

ADRIANA SCHETZ\*

LUKA EKSPLANACYJNA  
A POJĘCIA WŁASNOŚCI FENOMENALNYCH  
STANOWISKO JOSEPHA LEVINE’A<sup>1</sup>

Słowa kluczowe: Joseph Levine, Brian Loar, pojęcia fenomenalne, własności fenomenalne, luka eksplanacyjna, reprezentacjonizm, internalizm, eksternalizm, charakter fenomenalny

Keywords: Joseph Levine, Brian Loar, phenomenal concepts, phenomenal properties, explanatory gap, representationalism, internalism, externalism, phenomenal character

Problem logicznej niezależności koncepcji stanów fenomenalnych (takich jak wrażenia zmysłowe czy bardziej złożone stany percepcyjne

---

\* Adriana Schetz – adiunkt w Instytucie Filozofii Uniwersytetu Szczecińskiego. Do jej obszarów kompetencji należą: filozofia analityczna, filozofia umysłu, filozofia psychologii. Publikacje dotyczą głównie zagadnień związanych z filozoficznymi, psychologicznymi i kognitywistycznymi ujęciami percepcji, świadomości, pamięci oraz ogólnie kategorii reprezentacji umysłowej. Publikuje w takich naukowych czasopismach, jak m.in.: „Analiza i Egzystencja”, „Filozofia Nauki”, „Przegląd Filozoficzny”, „Roczniki Psychologiczne”, „Studia z Kognitywistyki i Filozofii Umysłu”. E-mail: [adriana.schetz@gmail.com](mailto:adriana.schetz@gmail.com).

<sup>1</sup> Artykuł ten powstał w czasie mojego pobytu w Ohio State University w ramach stypendium Fulbright Junior w 2006 r. Choć przeleżał cztery lata w szufladzie, to nadal szczególne podziękowania należą się profesorowi Josephowi Levine’owi za inspirujące dyskusje oraz cenne uwagi do wcześniejszej wersji opracowania.

o charakterze intencjonalnym) od ustaleń dotyczących fizjologii, czy po prostu fizycznej konstytucji organizmu, skłania wielu filozofów do wniosku, że istnieje luka pomiędzy wyjaśnieniem w kategoriach pojęć fenomenalnych a wyjaśnieniem w kategoriach pojęć z zakresu fizjologii czy fizyki nazywana luką eksplanacyjną. Pojęcie luki wiąże się z podejściem nieredukcyjnym do dziedziny mentalnej i tym samym z brakiem możliwości wyjaśnienia, jak konstytuowane są stany fenomenalne przez procesy naturalne. Stanowisko takie jest groźne dla wszystkich naturalistycznych koncepcji relacji psychofizycznej, gdyż prowadzi do podważenia wyjściowego założenia naturalizmu o jedności metodologicznej dziedzin traktujących o procesach i stanach psychofizycznych.

Joseph Levine jest jednym z najbardziej konsekwentnych obrońców idei, w myśl której istnieje niedająca się w żaden sposób uzupełnić luka w rozwiązaniach problemu umysł–ciało, bazujących na wyjaśnieniu naturalistycznym (Levine 1983)<sup>2</sup>. W pierwszej części artykułu zaprezentuję założenia wyjaśnienia epistemologicznego, którego potrzebujemy zdaniem Levine'a, aby wypełnić przepaść pomiędzy koncepcją stanów fenomenalnych a koncepcją odpowiednich stanów fizykalnych. W drugiej części przyjrę się dokładnie czynnikom konstytuującym lukę eksplanacyjną. Podejmę próbę odsłonięcia przyjmowanej przez Levine'a metodologii oraz koncepcji pojęć fenomenalnych prowadzących w konsekwencji do przekonania o istnieniu luki eksplanacyjnej. Trzecia i czwarta część artykułu podejmuje nasuwające się wątpliwości i problemy związane z proponowanym przez Levine'a ujęciem kategorii pojęć fenomenalnych.

## 1. Wyjaśnienie epistemologiczne

Redukcja teoretyczna jest bardzo popularna na gruncie fizyki i chemii. Polega ona na logicznym sprowadzeniu koncepcji substancji o pewnych niewytłumaczalnych dotąd własnościach do koncepcji innej substancji, której własności tłumaczą problematyczne dotąd zjawiska. Tak było na przykład z wodą, substancją, która służy dziś filozofom za pokazowy przykład takiej redukcji. Zdanie (1) „Woda jest H<sub>2</sub>O”, redukuje pojęcie wody, a więc substancji o takich znanych nam z życia codziennego własnościach, jak

---

<sup>2</sup> Joseph Levine jest autorem blisko trzydziestu opracowań poświęconych problematyce luki eksplanacyjnej. Do najważniejszych należą: (Levine 1983), (Levine 1998), (Levine 2001), (Levine 2006), (Levine 2010).

płynność w temperaturze pokojowej, zdolność do zamarzania w temperaturze ujemnej, przejrzystość itp., do pojęcia substancji chemicznej złożonej z dwóch atomów wodoru i jednego atomu tlenu, a więc do pojęcia  $H_2O$ . Redukcja teoretyczna w przypadku wody i  $H_2O$  jest uzasadniona, gdyż istnieją bezsporne dane obserwacyjne dowodzące systematycznych korelacji między wodą a  $H_2O$ . Redukcja pojęcia wody do pojęcia  $H_2O$  ma również niebagatelne znaczenie eksplanacyjne, gdyż prawdziwość zdania (1) pozwala zrozumieć, dlaczego woda ma takie a nie inne własności.

Przyjrzyjmy się teraz zdaniu (2) „Ból jest aktywnością określonych włókien nerwowych”. Wydaje się, że, podobnie jak w przypadku zdania (1), jesteśmy w pełni usprawiedliwieni w przypisywaniu mu statusu zdania prawdziwego, gdyż istnieją wiarygodne obserwacje, które pozwalają wnioskować o koniecznej korelacji pomiędzy występowaniem bólu oraz występowaniem określonej aktywności włókien nerwowych. Czy jest jednak tak, że zdanie mówiące, iż ból jest identyczny z aktywnością danych włókien nerwowych pozwala zrozumieć, że ból ma takie własności, jakie się mu przypisuje? Joseph Levine udziela na to pytanie negatywnej odpowiedzi (Levine 2006). Według niego to, co stanowi rdzeń stanu mentalnego, jakim jest ból, mianowicie jego charakter fenomenalny (to, że jest coś takiego, jak *jak to jest odczuwać* ból), nie staje się zrozumiałe dzięki odwołaniom do faktów fizjologicznych. Fakty te bowiem wskazują jedynie, że gdy ktoś odczuwa ból, to w jego mózgu aktywują się określone włókna nerwowe. Zdaniem Levine’a ta określona aktywność włókien nerwowych ewidentnie jest tożsama z posiadaniem bólu. Metafizycznie rzecz biorąc, ból jest aktywnością mózgu. Niemniej pojawia się problem epistemologiczny, jak za pomocą metod wykorzystywanych w fizjologii wyjaśnić to, że podmiot ma doznanie bólu? (Levine 2001, s. 61–63, 70–76; 2006, s. 145–149)?

Oczywiście nie jest tak, że wyjaśnienie odwołujące się do faktów metafizycznych jest niezależne od wyjaśnienia odwołującego się do faktów epistemologicznych. Mówiąc słowami Levine’a, stanowią one dwie strony tego samego zjawiska. Niemniej jednak, kiedy podajemy metafizyczne wyjaśnienie dla analizowanego fenomenu, jesteśmy zobowiązani wskazać mechanizm odpowiedzialny za powstanie czy zajście tego fenomenu. Innymi słowy, *A* wyjaśnia metafizycznie *B* wtedy i tylko wtedy, gdy *A* jest odpowiedzialne za zajście *B*. Tego rodzaju określenie zależności między dwoma zjawiskami jest bardzo ogólne i zwykle znajduje swoją konkretyzację w regule zależności przyczynowej. Istotne jest jednak to, że w odpowiedzi

na metafizycznie rozumiane pytanie: dlaczego  $B$ ?, odpowiedź brzmi: dlatego, że zaszło  $A$ ; dlatego, że spowodowało to  $A$ ; na mocy  $A$  itd. Łącznie z założeniem przyczynowego domknięcia dziedziny zjawisk naturalnych takie postawienie problemu daje w efekcie redukcję sfery mentalnej do sfery naturalnej (Levine 2001, s. 11–26).

Inaczej jest w przypadku epistemologicznie pojętego pytania: dlaczego zaszło  $B$ ? Levine podkreśla, że aby dowiedzieć się dlaczego, gdy patrzy na czerwony przedmiot doznaje czerwieni, musiałby znać czynniki sprawiające, że utożsamienie stanu mentalnego ze stanem neuronalnym mózgu pozwala *zrozumieć*, dlaczego stan fenomenalny ma takie a nie inne własności. Innymi słowy musiałby się dowiedzieć, dlaczego występowanie  $A$  czyni inteligibilnym posiadanie przez  $B$  własności dla niego specyficznych (Levine 2001, s. 7 oraz 75–76). W przypadku bólu musiałby pokazywać, jak pobudzenie danego obszaru neuronów pozwala pojąć to, że ból jest nieprzyjemny, pulsujący lub kłujący itp. Przyjrzyjmy się, jakimi metodami wyjaśniania dysponuje filozofia i spróbujmy ustalić, czy któraś z nich może pomóc rozwiązać problem wskazany przez Levine’a.

Carl Hempel posługiwał się dedukcyjno-nomologicznym (D-N) oraz indukcyjno-statystycznym (I-S) modelem wyjaśniania zachodzenia danego zjawiska. W przypadku modelu D-N wyjaśniamy zachodzenie zjawiska  $z$ , gdy budujemy argumentację, która spełnia następujące warunki. Po pierwsze, jedna z jej przesłanek ma charakter prawa. Po drugie, jedna z jej przesłanek opisuje specyficzne warunki, które miały miejsce w danym czasie i miejscu, w którym zaszło  $z$ . Po trzecie, przesłanki wzięte łącznie tworzą *eksplanans* oraz prowadzą do konkluzji, tj. *eksplanandum* (Levine 2001, s. 70). Na przykład, aby wyjaśnić dlaczego jabłko spada z drzewa, należy przedstawić argumentację, która będzie zawierała prawo grawitacji oraz będzie opisywała specyficzne warunki, w których zachodziło spadanie konkretnego jabłka itd.

Model D-N nie jest jednakże wydolny w wyjaśnianiu wszelkiego rodzaju niezrozumiałych zjawisk. Aby poszerzyć zakres wyjaśnianych zjawisk, Hempel wprowadził model I-S, który charakteryzuje się tym, że w przeciwieństwie do modelu D-N wnioski nie wynikają koniecznie z przesłanek, ale jest jedynie prawdopodobny czy oczekiwany (Levine 2001, s. 71). Na przykład to, że ktoś został zarażony wirusem HIV można wyjaśnić przez wskazanie zachowań seksualnych tej osoby. Jednocześnie pozostaje rzeczą oczywistą, iż nie każdy zostaje zakażony wirusem HIV na drodze

plciowej. Jednak statystycznie rzecz biorąc, najczęściej do tego rodzaju zakażenia dochodzi w taki właśnie sposób.

Wskazane przez Hempla metody można uzupełnić o tzw. teorię unifikującą (TU), do której odwołuje się m.in. Michael Friedman, a którą rozwija Philip Kitcher. Metoda Kitchera polega na pokazaniu, w jaki sposób schemat argumentacji przewidujący zajście danego zjawiska czy zdarzenia z może zostać zunifikowany ze schematem argumentacji przewidującej zajście innych zjawisk czy zdarzeń. Wyjaśnienie unifikujące zostaje osiągnięte wówczas, gdy wykażemy, że zajście z jest jednostkowym przypadkiem pewnej generalnej prawidłowości, która tworzy spójną całość z koncepcjami i prawami dotyczącymi innych zdarzeń i zjawisk. Analizując dość szczegółowo przywołane tu trzy modele wyjaśnienia, Levine dochodzi do wniosku, że żaden z nich nie nadaje się do zastosowania w celu rozwiązania problemu psychofizycznego, gdy bierze się pod uwagę *qualia*.

Model D-N jest powszechnie krytykowany przede wszystkim za eksplanacyjną niesymetryczność. Chodzi o to, że skoro eksplanacja ma charakter dedukcji wniosku z przesłanek będących prawami oraz deskrypcjami specyficznych warunków, to powinna istnieć możliwość wydedukowania z tych aktualnie obowiązujących praw i deskrypcji zdarzenia, które miało miejsce w przeszłości. Jednak nie praktykuje się takiego postępowania. Analogii dostarcza następujący przykład. Wiadomo, że możemy obliczyć pole powierzchni flagi na podstawie przekształceń informacji na temat pola powierzchni rzutowanego przez nią cienia. Mimo tego tylko odwrotna sytuacja jest uznana za wyjaśnienie we właściwym sensie, tj. wyprowadzenie danych na temat pola cienia flagi z informacji na temat rozmiaru pola powierzchni flagi. Innymi słowy, wielkość pola flagi wyjaśnia rozmiar cienia, a nie odwrotnie – rozmiar cienia nie wyjaśnia wielkości pola powierzchni flagi<sup>3</sup>. Zanim przekonany się, jakie „problem flagi” niesie ze sobą konsekwencje dla zagadnienia relacji umysł–ciało, przyjrzyjmy się krytyce modelu D-N Roberta Cummins’a.

Otóż wychodzi on od spostrzeżenia, że w wyjaśnieniach zjawisk z zakresu psychologii stawia się na przykład pytanie: dlaczego podmiot posiada

---

<sup>3</sup> Jest to znany w literaturze przykład podawany przez Sylvaina Brombergera. Przytaczam go za Salmonem (Salmon 1989, s. 47), który zaznacza, iż prawdopodobnie nie istnieje żadna opublikowana forma wypowiedzi Brombergera, w której przywoływałby on ten przykład.

jakaś umiejętność? Wówczas gra toczy się nie o to, dlaczego występuje jakieś  $z$  (gdzie  $z$  to dana umiejętność podmiotu), ale dlaczego *ten* podmiot ma *tę* umiejętność? Zatem istotna jest tu własność i to, w jaki sposób jest ona konstytuowana, a nie typ przedmiotu i to, że występuje on zgodnie z pewną prawidłowością. Jak zauważa Cummins, tylko teoria własności byłaby w stanie odpowiedzieć na to pytanie. Teoria oparta na modelu dedukcyjno-nomologicznym pokazuje relacje między stanami, a nie między własnościami.

Levine podobnie nie widzi, w jaki sposób model D-N mógłby posłużyć za wyjaśnienie faktu posiadania przez podmiot własności fenomenalnych. Jego ustalenia w tej sprawie można zrekonstruować w następujący sposób. Pomimo że w oparciu o metodę dedukcyjno-nomologiczną możemy stwierdzić, że dla każdego stanu fizycznego o danej charakterystyce podmiot mający ten stan ma zarazem własność mentalną o charakterystyce fenomenalnej, to mimo wszystko zawsze będzie pojawiała się kwestia, skąd wiemy, że tak jest. Model D-N po prostu nie działa w przypadku związku psychofizycznego ani na płaszczyźnie logicznej, ani na płaszczyźnie empirycznej. Nie działa na płaszczyźnie logicznej, gdyż nie mamy uzasadnienia dla przejścia od przesłanki mówiącej, że podmiot mający pewne własności fizjologiczne mózgu ma zarazem dane własności o charakterystyce fenomenalnej, do wniosku, iż pewien konkretny podmiot ma określoną własność fenomenalną. Przejście od przesłanek do wniosku w przypadku dedukcji stanów czy własności fenomenalnych jest niemożliwe dlatego, że predykaty odnoszące się do stanów fizjologicznych czy fizycznych nie są logicznie powiązane z predykatami odnoszonymi do stanów fenomenalnych. Nie jest tak, że wiedząc wszystko na temat semantyki pojęcia stanu fizycznego, na przykład danej aktywności określonych włókien nerwowych, możemy dowiedzieć się czegoś na temat bólu, albo choćby czegoś na temat semantyki pojęcia bólu. Chyba że tą wiedzą okaże się informacja o tym, że pojęcie aktywności włókien nerwowych oraz pojęcie bólu w żaden sposób się logicznie *nie* implikują. W tym miejscu nasuwa się spostrzeżenie Thomasa Nagela, iż taka negatywna wiedza na temat relacji między pojęciem własności fenomenalnej i odpowiadającym jej pojęciem własności fizycznej, tłumaczy brak możliwości wyjaśnienia relacji umysł–ciało w całej rozciągłości, tj. także na płaszczyźnie empirycznej. W przypadku pojęcia na przykład liczby oraz pojęcia bycia potomkiem, jest faktem semantycznym czy logicznym to, że liczba nie może być niczym potomkiem. Podobnie,

konkluduje Nagel, z faktu niezależności semantycznej czy logicznej pomiędzy pojęciem własności fizycznej a pojęciem własności fenomenalnej możemy co najwyżej wnioskować, że pojęcie własności fenomenalnych nie jest logicznie implikowane przez pojęcie własności fizycznych i *vice versa* (Nagel 2002/2008, s. 404).

Przejdźmy do modelu I-S. Wydaje się, że skoro nie jesteśmy w stanie zagwarantować koniecznego charakteru naszemu wnioskowaniu, to jesteśmy przynajmniej w stanie wykazać, że jest to wnioskowanie indukcyjno-statystyczne. Teza o tym, że dla każdego podmiotu, który znajduje się w danym stanie fizycznym zachodzi to, że zarazem znajduje się on w danym stanie fenomenalnym jest wysoce prawdopodobna, statystycznie potwierdzona, indukcyjnie zagwarantowana itd. Niestety za modelem I-S ciągnie się poważny zarzut, który sprawia, że tę metodę Levine widzi również jako nieobiecującą na gruncie filozofii umysłu. Otóż, zauważmy, że możemy stwierdzić na przykład, że prawdopodobieństwo zajścia zjawiska  $z$  w warunkach  $w$  wynosi 90%. Wydaje się więc, że mamy prawo sądzić, że tam, gdzie wystąpią warunki  $w$ , zjawisko  $z$  również wystąpi. Levine wątpi jednak w tego typu wyjaśnienie. Mówi on, że tak, jak  $w$  daje 90% szansy na to, że zajdzie  $z$ , to tak samo  $w$  daje 10% szansy na to, że  $z$  nie zajdzie. Zatem  $w$  wyjaśnia zachodzenie  $z$ , tak samo, jak  $w$  wyjaśnia niezachodzenie  $z$ . W przełożeniu na problem psychofizyczny wygląda to w ten sposób, że jeśli mamy nawet 99% szansy, że stan fenomenalny  $z$  wystąpi wówczas, gdy zachodzi stan fizyczny  $w$ , to jednocześnie istnieje ten 1% szansy, że ów stan fenomenalny się nie pojawi, pomimo że zajdzie  $w$ . Zatem  $w$  nie wyjaśnia zachodzenia  $z$  w sposób filozoficznie satysfakcjonujący.

Największą szansę na wyjaśnienie korelacji psychofizycznej Levine widzi w modelu TU, który stanowi metodę łączącą wspomnianą przez Cumminsa próbę analizy własności z jednoczesnym zorientowaniem na nomologiczny charakter wyjaśnienia. Próba wyjaśnienia, dlaczego dany podmiot ma pewną umiejętność, może przebiegać w ten sposób, że wskazuje się najpierw na pewną rolę funkcjonalną, z którą ta umiejętność jest utożsamiana. Następnie można pokazać, że umysł stanowiący systemem funkcjonalny spełniający rolę określoną dla pewnej umiejętności, ma tę umiejętność. Tak rozumiany model TU jest po prostu rodzajem „argumentacji funkcjonalnej”. Problem niesymetryczności („problem flagi”) zostaje tu rozwiązany w ten sposób, że niepożądane wyjaśnienia (np. tłumaczące pole powierzchni flagi poprzez odwołanie się do pola powierzchni jej cienia)

nie dają się zunifikować za pomocą wspólnej dla wartościowych wyjaśnień zasady. Jest tak dlatego, że gdy można mówić o tym, że cieniem flagi jest coś, co pełni określoną rolę funkcjonalną w stosunku do flagi, to nie istnieje rola funkcjonalna flagi w stosunku do jej cienia. Nawet jeśli ktoś wymyśli charakterystykę takiej roli, to nie będzie ona dawała się zunifikować z teoriami dotyczącymi zajścia innych zjawisk o naturze zbliżonej do zjawiska wielkości pola powierzchni flagi. Co jednakże sprawia, że i ten model eksplanacyjny Levine odrzuca jako niezadowolający?

Zauważmy, że aby wyjaśnić, dlaczego ktoś ma daną własność fenomenalną trzeba tak skonstruować *eksplanans*, aby wynikało z niego logicznie *eksplanandum* lub aby wynikała z niego alternatywa, której jeden z członów jest naszym *eksplanandum*. Innymi słowy wyjaśnienie powinno mieć charakter dedukcyjny. W przypadku wyjaśniania występowania danej własności fenomenalnej powinniśmy potrafić wywnioskować ze zbioru twierdzeń dotyczących własności fizykalnych (zasad funkcjonalnych, zasad fizyki, zasad biologii, zasad neurologii itp.) twierdzenie zawierające pojęcia ze słownika mentalnego, tj. zawierające pojęcia fenomenalne. Rzecz polega więc na tym, aby, mówiąc słowami Levine'a, „uczynić oczekiwaną” relację konstytuowania przez warstwę niższą – fizykalną czy naturalną – warstwy wyższej – fenomenalnej (Levine 2001, s. 74, 76). Tak jak w przypadku wody i H<sub>2</sub>O potrafimy tego dokonać, tak w przypadku stanów fenomenalnych i odpowiednich stanów fizykalnych czy naturalnych nie jesteśmy w stanie tego zagwarantować. Gdy wiadomo, że między atomami wodoru a atomem tlenu istnieją określone połączenia, to staje się *oczekiwane*, że woda jest substancją płynną w temperaturze pokojowej. Oczekiwanie oznacza tu po prostu logiczny związek pomiędzy charakterystyką H<sub>2</sub>O a charakterystyką wody. W przypadku aktywności włókien nerwowych problem polega na tym, że nie istnieje logiczne powiązanie charakterystyki w rodzaju: „Przez włókna C przepłynął w czasie *t* impuls elektryczny o takim-a-takim wzorcu pobudzenia” a charakterystyką w rodzaju: „W czasie *t* miałam doznanie palącego bólu w dużym palcu u nogi”. Informacja na temat struktury chemicznej H<sub>2</sub>O jest zarazem informacją na temat stanu skupienia wody. Informacja na temat aktywności włókien nerwowych nie jest zarazem informacją na temat subiektywnych doznań podmiotu. Jeśli wyjaśnienie funkcjonalne cokolwiek zmienia to tylko tyle, że ma ono w najlepszym razie charakter analityczny. Własność palącego bólu jest tożsama z taką-a-taką rolą funkcjonalną włókien nerwowych lub innych czynników naturalnych. Levine'owi zależy co



prawda na pokazaniu, że identyczność między własnościami fenomenalnymi a własnościami naturalnymi jest konieczna, chce jednak, by konieczność tę odkrywać na drodze aposteriorycznej, a nie konstruować apriorycznie. Model unifikacyjny nie przynosi oczekiwanego w tym miejscu rezultatu z jednego powodu. Nie wiemy mianowicie, z jakimi prawami mogłaby zostać zunifikowana teza wyrażająca prawdopodobieństwo, że ból jest tożsamy z aktywnością włókien nerwowych. W przypadku wyjaśnienia funkcjonalnego nie wiadomo, jak mogłoby ono zostać zunifikowane z zasadami fizyki (biologii, fizjologii itp.).

## 2. Co konstituuje lukę eksplanacyjną?

Levine pisze:

Połączenie pomiędzy opisem neurologicznym a naszą standardową pierwszoosobową koncepcją tego, *jak to jest* [podkr. – A.S.] ] wydaje się całkowicie arbitralne. Można mieć poczucie, że dana neuronalna konfiguracja mogłaby zniknąć zwyczajnie tak samo wraz z pojawieniem się doświadczenia wzrokowego zarówno niebieskości, jak i czerwoności; lub równie dobrze mogłaby zniknąć wraz z pojawieniem się takiego stanu, że nie ma w ogóle nic takiego, jak *jak to jest* [podkr. – A.S.]. Oto ten kontrast – to poczucie arbitralności towarzyszące psychofizycznej redukcji jako przeciwieństwo poczucia rozumienia towarzyszącego teoretycznym redukcjom – stanowi rdzeń problemu, określanego mianem „luki eksplanacyjnej” (Levine 2006, s. 145).

Przytoczony fragment wypowiedzi Levine’a wskazuje, że epistemologiczny problem ze zdaniem utożsamiającymi własności czy stany fenomenalne z własnościami czy stanami naturalnymi stanowi sedno luki eksplanacyjnej, co w efekcie powoduje, iż zdanie typu (2), mianowicie „Ból jest aktywnością określonych włókien nerwowych”, nie powinno być rozumiane jako redukcja teoretyczna pojęcia fenomenalnego do pojęcia naturalnego. Levine zauważa, iż zdania typu (2) robią wrażenie identyczności przygodnej, a nie, jak zdania typu (1), mianowicie „Woda jest H<sub>2</sub>O”, identyczności o charakterze koniecznym. Co sprawia, że możemy pomyśleć o aktywności włókien nerwowych lub jakiegoś stanu funkcjonalnego, któremu nie towarzyszy żadne *jak to jest*?

Fizykaliści twierdzą, że redukcja teoretyczna ma zastosowanie zarówno na gruncie utożsamień typu (1), jak i na gruncie utożsamień typu (2). To, co

wyduje się problematyczne dla Levina, daje się w prosty sposób wyjaśnić poprzez wskazanie na nieaprioryczny charakter zdań identycznościowych typu (1) i (2). Wedle obrońców fizykalizmu nie istnieje możliwość wyprowadzenia zdania „ta istota ma doświadczenie wzrokowe czerwieni” ze zdania lub zbioru zdań sformułowanych wyłącznie za pomocą słownika fizykalnego czy naturalistycznego. Podobnie jest jednak na przykład ze zdaniem „w tym kubku jest woda”. Nie da się go wyprowadzić *a priori* ze słownika naturalisty. Jeżeli zatem niemożliwość inferencji powoduje, że możemy sobie pomyśleć istotę fizykalnie identyczną ze mną, ale pozbawioną stanów fenomenalnych, tj. zombi, to równie dobrze możemy pomyśleć substancję chemiczną, która składa się z innych pierwiastków niż odpowiednia kombinacja wodoru i tlenu, która jest odpowiednikiem wody, tj. Putnamowskie XYZ. Fakt, że jest możliwe *a priori* istnienie zombi oraz istnienie XYZ nie podważa faktu, iż zdania (1) i (2) są zrozumiałe. Ból jest aktywnością włókien nerwowych tak, jak woda jest określoną substancją chemiczną. Jeśli chcemy wiedzieć, dlaczego ból jest taką aktywnością, nie otrzymamy odpowiedzi, tak samo jak w przypadku pytania o to, dlaczego woda jest  $H_2O$ . Nie ma sensu zadawać pytania „dlaczego?” w przypadku stwierdzenia identyczności dwóch substancji, zdarzeń, przedmiotów czy własności.

Levine jest świadomy zasadności tego rozumowania. Jego zdaniem nie ma jednak realnego podobieństwa między pytaniem „dlaczego woda jest  $H_2O$ ?” a pytaniem „dlaczego ból jest aktywnością włókien nerwowych?”. Jeśli bowiem ktoś ma pełną wiedzę z zakresu chemii i fizyki, wie wszystko na temat substancji chemicznych i ich fizycznych własności, pytanie o to, dlaczego dany układ cząsteczek  $H_2O$  jest tożsamy z płynnością nie ma po prostu sensu. Inaczej jest w przypadku utożsamienia fenomenalnego charakteru bólu z aktywnością pewnych włókien w mózgu lub określonym stanem funkcjonalnym mózgu. Levine przywołuje znany przykład argumentu modalnego pochodzący od Franka Jacksona, opowiadający historię Mary, naukowca z przyszłości mającego kompletną wiedzę fizykalną o kolorach, zamkniętą przez całe życie w biało-czarnym pokoju (Jackson 1986). Na pytanie, czy po pokazaniu Mary czerwonego przedmiotu dowiedziałaby się ona czegoś nowego na temat kolorów, Levine odpowiada, że tak. Levine nie twierdzi jednak, w przeciwieństwie do wielu innych filozofów, że z faktu, iż Mary dowiedziała się czegoś nowego wynika, że istnieje niefizyczna własność odpowiadająca czerwieni fenomenalnej. Jego zdaniem fakt poszerzenia wiedzy Mary dowodzi raczej istnienia luki eksplanacyjnej pomiędzy

pojęciem naturalistycznie rozumianego koloru czerwieni jako długości fali świetlnej, a pojęciem fenomenalnie rozumianego koloru, jako tego, *jak to jest* doznawać czerwieni. I choć proponuje odmienne nazwy na oznaczenie koloru w sensie fizycznym – „czerwień” (*red*) – oraz na oznaczenie koloru w sensie fenomenalnym – „czerwoność” (*reddish*) – nie przeczy, że mamy wystarczające racje, by przyjmować, że metafizycznie czerwień i czerwoność są tym samym (Levine 2001, s. 6). Tym niemniej to, iż najbardziej nawet detaliczna i zupełna wiedza z zakresu fizyki i fizjologii nie tłumaczy, dlaczego Mary nie wiedziała, jak to jest doznawać czerwieni dopóki jej nie zobaczyła, sprawia, że pojmowalne jest zachodzenie odpowiedniej aktywności fizjologicznej mózgu bez obecności danego doznania, które powinno, zgodnie z redukcją typu (2), mieć miejsce. Jakie konsekwencje niesie ze sobą ta konkluzja dla teorii stanów fenomenalnych?

Saul Kripke (Kripke 1972/2001), a za nim David Chalmers (Chalmers 1996) zauważyli, że pomimo, iż pewne twierdzenia wydają się nie mieć charakteru analitycznego, to mimo wszystko są konieczne w sensie metafizycznym. Stosując pojęcia semantyki dwuwymiarowej, idea sformułowana przez Kripkego daje się ująć na nowo w następujący sposób. Twierdzenie, że na przykład wrażenie czerwieni jest aktywnością pewnych obszarów mózgu (lub stanem funkcjonalnym o danej charakterystyce itd.) wydaje się współwystępować z jednoczesną możliwością pomyślenia sobie sytuacji, w której ktoś ma wrażenie czerwieni, choć w jego mózgu nie zachodzi odpowiedni stan naturalny. Zatem twierdzenie utożsamiające wrażenie czerwieni ze stanem mózgu nie ma charakteru apriorycznego. Jego pierwotna intensja – gdzie przez pierwotną intensję rozumie się funkcję wyznaczników referencji ze świata możliwego do ekstensji wyrażenia – wydaje się mieć charakter przygodny. Wyznaczniki referencji wyrażenia „wrażenie czerwieni” wydają się nie wskazywać w sposób konieczny na określoną aktywność w mózgu. Jednakże intensja wtórna – funkcja wyznaczników referencji w naszym świecie aktualnym, zastosowanych jako sztywno desygnujących do każdego możliwego świata, w którym używamy wyrażenia – ma charakter konieczny. Nie można pomyśleć sytuacji, w której to, co w naszym aktualnym świecie jest referentem wyrażenia „wrażenie czerwieni”, mogłoby nie być nim w jakimś możliwym świecie. Intensja wtórna ustala bowiem warunek identyczności dla przedmiotu, do którego się odnosi. Twierdzenie, że wrażenie czerwieni jest pewną aktywnością mózgu wydaje się zatem mieć charakter aposterioryczny i konieczny za-

razem, a zatem taki, jak wszelkie inne twierdzenia aposterioryczne tożsamościowe (np. woda jest  $H_2O$ ). Jednakże Kripke nie ulega temu wrażeniu. Wykazuje on mianowicie, że już to, co za Chalmerssem możemy tu nazwać pierwotną intensją, w przypadku zdań utożsamiających stany fenomenalne ze stanami naturalnymi ma charakter konieczny, choć nie aprioryczny.

Przyjrzyjmy się ponownie twierdzeniu, że wrażenie czerwieni jest pewną aktywnością mózgu. Przyjmijmy zatem, że dla każdego  $x$  to, że  $x$  jest egzemplifikowane przez jakieś  $C$ , jest identyczne z tym, że  $x$  jest egzemplifikowane przez jakieś  $M$  (gdzie  $x$  to stan fenomenalny,  $C$  to wrażenie czerwieni, a  $M$  to odpowiedni stan mózgu). Wyznacznik referencji dla egzemplifikacji  $x$ -a przez  $C$  ma postać deskrypcji „to, co jest fenomenalnym odczuwaniem dla podmiotu, kiedy patrzy na czerwony przedmiot”, nazwijmy ją  $D_C$ . W naszym aktualnym świecie deskrypcja ta nie wskazuje wcale aktywności mózgu, a wręcz samonarzucające jest spostrzeżenie, że sztywno desygnuje ona właśnie to, o czym mówi: stan fenomenalny. Podobnie drugi człon naszej identyczności, tj. egzemplifikacja  $x$ -a przez  $M$ , z deskrypcją „taka-a-taka aktywność mózgu” już przy pierwotnej intensji sztywno desygnuje taką-a-taką aktywność mózgu. Zatem nasze zdanie identycznościowe wydaje się mieć charakter konieczny zarówno w odniesieniu do pierwotnej, jak wtórnej intensji, a to wydaje się rzeczą niemożliwą. Wiemy bowiem, że o tym, iż doznanie czerwieni jest aktywnością mózgu, dowiadujemy się aposteriorycznie, a o aposterioryczności twierdzenia decyduje to, że ma ono przynajmniej w odniesieniu do jednego terminu utożsamianego przygodny wyznacznik referencji (Loar 1997, s. 600).

Argumentacja Kripkego jest powszechnie znana w filozofii umysłu i nie ma sensu wymieniać wszystkich kluczowych reakcji, jakie wywołała. Levine zauważa, że nie sposób jej obejść nie wykazując, że materializm jest stanowiskiem fałszywym. Jest możliwe pojęciowo czy logicznie, że występuje stan fenomenalny doznawania czerwieni, gdy nie występuje pewna aktywność mózgu. Nie znaczy to jednak – jak wskazuje intensja wtórna – że jest możliwe metafizycznie występowanie takiego stanu fenomenalnego, które nie byłoby zarazem pewną aktywnością mózgu. Wydaje się, że można by się powołać na dystynkcję pojęcie–własność i argumentować, iż choć istnieje „logiczna niemoc” przy próbie sprowadzenia pojęć fenomenalnych do pojęć naturalnych, to jednak własność, która te pojęcia spełnia, jest dla nich ta sama. Levine nie widzi jednakże, w jaki sposób strategia ta mogłaby zadziałać na korzyść materializmu. Ostatecznie bowiem wykazanie, że

dwa niesprowadzalne do siebie pojęcia są spełniane przez tę samą własność polega na wykazaniu, iż mają one jakiś wspólny, bardziej podstawowy wyznacznik referencji. W przypadku wody i  $H_2O$  byłaby to deskrypcja na przykład o postaci „substancja chemiczna złożona z odpowiednich pierwiastków połączonych ze sobą w odpowiednich proporcjach o danych cechach makroskopowych”. Nie istnieje jednak, konkluduje Levine, taki wyznacznik dla pojęcia fenomenalnego oraz dla pojęcia naturalnego, jak bowiem miałyby wyglądać wspólna deskrypcja wyznaczająca zarówno referencję doznania, jak i stanu fizjologicznego? (Levine 2001, s. 48–49). W związku z tym, zdaniem Levine’a, nie potrafimy zrozumieć, jak mógłby przebiegać pomost pomiędzy koncepcją stanów fenomenalnych a koncepcją stanów fizykalnych.

### 3. Krytyka pojęcia luki eksplanacyjnej

Można się zgodzić z Levinem, że istnieje luka eksplanacyjna w przypadku problemu umysł–ciało, zastrzegając jednak, że jest to przede wszystkim luka powstała ze względu na logiczną niesprowadzalność poziomów opisu faktów składających się na zjawiska, które mają zostać wyjaśnione. Nie możemy literalnie czy wprost obserwować, jak stany fizjologiczne czy fizyczne konstituują umysł, w przeciwieństwie do tego, jak możemy obserwować, zmieniając stopień powiększenia porcji wody na szkiełku mikroskopu, jak dana struktura chemiczna konstituuje płynność wody. W związku z tym zasadna wydaje się potrzeba posłużenia się inną metodą wyjaśnienia natury stanów fenomenalnych niż metoda obserwacyjna. Nic dziwnego, że potrzebę tę dostrzegli emergentyści, twierdząc, że jakości fenomenalne doświadczenia należy opisywać i wyjaśniać w oparciu o prawidłowości postulowane dla tych własności, a nie dla własności realizowanych na poziomie naturalnym rzeczywistości rządzonej prawami fizyki. Zastanówmy się, czy postulowane przez Levine’a rozumienie luki eksplanacyjnej można sprowadzić do pojęcia luki między poziomami opisu.

Przyjrzyjmy się ponownie zdaniu stwierdzającemu identyczność aposterioryczną: (1) „Woda jest  $H_2O$ ”. Obserwacja dostarcza, wedle Levine’a, wystarczających racji, aby uznać metafizyczną tożsamość pomiędzy substancją, która towarzyszy nam w życiu codziennym, a substancją, która jest poddawana analizom w laboratoriach chemicznych. Przypomnijmy, że fakt, iż zdanie (1) jest prawdziwe, sprawia, że nawet mówienie o identyczności

między wodą i  $H_2O$  brzmi zupełnie nieswojo. Chce się od razu powiedzieć, że w zdaniu (1) nie utożsamia się żadnych substancji, tylko wyraża coś na temat jednej i tej samej substancji. Fakt ten suponuje – jak była już o tym mowa – że zdania o postaci (1), jeśli tylko są dostatecznie umotywowane, nie domagają się żadnych dalszych wyjaśnień. Gdy już wiemy, że woda jest  $H_2O$ , to potrafimy zrozumieć, dlaczego woda ma postać płynną w temperaturze pokojowej, dlaczego może zamarzać, dlaczego jest przejrzysta itp.

Weźmy teraz zdanie: (2) „Ból jest aktywnością włókien nerwowych”. Wedle Levine’a zdanie to, stwierdzając identyczność między bólem i pewnym stanem fizjologicznym mózgu, jest równie metafizycznie uzasadnione, co zdanie (1). Niemniej Levine twierdzi, iż w przeciwieństwie do zdań, w których dokonuje się redukcji teoretycznej, zdania takie, jak (2) nie dostarczają wyjaśnienia, dlaczego pewne własności stanów fenomenalnych takie, jak na przykład nieprzyjemny charakter bólu są tym, czym są. Odwołanie się do własności włókien nerwowych nie sprawia, że fakt, iż ból jest odczuwany staje się zrozumiałym. Pamiętajmy jednak, że zarazem twierdzi on, że jest faktem metafizycznie bezspornym, iż zachodzi (2). Jak zatem rozumieć to, że istnieje problem epistemologiczny związany z utożsamieniem stanów fenomenalnych ze stanami naturalnymi?

Twierdzenie to jest uzasadnione, gdyż zdaniem Levine’a mamy wystarczające racje, by wierzyć, iż stany fenomenalne są stanami naturalnymi. Właściwie gwarancja, że nie pojawi się luka metafizyczna między dziedziną bytów mentalnych i dziedziną bytów fizycznych jest pochodną założenia, które Levine zawsze czyni na wstępie swoich rozważań, że dziedzina fizyczna jest zupełna i przyczynowo domknięta. Zatem, tak jak nie mamy podstaw, by twierdzić, że substancja, którą jest woda nie jest tą samą substancją, którą jest  $H_2O$ , tak też rzecz się ma z metafizycznie rozumianą tezą stwierdzającą identyczność psychofizyczną.

Przejdźmy teraz do dystynkcji, jaką posługuje się Levine, na dwa niesprowadzalne do siebie sposoby, na jakie podmiotowi dane są pojęcia (*modes of presentation*). Idea sposobu dania pojęć stanowi nawiązanie do pomysłu Gottloba Fregego, by oprócz sensu i znaczenia wyrażań językowych mówić również w miarę potrzeby o tym, że dla danego podmiotu, jeśli dysponuje on jakimś pojęciem, to jest coś takiego, jak sposób, na jaki to pojęcie jest mu dane. Wedle Levine’a pojęcia naturalne są dane w sposób niejako ubogi (*presentationally thin*), co oznacza, że pozostawanie w określonego rodzaju kognitywnej relacji pomiędzy podmiotem a tymi pojęciami nie ma

znaczenia dla zasadności przypisywania temu podmiotowi dysponowania tymi pojęciami. Nie ma mianowicie znaczenia to, czy pojęcia te dane są podmiotowi przez kognitywny kontakt ze stowarzyszonymi z nimi *qualiami* (Levine 2001, s. 8). Na przykład nie ma znaczenia to, że może nigdy nie widziałam, nie słyszałam itd. ptaka kiwi dla faktu, że mimo wszystko zasadne będzie przypisywanie mi dysponowania pojęciem ptaka kiwi, jeśli tylko znam istotne cechy tego stworzenia.

W przeciwieństwie do tego, sposób na jaki dane są pojęcia fenomenalne, Levine określa mianem bogatego (*presentationally thick*). Oznacza to, że kognitywna relacja pomiędzy *qualiami* stowarzyszonymi z nimi a podmiotem konstituuje stan dysponowania takimi pojęciami. Pojęcie bieli, gdy odnoszę je do koloru kwiatów magnolii, nie tylko mówi coś na temat kwiatostanu tego drzewa, ale jednocześnie „niesie ze sobą biel”. Inaczej mówiąc, nie mogę wiedzieć, co to znaczy, że kwiat jest biały, jeśli nie pozostaję w relacji jakiegś kognitywnej imaginacji do bieli. Dystynkcja na pojęcia, których sposób, na jaki dane są podmiotowi (jak mu się prezentują), jest ubogi lub bogaty ma tłumaczyć na gruncie teorii Levine’a, gdzie leży czynnik decydujący o kognitywnej wyjątkowości pojęć fenomenalnych. Levine znajduje go w fakcie, który można zinterpretować następująco: dla możliwości przypisania podmiotowi dysponowania pojęciem fenomenalnym istotne jest to, by zachodziła relacja jakiegś kognitywnej imaginacji pomiędzy tym podmiotem a przypisywanym mu pojęciem (Levine 2001, s. 8 oraz np. s. 164–165).

Fakt, że sposób, na jaki dane są podmiotowi pojęcia fenomenalne jest bogaty w sensie Levine’a, powoduje, że nie potrafimy zrozumieć, jak zdanie „Ból jest aktywnością włókien nerwowych” mogłoby być dla tych pojęć prawdziwe. W myśl rozumowania Levine’a bowiem nie możemy wyjaśnić pojęcia o bogatym sposobie prezentacji podmiotowi przez pojęcie o ubogim sposobie prezentacji podmiotowi. Przyjrzyjmy się jeszcze, co dokładnie decyduje o tym, że pojęcia fenomenalne nie są redukowalne do pojęć fizykalnych, a następnie sprawdźmy, czy fakt ten ma takie konsekwencje, na jakie wskazuje Levine, tj. konstituuje lukę epistemologiczną a nie metafizyczną.

Levine twierdzi, że zdanie o postaci „Doznanie bólu jest tożsame z aktywnością danego obszaru mózgu” jest prawdziwe *a posteriori*, tylko dlatego, że obserwujemy korelacje zachodzące pomiędzy stanami fenomenalnymi a stanami fizjologicznymi oraz dlatego, że fizyka (czy „nauki

naturalne”) jest lub będzie zupełna w swoich wyjaśnieniach, zaś postulowana przez nią przyczynowość fizyczna ma charakter domknięty. Uzasadnienie podane przez Levine’a można jednak uznać nie tylko za niewystarczające, ale nawet za fałszywe. Może nie być ono wystarczające, gdyż nawet jeśli okazałoby się, że rzeczywiście obserwujemy stosowne korelacje, to z tezy o zachodzeniu systematycznych korelacji nie wynika teza o tym, że korelacja ta ma charakter konieczny metafizycznie. Uzasadnienie Levine’a dla tezy o metafizycznie koniecznej identyczności psychofizycznej można natomiast uznać za fałszywe podważając to, że faktycznie obserwujemy cudze stany fenomenalne. Nie jest tak, że obserwuję czyjś ból czy czyjes doznanie bieli, tak jak obserwuję skład chemiczny wody. Tylko przy bardzo liberalnym rozumieniu pojęcia obserwacji możemy powiedzieć, że obserwujemy korelacje psychofizyczne. Takim liberalnym rozumieniem obserwacji byłaby koncepcja, w myśl której na przykład postulowanie pewnych własności cząsteczek materii jest tłumaczone przez obserwację zachowań elektronów atomów, z których się one składają. Przy czym śledzenie ruchu elektronów odbywa się poprzez bezpośrednią obserwację na przykład śladów, jakie zostawiają na specjalnie przygotowanym ekranie. Zatem nie byłoby tak, że obserwujemy bezpośrednio ruch elektronów, niemniej mimo wszystko mówiłoby się o ich obserwacji.

Można by zatem co najwyżej zaszcześcić ten model rozumienia pojęcia obserwacji na grunt filozofii umysłu i twierdzić, że w analogiczny sposób obserwujemy stany fenomenalne, tj. poprzez bezpośrednią obserwację szeroko pojmowanego zachowania organizmu oraz, w przypadku człowieka, w oparciu o raporty introspekcyjne. Fakt, że tak liberalne pojęcie obserwacji przyjęło się na gruncie fizyki znajduje motywację w tym, iż przyjmuje się, że pomiędzy śladami elektronów na ekranie a rzeczywistym ruchem elektronów zachodzi związek konieczny. W przełożeniu na problem umysł–ciało znaczyłoby to, że Levine musiałby przyznać, że pomiędzy zachowaniem, czy danymi introspekcyjnymi, a stanami fenomenalnymi zachodzi związek konieczny. Z oczywistych, historycznych już racji nie może tego jednak zrobić. W związku z powyższym, wbrew powszechnym praktykom filozofów, trudno jest zachować tezę, iż obserwujemy konieczne relacje współwystępowania pomiędzy stanami fenomenalnymi a stanami fizjologicznymi czy naturalnymi. Można zaś bez większego trudu przystać na słabszą tezę, mianowicie, że wnioskujemy na podstawie rozmaitych świadectw i przesłanek o tym, że zachodzą konieczne korelacje między umysłem



i ciałem. Należy jednak mieć świadomość, że wnioskowanie to może mieć co najwyżej charakter wnioskowania do najlepszego wyjaśnienia.

#### 4. Dalsze trudności stanowiska Levine'a

Proponuję rozpatrzyć dwa spośród nasuwających się zastrzeżeń adresowanych do pojęcia luki eksplanacyjnej w rozumieniu Levine'a. Pierwsze z nich wypływa z koncepcji treści doświadczenia percepcyjnego, którą zaszczerpia na grunt swoich rozważań Levine, drugie dotyczy kategorii bogatego sposobu, na jaki dane są pojęcia mentalne.

Poniżej spróbuję wykazać, że problem z utożsamieniami typu: „Ból jest aktywnością określonych włókien nerwowych” polega ostatecznie, jeśli przyjąć podział na dwa sposoby prezentacji pojęć, na niewspółmierności treściowej pomiędzy *definiendum* a *definiensem*. Jeśli tak jest, to zanika główna przesłanka, na mocy której Levine opiera swoją argumentację na rzecz istnienia luki epistemologicznej. Jest on natomiast zobowiązany, w pierwszej kolejności, do postulowania istnienia luki metafizycznej.

Pomimo że Levine zastrzega, iż metafizyczny problem luki eksplanacyjnej nie występuje, deklaracja ta wydaje się na wyrost w świetle dystynkcji na ubogi i bogaty sposób prezentacji pojęć. Przeanalizujemy, czym dokładnie jest kognitywnie bogata prezentacja pojęcia podmiotowi.

Levine rozpatruje rolę różnych wersji reprezentacjonizmu, w tym reprezentacjonizmu relacyjnego, w objaśnianiu natury relacji między treścią pojęć odnoszącą się do elementów świata a jakością fenomenalną towarzyszącą tej treści. W odróżnieniu od reprezentacjonizmu internalistycznego, zgodnie z którym fenomenalny charakter doświadczenia jest egzemplifikowany (*instantiated*) przez podmiot tego doświadczenia, jest to stanowisko przyznające rację teoriom świadomości wyższego rzędu (HO – *high order theories*). Albowiem reprezentacjonalista relacyjny wskazuje na kognitywną relację między podmiotem doświadczenia a tym, co jest doświadczane, czyli na fakt reprezentowania. Pomimo że teorie świadomości wyższego rzędu są zasadniczo internalistyczne (doświadczenie konstytuowane jest przez zależność między podmiotem a jego mentalnymi stanami wewnętrznymi), to z reprezentacjonizmem relacyjnym dzielą wspólne założenie, że doświadczać czegoś oznacza dla podmiotu coś reprezentować (Levine 2010, s. 211).

Zgodnie z ustaleniami Levine'a powoływanie się na eksternalistycznie rozumianą relację percepcyjnego doświadczania daje szansę, z jednej

strony, ocalenia intuicyjnego przekonania o tym, że treścią doświadczenia jest jakiś fragment świata zewnętrznego ujęty mentalnie w dany sposób (reprezentowany), z drugiej zaś strony pozwala zachować fenomenalny czy jakościowy charakter tego aktu. Przy czym charakter fenomenalny jest przez niego rozumiany następująco:

Wykorzystując tak zwany argument „przezroczystości” reprezentacji utrzymują, że doświadczanie jest pewnym sposobem reprezentowania świata. Jest tym, co faktycznie czyni je szczególnie mentalnym, a zatem tym, co odróżnia egzemplifikowanie charakteru fenomenalnego od egzemplifikowania wagi lub wysokości – tego, co odróżnia, powiedzmy, doświadczenie percepcyjne od procesów trawiennych (Levine 2010, s. 211).

Ostatecznie Levine podważa jednak reprezentacyjny charakter doświadczenia (Levine 2010, s. 213). Rysując ostrą granicę między reprezentacjonizmem relacyjnym a reprezentacjonizmem internalistycznym stara się ustalić, które z tych dwóch podejść do natury doświadczenia percepcyjnego lepiej współgra z teorią świadomości wyższego rzędu (HO). Nie można jej pominąć w związku z tym, że według niego tylko ta teoria świadomości nadaje się do wyjaśnienia na gruncie reprezentacjonizmu zjawiska, które stanowi wyzwanie i swoisty test mocy eksplanacyjnej dla każdej teorii percepcji, tj. zjawisko doświadczania halucynacji, o którym nieco więcej przy okazji omawiania drugiego zarzutu.

Przez reprezentacjonizm relacyjny rozumie Levine takie podejście do doświadczenia percepcyjnego, którego centralną kategorią jest relacja reprezentowania. Inaczej rzecz się ma w przypadku reprezentacjonizmu internalistycznego, gdzie, jak podkreśla Levine, akcentuje się rolę tego aspektu reprezentacji, który przesądza o jej charakterze fenomenalnym. Jak wyjaśnia Levine:

To, co nazywam „internalistycznym” reprezentacjonizmem jest zgodne ze stanowiskiem, że dane treści reprezentacyjne są treściami szerokimi dotyczącymi dystalnych własności przedmiotów, a nie wąskich treści lub wrażeń. Treść reprezentacji jest eksternalna, lecz to, co jest internalne, to sam charakter fenomenalny jako utożsamiony z własnością reprezentacji, a nie z własnością przedmiotu dystalnego (Levine 2010, s. 224, przypis 4).

Cały kłopot polega na tym, że nie wiadomo, co dokładnie ma znaczyć internalistyczny w zestawieniu terminów, które proponuje Levine, a zatem, co ostatecznie rozumie on przez treść reprezentacji. Przywołując klasyczną interpretację internalizmu co do treści oraz jego przeciwieństwa eksternalizmu oraz podział na wąską i szeroką treść, nie jest oczywiste, że idee te można by ze sobą połączyć w sposób zaproponowany przez Levine'a: internalizm z kategorią wąskiej treści, zaś eksternalizm z kategorią szerokiej treści.

Zgodnie z internalizmem treść doświadczenia jest w całości determinowana czynnikami wewnętrznymi, tj. umysłowymi. Eksternaliści zaprzeczają temu twierdzeniu, starając się wykazać, że przynajmniej niekiedy treść doświadczenia jest determinowana przez sam świat zewnętrzny. Internalizm i eksternalizm w odniesieniu do treści mentalnej są stanowiskami dotyczącymi tego, jakie czynniki – wewnętrzne czy zewnętrzne w stosunku do umysłu – determinują treść doświadczenia. Internaliści będą się spierać z eksternalistami o to, czy jest to relacja przyczynowa, czy raczej jakiś rodzaj umysłowej konstrukcji itp.

Koncepcja wąskiej i szerokiej treści dotyczy innego zagadnienia. Powstała jako odpowiedź na pytanie o to, co jest przedmiotem doświadczenia. W przypadku treści wąskiej przedmiotem tym jest jakiś byt umysłowy, zaś w przypadku treści szerokiej jest nim bądź obiektywnie rozumiane znaczenie wyrażenia językowego, bądź przedmiot w świecie.

Levine twierdząc, że najbliższy wyjaśnienia fenomenalnego charakteru doświadczenia percepcyjnego jest internalistyczny reprezentacjonizm – rozumiany jako połączenie ze sobą eksternalizmu z internalizmem, a jeszcze na dodatek wąskiego i szerokiego podejścia do treści doświadczenia więcej zaciemnia, niż wyjaśnia. Być może chodzi mu o to, że treść doświadczenia składa się z elementu konstytuowanego przez czynniki zewnętrzne, przez które rozumie on sposób oddziaływania przedmiotu dystalnego na receptory. Jednakże, jak zastrzega, aby zachować fenomenalny charakter doświadczenia, trzeba przyjąć, że obok tego elementu znajduje się również taki, który jest konstytuowany przez jakieś czynniki wewnętrzne, tj. własności samej reprezentacji, które są efektem, jak można się domyślać, działania świadomości wyższego rzędu. Niejasne jest to, w jakim celu wykorzystuje on do określenia tej idei zarówno stanowisko eksternalizmu, jak i internalizmu. Przecież to, co twierdzi Levine, jest zawarte w postulatach eksternalizmu umiarkowanego: część treści doświadczenia percepcyjnego jest konstytu-

owana jakąś relacją (przyczynową, ujmowania, obejmowania, sondowania itp.) między umysłem a światem (por. teoria przyczynowa). Eksternalista nie jest zobowiązany do akceptacji twierdzenia, że cała treść mentalna jest konstituowana czynnikami zewnętrznymi.

Levine'owi zależy z pewnością na tym, by podkreślić rolę jakościowego aspektu doświadczenia, tj. faktu, że treść doświadczenia może być przedmiotem aktu świadomości. Może dlatego posługuje się kategoriami wąskiej i szerokiej treści. Obok przezroczystej treści odnoszącej się do elementów świata, jest jeszcze treść sprowadzająca się do tego, *jak to jest mieć* treść pierwszego rodzaju. Do ich określenia Levine używa nazw „wąska” i „szeroka” treść. Spory między zwolennikami wąskiej treści oraz ich przeciwnikami argumentującymi na rzecz treści szerokiej pokazują, że jest czymś niesłychanie skomplikowanym, być może nawet niemożliwym, połączenie ze sobą tych podejść w jedno spójne stanowisko. Spór dotyczy bowiem tego, co jest przedmiotem doświadczenia i karkołomną próbą wydaje się opowiadanie, że jest nim zarazem coś zewnętrznego w stosunku do aktu psychicznego, jak i coś w stosunku do niego wewnętrznego. Ostatecznie bowiem przez to, co jest zewnętrznym przedmiotem doświadczenia rozumie się obiektywnie, tj. niezależnie od aktów poznawczych podmiotu, istniejącą rzeczywistość, w przeciwieństwie do przedmiotu wewnętrznego w stosunku do umysłu, który nie stanowi elementu obiektywnego umeblowania świata. Jaki przedmiot spełniałby warunek Levine'a, tj. byłby zarazem przedmiotem obiektywnym, jak i subiektywnym doświadczenia?

Podsumowując, Levine posługuje się niejasną kategorią treści doświadczenia, co utrudnia zorientowanie się, jakiego dokładnie broni poglądu w kwestii relacji umysł–ciało.

Przejdźmy teraz do trudności związanej ze sposobem, w jaki rozumie on kategorię bogatego sposobu dania podmiotowi pojęć własności fenomenalnych. Problem z percepcją nieweredywną (iluzje, halucynacje, złudzenia zmysłowe), jaki napotyka reprezentacjoniści wynika z przyjmowanego przez nich założenia o przezroczystości doświadczenia percepcyjnego. Jeśli bowiem treść reprezentacji jakiegoś przedmiotu wyczerpuje się w cechach tego przedmiotu na takiej zasadzie, że ostatecznie treść tę utożsamia się z owym przedmiotem, to nie do końca wiadomo, w jaki sposób możliwe są doświadczenia, w których podmiotowi coś się jawi, pomimo że nie stanowi to aktualnego przedmiotu jego percepcji. Jeśli treść reprezentacji percep-

cyjnej jest przezroczysta, to w jaki sposób może niekiedy ukazywać coś, co aktualnie nie jest dane w akcie percepcji?

Wprowadzając do podejścia reprezentacjonistycznego teorię świadomości wyższego rzędu otrzymujemy, zdaniem Levine'a, proste rozwiązanie tego problemu. W akcie percepcji nieweredycyjnej dochodzi do sytuacji, w której świadomość pewnej treści jakiegoś stanu percepcyjnego staje się doświadczeniem, nie zaś świadomość skierowana niejako na zewnątrz umysłu, tj. do przedmiotów w świecie. Innymi słowy, zdaniem Levine'a połączenie reprezentacjonizmu z teorią świadomości wyższego rzędu wyjaśnia zjawisko percepcji nieweredycyjnej. Koncepcja powstała z połączenia tych dwóch składników określana jest przez niego teorią świadomości wyższego rzędu reprezentacjonizmu internalistycznego (HOIR), (Levine 2010, s. 213). Teoria HOIR ma również, według Levine'a, zastosowanie w wyjaśnieniach wszelkich fenomenów percepcji, a nie wyłącznie w wyjaśnieniach doświadczenia nieweredycyjnego.

W miejsce reprezentacjonizmu proponuje on ostatecznie koncepcję tego, jak świat jawi się podmiotowi w percepcji, w skrócie koncepcję jawienia się (*appearance theory*). Levine pisze:

Charakter fenomenalny zatem – sposób jawienia się [przedmiotu – A.S.] – jest rozumiany jako determinant relacji jawienia się. Na przykład, gdy wzrokowo doświadczam zaśniewzonego pola na zewnątrz, to stoję w relacji świadomości [*awareness*] do zaśniewzonego pola, zaś białawy charakter fenomenalny jest determinująca formą tej relacji. W przeciwieństwie do HOIR charakter fenomenalny nie stanowi cechy przedmiotu świadomości, ale samej relacji świadomości (Levine 2010, s. 220).

Jakie konsekwencje dla proponowanej przez Levine'a koncepcji pojęć fenomenalnych ma opowiedzenie się przez niego po stronie koncepcji jawienia się? Zauważmy, że Levine jest zobowiązany do twierdzenia, że sposób prezentacji czy tego, jak dane są pojęcia fenomenalne, wymaga od podmiotu, aby potrafił on dokonać imaginatywnej rekonstrukcji (wizualizacji) własności spełniającej to pojęcie. W przypadku barwy czerwonej powiemy, że ktoś dysponuje pojęciem czerwieni, jeśli jest w stanie odtworzyć nie tylko deskryptywną treść pojęcia czerwieni (czyli wszystkie istotne dla niej cechy), ale również *przedstawić to sobie* w formie wyobrażenia doznawania czerwieni. Skoro bowiem: „Charakter fenomenalny jest zasadniczo sposobem jawienia się, cechą, która jest egzemplifikowana tylko wówczas, gdy ma

miejsce jawienie się [przedmiotu – A.S.]” (Levine 2010, s. 220) i „nie możemy uczynić współmiernymi pierwszoosobowej reprezentacji [własności fenomenalnych – A.S] z trzecioosobowymi, teoretycznymi [własnościami – A.S.]” (Levine 2001, s. 164), to przypisywanie danemu podmiotowi dysponowania pełnym pojęciem własności wyposażonej w charakter fenomenalny (np. czerwieni) zakłada konieczność reprezentowania sobie przez ten podmiot, obok cech deskrypcyjnych przedmiotu, również charakteru fenomenalnego, czyli tego, jak on się percepcyjnie podmiotowi jawi.

Wydaje się więc, że Levine przesądza, iż istnieje luka metafizyczna i tym samym podważa swoje wcześniejsze stanowisko w tej sprawie. Jeżeli przyjmiemy założenie, że sposób prezentacji danego pojęcia wymaga od dysponującego nim podmiotu umiejętności wizualizacji desygnatu, to trzeba także przyjąć, że istnieje coś takiego, jak własność bycia doznawanym dla tego desygnatu. Levine zresztą otwarcie zgadza się na ten krok już w chwili, w której opowiada się po stronie koncepcji jawienia się. Jeśli teraz dla pojęcia barwy czerwonej trzeba umieć sobie wizualizować, jak to jest doznawać czerwieni, to musi istnieć taka własność czerwieni, która jest własnością *bycia doznawaną jako czerwień*. Jest tak dlatego, że to, co jest dane w wizualizacji, stanowi własność definiującą czerwień obok wszelkich własności deskryptywnych czy teoretycznych (trzecioosobowych). Zatem własność przesądzająca, że mamy do czynienia z pojęciem danym podmiotowi na sposób bogaty byłaby tożsama z niefizyczną własnością *bycia doznawanym jako to-a-to*. Następnie, z faktu, że cechą definicyjną jakiegoś przedmiotu lub własności jest to, że jest on doznawany w pewien sposób, wynika, że ma on własność bycia doznawanym w ten sposób.

Można jednak w tym miejscu podnieść następującą trudność. Sam przedmiot ma co najwyżej własność *bycia doznawanym jako  $\phi$* , zaś własność doznawania (czyli *jak to jest doznawać, jako...*), jest własnością relacyjną, wiążącą podmiot z przedmiotem w akcie percepcji. Zauważmy jednak, że własność *bycia doznawanym jako  $\phi$*  jest mimo wszystko tożsama z własnością *jak to jest doznawać jako  $\phi$* . Własność bycia doznawanym jako czerwony jest treściowo identyczna z własnością doznawania czerwieni. Pierwsza jest, wbrew zabiegom Levine’a, elementem świata zewnętrznego, a nie samej relacji między podmiotem a światem. Jako niefizyczna część umeblowania świata przemawia ona na rzecz istnienia luki metafizycznej.

Podsumowując, wprowadzenie przez Levine'a kategorii bogatego sposobu dania pojęć fenomenalnych podmiotowi, w połączeniu jeszcze z koncepcją jawienia się przedmiotów w percepcji oraz z zastrzeżeniem do nieredukowalności pojęć prezentujących się podmiotowi na sposób bogaty do pojęć danych podmiotowi na sposób ubogi, prowadzi do konieczności akceptacji istnienia luki metafizycznej, idei, którą Levine odrzuca bez przekonującego wyjaśnienia.

## Literatura

- Chalmers D. (1996), *The Conscious Mind*, Oxford: Oxford University Press.
- Jackson F. (1986), *What Mary didn't know?*, „Journal of Philosophy”, 83, s. 291–295.
- Kitcher P. (1989), *Explanatory Unification and the Causal Structure of the World*, [w:] P. Kitcher, W.C. Salmon (eds.), *Minnesota Studies in the Philosophy of Science*, Vol. XIII: *Scientific Explanation*, Minneapolis: University of Minnesota Press.
- Kripke S. (1972/2001), *Nazywanie a konieczność*, przeł. B. Chwedeńczuk, Warszawa: Fundacja Aletheia.
- Levine (1983), *Materialism and Qualia: The Explanatory Gap*, „Pacific Philosophical Quarterly”, 64, s. 354–361.
- Levine J. (1998), *Conceivability and the Metaphysics of Mind*, „Noûs”, 32 (4), s. 449–480.
- Levine J. (2001), *Purple Haze. The Puzzle of Consciousness*, Oxford: Oxford University Press.
- Levine J. (2006), *Phenomenal Concepts and the Materialist Constraint*, [w:] T. Alter, S. Walter (eds.), *Phenomenal Concepts and Phenomenal Knowledge: New Essays on Consciousness and Physicalism*, Oxford: Oxford University Press, s. 145–166.
- Levine J. (2010), *Phenomenal Experience: A Cartesian Theater Revival*, „Philosophical Issues”, 20, s. 209–225.
- Loar (1997), *Phenomenal States*, [w:] N. Block, O. Flanagan, G. Güzeldere (eds.), *The Nature of Consciousness: Philosophical Debate*, Cambridge: MIT Press.

- Nagel T. (2002/2008), *Związek psychofizyczny*, [w:] M. Miłkowski, R. Poczobut (red.), *Analityczna metafizyka umysłu*, Warszawa: Wydawnictwo Instytutu Filozofii i Socjologii Polskiej Akademii Nauk, s. 394–441 [oryginał: T. Nagel, *The Psychophysical Nexus*, [w:] tenże, *Concealment and Exposure*, Oxford: Oxford University Press 2002, s. 194–235].
- Salmon W.C. (1989), *Four Decades of Scientific Explanation*, [w:] P. Kitcher, W.C. Salmon (eds.), *Minnesota Studies in the Philosophy of Science*, vol. XIII: *Scientific Explanation*, Minneapolis: University of Minnesota Press.

## EXPLANATORY GAP AND THE CONCEPT OF PHENOMENAL PROPERTIES: JOSEPH LEVINE'S VIEW

### Summary

The explanatory gap problem arises in the context of the mind-body relation, and especially the phenomenal concepts-physical concepts relation. It is posed as a question about the method which is appropriate for the reduction of consciousness to physical or neural states of the brain. Therefore it concerns – as the paper suggests – theoretical incapability of naturalistic attempts to explain what phenomenal concepts are about in terms of what physical or natural concepts are about. The paper discusses the argument of Joseph Levine, one of the best-known critics of the reductive attempts to close the explanatory gap. The bottom line of the argument is that since phenomenal concepts are theoretically thick, and physical concepts are theoretically thin, there is no way to reduce the former to the latter.