającą, a świat fizyczny nie jest przyczynowo domknięty. Stoi za tym intuicja, że jeśli skutek fizyczny nie może nigdy wystąpić bez przyczyny mentalnej, to przyczyna fizyczna nie może być wszystkim, czego potrzeba do zajścia tego skutku.

Jeśli zarzut ten jest zasadny, to kompatybilizm jest wewnętrznym sprzecznym: jednocześnie zakłada zasadę domknięcia przyczynowego świata fizycznego i można z niego wywieść zaprzeczenie tej zasady. Kompatybiliści mogą jednak próbować bronić swojego stanowiska, twierdząc, że przedstawiona argumentacja w nieuprawniony sposób zrównuje posiadanie mocy przyczynowych z ich ujawnianiem. Otóż f może być w stanie samodzielnie spowodować f^* i jednocześnie nigdy tego nie robić z przyczyn, które nie mają żadnego związku z mocami przyczynowymi f . Przyczyny fizyczne mają niejako wszystko, co potrzebne, żeby poradzić sobie ze spowodowaniem swoich fizycznych skutków. Nie możemy tylko tego udowodnić, ponieważ tak się składa, że w każdym świecie możliwym f występuje razem ze zdarzeniem mentalnym m . Można tutaj zaproponować nieformalną analogię do sytuacji, w której matka samotnie opiekuje się kilkuletnim synem. Syn jest za mały, żeby zostawić go samego w domu, a więc wszędzie, gdzie pojawia się matka, tam też koniecznie pojawia się syn. Gdy matka idzie do sklepu po piwo, zawsze pozwala synowi wziąć pieniądze z jej portfela i zapłacić. Nigdy nie uświadczymy sytuacji, w której matka kupuje piwo bez obecności syna. Jej obecność jest jednak w oczywisty sposób wystarczająca do sprzedaży, nawet jeśli zawsze pozwala synowi wziąć pieniądze z jej portfela i zapłacić, ponieważ, jako osoba pełnoletnia, bez obecności syna też mogłaby kupić sobie piwo. Nie jest to oczywiście pełna analogia do sytuacji przyczynowania mentalnego, ale pozwala uświadomić sobie, że pewne zdarzenie może mieć moce przyczynowe uprawniające je do bycia przyczyną wystarczającą innego zdarzenia, ale jednocześnie nigdy tych mocy w pełni i samodzielnie nie okazywać. Takie przedstawienie problemu sprowadza go do jego pierwotnej postaci: kompatybilizm nie pozwala na ocenę prawdziwości założenia o domknięciu przyczynowym świata fizycznego, ale już nie jest wewnętrznym sprzecznym.

3.2. WĄTPLIWE PRZYZCYNOWANIE MENTALNE

Na problemy nieco innej natury zwraca uwagę Kallestrup (2006: 475-478), który wskazuje, że nierzeczywiste okresy warunkowe mogą nie być najlepszym narzędziem dla kompatybilistów, ponieważ nie są w stanie zapewnić zdarzeniom mentalnym mocy przyczynowych. Pierwszą przeszkodą jest za-

groźenie epifenomenalizmem. Otóż kompatybiliści wykazują, że zdarzenie mentalne m jest przyczyną zdarzenia fizycznego f^* , za pomocą prawdziwości nierzeczywistego okresu warunkowego $\neg m \Box \rightarrow \neg f^*$. Prawdziwość ta wynika z przesłanek $\neg m \Box \rightarrow \neg f$ (na mocy superwencji psychofizycznej) i $\neg f \Box \rightarrow \neg f^*$ (na mocy domknięcia przyczynowego świata fizycznego) oraz z tego, że relacja zależności warunkowej jest przechodnia¹¹. Te same okresy warunkowe są jednak prawdziwe również według epifenomenalistów, a przecież ich zdaniem zdarzenia mentalne nie są przyczynami żadnych innych zdarzeń. Sama analiza okresów warunkowych nie pozwala więc odróżnić kompatybilizmu od epifenomenalizmu.

Sposobem na odróżnienie kompatybilizmu od epifenomenalizmu może być według Kallestrupa wprowadzenie pojęcia eliminacji przyczyny (*pre-emption*). Otóż przyczyna p eliminuje przyczynę q w stosunku do ich skutku e wtedy, gdy e jest powiązane nierzeczywistym okresem warunkowym z p , a nie jest w ten sposób powiązane z q dopóty, dopóki zachodzi p (Loewer 2001). W epifenomenalizmie f eliminuje m w stosunku do f^* , czyli, zgodnie z podaną definicją, prawdziwe są nierzeczywiste okresy warunkowe $\neg f \Box \rightarrow \neg f^*$ oraz $(f \wedge \neg m) \Box \rightarrow f^*$. W kompatybilizmie jednak ten ostatni okres warunkowy jest, jak już ustaliliśmy, jałowy, a zatem warunki eliminacji nie zachodzą i przyczyna mentalna nie jest wyeliminowana. W ten sposób można za pomocą nierzeczywistych okresów warunkowych wykazać, że m jest przyczyną f^* .

Kallestrup widzi w tym jednak pewien kłopot: otóż tą metodą można też wykazać, że m eliminuje f w stosunku do f^* . Jest tak, jeśli prawdziwe są okresy warunkowe $\neg m \Box \rightarrow \neg f^*$ oraz $(m \wedge \neg f) \Box \rightarrow f^*$. Pierwszy z nich, jak widzieliśmy w poprzednim akapicie, powinien być prawdziwy, natomiast drugi, jak pokazaliśmy w poprzedniej części, może być prawdziwy, w zależności od przyjętych założeń. Możliwe jest zatem, że w kompatybilizmie przyczyny mentalne eliminują przyczyny fizyczne. Takie stanowisko brzmi jednak kompletnie niefizykalistycznie: eliminacja przyczyn fizycznych w stosunku do skutków fizycznych wygląda na większe zagrożenie dla domknięcia przyczynowego świata fizycznego niż uwagi Moore'a. Jeśli utrzymujemy więc kryteria eliminacji podane przez Loewera, to nie powinniśmy uznawać prawdziwości okresu warunkowego (2).

Możliwe jednak, że widmo epifenomenalizmu można odegnąć bez odwoływania się do eliminacji przyczyn. Kroedel (2015: 845) przedstawia sposób oceny nierzeczywistych okresów warunkowych, w którym zdanie $\neg m \Box \rightarrow \neg f^*$ jest prawdziwe dla kompatybilistów, a fałszywe dla epifenomenalistów:

¹¹ Chociaż przechodność nierzeczywistych okresów warunkowych bywa podważana, to w tym konkretnym przypadku można się na nią zgodzić z uwagi na silną naturę relacji między f a m (por. Kallestrup 2006: 482).

w przypadku okresów warunkowych, których poprzedniki mówią o niezachodzeniu pewnego zdarzenia zachodzącego w świecie rzeczywistym, za najbliższe światy możliwe powinniśmy uważać takie, w których poprzednik jest prawdziwy (czyli owo zdarzenie nie zachodzi), ale przeszłość jest dokładnie taka sama jak w świecie rzeczywistym do momentu zajścia wspomnianego zdarzenia. Stosując to do zdania $\neg m \Box \rightarrow \neg f^*$, za najbliższe światy możliwe powinniśmy uznać te, w których m nie zachodzi, a cała przeszłość jest taka sama jak w świecie rzeczywistym (w którym m zachodzi). Oznacza to, że zachodzi między innymi jakieś zdarzenie fizyczne f_0 , które jest przyczyną zdarzenia f , na którym superweniuje m . Zgodnie z epifenomenalizmem, jak pisze Kallestrup (2006: 476), superweniencja psychofizyczna jest konieczna zaledwie nomologicznie: to tylko prawo naszego świata, które w pewnych innych światach możliwych może nie zachodzić¹². Nic więc nie stoi na przeszkodzie, by w najbliższym możliwym świecie bez m zaszło f . Co więcej, powinno ono zajść, ponieważ zachodzi jego przyczyna f_0 . Skoro zaś zachodzi f , które jest przyczyną f^* , to zachodzi też samo f^* . Doszliśmy do tego, że według epifenomenalizmu najbliższym światem możliwym, w którym nie zachodzi m , jest świat, w którym zachodzi f^* , a zatem okres warunkowy $\neg m \Box \rightarrow \neg f^*$ jest fałszywy. W kompatybilizmie zachowuje on swoją prawdziwość, ponieważ bez m niemożliwe jest f , a bez f niemożliwe jest f^* . Da się więc odróżnić kompatybilizm od epifenomenalizmu za pomocą analizy nierzeczywistych okresów warunkowych.

Nawet jeśli poradzimy sobie z problemem epifenomenalizmu, to Kallestrup ma w zanadru jeszcze jeden zarzut, bardziej ogólny i, jak się wydaje, poważniejszy. Problematyczny jest okres warunkowy $\neg m \Box \rightarrow \neg f^*$, który według kompatybilistów powinien zapewniać przyczynowanie mentalne. Czy jednak faktycznie to robi? Jak już wspomniałem, prawdziwość tego zdania zapewniana jest przez prawdziwość dwóch innych okresów warunkowych, $\neg m \Box \rightarrow \neg f$ i $\neg f \Box \rightarrow \neg f^*$, a także przez ich przechodniość. Jednak żeby $\neg m \Box \rightarrow \neg f^*$ było zależnością przyczynową, to przyczynowe muszą być też przesłanki, z których ją wywodzimy. Zależność $\neg f \Box \rightarrow \neg f^*$ jak najbardziej jest przyczynowa, natomiast $\neg m \Box \rightarrow \neg f$ jest prawdziwe na mocy superweniencji psychofizycznej. Relacja superweniencji nie jest rodzajem relacji przyczynowania: przyczyna i skutek muszą zajść w różnych momentach, natomiast zdarzenia

¹² Pisząc o epifenomenalizmie, zazwyczaj mamy na myśli jego klasyczną, dualistyczną wersję. Dualiści mogą zaś utrzymywać co najwyżej nomologiczną konieczność superweniencji psychofizycznej (Crane 2000). Horowitz (1999) opisuje epifenomenalizm fizykalistyczny, który przyjmuje fizykalizm. Nie precyzuje jednak, czy jest to (według wprowadzonej tu wcześniej terminologii) fizykalizm mocny, czy słaby. Należy więc nadmienić, że opisywany tu kontrargument nie musi stosować się do epifenomenalizmu fizykalistycznego, ten bowiem może zakładać metafizyczną konieczność warunku $f \rightarrow m$.

superwenujące i bazowe zachodzą dokładnie w tym samym czasie. Raczej nikt nie próbowałby twierdzić, że nasze uczucie bólu pojawia się przed stanem mózgu odpowiadającym bólowi i jest jego przyczyną (a tak by było, gdyby superwenujenca zakładała przyczynowość). Brak przyczynowego charakteru zależności m i f sprawia, że nie jesteśmy w stanie wskazać łańcucha przyczynowego od m do f^* , a zatem nie możemy udowodnić przyczynowego charakteru ich zależności. Co więcej, trudno jest wierzyć w możliwość istnienia takiego łańcucha, skoro wykluczyliśmy przyczynowanie między m a f . Taki łańcuch musiałby biec z m do f^* z pominięciem f , co wygląda na niezgodne z fizycznym charakterem świata. Kompatybiliści nie są zatem w stanie udowodnić przyczynowania mentalnego za pomocą analizy nierzeczywistych okresów warunkowych. Ten problem nie przekreśla szans na powodzenie ich projektu, ale do kompletności (i idącej za nią atrakcyjności) stanowiska brakuje im takiego opisu przyczynowania mentalnego, który byliby w stanie udowodnić w ramach swojej teorii.

4. WNIOSKI

Problem wykluczenia przyczynowego określony jest przy użyciu pojęcia przedeterminowania, a zatem wymaga ono odpowiedniej definicji. Najlepsza z podanych definicji sformułowana jest za pomocą nierzeczywistych okresów warunkowych. Kompatybiliści chcący odrzucić zasadę wykluczenia przyczynowego powinni więc zrobić to za pomocą analizy tych okresów warunkowych. Dzięki odpowiednio mocnej relacji superwenujenca psychofizycznej są w stanie wykazać, że przedeterminowanie nie zachodzi w przyczynowaniu mentalnym. Co więcej, mogą to zrobić na różne sposoby, w zależności od preferowanych relacji między zdarzeniami mentalnymi a fizycznymi oraz od używanych metod wyznaczania najbliższych światów możliwych. Kompatybiliści mogą też odeprzeć oskarżenia o milczące odrzucenie domknięcia przyczynowego świata fizycznego lub przyczynowania mentalnego. Nie oznacza to jednak, że ich stanowisko jest wolne od wad: korzystając z analizy nierzeczywistych okresów warunkowych, nie są oni w stanie wykazać prawdziwości zasady domknięcia przyczynowego ani przyczynowania mentalnego. Nie są to fundamentalne błędy w ich systemie, lecz raczej luki, które trzeba wypełnić, żeby system był kompletny i atrakcyjny. W tym celu kompatybiliści powinni jednak szukać innych narzędzi niż nierzeczywiste okresy warunkowe, ponieważ ich użyteczność jest ograniczona.

BIBLIOGRAFIA

- Bechtel W., Mundale J. (1999), *Multiple Realizability Revisited. Linking Cognitive and Neural States*, „Philosophy of Science” 66(2), 175-207.
- Bennett K. (2003), *Why the Exclusion Problem Seems Intractable, and How, Just Maybe, to Tract It*, „Noûs” 37(3), 471-497.
- Carey B. (2011), *Overdetermination and The Exclusion Problem*, „Australasian Journal of Philosophy” 89(2), 251-262.
- Chalmers D. (1996), *The Conscious Mind. In Search of a Fundamental Theory*, Oxford: Oxford University Press.
- Crane T. (2000), *Dualism, Monism, Physicalism*, „Mind & Society” 1(2), 73-85.
- Davidson D. (2001), *Mental Events [w:] Essays on Actions and Events*, Oxford: Clarendon Press, 207-225.
- Francescotti R. (2014), *The Problem of Extras and the Contingency of Physicalism*, „Philosophical Explorations” 17(2), 241-254.
- Horowitz A. (1999), *Is There a Problem in Physicalist Epiphenomenalism?*, „Philosophy and Phenomenological Research” 59(2), 421-434.
- Hyslop A. (1998), *Methodological Epiphenomenalism*, „Australasian Journal of Philosophy” 76(1), 61-70.
- Jackson F. (1994), *Armchair Metaphysics [w:] Philosophy in Mind*, M. Michael, J. O’Leary-Hawthorne (eds.), Dordrecht: Kluwer, 23-42.
- Kallestrup J. (2006), *The Causal Exclusion Argument*, „Philosophical Studies” 131(2), 459-485.
- Kim J. (2001), *Superweniencja jako pojęcie filozoficzne*, tłum. R. Poczobut, „Roczniki Filozoficzne” 49(1), 195-227.
- Kim J. (2002), *Umysł w świecie fizycznym*, tłum. R. Poczobut, Warszawa: IFiS PAN.
- Kim J. (2008), *Mit nieredukcyjnego materializmu*, tłum. P. Gutowski, T. Szubka [w:] *Analityczna metafizyka umysłu*, M. Miłkowski, R. Poczobut (red.), Warszawa: IFiS PAN, 76-97.
- Kroedel T. (2015), *A Simple Argument for Downward Causation*, „Synthese” 192(3), 841-858.
- Kroedel T., Schulz M. (2016), *Grounding Mental Causation*, „Synthese” 193(6), 1909-1923.
- Lewis D. (1973), *Causation*, „The Journal of Philosophy” 70(17), 556-567.
- Lewis D. (1983), *New Work for a Theory of Universals*, „Australasian Journal of Philosophy” 61(4), 343-377.
- Lewis D. (2000), *Causation as Influence*, „The Journal of Philosophy” 97(4), 182-197.
- Loewer B. (2001), *Book Review: «Mind in a Physical World»*, „The Journal of Philosophy” 98(6), 315-324.
- McLaughlin B., Bennett K. (2014), *Supervenience [w:] The Stanford Encyclopedia of Philosophy* (Spring 2014 Edition), E. N. Zalta (ed.), <https://stanford.io/2SkuxMn>.
- Merricks T. (2001), *Objects and Persons*, Oxford: Oxford University Press.
- Moore D. (2017), *Mental Causation, Compatibilism, and Counterfactuals*, „Canadian Journal of Philosophy” 47(1), 20-42.
- Pereboom D. (2002), *Robust Nonreductive Materialism*, „Journal of Philosophy” 99(10), 499-531.
- Polger T., Shapiro L. (2016), *The Multiple Realization Book*, Oxford: Oxford University Press.
- Schwitzgebel E. (2014), *The Crazyist Metaphysics of Mind*, „Australasian Journal of Philosophy” 92(4), 665-682.

- Sendlak M. (2014), *Realizm modalny i okresy warunkowe z niemożliwymi poprzednikami*, „Filozofia Nauki” 22(4) [88], 91-107.
- Sider T. (2003), *What's So Bad about Overdetermination?*, „Philosophy and Phenomenological Research” 67(3), 719-726.
- Tse P. U. (2015), *Ontological Indeterminism Undermines Kim's Exclusion Argument*, <https://philarchive.org/archive/TSEOIU>.
- Yablo S. (1992), *Mental Causation*, „The Philosophical Review” 101(2), 245-280.
- Wilson J. (2005), *Supervenience-Based Formulations of Physicalism*, „Noûs” 39(3), 426-459.
- Zhong L. (2017), *Semantic Normativity and Semantic Causality*, „Philosophy and Phenomenological Research” 94(3), 626-645.