

Anna Brożek, Jacek Jadacki

## **Eksperymenty myślowe w nauce**

1. Uwagi wstępne. 2. Eksperyment rzeczywisty. 3. Warianty definicji „eksperymentu rzeczywistego”. 4. Typy eksperymentów rzeczywistych. 5. Rzeczywisty eksperyment heurystyczny a rzeczywisty eksperyment testujący. 6. Eksperyment myślowy. 7. Typy eksperymentów myślowych. 8. Myślowy eksperyment heurystyczny i testujący. 9. Rzeczywisty a myślowy eksperyment testujący. 10. Eksperyment udany i nieudany. 11. Eksperymenty myślowe w fizyce. 11.1. Spadające kule. 11.2. Światło w pociągu. 12. Podsumowanie.

**Tomasz Bigaj: Komentarz do opisu eksperymentu ze światłem w pociągu**

Bibliografia

### **1. UWAGI WSTĘPNE**

Trudno sobie wyobrazić naukę bez eksperymentu myślowego. Po pierwsze, towarzyszy on nauce od starożytności,<sup>1</sup> poprzez średniowiecze<sup>2</sup> i nowożytność,<sup>3</sup> do czasów nam współczesnych.<sup>4</sup> Niektórzy utrzymują nawet, że pewne spośród kluczowych dla historii nauki eksperymenty — które zmieniły bieg tej historii — były właśnie eksperymentami myślowymi, przeprowadzonymi w umysłach genialnych

<sup>1</sup> Por. np. Rescher (1991).

<sup>2</sup> Por. np. King (1991).

<sup>3</sup> Por. np. Alanen (1991), George (1991), Lennox (1991) i Wilson (1991). I tak np. George pisze: „It [...] seems that some of the grander conclusions of the philosophers [...], their idealism, or scepticism, are comments or glosses on, or lessons taken from, constructions that we may reasonably describe as thought experiments” (1991: 282).

<sup>4</sup> Por. np. Horowitz (1991), Mohanty (1991), Norton (1991) i Thomason (1991).

fizyków. Po drugie, w niektórych dyscyplinach — np. matematyce<sup>5</sup> i filozofii<sup>6</sup> — nie stosuje się eksperymentów innych niż myślowe.

Stawiamy tu sobie za cel: skonstruowanie definicji „eksperymentu rzeczywistego” i „eksperymentu myślowego”, które w maksymalnym stopniu odpowiadają praktyce naukowej,<sup>7</sup> przedstawienie pewnej typologii eksperymentów rzeczywistych i myślowych oraz wskazanie ich funkcji w naukach przyrodniczych.

Realizując te cele — mamy świadomość ograniczeń, które tę realizację obciążają. Po pierwsze — termin „eksperyment myślowy” jest używany w sposób chwiejny i to nie tylko przez dyletantów, lecz także przez fachowców.<sup>8</sup> Zaproponowana przez nas definicja będzie więc miała rekonstrukcyjny — w sensie Łukasiewicza — charakter. Po drugie — analiza funkcji z konieczności oprze się tylko na kilku przykładach, spostrzeżeniach innych autorów oraz naszych (czasem mglistych) intuicjach.

W naszym tekście zasadniczo zdajemy sprawę z wyników naszej rekonstrukcji (gdziekolwiek dajemy jedynie rozwiązania alternatywne). Uwagi o charakterze historycznym umieszczamy w przypisach.

## 2. EKSPERYMENT RZECZYWISTY

Zakładamy, że pomiędzy pojęciami eksperymentu rzeczywistego i eksperymentu myślowego zachodzi pewna analogia (innymi słowy — że eksperymenty rzeczywiste myślowe należą do wspólnej klasy procedur naukotwórczych). Rozpocniemy w związku z tym od rekonstrukcji pojęcia eksperymentu rzeczywistego.

Przypuśćmy, że osoba *O* zażyła truciznę. To, że osoba *O* zażyła truciznę, jest pewnym stanem rzeczy. Kiedy dwa stany rzeczy są takie, że zajście pierwszego pociąga za sobą zajście drugiego, będziemy mówili wtedy, że drugi jest zależny od pierwszego. Zależność taka jest również pewnym stanem rzeczy — można powiedzieć: stanem rzeczy wyższego rzędu.

Ogólnie rzecz ujmując: zależność taka może mieć m.in. charakter kauzalny, funkcjonalny lub semantyczny (a w szczególności logiczny). Będziemy mówili wtedy, że drugi stan rzeczy jest odpowiednio efektem, korelatem lub konsekwencją — a ogólnie: rezultatem — pierwszego stanu rzeczy. Przykładowo: efektem zażycia tru-

<sup>5</sup> Zob. Anapolitanos (1991: 87). Dodajmy jednak, że za rodzaj eksperymentu rzeczywistego można uznać pomiar, który mógł stać u początków uogólniających hipotez geometrycznych.

<sup>6</sup> Mohanty podkreśla, że eksperymenty myślowe zaczęto stosować na szeroką skalę we współczesnej filozofii analitycznej zwłaszcza po upowszechnieniu się pojęcia możliwego świata, podkreślając zarazem (nie jesteśmy pewni, czy słusznie), że możliwość jest w tym wypadku rozumiana jako możliwość logiczna (1991).

<sup>7</sup> Ponieważ są to definicje regulujące względem co najmniej niektórych definicji znanych w literaturze, trudno tu mówić o adekwatności.

<sup>8</sup> Zwraca na to uwagę m.in. Mohanty: „The term «thought experiment» does not [...] have any «ordinary extension», and, even as used in the context of the physical sciences, it has conflicting uses” (1991: 271).

cizny przez osobę  $O$  jest śmierć osoby  $O$ ; korelatem zatrucia u osoby  $O$  jest u osoby  $O$  gorączka; konsekwencją tego, że osoba  $O_1$  jest mężem osoby  $O_2$ , jest to, że osoba  $O_2$  jest żoną osoby  $O_1$ .

Będziemy mówili, że:

Osoba  $O$  obserwuje  $X$ -a, gdy osoba  $O$  uważnie przygląda się  $X$ -owi po to, aby dostrzec części  $X$ -a<sup>9</sup> lub zachodzące w  $X$ -ie zmiany.<sup>10</sup>

Zauważmy, że  $X$  może być w szczególności zbiorem pewnych stanów rzeczy połączonych różnymi zależnościami: obserwować można zarówno zażywanie trucizny, jak i ciąg, którego pierwszym elementem jest zażywanie trucizny, a drugim -- efekt zażycia trucizny, tj. śmierć.

Kiedy mamy do czynienia z rzeczywistym eksperymentem?

Przypuśćmy, że stawiamy osobie  $O_1$  pytanie:

(1) Co jest efektem zażycia trucizny przez osobę  $O_2$ ?

Ogólnie:

(2) Co jest efektem  $X$ -a?

*Datum quaestionis* pytania (3) brzmi:

(3) Efektem  $X$ -a jest  $Y$ .

Jeżeli  $X$  nie był wywołany przez osobę  $O_1$  (gdyż np. osoba  $O_2$  z własnej woli zażyła truciznę -- z punktu widzenia osoby  $O_1$  zażycie trucizny było więc «samoistne», a nie «sztuczne»), to aby odpowiedzieć na pytanie (1) wystarczy, jeżeli osoba  $O_1$  będzie obserwowała efekt  $X$ -a.

Jeśli  $X$  «samoistnie» (a więc bez udziału osoby  $O_1$ ) nie zachodzi -- i jeśli osoba  $O_1$  ma odpowiedzieć na pytanie (2), a nie zna lub nie uznaje prawa, z którego taka odpowiedź by wynikała -- to musi wywołać  $X$ .<sup>11</sup> Właśnie wtedy będziemy mieli do czynienia z eksperymentem rzeczywistym:

<sup>9</sup> Jak to sformułował Mill -- „obserwatorem [...] jest nie ten, kto po prostu widzi rzecz, jaka jest przed jego oczami, lecz ten, kto widzi, z jakich części dana rzecz się składa” (1843: 587).

<sup>10</sup> Tak Podsiad definiuje „obserwację” w potocznym rozumieniu: jako „czynność (lub wynik) uważnego spostrzegania, mającą na celu dokładniejsze poznanie jakiejś rzeczy lub zjawiska” (2000: 576). Za „obserwację” w sensie metodologicznym uważa on natomiast „planowe (o ustalonej kolejności), systematyczne (w określonych odstępach czasu) i selektywne (wybierające elementy ważne) poznawanie jakiegoś fragmentu rzeczywistości, mające niekiedy na celu sprawdzenie postawionej hipotezy” (2000: 576). Podobną ideę wyraził wcześniej Kotarbiński, pisząc, że „obserwowanie jest to postrzeganie planowe” (1929: 352). Nie negując tego, że termin „obserwacja” bywa tak używany, tutaj ograniczamy się do rozumienia nazwanego przez Podsiad „potocznym”. Idziemy tu za Stępnem, według którego „OBSERWACJA to doświadczenie z uwagą, w określonym celu” (2007: 119).

<sup>11</sup> Jak pisze Ajdukiewicz -- z eksperymentem mamy do czynienia, gdy „wpływamy lub próbujemy wpłynąć na naturalny tok zdarzeń lub stanu rzeczy, zmieniając w sposób dowolny i dobrze nam wiadomy warunki, w jakich on przebiega, i to w tym celu, by zaobserwować, czy i w jaki sposób wraz ze zmianą tych warunków zmienił się tok tych zdarzeń” (1965: 229). W ujęciu Sucha: Eksperymentem nazywa się w nauce procedurę doświadczalną, mającą na celu rozstrzygnięcie jakiegoś problemu teoretycznego, w której toku bądź wywołuje się samo badane zjawisko, bądź wpływa -- poprzez modyfikację warunków -- na jego przebieg” (1987: 120). Eksperymenty prze-

(4) Osoba  $O$  eksperymentuje rzeczywiście (przeprowadza eksperyment rzeczywisty) z  $X$ -em, gdy:

(a) osoba  $O$  wywołuje  $X$ -a

i

(b) osoba  $O$  ustala — za pomocą obserwacji — co jest efektem  $X$ -a.<sup>12</sup>

### 3. WARIANTY DEFINICJI „EKSPERYMENTU RZECZYWISTEGO”

Chociaż za punkt wyjścia do dalszych rozważań przyjmiemy definicję (4), rozważymy pewne definicje alternatywne.

Przypuśćmy, że  $X$  jest zażyciem chininy przez osobę  $O_2$ . Eksperyment z zażyciem chininy przez osobę  $O_2$  polegałby więc na wywołaniu zażycia chininy przez osobę  $O_2$  — czyli na podaniu chininy osobie  $O_2$  — i uważnym przyglądaniu się, co jest efektem podania chininy osobie  $O_2$ .

Rzecz jasna w eksperymencie takim mogłoby chodzić o osobę zarażoną malarią. Sytuację tę opisuje się bądź mówiąc, że chodzi o zażycie chininy przez osobę  $O_2$ , która jest zarażona malarią — bądź mówiąc, że chodzi o osobę  $O_2$  (znajdującą się) w warunkach zarażenia malarią.

Przy tym drugim ujęciu definicja „eksperymentu rzeczywistego” przybrałaby postać:

(1) Osoba  $O$  eksperymentuje rzeczywiście (przeprowadza eksperyment rzeczywisty) z  $X$ -em, gdy są takie warunki  $W$ , że:

(a) osoba  $O$  wywołuje  $X$ -a w warunkach  $W$ .<sup>13</sup>

i

ciwstawia się przy tym obserwacji, zwracając uwagę na „czynny” charakter eksperymentu w porównaniu z bierną obserwacją oraz jego „teoretyczne ukierunkowanie”. Jak pisze Grobler: „Eksperyment nie jest prostą obserwacją tego, co doświadczenie przynosi. Eksperyment polega na zaaranżowaniu sytuacji stwarzającej okazję do obserwacji, która bez planowego działania uczonego mogłaby się nie nadarzyć. Ujmując rzecz nieco frywolnie, eksperyment różni się od obserwacji tym, czym przesłuchiwanie od podsłuchiwanie w śledztwie prowadzonym w sprawie przyrody” (2006: 24). Granica między eksperymentem a obserwacją nie jest jednak wyraźna: wskazuje się m.in. na fakt, iż każda obserwacja jest w pewnym stopniu uteoretyzowana, a pojęcie ingerencji w przebieg zdarzeń jest nieostre. Tutaj tego zagadnienia nie rozwijamy, zwłaszcza że nie ma ono zasadniczego znaczenia dla omawianych kwestii.

<sup>12</sup> Jak piszą Nowaczyk i Żołnowski — „eksperymentowanie [...] służy przede wszystkim stwierdzaniu zależności między zjawiskami, z którym przynajmniej jedno zostało celowo spowodowane przez badacza” (1974: 152).

<sup>13</sup> Twardowski utożsamiał eksperyment z działaniem opisanym w (a). Pisał: „Dowolne wywołanie zjawiska w warunkach dowolnych w celu przeprowadzenia nad nim obserwacji nazywa się «eksperymentem»” (1901: 90). Podobnie stawia sprawę Hajduk, pisząc, że eksperyment jest „dowolnie powtarzalnym wywoływaniem zjawiska (zdarzenia, stanu rzeczy) przyrodniczego, psychicznego, społecznego” (2002: 83). Takie rozumienie „eksperymentu” odnotowuje również — ale z dezaprobatą — Kotarbiński, pisząc: „Częstokroć [...] przez eksperyment rozumie się wszelkie wytworzenie «sztucznych» warunków obserwacji wraz z samą odpowiednią obserwacją” (1929: 352).

(b) osoba  $O$  ustala — za pomocą obserwacji — co jest efektem  $X$ -a w warunkach  $W$ .<sup>14</sup>

Obserwacja dotyczy zatem w tym wypadku nie tylko efektu  $X$ -a, lecz także warunków  $W$ . Dla prawomocności eksperymentu może bowiem być istotne, czy warunki  $W$  pozostają niezmiennione w ciągu eksperymentu i czy nie wpływają na efekt  $X$ -a.

Zauważmy, że  $X$  — np. wspomniane zażycie chininy przez osobę  $O_2$  — można opisać również jako wpływanie na zdrowie osoby  $O_2$  lub jako zmianę osoby  $O_2$  (z niemającej chininy — na mającą chininę w organizmie). Dlatego terminu „eksperyment rzeczywisty” używa się czasem w sposób, który rejestrują definicje sprowadzalne do postaci następujących:

(2) Osoba  $O$  eksperymentuje rzeczywiście (przeprowadza eksperyment rzeczywisty) z  $X$ -em, gdy są takie warunki  $W$ , że:

(a) osoba  $O$  wpływa na stan  $X$ -a w warunkach  $W$

i

(b) osoba  $O$  ustala — za pomocą obserwacji — co jest efektem wpływania na stan  $X$ -a w warunkach  $W$ .<sup>15</sup>

(3) Osoba  $O$  eksperymentuje rzeczywiście (przeprowadza eksperyment rzeczywisty) z  $X$ -em, gdy:

(a) osoba  $O$  wywołuje zmianę w  $X$ -ie

i

(b) osoba  $O$  ustala — za pomocą obserwacji — co jest efektem zmiany w  $X$ -ie.<sup>16</sup>

<sup>14</sup> Według Kotarbińskiego „eksperyment [...] to tyle, co zabieg polegający na wywołaniu czegoś w takich właśnie, a nie innych warunkach, po to, by można było zaobserwować, czy w tych warunkach towarzyszy temu czemuś coś takiego a takiego” (1963: 352). Sformułowanie to ma dwie wady: (1) definicja jest dystynktywna (tj. *definiendum* jest identyczne z definiowanym terminem „eksperyment”), co sugeruje pełną wymienialność terminu „eksperyment” na *definiens*, podczas gdy w rzeczywistości taka wymienialność nie ma miejsc; (2) definicja jest za wąska, gdyż ogranicza eksperymenty do eksperymentów testujących.

<sup>15</sup> Tak można sparafrazować alternatywną definicję Hajduka, zgodnie z którą eksperyment polega na „wpływniu na [...] przebieg [zjawiska] w kontrolowalnych warunkach i poddawanie go systematycznej obserwacji” (Hajduk 2002: 83). Podobnie jest u Irvine’a: „Typically, a physical experiment can be defined as a designed intervention in nature or the laboratory whose observable consequences serve to test some hypothesis previously developed in light of both past observations and a particular theoretical context. Typically, too, the experiment is carried out under controlled, and hence reproducible, conditions. These conditions are associated with particular events or states of affairs from which more general conclusions regarding the nature of the world can then be generated. The observations made then often lead to a supplementing or revising of the original theoretical context” (1991: 151).

<sup>16</sup> Jak pisze Kotarbiński — „niektórzy [...] uważają za znamienne przy eksperymentowaniu wprowadzanie zmian w czymś jednym i badanie zmian towarzyszących temu w czymś innym lub w tym samym przedmiocie pod innym względem” (1929: 353). Tak ujmują eksperymentowanie m.in. Nowaczyk i Żołnowski: „Eksperyment jest badaniem układu celowo zmienianego; wpływamy celowo na tok zdarzeń lub stan rzeczy, by móc obserwować to, co jest skutkiem spowodowanych zmian. Skutkiem eksperymentu może być zjawisko nowego zjawiska lub zmiana przebiegu



Ujęcia powyższe uważamy za równoważne, tyle że są one w różny sposób uschematyzowane.

Ujęciem nierównoważnym z powyższymi jest uznanie eksperymentu rzeczywistego za odmianę obserwacji. W tym wypadku stawiamy sobie nie pytanie (2) z § 2, lecz pytanie:

(4) Jak przebiega zajście  $X$ -a?

Wtedy mamy:

(5) Osoba  $O$  eksperymentuje rzeczywiście (przeprowadza eksperyment rzeczywisty) z  $X$ -em, gdy:

(a) osoba  $O$  wywołuje zajście  $X$ -a

i

(b) osoba  $O$  ustala — za pomocą obserwacji — jak przebiega zajście  $X$ -a.<sup>17</sup>

Zauważmy, że w odróżnieniu od definicji (4) z § 2 chodzi tu o ustalenie *przebiegu*  $X$ -a, który w tym wypadku jest pewnym procesem.

Eksperyment rzeczywisty, w takim ujęciu, byłby więc po prostu obserwacją w sensie definicji z § 2 — tyle że obserwacją  $X$ -a wywołanego przez osobę  $O$ .

Warto zauważyć, że szczególnym przypadkiem  $X$ -a może być ciąg przyczynowo-skutkowy — np. zażycie trucizny przez osobę  $O_2$  i następująca po nim śmierć osoby  $O_2$ . Wtedy poprzednie ujęcia — (4) z § 2 i (1)-(3) z § 3 — byłyby szczególnymi przypadkami ujęcia (2) z tego paragrafu.

Jako tło dalszych rozważań przyjmijmy ujęcie (4) z § 2.

#### 4. TYPY EKSPERYMENTÓW RZECZYWISTYCH

W zależności od pewnych szczegółów — można wyodrębnić wśród eksperymentów następujące typy:<sup>18</sup>

zjawisk już istniejących" (1974: 153). Tak też stawia sprawę np. Podsiad w swojej definicji „eksperymentu”. Jest to — według niego — „zabieg badawczy polegający na wywołaniu określonego zjawiska w określonych warunkach, w celu stwierdzenia związków, np. przyczynowych, między zmianami warunków (*zmienna niezależna*) i zmianami zjawiska (*zmienna zależna*)” (2000: 211). Uszczegółowienie to nie jest skądinąd zgodne z jego definicją „eksperymentu diagnostycznego”, tj. „takiego, który dzięki nowym obserwacjom pozwala bliżej określić pewne cechy badanego przedmiotu” (2000: 211-212).

<sup>17</sup> Horowitz i Massey piszą wprost: „Observation itself can be passive and even unplanned, or active and artfully contrived. Observation of this second, deliberate sort is commonly known as *experimentation*” (1991: 1). To, że niekiedy eksperyment utożsamia się „obserwacją czynną (sprowokowaną)”, sugeruje Hajduk (2002: 83). Tak jest np. u Wojtysiaka, według którego eksperyment jest to „obserwacja zjawiska celowo wywołanego przez obserwatora w sytuacji zaplanowanej” (2007: 344). Wojtysiak idzie tu zapewne za Słepniem, który pisze jeszcze dobitniej: „Eksperyment [...] to obserwacje (a więc spostrzeżenia) określonego stanu, zdarzenia lub procesu, wywołanego przez eksperymentatora (ewentualnie przy użyciu odpowiedniej aparatury) w sytuacji zaplanowanej” (2007: 78).

<sup>18</sup> Nie podejmujemy się przeprowadzenia klasyfikacji eksperymentów rzeczywistych, która spełniałaby tradycyjne kryteria jej logicznej poprawności.

(a) eksperymenty teoretyczne (których celem jest heureka, test lub selekcja — zob. w tej sprawie niżej) i dydaktyczne (których celem jest tylko demonstracja odpowiednich procedur<sup>19</sup>);

(b) eksperymenty orientacyjne (próbne), konstytutywne (główne) i kontrolne (w stosunku do konstytutywnych?);<sup>20</sup>

(c) eksperymenty autentyczne (przeprowadzane na obiektach «oryginalnych») i modelowane (przeprowadzane na modelach owych obiektów<sup>21</sup>);

(d) eksperymenty laboratoryjne (które przebiegają w warunkach laboratoryjnych) i naturalne (które przebiegają w warunkach naturalnych — zmodyfikowanych jedynie interwencją eksperymentatora polegającą na wywołaniu danego stanu rzeczy);<sup>22</sup>

(e) eksperymenty heurystyczne (mające na celu ustalenie, jakie są rezultaty pewnych stanów rzeczy);<sup>23</sup> testujące (mające na celu sprawdzenie wysuniętej hipotezy)<sup>24</sup> i selekcyjne (mające na celu wybór jednej spośród konkurencyjnych hipotez);<sup>25</sup>

(f) eksperymenty konstruktywne (tj. eksperymenty testujące mające potwierdzić testowaną hipotezę)<sup>26</sup> i destruktywne (tj. mające ją obalić).<sup>27</sup>

Zauważmy, że typologie (b)–(f) są względem typologii (a) niezależne. Niezależne względem siebie są także typologie (b) i (c), (b) i (d), (b) i (e), (c) i (e). Jeśli chodzi

<sup>19</sup> Zob. np. Hajduk (2002: 83). Eksperymenty dydaktyczne przeprowadza się jedynie w celu zobrazowania teorii (np. na szkolnych zajęciach fizyki). Różni je od eksperymentów teoretycznych duży stopień subiektywnej pewności co do uzyskania przewidzianego rezultatu. Przypomina to sytuację ćwiczenia polegającego na dowodzeniu skądinąd dowiedzionych już też jakiejś teorii aksjomatyzowanej.

<sup>20</sup> Por. Hajduk (2002: 83).

<sup>21</sup> Jako przykład takiego „zastępczego” eksperymentu — Hajduk podaje — „symulacje komputerowe” (2002: 83).

<sup>22</sup> Jak pisze Hajduk — eksperyment laboratoryjny „przebiega w warunkach kontrolowalnych przez eksperymentatora”, natomiast naturalny — „w warunkach możliwie najmniej modyfikowanych, zaś dokonywane obserwacje dotyczą zmian, o ile posiadają one wyraźne przyczyny oraz klarowny przebieg” (2002: 83).

<sup>23</sup> Według Hajduka chodzi w nich o ustalenie „nowych faktów o własnościach badanych obiektów oraz zależności między nimi” (2002: 83).

<sup>24</sup> Według Hajduka — eksperyment testujący (czyli sprawdzający) — to eksperyment „pozwalający ustalić nowe fakty o własnościach badanych obiektów oraz zależności między nimi” (2002: 83).

<sup>25</sup> Nowaczyk i Żołądowski piszą, że eksperymenty takie są „podejmowane wtedy, gdy chodzi o uzasadnienie tezy, że z dwóch tych możliwych odpowiedzi na dane zagadnienie pierwsza a nie druga (lub odwrotnie) jest odpowiedzią prawdziwą” (1974: 153). W ujęciu Hajduka — eksperyment selekcyjny lub inaczej rozstrzygający, krzyżowy (*experimentum crucis*) „pozwala rozstrzygnąć na rzecz jednej z konkurencyjnych hipotez” (2002: 38). Podobnie charakteryzuje go Wojtyśiak — jako eksperyment „rozstrzygający o wyborze jednego z dwóch konkurencyjnych rozwiązań określonego problemu” (2007: 344).

<sup>26</sup> Jak pisze Kotarbiński — „robiąc [taki] eksperyment [...] idzie nam o poparcie [pewnej] tezy” (1929: 353).

<sup>27</sup> W wypadku takich eksperymentów — „idzie nam o to, by [pewną] tezę odrzucić” (Kotarbiński 1929: 353).

o typologii (b) i (e), to eksperymenty orientacyjne i konstytutywne mogą być zarówno heurystyczne, jak i testujące lub selekcyjne, a eksperymenty kontrolne można uważać za rodzaj eksperymentów testujących lub selekcyjnych (w odniesieniu do rezultatów odpowiednich eksperymentów orientacyjnych i konstytutywnych). Natomiast w wypadku typologii (c) i (d) — eksperymenty modelowane są typem eksperymentów laboratoryjnych. Wreszcie w wypadku typologii (e) i (f) — ta druga jest oczywiście typologią eksperymentów testujących, a nie klasy eksperymentów wziętej *in toto*.

Ze względu na wagę pojęć rzeczywistego eksperymentu heurystycznego i testującego<sup>28</sup> — dla rekonstrukcji pojęcia eksperymentu myślowego — zajmiemy się teraz krótko tymi pojęciami.

## 5. RZECZYWISTY EKSPERYMENT HEURYSTYCZNY A RZECZYWISTY EKSPERYMENT TESTUJĄCY

Z rzeczywistym eksperymentem heurystycznym — przypomnijmy — mamy do czynienia wtedy, gdy chcemy ustalić, co jest efektem stanu rzeczy określonego przez nas typu. Dokładniej:

(1) Jeśli osoba  $O$  chce ustalić w drodze eksperymentu rzeczywistego odpowiedź na pytanie:

(a)  $\wedge x (Px \Rightarrow ?x)$ ,<sup>29</sup>

to osoba  $O$  powinna dla pewnego  $a$  wywołać to, że:

(b)  $Pa$ ,

i ustalić — w drodze obserwacji — przy jakim  $X$  jest tak, że:

(c)  $Pa \Rightarrow Xa$ .

Od razu zaznaczmy, że nie jest to łatwe zadanie — m.in. dlatego, że:

(a) związek pociągania nie jest (nigdy? na ogół?) spostrzegalny;

(b) niezbędnym warunkiem tego, że dla pewnego  $X$  zachodzi (c), jest to, że to, że  $Xa$ , jest późniejsze od tego, że  $Pa$ ;

(c) to, że  $Pa$ , pociąga (zawsze? na ogół?) więcej niż jeden stan rzeczy.

Jest jasne, że zaryzykowanie hipotezy, że to  $Q$  jest owym poszukiwanym  $X$ , a więc, że odpowiedź na pytanie (a) brzmi:

(d)  $\wedge x (Px \Rightarrow Qx)$ ,

wymaga zaobserwowania, że dla wielu przedmiotów zachodzi (d). Zwykle dopiero taki ciąg pojedynczych eksperymentów uważa się za eksperyment heurystyczny.

Rozważmy z kolei strukturę eksperymentu testującego.

<sup>28</sup> Pomijamy sprawę eksperymentu selekcyjnego, gdyż da się on zinterpretować w kategoriach eksperymentu testującego.

<sup>29</sup> Jest to symboliczny zapis pytania, którego quasi-naturalnojęzykowy odpowiednik brzmi: „Jaki stan przedmiotu należącego do pewnej określonej klasy jest rezultatem tego, że przedmiot  $x$  jest  $P$ ?”



Z rzeczywistym eksperymentem testującym mamy do czynienia wtedy, gdy chcemy sprawdzić pewną wysuniętą uprzednio hipotezę. Dokładniej:

(2) Jeśli osoba  $O$  chce sprawdzić w drodze eksperymentu rzeczywistego, czy:

(a)  $\wedge x (Px \Rightarrow Qx)$ ,

to osoba  $O$  powinna dla pewnego  $a$  wywołać to, że:

(b)  $Pa$ ,

i ustalić — w drodze obserwacji — czy:

(c)  $Pa \Rightarrow Qa$ .

Jeśli okaże się, że  $Pa \Rightarrow Qa$ , to osoba  $O$  potwierdziła (w pewnym stopniu) hipotezę (a). Jeśli natomiast okaże się, że nieprawda, że  $Pa \Rightarrow Qa$ , to osoba  $O$  obaliła hipotezę (a).

## 6. EKSPERYMENT MYŚLOWY

W jakim sensie przeprowadzenie eksperymentu myślowego jest analogonem przeprowadzenia eksperymentu rzeczywistego?

Otóż odpowiednikiem wywołania pewnego stanu rzeczy w eksperymentcie rzeczywistym jest w eksperymentcie myślowym *założenie* lub *wyobrażenie sobie* zajścia pewnego stanu rzeczy (*scil.* że pewien stan rzeczy zachodzi)<sup>30</sup> W braku lepszego terminu za *genus proximum* założenia i wyobrażenia sobie czegoś przyjmiemy przedstawienie sobie czegoś.

Natomiast odpowiednikiem obserwacji rezultatu eksperymentu jest w wypadku eksperymentów myślowych odpowiednie rozumowanie.

Powiemy więc — parafrazując odpowiednio definicję „eksperymentu rzeczywistego” z § 2 — że:

Osoba  $O$  eksperymentuje myślowo (przeprowadza eksperyment myślowy) z  $X$ -em, gdy:

(a) osoba  $O$  przedstawia sobie  $X$ -a

i

(b) osoba  $O$  ustala — za pomocą rozumowania — co jest efektem  $X$ -a.

Nazwijmy  $X$ -a „bazą eksperymentu myślowego”, a zależność między  $X$ -em a efektem  $X$ -a „rezultatem eksperymentu myślowego”.

Kluczowe jest tutaj, na jakiej podstawie osoba  $O$  z przedstawienia sobie  $X$ -a wprowadza przedstawienie sobie efektu  $X$ -a — a więc na jakiej podstawie ustala re-

<sup>30</sup> Tak stawia sprawę m.in. Brown (1991a: 1) — pisząc, że eksperyment myślowy to coś, co przypomina eksperyment empiryczny, tyle że odbywa się w „laboratorium naszego mózgu”. Hajduk z kolei pisze, że eksperyment myślowy — to eksperyment „odbywający się bez jakichkolwiek manipulacji fizycznych (dotyczy sytuacji, która może mieć miejsce, jeśli zostanie zrealizowany proponowany scenariusz zdarzeń)” (2002: 83). Tak jest też u Wojtyśiaka, według którego eksperyment myślowy polega na „pojęciowym lub wyobrażeniowym przedstawieniu i analizie pewnej możliwej lub fikcyjnej sytuacji” (2007: 344).

zultat swojego eksperymentu myślowego. Innymi słowy — co przesądza o tym, że osoba przeprowadzająca eksperyment przedstawia sobie jako rezultat bazy taki, a nie inny stan rzeczy? Co steruje jego umysłem, a w szczególności wyobraźnią? Analiza wielu procedur określanych mianem „eksperymentu myślowego” każe nam sądzić, że podstawa ta musi być jakieś uznane już, w każdym razie przez osobę przeprowadzającą eksperyment myślowy, prawo lub ciąg praw, z których w koniunkcji ze zdaniem stwierdzającym bazę wynika zdanie stwierdzające rezultat eksperymentu myślowego.

Będziemy nazywać to prawo — *resp.* te prawa — „horyzontem eksperymentu myślowego”.

## 7. TYPY EKSPERYMENTÓW MYŚLOWYCH

Wśród eksperymentów myślowych da się wyróżnić typy podobne do tych, które wyodrębniliśmy wśród eksperymentów rzeczywistych (z wyjątkiem pary: eksperymenty laboratoryjne i naturalne); można więc przeprowadzać eksperymenty myślowe: (a) teoretyczne i dydaktyczne; (b) orientacyjne, konstytutywne i kontrolne; (c) autentyczne i modelowane; (d) heurystyczne, testujące i selekcyjne;<sup>31</sup> (e) konstruktywne i destruktywne.<sup>32</sup>

Poza tym są wśród nich typy specyficzne, a mianowicie:

(f) eksperymenty supozycyjne, w których baza jest założona,<sup>33</sup> i eksperymenty imaginacyjne, w których baza jest wyobrażona;<sup>34</sup>

<sup>31</sup> Niewykluczone, że poza testującymi (i selekcyjnymi) — eksperymenty myślowe miewają cele perswazyjne. O takich eksperymentach perswazyjnych w lingwistyce pisze Thomason (1991: 247-248). Zauważmy przy okazji, że odwoływanie się do intuicji językowych przy testowaniu hipotez lingwistycznych — to w istocie przeprowadzanie eksperymentów rzeczywistych.

<sup>32</sup> Brown jako przykład pierwszych podaje eksperyment Newtona z obracającym się czerpakiem, wykorzystywany dla poparcia Newtonowskiej teorii absolutnej przestrzeni i czasu, a drugich — z tzw. kotem Schrödingera, wykorzystywany dla obalenia kopenhaskiej interpretacji mechaniki kwantowej. Brown zwraca uwagę, że niektóre eksperymenty myślowe służą zarazem do potwierdzenia jednej i obalenia drugiej, konkurencyjnej hipotezy; jako przykład podaje eksperyment z kulą armatnią i muszkietową Galileusza, wykorzystywany zarazem dla potwierdzenia teorii Galileusza i dla obalenia teorii Arystotelesa (Brown 1991b).

<sup>33</sup> Rescher utożsamia myślowe eksperymenty supozycyjne z myślowymi eksperymentami *tout court*: „A „thought experiment” is an attempt to draw instruction from a process of hypothetical reasoning that proceeds by eliciting the consequences of a hypothesis which, for aught that one actually knows to the contrary, may well be false. It consists in reasoning from a supposition that is not accepted as true — perhaps is even known to be false — but is assumed provisionally in the interests of making a point of resolving a conclusion” (1991: 31). Podobnie jest u Nortona, który tak charakteryzuje warunki niezbędne bycia eksperymentem myślowym: „Thought experiments [w fizyce] are arguments which: (i) posit hypothetical or counterfactual states of affairs [inaczej nie byłby eksperymentem myślowym], and (ii) invoke particulars irrelevant [pod większością względów?] to the generality of the conclusion [inaczej nie byłby eksperymentem]” (1991: 129). Nieco inaczej ujmują

(g) eksperymenty faktualne, tj. takie, w których baza jest czymś, co nie zachodzi, ale może zajść w rzeczywistości,<sup>35</sup> i kontrfaktualne, tj. takie, które nie mogą stać się eksperymentami rzeczywistymi.<sup>36</sup>

Te ostatnie dwie pary typów — (f) i (g) — wymagają trzech komentarzy.

Po pierwsze — eksperymenty, zwane przez niektórych „eksperymentami introspekcyjnymi”, są w istocie eksperymentami rzeczywistymi, tylko dotyczącymi psychicznych stanów rzeczy.<sup>37</sup>

to Lennox: „Thought experiments are: [a] tests of a theory's explanatory potential which [b] posit hypothetical or counterfactual test conditions and [c] invoke particulars which are irrelevant to the generality of the theory, and which [d] are selected to instantiate feature of the theory under special consideration” (1991: 236). Według Nortona cechą charakterystyczną eksperymentów myślowych jest to, że są one zasadniczo eliminowalne (1991: 131).

<sup>34</sup> Eksperyment imaginacyjny ma na myśli Kotarbiński, kiedy pisze, że w wypadku eksperymentu myślowego ktoś „uprzytamnia [...] sobie tylko, co by było przy zmianie pewnych okoliczności”, a więc „wprawdzie tylko w myśli — wprowadza się [...] zmiany i bada, co im towarzyszy” (Kotarbiński 1929: 353); podobnie Podsiad uważa eksperyment myślowy za „uprzytomnienie sobie ewentualnego przebiegu eksperymentu rzeczywistego i wysnuwanie na tej podstawie odpowiednich wniosków” (2000: 212). Podsiad dodaje że „postępowanie takie nie ma wprawdzie wartości dowodowej, ale może mieć wartość heurystyczną” (2000: 212).

<sup>35</sup> Czasem uważa się, że właściwymi eksperymentami myślowymi są eksperymenty zasadniczo faktualne. Janis np. pisze: „There are two characteristics of a thought experiment [...] [in physics]. First, a thought experiment as a description of an experimental procedure as well as its outcome or possible outcomes. (In referring to an experimental procedure, I do not mean to imply that the experiment could actually be carried out. The reasons could range from technological difficulties that could, in principle be overcome to difficulties of principle; for example, it may be desirable in a thought experiment to suppose that the universe were quite different from the way it actually is.) Second, the outcome or possible outcomes must be deduced by reasoning consistent with a given theoretical framework” (1991: 113). Podobne jest stanowisko Browna: „We recognize them [*scil.* thought experiments] when we see them: they are visualizable; they involve mental manipulation; they are not the mere consequence of a theory-base calculation; they are often (but not always) impossible to implement as real experiments, either because we lack the relevant technology or because they are simply impossible in principle” (1991b: 122). Za dopuszczeniem uznania eksperymentów faktualnych za rodzaj eksperymentów myślowych — a może nawet za ich utożsamienie — opowiada się także Laymon (1991).

<sup>36</sup> Do nieurzeczywistnialnych eksperymentów myślowych należy tzw. analiza imaginatywna — tak jak ją charakteryzuje Szubka: „ANALIZA IMAGINATYWNA przyjmuje postać wyobrażenia lub przedstawiana sobie, jak stosowalibyśmy nasze pojęcia, kategorie i rozróżnienia w sytuacjach innych od tych znanych z rzeczywistego świata. Wydaje się ona atrakcyjnym sposobem ustalania stałych, nieprzygodnych związków i wyszukiwania rozmaitych kontrprzykładów dla zastanych rozwiązań. Wykorzystanie jej do ustalania tez metafizycznych wymaga założenia prawomocności przejścia od tego, co pojmowalne (lub niepojmowalne) do tego, co metafizycznie możliwe (lub niemożliwe)” (2000: 207). *Nb.* Eksperymenty kontrfaktualne opisywane bywają w literaturze jako paradoksy, gdyż mają strukturę podobną do rozumowań antynomialnych.

<sup>37</sup> Chodzi np. o sprawdzanie hipotez lingwistycznych, a w szczególności gramatycznych czy semantycznych, przez odwoływanie się do odpowiednich intuicji użytkowników danego języka.

Po drugie — kontrfaktualność może mieć charakter ontyczny, ale także techniczny i etyczny. Bywa, że pewien stan rzeczy nie może zajść, gdyż tak jest urządzony świat, ale bywa, że dlatego, iż my (jeszcze) nie możemy go wywołać — albo uważamy, że nie powinien zostać wywołany, gdyż byłoby to pogwałcenie jakichś norm moralnych.<sup>38</sup>

Po trzecie — wielu badaczy stoi na stanowisku, że termin „eksperyment myślowy” należy zarezerwować wyłącznie dla eksperymentów imaginacyjnych i kontrfaktualnych.<sup>39</sup> Uzasadnia się to tym, że tzw. eksperymenty supozycyjne są jedynie pewnymi szczególnymi przypadkami «zwykłych» rozumowań,<sup>40</sup> a eksperymenty faktualne — to nic innego, jak «projekty» eksperymentów rzeczywistych.<sup>41</sup> Nie chcemy jednak tutaj tej sprawy przesądzać.

Powiedzmy jednak wyraźnie, jakich procedur za eksperymenty myślowe nie uważamy.

Jako odmianę eksperymentów myślowych wymienia się np. tzw. eksperymenty eksplanacyjne.<sup>42</sup> Ich opis pozwala utożsamić je z zakładaniem prawdziwości pewnego zdania, gdy nie jest ono jawnie fałszywe, a stanowi oczywistą rację pewnego innego zdania, dla którego jeszcze nie mamy «lepszego» racji.<sup>43</sup> Jest to zatem nic innego

Por. w tej sprawie np. Gale (1991) i Thomason (1991).

<sup>38</sup> O tym ostatnim wypadku wspomina m.in. Thomason (1991: 250).

<sup>39</sup> Do obrońców stanowiska, że każdy eksperyment myślowy jest eksperymentem kontrfaktualnym (lub uważanym za taki przez eksperymentatora), należy Rescher: „This is in fact only one, particularly strong form of thought experiment. When the detective reasons, «Now suppose that the buller did it...», at some early stage of the investigation, his reasoning is clearly not unraveled as a thought experiment if it eventually turns out that he indeed did so” (1991: 40-41). Podobne stanowisko zajmuje m.in. Mohanty, który pisze: „A genuine thought-experiment — if our talk about thought-experiments is to be significant — must be a process which cannot be reiterated physically” (1991: 263). Warunek ten spełniają — według niego — eksperymenty myślowe Kanta z *Krytyki czystego rozumu* oraz fenomenologiczna wariacja imaginacyjna (ejdetyczna). Przy okazji Mohanty ustosunkowuje się do zarzutów kierowanych wobec metody wariacji imaginacyjnej, krytykowanej za *circulus* eksplikacyjny. Krytycy tej metody zwracają uwagę na to, że aby ustalić na podstawie analizy egzemplifikacji *X-a*, co jest istotą *X-a* (przez ustalenie, jakich własności można tę egzemplifikację pozbawić bez utraty tego, że przestanie być taką egzemplifikacją), trzeba już przedtem wiedzieć, co jest istotą *X-a*, gdyż bez tego nie da się rozstrzygnąć, czy analizowany przedmiot jest rzeczywiście egzemplifikacją *X-a*. Mohanty zwraca uwagę na to, że na zarzut taki (lub — dodajmy — zarzut regresu *ad infinitum*) narażona jest każda procedura eksplikacyjna. Innego zdania w omawianej sprawie jest Irvine, który zalicza do eksperymentów myślowych zarówno eksperymenty kontrfaktualne, jak i faktualne (1991: 150-151). Podobnie chyba jest u Nortona (1991).

<sup>40</sup> Por. też np. Norton (2004). Za rodzaj rozumowania — jak się zdaje — uznaje eksperyment myślowy wielu matematyków. Por. Anapolitanos (1991: 94).

<sup>41</sup> Z tego zapewne powodu Thomason nazywa je „eksperymentami inicjującymi” (1991).

<sup>42</sup> Przykłady eksperymentów eksplanacyjnych, apagogicznych (*negatively demonstrative experiments* lub *refutatory experiments*), przez *reductio ad absurdum*, sceptycznych, *per analogiam* i *a contrario* — stosowanych przez presokratyków — podaje Rescher (1991).

<sup>43</sup> W sformułowaniu Reschera wygląda to tak: „*X* is hard to account for, but if we assume that *P*,



jak wyjaśnianie, czyli rozumowanie, w którym dla uznanego zdania wskazujemy hipotetyczną rację.

Inną odmianą eksperymentów myślowych miałyby być tzw. eksperymenty apagogiczne i przez *reductio ad absurdum*. Okazuje się, że można je utożsamić z wykazaniem prawdziwości pewnego zdania poprzez wyprowadzenie fałszu — lub odpowiednio fałszu logicznego — z założenia o fałszywości owego zdania. Jest to zatem nic innego jak dowodzenie.<sup>44</sup>

Coś podobnego można powiedzieć o tzw. eksperymentach *per analogiam* i *a contrario*. Przypominają one — sądząc znowu z opisu — wnioskowania zawodne funkcjonujące jako znane chwytty erystyczne: z tego, że *A* ma cechę *C*, a *B* jest podobne do *A*, wnioskuje się, że i *B* ma cechę *C*; z kolei z założenia, że nic nie ma cechy *C* (np. nie jest gorące albo nie jest niczego przyczyną), wnioskuje się, że nic nie mogłoby też mieć korelatywnej cechy *D* (w tym wypadku nie mogłoby być zimne albo odpowiednio nie mogłoby być niczego skutkiem), a tymczasem cecha *D* pewnym przedmiotem przysługuje — zatem musi też czemuś przysługiwać cecha *C*.

## 8. MYŚLOWY EKSPERYMENT HEURYSTYCZNY I TESTUJĄCY

Rozważmy teraz, jaką strukturę mają myślowe odpowiedniki heurystycznych i testujących eksperymentów rzeczywistych.

(1) Jeśli osoba *O* chce ustalić w drodze eksperymentu myślowego odpowiedź na pytanie:

(a)  $\wedge x (Px \Rightarrow ?x)$ ,

to osoba *O* powinna dla pewnego *a* przedstawić sobie, że:

(b)  $Pa$ ,

i ustalić — w drodze rozumowania — przy jakim *X* jest tak, że:

(c)  $Pa \Rightarrow Xa$ .<sup>45</sup>

Co jest horyzontem takiego eksperymentu?

Nie może być nim oczywiście prawo o postaci:

(d)  $\wedge x (Px \Rightarrow Qx)$ ,

ponieważ (d) jest po prostu odpowiedzią na (a). Zdanie (d) będziemy nazywać „obiektem eksperymentu myślowego”.

which we certainly don't know but which is not inherently implausible, then we obtain a perfectly good explanation of *X*” (1991: 32).

<sup>44</sup> Mówi się także o tzw. eksperymentach sceptycznych. Uznajemy w nich to, że *p*, oraz że z tego, że *p*, wynika to, że *q* — a zarazem co prawda wydaje się nam, że *q*, ale nie wykluczamy, że nie-*q*; skoro zaś tak — to nie jest też wykluczone, że nie *p*. Jest to — jak widać — nic innego jak wnioskowanie o szczególnych przesłankach.

<sup>45</sup> Szczególny wypadek polega tutaj na tym, że formułę (c) spełnia więcej niż jeden predykat będący podstawieniem '*X*-a'. Rezultatem eksperymentu byłaby więc pewna formuła alternatywna (na którą nakłada się na ogół co najmniej warunek niesprzeczności wewnętrznej).



Horyzontem musi być zatem coś innego nie — obiekt eksperymentu. Najprostszy wypadek polegałby na tym, że horyzont składałby się tutaj np. z dwóch praw:

$$(e) \wedge x (Px \Rightarrow Rx)$$

i

$$(f) \wedge x (Rx \Rightarrow Qx).$$

Na podstawie (b) oraz (c) i (f) (a także założenia, że  $a$  należy do zakresu zmienności zmiennej ' $x$ ') osoba  $O$  może — rzecz jasna — wyprowadzić to, że:

$$(g) Qa.$$

Powstaje jednak pytanie, po co zakładać lub wyobrażać sobie, że (b), skoro to, że:

$$(h) Pa \Rightarrow Qa$$

okazuje się podstawieniem prawa (d), które przecież wynika z koniunkcji (e) i (f)?

Jeśli rekonstrukcja powyższa jest adkwatna, to wyglądałoby na to, że myślowy eksperyment heurystyczny nie odkrywa żadnej zależności — tylko co najwyżej jest jej ilustracją: wskazuje przykład pewnych już znanych zależności.

Trzeba jednak pamiętać, że zachodzą wypadki, w których jest czymś więcej niż tylko taką ilustracją.

Po pierwsze, może być tak, że zależność (d) wynika z horyzontu, ale jest to wynikanie «skomplikowane» (gdyż np. horyzont jest koniunkcją bardzo wielu zdań) i dopiero eksperyment uzmysławia zachodzenie tego wynikania. Eksperyment spełnia wtedy funkcję symplifikacyjną.

Po drugie, może być tak, że horyzont nie należy do tzw. wiedzy powszechnie uznanej, lecz do zestawu hipotez autora eksperymentu; eksperyment ma wtedy na celu przekonanie społeczności uczonych do tej hipotezy. Ma on wtedy wartość jedynie perswazyjną.

Po trzecie, może być tak, że zależność (d) pojawia się w umyśle osoby przeprowadzającej eksperyment w momencie wyobrażenia sobie bazy, które w niepoddający się dyskursywnemu opisowi narzuca «nieodparcie» tej osobie to, że rezultatem eksperymentu jest to, że  $Qa$ . Eksperyment spełnia wtedy funkcję «inkubacyjną».

Jeśli z kolei chodzi o myślowy eksperyment testujący, to zauważmy najpierw, że nie można przedstawić procedury testującej, odwołującej się do eksperymentu myślowego — analogicznie do ujęcia rzeczywistego eksperymentu testującego, a więc formuły z § 7, i powiedzieć że:

(2) Jeśli osoba  $O$  chce sprawdzić w drodze eksperymentu myślowego, czy:

$$(a) \wedge x (Px \Rightarrow Qx),$$

to osoba  $O$  powinna dla pewnego  $a$  przedstawić sobie, że:

$$(b) Pa,$$

i ustalić — w drodze rozumowania — czy:

$$(c) (Pa \Rightarrow Qa).$$

Jest tak dlatego, że osoba  $O$  może — w drodze rozumowania — ustalić, czy (c), bądź na podstawie (a) — co nie wchodzi w rachubę, gdyż (a) ma właśnie podlegać sprawdzeniu — bądź jakichś innych praw (wykluczających np. że w ogóle możliwe jest, że  $Qa$ ).

Schemat sprawdzania w drodze eksperymentu myślowego musi więc mieć postać następującą:

(3) Jeśli osoba  $O$  chce sprawdzić w drodze eksperymentu myślowego, czy:

(a)  $\wedge x (Px \Rightarrow Qx)$ ,<sup>46</sup>

to osoba  $O$  powinna dla pewnego  $a$  przedstawić sobie, że:

(b)  $Pa$ .

Teraz możliwe są dwa typy sytuacji.

(A) Załóżmy, że osoba  $O$  uznaje prawa:

(c)  $\wedge x (Px \Rightarrow Rx)$

i

(d)  $\wedge x (Rx \Rightarrow \sim Qx)$

które — jako horyzont tego eksperymentu — wraz z bazą (b) pociągają:

(e)  $\sim Qa$ .

Podkreślmy raz jeszcze, że sprawdzane prawo (a) do horyzontu tego eksperymentu należeć nie może — jako obiekt tego eksperymentu.

Otóż powstałą sytuację osoba  $O$  może zinterpretować dwojako: albo osoba  $O$  podtrzymuje prawa (c) i (d) — i uznaje, że rezultat eksperymentu falsyfikuje sprawdzane prawo (a); albo osoba  $O$  podtrzymuje prawo (a) — traktując eksperyment za jego weryfikację — i uznaje, że eksperyment falsyfikuje jego horyzont, a więc prawo (c) lub (d).

To, czy osoba przeprowadzająca eksperyment podtrzyma (w całości) horyzont, czy sprawdzane prawo, zależy od tego, czy jako rezultat eksperymentu «narzuci się» jej wyobraźni to, że  $Qx$ , czy też to, że  $\sim Qx$ .

(B) Osoba  $O$  nie uznaje żadnych praw, z których dałoby się wyprowadzić sprawdzane prawo (a) lub jego negacja.

W tej sytuacji osoba  $O$  z (b) nie może wyprowadzić ani tego, że  $Qa$ , ani tego, że  $\sim Qa$  — a więc eksperyment nie nadaje się do sprawdzenia prawa (a).

Zauważmy, że i w tym wypadku — podobnie jak w wypadku myślowego eksperymentu heurystycznego — jego konkluzje można wyprowadzić z pominięciem zakładania lub wyobrażania sobie, że (b), gdyż w wypadku (A) sam horyzont eksperymentu pociąga odrzucenie hipotezy (a), natomiast w wypadku (B) horyzont niczego o hipotezie (a) nie rozstrzyga.

Znowu więc, jeśli rekonstrukcja powyższa jest adekwatna, to myślowy eksperyment testujący stanowi również co najwyżej ilustrację pewnych już znanych zależności albo — pomaga wydobyć «tkwiące» już w teorii, ale nieświadomione zależności.

<sup>46</sup> Warto dodać, że według niektórych sprawdzanie za pomocą eksperymentu rzeczywistego różni się od sprawdzania za pomocą eksperymentu myślowego m.in. tym, że w wypadku pierwszego chodzi o sprawdzenie prawdziwości hipotezy, a w drugim — jej mocy wyjaśniającej (czyli stopnia ogólności). Według Lennexa np.: „Thought experiments are intended as tests, not of the truth of the statements comprising the theory, but of the *explanatory potential* of the theory. They are designed either to display, or to challenge, a theory's *ability* to explain the full range of phenomena it claims for its domain” (1991: 223).

## 9. RZECZYWISTY A MYŚLOWY EKSPERYMENT TESTUJĄCY

Definicje „eksperymentu testującego” — i w wypadku eksperymentu rzeczywistego, i w wypadku eksperymentu myślowego — wymagają kilku komentarzy.

(A) Definicje te są podwójną idealizacją. Po pierwsze — nie uwzględniają one hipotez statystycznych, dla których nie istnieją rezultaty eksperymentów, co do których można mieć pewność, że potwierdzają one lub obalają takie hipotezy, gdyż żadne zdanie o rezultacie eksperymentu nie wynika w tym wypadku ze sprawdzanej hipotezy. Po drugie — faktycznie wykonywane eksperymenty polegają na ogół na wywołaniu lub przedstawieniu sobie nie pojedynczego stanu rzeczy, lecz ciągu takich stanów.

(B) Przeprowadzanie zarówno eksperymentów rzeczywistych, jak i myślowych, to działania celowe. Nie jest eksperymentem rzeczywistym wywołanie np. swobodnego spadania szklanki na ziemię, dopóki nie ma ono na celu potwierdzenia prawa powszechnego ciężenia. Nie jest eksperymentem myślowym wyobrażenie sobie stanu rzeczy polegające np. na tym, że upuszczona szklanka «zawisa» w powietrzu, jeśli nie ma się na celu obalenia prawa powszechnego ciężenia.

(C) W wypadku konfliktu wyników eksperymentu rzeczywistego i myślowego, (prawie?) zawsze rewizji podlega wynik eksperymentu myślowego.<sup>47</sup>

(D) Stany rzeczy wywoływane w eksperymentach rzeczywistych są dostępne intersubiektywnie, podczas gdy bazy eksperymentów myślowych (a w każdym razie stany rzeczy wyobrażone) dostępne są tylko osobie eksperyment przeprowadzającej. Kluczowy okazuje się w tym wypadku opis przeprowadzanego eksperymentu myślowego, dzięki któremu inni mogą przeprowadzić eksperyment podobny.

## 10. EKSPERYMENT UDANY I NIEUDANY

Eksperyment rzeczywisty może się nie udać co najmniej w następujących sytuacjach:<sup>48</sup>

(A) Eksperymentatorowi nie udało się wywołać stanu rzeczy, którego efekt chciał zbadać. Chciał np. zbadać efekt zażycia chininy przez daną osobę, ale nie udało mu się tej osoby do tego skłonić — albo przez pomyłkę podał jej nie chininę, lecz strychninę.

(B) Eksperymentatorowi nie udało się zaobserwować, co jest efektem tego, co wywołał. Podał np. chininę danej osobie, ale osoba ją zaraz potem zwymiotowała lub ktoś ją pozbawił życia.

Kiedy natomiast nie udaje się eksperyment myślowy?

Należy to oddzielnie rozważyć dla eksperymentu supozycyjnego i imaginacyjnego — oraz dla faktualnego i kontrfaktualnego.

<sup>47</sup> Por. Irvine (1991: 161).

<sup>48</sup> Więcej takich sytuacji rozważa Janis (1991: 114, 115).

Analogenicznie jak w wypadku (A) — eksperymentatorowi może nie udać się przedstawić sobie bazy, tj. odpowiednio założyć jej lub sobie wyobrazić. Trzeba przyznać, że o wiele łatwiej jest podać kryteria «(nie)zakładalności» niż kryteria (nie)wyobrażalności.<sup>49</sup> Jest tak być może dlatego, że problem wyobrażalności jest w trudny do uchwycenia sposób powiązany z problemem racjonalności.

Mówimy więc:

(1) Założmy, że *p*.

Czy możemy czegoś nie móc założyć?

Jeśli czegoś uprzednio nie założyliśmy, to właściwie wszystko jest «zakładalne».

W szczególności — jest tak w wypadku, gdy dopuszczamy kontrfaktyczną bazę eksperymentu. Możemy jednak założyć, że baza ta — mimo że jest kontrfaktualna — powinna być możliwie podobna do bazy faktualnej (założony stan rzeczy powinien możliwie blisko przypominać stan rzeczy rzeczywistalny).<sup>50</sup> Może nam np. chodzić o to, abyśmy byli w stanie chociaż «w przybliżeniu» wykorzystać eksperyment do opisu rzeczywistego świata. Wtedy oczywiście nie wszystko jest «zakładalne»: nie wszystko wolno założyć. Nie wolno np. założyć, że:

(2) We wszechświecie jest tylko jeden przedmiot.

choć przy takim założeniu moglibyśmy łatwo wykazać, że nie ma własności relacyjnych.<sup>51</sup>

W wypadku, gdy eksperyment ma charakter faktualny, wolno nam dopuszczać jedynie taką bazę, która mogłaby się urzeczywistnić (w szczególności w przyszłości). Nie wolno byłoby nam wtedy założyć takiego '*p*', dla którego prawdą byłoby, że:

(3) Niemożliwe, że *p*.<sup>52</sup>

Mówimy:

(4) Wyobraźmy sobie, że *p*.

Czy wszystko możemy sobie wyobrazić? Czy jest jednak coś, co jest niemożliwe do wyobrażenia sobie?<sup>53</sup>

Wysuwane są w tym wypadku bądź kryteria ontologiczne, bądź epistemologiczne. Kryteria ontologiczne wskazują następujące zasady:

(5) Niewyobrażalne jest wszystko i tylko to, co jest niemożliwe.

(6) Niewyobrażalne jest wszystko to, co jest niemożliwe.

<sup>49</sup> Jest to o tyle sytuacja niedobra, że — jak się niekiedy sądzi (Massey 1991) — we współczesnej filozofii (zwłaszcza analitycznej) mamy do czynienia przede wszystkim z eksperymentami imaginacyjnymi. Pisze on: „Conceivability appears to furnish then, if not the philosopher's stone, at least a powerful philosophical tool” (1991: 290). I dalej: „The philosophical art is the art of counter-exampling” (1991: 292) „At bottom analyticity and conceivability are the same notion” (1991: 295).

<sup>50</sup> Niektórzy — np. Forge (1991: 216) — uważają, że założenie takie powinien robić każdy eksperymentator.

<sup>51</sup> Podobny przykład podaje Forge (1991).

<sup>52</sup> Jak zauważa Leymon — jeśli eksperyment myślowy ma mieć zastosowanie do świata, to (b) z (1)–(3) w § 8 powinien mieć postać: Możliwe, że *Pa* (1991: 177).

<sup>53</sup> Znamienne, że takie pytania zadawano sobie nawet w odniesieniu do Boga.

(7) Niewyobrażalne jest tylko to, co jest niemożliwe.<sup>54</sup>

Na ogół przy tym za niemożliwe uważa się wszystko i tylko to, czego opis jest wewnętrźnie sprzeczny.<sup>55</sup>

Czasami zasady te osłabia się do postaci «subiektywistycznej»:

(8) Osoba *O* jest nie jest w stanie wyobrazić sobie, że *p*, gdy osoba *O* nie ma podstaw wierzyć, że to, że *p*, jest możliwe.

(9) Jeżeli osoba *O* nie ma podstaw uważać, że to, że *p*, jest możliwe, to osoba *O* nie jest w stanie wyobrazić sobie, że *p*.

(10) Jeżeli osoba *O* nie jest w stanie wyobrazić sobie, że *p*, to osoba *O* nie ma podstaw uważać, że to, że *p*, jest możliwe.<sup>56</sup>

Kryteria epistemologiczne — to po prostu introspekcja:

(11) Niewyobrażalne jest wszystko i tylko to, o czym stwierdzamy na podstawie introspekcji, że nie jesteśmy w stanie sobie wyobrazić.

Odnajdujemy, że zasada (11) ma znamiona błędnego koła oraz potencjalnego źródła relatywizacji (nie)wyobrażalności.

Wydaje się, że o wiele łatwiej jest o operatywne kryteria pozwalające rozstrzygnąć, który z dwóch przedmiotów trudniej sobie wyobrazić. Otóż:

(12) *A* jest trudniej wyobrażalne dla osoby *O* niż *B*, gdy:

(a) *B* jest prostsze od *A*;

(b) *B* jest podobniejsze do tego, co jest osobie *O* już znane.<sup>57</sup>

Pewnego rodzaju analogonem wypadku (B) dla eksperymentów rzeczywistych jest w zakresie eksperymentów myślowych sytuacja, w której eksperymentatorowi nie udało się niczego poprawnie wywnioskować ze zdania stwierdzającego zajście przedstawionego sobie stanu rzeczy zdania stwierdzającego zajście rezultatu — albo jego wnioskowanie okazało się błędne formalnie (*non sequitur*).<sup>58</sup>

<sup>54</sup> Za (5) miał się opowiadać Hume, a za (6) — Descartes (Massey 1991).

<sup>55</sup> Zob. np.: Things are possible or thinkable (conceivable) insofar as their notion or essential descriptions are non-contradictory (Alanen 1991: 75). Na takim poglądzie na (nie)możliwość opierała się koncepcja możliwych światów u Dunsa Szkota: „Because of the relations of compatibility and incompatibility between things or their possible properties and existence, possibilities are portioned into classes of compossible and impossible states of affairs, forming several alternative «possible worlds»” (Alanen 1991: 76).

<sup>56</sup> Por. Hetherington (1991).

<sup>57</sup> Por. Anapolitanos (1991: 94-96).

<sup>58</sup> Jak pisze King: „The obvious question remains: how does one determine what happens in a thought-experiment, which sentence follow from which descriptions — and indeed what sense of «following from» is relevant here” (1991: 51). Za błąd, który występuje szczególnie często, Camp uważa rozumowanie według następującego schematu: „An object of thought *x* is contemplated in a context so impoverished that some one feature *F* is necessary for retaining *x* as an object of thought. The thought-experimenter then concludes that *x* must be *F*, and goes on to interpret this to mean that it is absolutely certain that *x* is *F*” (Horowitz & Massey 1991: 28).



## 11. EKSPERYMENTY MYŚLOWE W FIZYCE<sup>59</sup>

Wydaje nam się, że eksperymenty myślowe w fizyce (i innych naukach przyrodniczych) mają z intencji<sup>60</sup> charakter eksperymentów testujących. Dlatego ograniczymy się do rekonstrukcji przykładów dwóch takich właśnie eksperymentów.

### 11.1. Spadające kule<sup>61</sup>

Eksperymentem faktualnym był eksperyment ze spadającymi kulami opisany przez Galileusza w następujący sposób:

Jest niemożliwe, aby większy ciężar poruszał się prędzej niż mniejszy. [...] Gdybyśmy więc mieli dwa ciała, których naturalne prędkości są różne, to jasnym jest, że gdy powolniejsze złączymy z przedszym, to wtedy to ostatnie będzie opóźniane przez tamto, a powolniejsze będzie przyspieszone przez przedsze. [...] Jeżeli więc jest prawdą, że wielki kamień porusza się z prędkością ośmiu stopni, a mniejszy z prędkością czterech stopni, to te dwa kamienie złączone razem będą miały prędkość mniejszą od ośmiu stopni. Ale połączone dwa kamienie są większe od wielkiego kamienia, który miał prędkość ośmiu stopni, więc w tym razie ciężar większy poruszałby się wolniej od mniejszego, co sprzeciwiałoby się Waszemu założeniu. Widzicie więc jak z założenia, że większe ciało porusza się prędzej od mniejszego dochodzę do wniosku, że większy ciężar porusza się wolniej (Galileusz 1638: 53).

Oto jak mogłaby wyglądać rekonstrukcja tego opisu — tak, aby opisana procedura odpowiadała formule z § 6:

(1) Załóżmy, że:

(α) Ciało *A* spada z prędkością  $v_A$ , a ciało *B* — z prędkością  $v_B$ , przy czym  $v_A > v_B$ .

(β) Łączymy ciało *A* z ciałem *B*, tworząc ciało *C*.

Przy takich założeniach (oraz, dodajmy na wszelki wypadek, przy założeniu, że ciało *B* porusza się w kierunku, w którym porusza się ciało *A*):

(2) Prędkość ciała *A* się zmniejszy (ale — dodajmy — nie będzie mniejsza niż  $v_B$ ), a ciała *B* się zwiększy (ale — dodajmy — nie będzie większa niż  $v_A$ ).

A zatem:

(3) Ciało *C* będzie spadać z mniejszą prędkością niż samo ciało *A*.

Jednakże zarazem:

(4) Ciało *C* będzie miało większą masę niż samo ciało *A*.

<sup>59</sup> To, czy opisane eksperymenty myślowe przeprowadzone «w umysłach» Galileusza i Einsteina miały dokładnie taki przebieg, jak niżej opisujemy, nie ma szczególnego znaczenia — z punktu widzenia celów (nie-historycznych), które sobie tu stawiamy.

<sup>60</sup> Piszemy „z intencji”, gdyż — jak wskazaliśmy wyżej — eksperymenty takie w istocie nie sprawdzają zachodzenia jakiejś zależności, lecz co najwyżej ją ilustrują. W tym sensie ma rację Such, według którego „eksperyment myślowy — w odróżnieniu od rzeczywistego — bardziej nadaje się do pełnienia funkcji heurystycznych aniżeli weryfikacyjnych” (1987a: 123).

<sup>61</sup> Do przykładu tego jako pierwszy odwołał się Mach (1906). Zob. także analizę Czeżewskiego (1956: 212 i nn.).

Przypuśćmy teraz, że:

(5) Ciało o większej masie spada z większą prędkością od ciała o mniejszej masie.

Z (4) i (5) wynika, że:

(6) Ciało *C* będzie spadać z większą prędkością niż samo ciało *A*.

Ale (6) jest niezgodne z (3).

Jak wyglądałaby ta procedura ujęta jako eksperyment myślowy w wyeksplikowanym przez nas sensie?

Galileusz najpierw zakłada połączenie ciał *A* i *B*, scharakteryzowanych jak w założeniu (1). To jest baza jego eksperymentu myślowego.

Rezultatem eksperymentu byłyby wnioski (2), (3) i (4). Ale rezultat ten — a w szczególności wniosek (3) — obala hipotezę (5), będącą obiektem eksperymentu, gdyż z hipotezy tej w połączeniu z (4) wynika (6), a (6) jest niezgodne z (3).

Mielibyśmy więc tutaj do czynienia z eksperymentem supozycyjnym (niczego tu sobie wszak nie musimy wyobrażać). Co sądzić o jego wartości?

Kluczowa jest oczywiście z tego punktu widzenia baza (1). W wypadku jej odpowiednika w obrębie eksperymentu rzeczywistego — to, że rezultat bazy (1) jest taki a nie inny, jest efektem zachodzenia naturalnej prawidłowości. Natomiast wywnioskować (2), (3) i (4) z założenia (1) można jedynie wtedy, jeśli się dodatkowo założy, że zarazem:

(7) Połączenie dwóch ciał ma większą masę niż każde z połączonych ciał wziętych z osobna.

(8) Jeżeli połączy się dwa ciała poruszające się z różną prędkością w kierunku, w którym porusza się ciało mające większą prędkość, to połączenie ciał będzie poruszało się wolniej od pierwotnej prędkości tego ostatniego.

Wartość tego eksperymentu myślowego zależy więc w tym wypadku od zasadności hipotez (7) i (8).

Dodajmy, że eksperyment ten jest eksperymentem selekcyjnym. Z jednej strony obala on bowiem Arystotelesowską koncepcję swobodnego spadania ciał — czyli hipotezę (5) — z drugiej zaś potwierdza kontrkoncepcję Galileusza głoszącą, że:

(9) Ciała — bez względu na masę — spadają z tą samą prędkością.

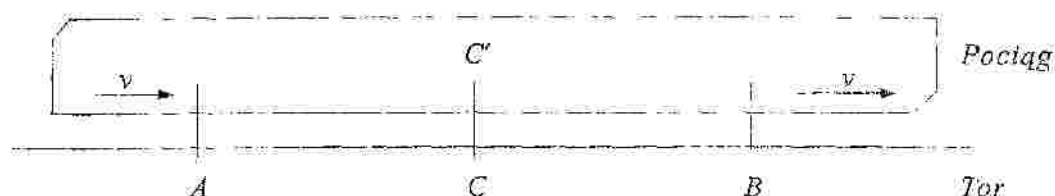
Założenie (1) jest bowiem założeniem kontrfaktualnym, gdyż na gruncie (9) nie ma ciał, które spełniałyby warunek ( $\alpha$ ) z założenia (1). Nie ma więc podstaw do wniosku (2) i (3), a kiedy zamiast (5) przyjmiemy (9) — nie będzie też podstaw do wniosku (6) niezgodnego z (3).

## 11.2. Światło w pociągu

Rozważmy z kolei eksperyment ze światłem w pociągu, opisany przez Einsteina:

Dotychczas odnosiliśmy nasze rozważania do pewnego określonego układu odniesienia. Był nim nasz tor kolejowy (ziemia). Dajmy na to teraz, że na tym torze jedzie bardzo długi pociąg ze stałą prędkością  $v$ , w kierunku podanym na [...] [rysunku]. Pasażerowie pociągu będą z na-

tury rzeczy obserwowali i notowali wszelkie zdarzenie w odniesieniu doń, czyli do połączonego z nim sztywnie „układu współrzędnych”. Każde zdarzenie, zachodzące w pewnym miejscu toru, zachodzi także w określonym punkcie pociągu i nawzajem. Określenie jednoczesności da się przy tym podać dokładnie w ten sam sposób w odniesieniu do pociągu, co i w odniesieniu do toru. Skoro jednakże zapytamy, czy dwa zdarzenia (np. oba uderzenia piorunów w  $A$  i  $B$ ) jednoczesne W ODNIESIENIU DO TORU, są także jednoczesne W ODNIESIENIU DO POCIĄGU, to, jak zaraz zobaczymy, musi odpowiedź wypaść przecząco.



Jeżeli bowiem powiadamy, że pioruny  $A$  i  $B$  są jednoczesne względem toru, to znaczy to, że promienie światła, sygnalizujące ich uderzenia i wychodzące z punktów  $A$  i  $B$ , spotykają się w środku odcinka toru  $AB$ . Ale obu zdarzeniom odpowiadają także miejsca  $A$  i  $B$  pociągu, pomiędzy którymi znajduje się w pociągu środek  $C'$ . Ten punkt schodzi się wprawdzie w chwili uderzenia piorunów z punktem  $C$  (dla obserwatora na torze), ale porusza się z prędkością  $v$  pociągu, wskutek czego zbliża się do idącego ku niemu sygnału świetlnemu z  $B$ , a oddala od takiego sygnału z  $A$ . Obserwator w  $C'$  zobaczy zatem najpierw piorun w  $B$ , a potem w  $A$ . Obserwatorzy, posługujący się pociągiem jako układem odniesienia, muszą przeto stwierdzić, w przeciwieństwie do obserwatorów na ziemi, że piorun  $B$  uderzył wcześniej niż  $A$ . Dochodzimy więc do następującego ważnego wyniku:

Zdarzenia jednoczesne w odniesieniu do toru nie są jednoczesne w odniesieniu do pociągu i nawzajem (względność równoczesności). KAŻDY UKŁAD ODNIESIENIA POSIADA SWÓJ OSOBNY CZAS; PODANIE CZASU MA OKREŚLONY SENS TYLKO WTEDY, GDY PODAJEMY ZARAZEM UKŁAD ODNIESIENIA, DO KTÓREGO ONO SIĘ ODNOSI.

Otóż przed teorią względności przyjmowała fizyka zawsze milcząco, że znaczenie podań czasu jest bezwzględne, tzn. niezależne od stanu ruchu układu odniesienia. Zobaczyliśmy właśnie, że to przyjęcie nie da się pogodzić z najbliższym określeniem jednoczesności: skoro je usuniemy, to znika [...] konflikt prawa rozchodzenia się światła z zasadą względności (Einstein 1917: 24-25).

Baza tego eksperymentu jest następująca (korzystamy z ilustracji Einsteina):

Założmy, że:

- (1) Ciało  $T$  (tj. pociąg) porusza się względem ciała  $E$  (tj. nasyp) z prędkością  $v$  w kierunku  $k$ .
- (2) Z punktu  $A$  wysłany zostaje promień świetlny w kierunku  $k$  (zdarzenie  $Z_A$ ).
- (3) A punktu  $B$  wysłany zostaje promień świetlny w kierunku przeciwnym  $k$  (zdarzenie  $Z_B$ ).
- (4) Do punktu  $M$  (na nasypie) promienie świetlne docierają jednocześnie: zdarzenia  $Z_A$  i  $Z_B$  obserwowane są jako jednoczesne.

Rezultat eksperymentu:

(5) Do punktu  $M$  promień świetlny wysłany z punktu  $B$  dociera przed promieniem świetlnym wysłanym z punktu  $A$ : zdarzenie  $Z_B$  obserwowane jest jako wcześniejsze niż zdarzenie  $Z_A$ .

Jakie rozumowanie prowadzi do takiego rezultatu?

Oznaczmy za pomocą:

- (a) ' $d_1$ ' — odległość od punktu  $A$  do  $M$  i od punktu  $B$  do  $M$ ,
- (b) ' $d_2$ ' — długość odcinka, o który w okresie  $o_1$  punkt  $M$  przemieszcza się względem punktu  $M$  w kierunku  $k$  przez (gdzie  $d_2 > 0$ );
- (c) ' $o_1$ ' — czas, który potrzebuje światło do przebycia drogi z  $A$  do  $M$  oraz z  $B$  do  $M$ ;
- (d) ' $t_1$ ' — moment, w którym promienie świetlne wysłane z punktów  $A$  i  $B$  docierają do  $M$ .

Otóż:

W momencie  $t_1$  promień światła wysłany z  $A$  nie dotrze jeszcze do  $M$ , gdyż odległość z  $A$  do  $M$  w  $t_1$  jest większa niż  $d_1$  (w szczególności wynosi  $d_1 + d_2$ ). Promień światła wysłany z punktu  $B$  dotrze do  $M$  zanim dotrze do  $M$ , gdyż odległość od  $B$  do  $M$  w czasie  $t_1$  wynosi mniej niż  $d_1$  (w szczególności  $d_1 - d_2$ ).

Dodajmy jeszcze, jakie twierdzenia należą do horyzontu tego eksperymentu. Jest to definicja kryterialna „jednoczesności” oraz teza o prędkości światła:

(6) Dwa zdarzenia (np. rozbłyśki światła) w danym układzie odniesienia są jednoczesne, gdy nieruchomy względem tego układu obserwator równo odległy od miejsc, w których zaszły te zdarzenia, zaobserwuje zajście owych zdarzeń jako jednoczesne.

(7) Prędkość światła jest stała.

Rozważany eksperyment jest — w intencji autora<sup>62</sup> — selekcyjny. Jego potwierdzonym obiektem jest teza o względności jednoczesności:

(8) Dwa zdarzenia jednoczesne w jednym układzie odniesienia mogą być niejednoczesne w innym.

Zarazem eksperyment obala tezę o bezwzględności jednoczesności:

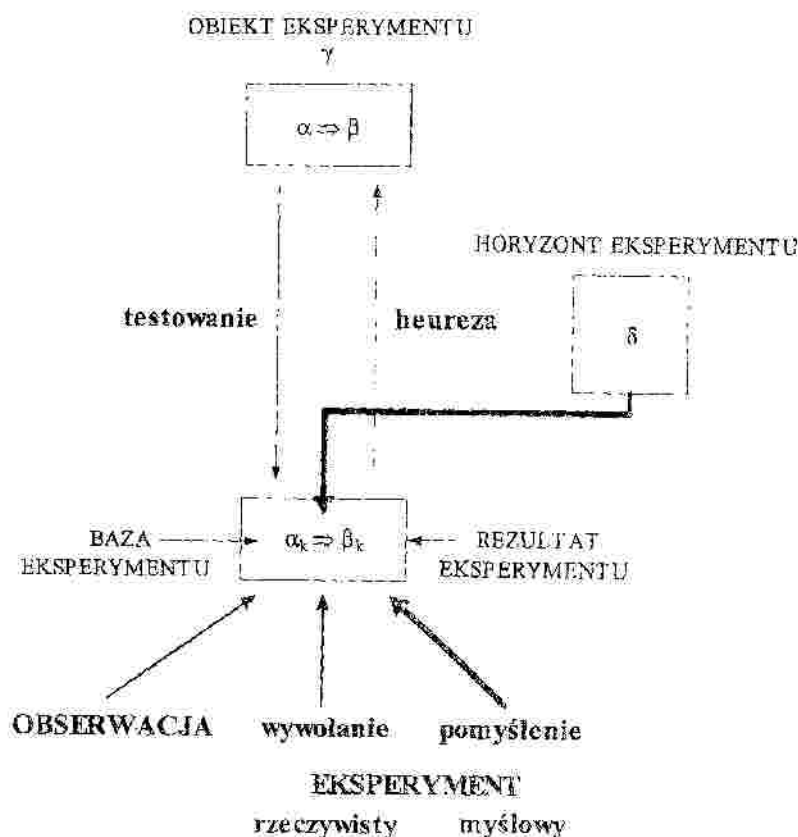
(9) Jeżeli dwa zdarzenia są jednoczesne w jednym układzie odniesienia, to są też jednoczesne w każdym innym układzie odniesienia.

<sup>62</sup> Nie wypowiadamy się tutaj na temat tego, czy intencja ta została w istocie urzeczywistniona; jak wiadomo, zgłaszano co do tego różne wątpliwości. Zauważmy tylko, że pojęcie jednoczesności (scilicet bycia-jednoczesnymi) zostaje w formule (6) określone za pomocą pojęcia bycia-zaobserwowanym-jako-jednoczesne, co ma znamiona błędnego koła w definicji. Jak zauważył Huber, polski tłumacz cytowanego tekstu Einsteina, „pojęcie równoczesności w jednym miejscu jest *a priori* jasne i nie nastręcza żadnych trudności” (1920: 12).

## 12. PODSUMOWANIE

Nasza koncepcja eksperymentu myślowego przedstawia się — w uproszczeniu — następująco.

Załóżmy, że chcemy dowiedzieć się, co pociąga za sobą zajście pewnego określonego stanu rzeczy  $\alpha_k$ . Oznaczmy przez ' $\beta_k$ ' stan rzeczy, którego zajście pociągane jest przez zajście stanu rzeczy  $\alpha_k$ . Jeżeli stan rzeczy  $\alpha_k$  jest «gotowy», to wystarczy jeśli dokonamy odpowiedniej obserwacji, dzięki której «zobaczymy» ów stan rzeczy  $\beta_k$ . W przeciwnym razie musimy wywołać stan rzeczy  $\alpha_k$ . Na tym właśnie polega eksperyment rzeczywisty. Stan rzeczy  $\alpha_k$  nazwaliśmy „bazą eksperymentu”, a to, że zajście stanu rzeczy  $\alpha_k$  pociąga za sobą zajście stanu rzeczy  $\beta_k$  — „rezultatem eksperymentu”.



Eksperyment może mieć cel heurystyczny: wtedy rezultat eksperymentu, a więc ustalenie, że zajście stanu rzeczy  $\alpha_k$  pociąga za sobą zajście stanu rzeczy  $\beta_k$ , staje się pierwszym krokiem do postawienia hipotezy — oznaczmy ją przez ' $\gamma$ ' — głoszącej, że zajście zjawiska  $\alpha$  pociąga za sobą zajście zjawiska  $\beta$  (gdzie zjawisko  $\alpha$  jest typem stanów rzeczy, do którego należy stan rzeczy  $\alpha_k$ , zjawisko zaś  $\beta$  — typem stanów



rzeczy, do którego należy stan rzeczy  $\beta_k$ ). Eksperyment taki może mieć również cel testujący: jeśli postawiliśmy już hipotezę  $\gamma$ , to ustalenie, że zajście stanu rzeczy  $\alpha_k$  pociąga za sobą zajście stanu rzeczy  $\beta_k$  potwierdza hipotezę  $\gamma$ ; gdyby to nie zachodziło — hipoteza  $\gamma$  zostałaby obalona.

Jeśli z jakichś powodów nie umiemy lub nie chcemy wywołać stanu rzeczy  $\alpha_k$ , możemy posłużyć się eksperymentem myślowym: zakładamy wtedy lub wyobrażamy sobie tylko, że stan rzeczy  $\alpha_k$  zachodzi. Pomyślenie tylko, że zachodzi stan rzeczy  $\alpha_k$ , nie wystarczy oczywiście do ustalenia, że zajście jakiego stanu rzeczy jest pociągane przez zajście stanu rzeczy  $\alpha_k$ , a więc że jest to (np.) stan rzeczy  $\beta_k$ . Musimy dysponować gotową hipotezą — oznaczmy ją przez ' $\delta$ ' — różną od hipotezy  $\gamma$ , taką, że z niej i ze zdania stwierdzającego zajście stanu rzeczy  $\alpha_k$  wynika zajście stanu rzeczy (powiedzmy)  $\beta_k$ . Nazwalismy hipotezę  $\gamma$  „horyzontem eksperymentu”. Niekiedy to, że istotnie uznajemy hipotezę  $\gamma$ , ujawnia się dopiero w trakcie eksperymentu — kiedy uświadamiamy sobie, dlaczego wyobrażamy sobie lub zakładamy, że nasz eksperyment ma taki, a nie inny rezultat.

Skoro rezultat eksperymentu myślowego wynika z horyzontu tego eksperymentu, to zakładanie lub wyobrażanie sobie bazy eksperymentu — z teoretycznego punktu widzenia — jest zbędne. Eksperyment myślowy sprowadza się bowiem do wywnioskowywania rezultatu z horyzontu i bazy. Eksperymenty myślowe pełnią natomiast ważne funkcje praktyczne: pozwalają na uświadomienie sobie elementów bazy, a także ułatwiają zilustrowanie lub zegzemplifikowanie odpowiednich zależności.

Zadajmy sobie pytanie, dlaczego fizycy odwołują się do myślowych eksperymentów kontrfaktualnych.

Otóż po pierwsze — dlatego, że różnice w alternatywnych opisach jakiegoś fragmentu rzeczywistości ujawniają się w sytuacjach granicznych, a sytuacji granicznych nie da się «wytworzyć» (niekiedy jest to co najmniej technicznie niemożliwe); można je sobie co najwyżej pomyśleć. Wszystkie obserwowalne przez nas ruchy ciał są np. «zgodne» zarówno z teorią Newtona, jak i teorią Einsteina — różnice widać dopiero przy ogromnych prędkościach.

Po drugie — dlatego, że myślowo możemy «stworzyć» idealizacyjne warunki, których nie jesteśmy w stanie «wytworzyć» w rzeczywistości.

Po trzecie zaś — zapewne dlatego, że egzemplifikacje ubrane w szaty eksperymentów myślowych bardziej przemawiają do wyobraźni, co jest nie bez znaczenia dla szybkości i powszechności recepcji nowych, «mało intuicyjnych» idei w nauce.

**Tomasz Bigaj**

## **Komentarz do opisu eksperymentu ze światłem w pociągu**

Analiza eksperymentu myślowego ze światłem w pociągu została przedstawiona przez Einsteina nieco nieprecyzyjnie i ten brak precyzji uwidacznia się również w interpretacji Autorów artykułu.

Przed wszystkim należy zwrócić uwagę na to, że termin „punkt” używany w analizie jest wieloznaczny. Może on znaczyć tyle, co „punkt przestrzenny” lub tyle, co „punkt czasoprzestrzenny”. Punkt czasoprzestrzenny jest to lokalizacja zdarzenia nierozciągniętego (punktowego), a bardziej formalnie — klasa abstrakcji na zbiorze zdarzeń od relacji koincydencji. Punkty czasoprzestrzenne są niezależne od układów odniesienia, ale odległości przestrzenne i czasowe między nimi są już zależne. Z kolei punkty przestrzenne mogą być zdefiniowane wyłącznie ze względu na dany układ odniesienia. Punkt przestrzenny, w którym zachodzi dane zdarzenie  $e$ , jest definiowany jako zbiór tych wszystkich zdarzeń, których lokalizacja przestrzenna w danym układzie jest tożsama z lokalizacją zdarzenia  $e$ . Punkty czasoprzestrzenne są bytami momentalnymi (chwilowymi), podczas gdy punkty przestrzenne trwają w czasie. Obrazowo można przedstawić punkt przestrzenny jako «kreskę» zaznaczoną np. na szybie wagonu pociągu, która porusza się z całym wagonem.

Nie jest do końca jasne, jaka jest interpretacja symboli  $A$  i  $B$  w opisie eksperymentu podanym przez Einsteina. Mogłyby być one interpretowane jako punkty czasoprzestrzenne będące lokalizacjami czasoprzestrzennymi obu uderzeń piorunów. Jednakże Einstein pisze w pewnym momencie „miejsca  $A$  i  $B$  pociągu”, co sugeruje interpretację przestrzenną. W takim jednak razie Einstein powinien użyć innych symboli, np.  $A''$  i  $B''$ , gdyż punkty przestrzenne uderzeń piorunów w układzie związanym z pociągiem są numerycznie różne od punktów w układzie związanym z torem. Punkty  $A$  i  $A'$  (oraz  $B$  i  $B'$ ) «koincydują» ze sobą tylko w momencie uderzenia piorunu, a następnie «się rozjeżdżają» (punkty  $A'$  i  $B'$  są sztywno powiązane z pociągiem).

Ważne jest, aby zauważyć, że punkty  $C$  i  $C'$  (później w tekście nieco myląc się one przemianowane na  $M$  i  $M'$ ) mogą być zdefiniowane tylko jako punkty przestrzenne. Są one wyznaczone geometrycznie jako środki odcinków  $AB$  i  $A'B'$ . Koincydencja między  $C$  i  $C'$  zachodzi tylko w jednym momencie (z punktu widzenia obserwatora na torze jest to moment uderzenia obu piorunów).

Zakładam, że symbole  $A$  i  $B$  oznaczają punkty przestrzenne zdefiniowane w układzie związanym z torem. Zatem wszystkie pojęcia zdefiniowane w punktach (a)-(d) muszą być zrelatywizowane do tego właśnie układu odniesienia (dotyczy to w szczególności pojęć czasowych z punktów (c) i (d) — czas przebycia drogi od  $A$  do  $M$  nie jest równy czasowi od  $A'$  do  $M'$ ). Ta relatywizacja może wydawać się zagrożeniem dla tezy końcowej, która przecież ma być twierdzeniem dotyczącym braku równoczesności w układzie związanym z pociągiem — a nie z torem. Jednak jest to problem pozorny. Z przyjętych założeń wynika bowiem teza absolutna (tj. prawdziwa we wszystkich układach odniesienia), iż oba promienie NIE SPOTYKAJĄ SIĘ w punkcie  $M'$  (ponieważ spotykają się w  $M$ , a  $M$  i  $M'$  nie koincydują w momencie spotkania promieni). To zaś wystarczy, aby pokazać, że w układzie związanym z pociągiem uderzenia piorunów nie są równoczesne zgodnie z kryterium Einsteina. W istocie argument dowodzi czegoś więcej: że uderzenie pioruna w  $B'$  jest w tym układzie wcześniejsze niż uderzenie w  $A'$ . Formalnie, ten wniosek wymaga przyjęcia dodatkowego kryterium: mianowicie kryterium wcześniejszości (późniejszości). Jest to kryterium czasikowe, stosowalne tylko dla zdarzeń odseparowanych interwałem przestrzennie-podobnym. Zdarzenie  $X$  jest wcześniejsze od zdarzenia  $Y$  w układzie  $u$ , gdy miejsce spotkania promieni wysłanych z  $X$  i  $Y$  wy-

pada bliżej punktu przestrzennego zajścia  $Y$ -a w  $u$  niż punktu zajścia  $X$ -a w  $u$ . Przyjmując to kryterium możemy znacznie uprościć całe rozumowanie (oraz pozbawić je kłopotliwej relatywizacji do układu związanego z torem). Wystarczy tylko zauważyć, że zdarzenie zetknięcia dwóch promieni (które z założenia wypada w  $M$ ) jest zlokalizowane pomiędzy punktami  $A'$  i  $M'$  w układzie związanym z pociągami, a zatem jest bliższe  $A'$  niż  $B'$ .

Nawiasem mówiąc – zarzut z przypisu 62 jest nietrafny. Relacja  $x$ -jest-zaobserwowany-jako-jednoczesny- $z$ - $y$  jest logicznie niezależna od relacji  $x$ -jest-jednoczesny- $z$ - $y$ . Aby jednak postawić sprawę na ostrzu noża, zauważmy, że kryterium Einsteina może (i powinno) być sformułowane bez odniesienia do obserwacji. Jednoczesność zdarzeń odległych przestrzennie definiuje się tutaj za pomocą pojęcia koincydencji, które z kolei traktowane jest jako pierwotne, a nie definiowalne jako jednoczesność połączona z kolokacją.

## BIBLIOGRAFIA

- Ajdukiewicz, Kazimierz  
(1965) *Logika pragmatyczna*, Warszawa, PWN.
- Alanen, Lilli  
(1991) Descartes, Conceivability, and Logical Modality, [w:] Horowitz & Massey (red.) 1991, s. 65-84.
- Anapolitanos, Dionysios A.  
(1991) Thought Experiments and Conceivability Conditions, [w:] Horowitz & Massey (red.) 1991, s. 87-97.
- Brown, James Robert  
(1991a) *The Laboratory of the Mind. Thought Experiments in the Natural Sciences*, London, Routledge.  
(1991b) Thought Experiments: A Platonic Account, [w:] Horowitz & Massey (red.) 1991, s. 119-128.
- Camp, Joseph L. Jr.  
(1991) The Ballad of Clyde the Moose, [w:] Horowitz & Massey (red.) 1991, s. 327-333.
- Czeżowski, Tadeusz  
(1956) Arystoteles, Galileusz, Bacon, [w:] *Odczyty filozoficzne*, Toruń 1958, Towarzystwo Naukowe w Toruniu, s. 208-217.
- Einstein, Albert  
(1917) *O szczególnej i ogólnej teorii względności*, Warszawa 1922, Książnica Polska TNSW.
- Forge, John  
(1991) Thought Experiments in the Philosophy of Physical Science, [w:] Horowitz & Massey (red.) 1991, s. 209-222.
- Gale, Richard M.  
(1991) On Some Pernicious Thought-Experiments, [w:] Horowitz & Massey (red.) 1991, s. 297-303.

- Galilei, Galileo [Galileusz]  
(1638) *Rozmowy i dowodzenia matematyczne w zakresie dwóch nowych umiejętności dotyczących mechaniki i ruchów miejscowych*, Warszawa 1930, Kasa im. Mianowskiego.
- George, Rolf  
(1991) The Tradition of Thought Experiments in Epistemology, [in:] Horowitz & Massey (red.) 1991, p. 273-283.
- Grobler, Adam  
(2006) *Metodologia nauk*, Kraków, Wydawnictwo Aureus & Wydawnictwo Znak.
- Hajduk, Zygmunt  
(2002) Eksperyment, [w:] Andrzej Maryniarczyk *et al.* (red.), *Powszechna encyklopedia filozofii*. Tom 3, Lublin, Polskie Towarzystwo Tomasza z Akwinu, s. 83-84.
- Hetherington, Stephen Cade  
(1991) Conceivability and Modal Knowledge, [w:] Horowitz & Massey (red.) 1991, s. 317-325.
- Horowitz, Tamara  
(1991) Newcomb's problem as a Thought Experiment, [w:] Horowitz & Massey (red.) 1991, s. 305-316.
- Horowitz, Tamara & Massey, Gerald J.  
(1991) Introduction, [do:] Horowitz & Massey (red.) 1991, s. 1-28.
- Horowitz, Tamara & Massey, Gerald J. (red.)  
(1991) *Thought Experiments in Science and Philosophy*, Savage (Maryland), Rowman & Littlefield Publishers, Inc.
- Huber, Maksymilian Tytus  
(1920) *Albert Einstein i jego teoria*, Lwów, Spółka Wydawnicza Słowa Polskiego.
- Irvine, Andrew D.  
(1991) On the Nature of Thought Experiments in Scientific Reasoning, [w:] Horowitz & Massey (red.) 1991, s. 149-165.
- Janis, Allen I.  
(1991) Can Thought Experiments Fail? [w:] Horowitz & Massey (red.) 1991, s. 113-118.
- King, Peter  
(1991) Mediaeval Thought-Experiments: The Metamethodology of Mediaeval Science, [w:] Horowitz & Massey (red.) 1991, s. 43-64.
- Kotarbiński, Tadeusz  
(1929) Obserwacja i eksperyment, [w:] *Elementy teorii poznania, logiki formalnej i metodologii nauk*, Wrocław 1961, Ossolineum, s. 351-355.
- Laymon, Ronald  
(1991) Thought Experiments of Stevin Mach and Gouy: Thought Experiments as Ideal Limits and as Semantic Domains, [w:] Horowitz & Massey (red.) 1991, s. 167-191.
- Lennox, James G.  
(1991) Darwinian Thought Experiments: A Function for Just-So Stories, [w:] Horowitz & Massey (red.) 1991, s. 223-245.

- Mach, Ernest  
(1906) Über Gedankenexperimente. [w:] *Erkenntnis und Irrtum*, Leipzig, Verlag von Johann Ambrosius Barth, s. 183-200.
- Massey, Gerald  
(1991) Backdoor Analyticity, [w:] Horowitz & Massey (red.) 1991, s. 285-296.
- Mill, John Stuart  
(1843) O obserwacji i eksperymencie, [w:] *System logiki dedukcyjnej i indukcyjnej*. T. 1, Warszawa 1962, PWN, s. 586-599.
- Mohanty, John N.  
(1991) Method of Imaginative Variation in Phenomenology, [w:] Horowitz & Massey (red.) 1991, s. 261-272.
- Norton, John D.  
(1991) Thought Experiments in Einstein's Work, [w:] Horowitz & Massey (red.) 1991, s. 129-148.  
(2004) On Thought Experiments: Is There More to the Argument? *Philosophy of Science* vol. 71, s. 1139-1151.
- Nowaczyk, Adam & Żołądowski, Zenobiusz  
(1974) Eksperyment, [w:] *Logika i metodologia badań naukowych dla lekarzy*, Warszawa, PZWL, s. 152-153.
- Podsiad, Antoni  
(2000) *Słownik terminów i pojęć filozoficznych*, Warszawa 2000, Instytut Wydawniczy Pax.
- Rescher, Nicholas  
(1991) Thought Experimentation in Presocratic Philosophy, [w:] Horowitz & Massey (red.) 1991, s. 31-41.
- Stepień, Antoni B.  
(2007) *Wstęp do filozofii*, Lublin, Towarzystwo Naukowe KUL.
- Such, Jan  
(1987) Eksperyment, [w:] Z. Cackowski, J. Kmita, K. Szaniawski & P. J. Smoczyński (red.), *Filozofia a nauka. Zarys encyklopedyczny*, Wrocław, Ossolineum, s. 120-131.
- Szubka, Tadeusz  
(2000) Analityczna filozofia, [w:] Andrzej Maryniarczyk *et al.* (red.), *Powszechna encyklopedia filozofii*. Tom 1, Lublin, Polskie Towarzystwo Tomasza z Akwinu, s. 204-208.
- Thomason, Sarah G.  
(1991) Thought Experiments in Linguistics, [w:] Horowitz & Massey (red.) 1991, s. 247-257.
- Twardowski, Kazimierz  
(1901) *Zasadnicze pojęcia dydaktyki i logiki*, Lwów, Towarzystwo Pedagogiczne.
- Wilson, Mark  
(1991) Reflections on Strings, [w:] Horowitz & Massey (red.) 1991, s. 193-207.
- Wojtyś, Jacek  
(2007) Słownik wybranych terminów filozoficznych, [w:] Stepień 2007, s. 327-438.



Anna Brożek, Jacek Jadacki

## **Eksperymenty myślowe w filozofii**

1. *Quasi-fizyczne eksperymenty myślowe*. 2. Bezludny wszechświat. 3. Wymiana «klepek». 4. Diagnostyczne eksperymenty myślowe. 5. Tor śmierci. 6. Semantyczne eksperymenty myślowe. 7. Statek Tezeusza. 8. Uzasadnione przekonanie. 9. Ziemia Bliźniacza. 10. Chiński pokój. 11. Mózg w naczyniu. 12. Konkluzja.

Bibliografia

### **1. QUASI-FIZYCZNE EKSPERYMENTY MYŚLOWE**

W tekście *Eksperymenty myślowe w nauce* zaproponowaliśmy pewną definicję „eksperymentu myślowego”. Przypomnijmy ją tutaj, gdyż będzie ona tłem dla rozważania funkcji eksperymentów myślowych w filozofii.

Otóż założyliśmy, że pomiędzy pojęciami eksperymentu rzeczywistego i eksperymentu myślowego zachodzi wyraźna analogia. Eksperyment rzeczywisty polega na wywołaniu pewnego stanu rzeczy po to, aby zaobserwować jego efekt.

(1) Osoba *O* eksperymentuje rzeczywiście (przeprowadza eksperyment rzeczywisty) z *X*-em, gdy:

(a) osoba *O* wywołuje *X*-a

i

(b) osoba *O* ustala — za pomocą obserwacji — co jest efektem *X*-a.

W eksperymentacie myślowym wywołujemy nie stan rzeczy, a — przedstawienie stanu rzeczy, tj. zakładamy lub wyobrażamy sobie, że dany stan rzeczy zachodzi — po to, aby przewidzieć, za pomocą wnioskowania, jego efekty:

(2) Osoba *O* eksperymentuje myślowo (przeprowadza eksperyment myślowy) z *X*-em, gdy:

(a) osoba *O* przedstawia sobie *X*-a

i

(b) osoba *O* ustala — za pomocą rozumowania — co jest efektem *X*-a.

Wyobrażony lub założony *X* nazwaliśmy „bazą eksperymentu myślowego”, a wywnioskowany efekt *X*-a — „rezultatem eksperymentu myślowego”.

Przyjęliśmy też, że osoba przeprowadzająca eksperyment myślowy wywnioskowuje, co jest efektem *X*-a, na podstawie pewnych uznawanych już (mniej lub bardziej świadomie) praw lub wysuwanych hipotez. Bywa też, że prawo pozwalające na wywnioskowanie rezultatu z bazy ujawnia się (*scil.* jest uświadamiane) dopiero w momencie «przeprowadzenia» eksperymentu. Owe prawa nazwaliśmy „horyzontem eksperymentu myślowego”.

Eksperymenty — zarówno rzeczywiste, jak i myślowe — miewają cele heurystyczne lub testujące.

W pierwszym wypadku ustalenie (w drodze obserwacji lub rozumowania), że zajście stanu rzeczy *a* pociąga za sobą zajście stanu rzeczy *b*, staje się pierwszym krokiem do postawienia hipotezy głoszącej, że zajście zjawiska *a* pociąga za sobą zajście zjawiska typu  $\beta$  (gdzie zjawisko *a* jest typem stanów rzeczy, do którego należy stan rzeczy *a*, zjawisko zaś  $\beta$  — typem stanów rzeczy, do którego należy stan rzeczy *b*).

Eksperymenty miewają także cel testujący: jeśli postawiliśmy już hipotezę  $\gamma$ , to ustalenie, że zajście stanu rzeczy *a* pociąga za sobą zajście stanu rzeczy *b* potwierdza tę hipotezę  $\gamma$ ; gdyby to nie zachodziło — hipoteza  $\gamma$  zostałaby obalona.

Zauważyliśmy, że skoro rezultat eksperymentu myślowego wynika z horyzontu tego eksperymentu, to zakładanie lub wyobrażanie sobie bazy eksperymentu z teoretycznego punktu widzenia jest zbędne. Eksperyment myślowy redukuje się do wywnioskowywania rezultatu z horyzontu i bazy. Wyobrażenie lub założenie bazy i rezultatu — spełniają tylko funkcję praktyczną: ilustrowania lub egzemplifikowania odpowiednich zależności.

Przedstawimy teraz analizę kilku filozoficznych eksperymentów myślowych, traktując powyższe rozróżnienia jako tło analizy.

Pierwsze dwa przykłady — to przykłady eksperymentów *quasi*-fizycznych.

Przypominają one eksperymenty fizyczne — a niezależnie od tego, że nie są rzeczywiste, duże trudności sprawia rekonstrukcja ich horyzontu.

## 2. BEZLUDNY WSZECHŚWIAT

W eseju Wierzbickiego *Roski umysł* (2011)<sup>1</sup> znajdujemy taki oto *passus*:

Proponujemy mały myślowy eksperyment. Spróbujmy wyobrazić sobie [...] świat w sytuacji, gdy nas jeszcze tu nie było i gdy nas znów na tej planecie zabraknie, świat nie oglądany przez

<sup>1</sup> Sam autor zastrzega, że jego esej ma raczej charakter literacki — nie zaś teoretyczny. Zauważmy więc, że przy pewnych założeniach co najmniej niektóre dzieła literatury pięknej (por. np. choćby *Podróż Guliwera* Jonathana Swifta czy *Mikołaja Doświadczyńskiego przypadki* Ignacego Krasickiego) wolno uważać za literackie sprawozdanie z eksperymentów myślowych.

niczyje oczy, nie dotknięty przez niczyją żywą myśl, świat, jak to mawiają „sam w sobie”, wolny od jakichkolwiek śladów naszych ja, ogołoceny z przypadłości (kształtów, kolorów, zapachów), które mu nadają nasze zmysły, pozbawiony widoków, świat, w którym w miejscu obiektów wyodrębnionych i uporządkowanych przez nasz umysł kotłuje się jakaś magma, zaś pajęczynę pojęć utrwalonych przez nasz język zastępuje tępe ani be, ani me, świat, o którym nikt nie słyszał i nie pomyślał. Wyteżamy wszystkie siły duszy, rozkręcamy imaginację, przerzucamy kolejne niewydarzone warianty wizerunku, któremu nie wolno mieć twarzy... Raptem coś to? Uświadamiamy sobie straszną rzecz: daremne próby, próżny trud, cel nasz jest nieosiągalny z założenia, świata nie oglądanego przez nikogo, świata nie dotkniętego niczym ja, świata, o którym nikt nie słyszał i nie pomyślał, nie zdołamy sobie nigdy wyobrazić, bo on zwyczajnie pod słońcem *nie istnieje*. Jakże by miało być inaczej? Przecież już w pierwszej sekundzie naszych zmagani, ba, jeszcze wcześniej, już wręcz w chwili, w której wpadł nam do głowy sam pomysł, złamaliśmy reguły przedsięwzięcia. Miał to być świat, o którym nie pomyślał nikt. Otóż *pomyślał*. Myśmy sami najpierw o nim *pomyśleli*, potem zaś *myśleli* w najlepsze. Że nie było innego wyjścia... Święta racja: w tym właśnie sęk, nie było wyjścia, nie ma wyjścia, nasze ja nie odstępuje światu nawet wtedy, gdy zdaje się nam, żeśmy go ogołocili z bagażu naszych zmysłów i umysłu.

Tak wychodzi na jaw najzaszczytniejszy tytuł naszych ja. Są współkreatorami bytu (Wierzbicki 2011: 18-19).

Wywód ten wolno w naszym przekonaniu zrekonstruować następująco.

Zadajmy sobie pytanie:

(1) Jaki byłby świat bezludny, tj. taki, w którym nie byłoby ludzi?

Na pytanie to nie da się odpowiedzieć inaczej niż wyobrażając sobie, jaki byłby ów świat. Zatem wyobraźmy sobie, że:

(2) Świat jest bezludny.

Stan rzeczy stwierdzany przez (2) jest bazą eksperymentu. Okazuje się, że zadanie wyobrażenia sobie, że (2), jest co najmniej dwuznaczne.

Pierwsza interpretacja uwzględniałaby to, że:

(3) Gdyby w świecie nie było ludzi, to nikt by go sobie nigdy m.in. nie wyobrażał.

Można więc (2) zinterpretować jako zadanie wyobrażenia sobie, że:

(4)  $\sim \forall x (x \text{ wyobraża sobie świat})$ .

Gdyby to się nam udało, to bylibyśmy kimś, kto wyobraża sobie świat, którego... nikt sobie nie wyobraża. W konsekwencji mielibyśmy sprzeczność:

(5)  $\forall x (x \text{ wyobraża sobie świat}) \wedge \sim \forall x (x \text{ wyobraża sobie świat})$ .

A więc zadanie wyobrażenia sobie, że (4), a zatem i zadanie wyobrażenia sobie, że (2) — jeśli do naszego horyzontu należy (3) — wyglądałoby na niewykonalne.

Nie jest to jednak rozumowanie poprawne, gdyż — dodajmy od siebie — zadanie wyobrażenia sobie, że (2), po dokładniejszej analizie, przybiera postać:

(6) Wyobraźmy sobie w chwili  $T$  świat bezludny w chwili  $T^*$  (różnej od  $T$ ).

Wtedy zamiast (5) mamy:

(7)  $\forall x (x \text{ wyobraża sobie świat w chwili } T) \wedge \sim \forall x (x \text{ wyobraża sobie świat w chwili } T^*)$ .

A tu oczywiście nie ma sprzeczności.

Ponadto — zakłada się, że w formułach zakres zmiennej 'x' ogranicza się do przedmiotów należących do wyobrażonego świata. Można sobie jednak... wyobrazić, że wyobrażamy sobie świat niejako «spoza niego»: zakres zmiennej 'x' w pierwszym członie formuły (5) byłby wówczas nadzbiorem zakresu zmiennej 'x' z drugiego członu.

A oto druga interpretacja formuły (2).

Nazwijmy (idąc za pewną tradycją) własności, które ludzie przypisują jakiemuś przedmiotowi dzięki temu, że spostrzegają ten przedmiot, „własnościami subiektywnymi”. Z kolei „własnościami obiektywnymi” nazwijmy własności, które dany przedmiot ma niezależnie od tego, czy jest spostrzegany przez ludzi.

Otóż:

(8) Gdyby w świecie nie było ludzi, to świat miałby wyłącznie własności obiektywne.

Założmy jednak, że:

(9) Ludzie są w stanie wyobrazić sobie tylko własności subiektywne świata.

(10) Własności obiektywne danego przedmiotu są różne od jego własności subiektywnych.

(11) Nie jesteśmy w stanie wyobrazić sobie niczego, co nie ma własności.

Zatem:

(12) Nie jesteśmy w stanie wyobrazić sobie świata bezludnego.

Podkreślmy: jest to niewykonalne nie dlatego, że w zadaniu tkwi rzekoma sprzeczność (jak w wypadku pierwszej interpretacji), lecz dlatego, że zakładamy (9), (10) i (11).

W tej interpretacji — horyzontem eksperymentu myślowego są zdania (10)-(11). Żadnego rezultatu tego eksperymentu nie da się natomiast — zgodnie z (12) — wskazać.

### 3. WYMIANA «KLEPEK»

O ile poprzedni eksperyment wolno uznać za argument na rzecz ontologicznego idealizmu, o tyle eksperyment Searle'a z wymianą «klepek» jest wykorzystywany w dyskusji nad dualizmem psychofizycznym. Oto jego brzmienie:

Wyobraź sobie, że twój mózg zaczyna psuć się w ten sposób, że stopniowo ślepniesz. Wyobraź sobie dalej, że zdesperowani lekarze, pragnący poprawić twój stan, próbują wszelkich metod przywrócenia ci wzroku. Jako ostatnią deskę ratunku postanawiają wszczepić układy scalone w część twojej kory mózgowej odpowiedzialną za widzenie. Ku twojemu i ich zaskoczeniu układy scalone przywracają ci normalne widzenie. Wyobraź sobie teraz, że twój mózg znów niestety nadal się psuje, a lekarze wszczepiają coraz to nowe układy scalone. Widać już, dokąd zmierza ten eksperyment myślowy: w końcu wyobrażamy sobie, że twój mózg zostaje całkiem zastąpiony przez układy scalone — tak, że gdy potrząsaszą głową, możesz słyszeć krzemowe płytki klekoczące w środku (Searle 1992: 67).

Skłonni jesteśmy uznać, że eksperyment z wymianą «klepek» ma następującą strukturę.

Obiektem eksperymentu jest hipoteza:

(1) Jeżeli przedmiot  $P_1$  złożony z części  $\{C_1, \dots, C_j\}$  pełni funkcje  $\{F_1, \dots, F_k\}$ , to przedmiot  $P_2$ , złożony z części  $\{D_1, \dots, D_j\}$ , takich, że  $D_1$  pełni taką samą funkcję jak  $C_1$ ,  $D_2$  taką samą jak  $C_2$  etc., także pełni funkcje  $\{F_1, \dots, F_k\}$ .

Bazę eksperymentu stanowi sytuacja polegająca na tym, że:

(2) Wszystkie części pewnego mózgu zostają wymienione na ich odpowiedniki w postaci układów scalonych.

Z kolei na horyzont eksperymentu składają się dwie tezy:

(3) Mózg pełni co najmniej dwie funkcje: steruje zachowaniem i jest ośrodkiem świadomości.

(4) Pewne układy scalone pełnią te same funkcje, co pewne części mózgu.<sup>2</sup>

Searle nie przesądza żadnego wyniku eksperymentu, a wymienia tylko różne możliwości:

(5) System układów scalonych pełni wszystkie funkcje pełnione przez mózg — w szczególności steruje zachowaniem i jest ośrodkiem świadomości.<sup>3</sup>

(6) System układów scalonych steruje zachowaniem, ale nie jest ośrodkiem świadomości.<sup>4</sup>

(7) System układów scalonych jest ośrodkiem świadomości, ale nie steruje zachowaniem.<sup>5</sup>

Zauważmy, że nie jest to pełna lista «czystych» możliwości; istnieje np. jeszcze możliwość, polegająca na tym, że po wymianie jednego z neuronów na układ scalony zanikają wszystkie funkcje mózgu.

Mamy tutaj w gruncie rzeczy inną wersję starożytnego paradoksu nieostrości. Załóżmy mianowicie, że:

(8) Wymiana jednego neuronu mózgu na odpowiedni układ scalony nie narusza funkcjonowania mózgu.

<sup>2</sup> Nie jest jasne, czy teza (4) należy do horyzontu, czy też jest już częścią wyniku eksperymentu. Chodzi tu o to, że po wszczepieniu układu scalonego dochodzi do odzyskania wzroku.

<sup>3</sup> „One logical possibility, not to be excluded on any a priori grounds alone, is surely this: you continue to have all of the sorts of thoughts, experiences, memories, etc. that you had that you had previously; the sequence of your mental life remains unaffected. In this case, we are imagining that the silicon chips have the power not only to duplicate the mental phenomena, conscious and otherwise that are normally responsible for your input-output functions.” Co ciekawe, Searle dodaje: „I don't for a moment think that such a thing is even remotely empirically possible. I think it is empirically absurd to suppose that we could duplicate the causal powers of neurons entirely in silicon. But that is an empirical claim of my part. It is not something that we could establish a priori. So the thought experiment remains valid as a statement of logical or conceptual possibility.”

<sup>4</sup> „The second possibility: your conscious experience slowly shrinks to nothing while your externally observable behavior remains the same.”

<sup>5</sup> „Your thoughts, feelings, experiences, memories etc. remains intact but your observable external behavior slowly reduces to total paralysis.”



(9) Dla każdego  $k$ : jeżeli wymiana  $k$  neuronów mózgu na odpowiednie układy scalone nie narusza funkcjonowania mózgu, to i wymiana  $k+1$  neuronów mózgu na układy scalone nie narusza funkcjonowania mózgu.

Przy założeniach (8) i (9) wymiana wszystkich neuronów mózgu na odpowiednie układy scalone nie narusza funkcjonowania mózgu; jeśli zatem mózg jest ośrodkiem świadomości, to i powstały w wyniku takiej wymiany obiekt jest ośrodkiem świadomości.

Oczywiście jest, że oba założenia — (8) i (9) wymagają faktycznego sprawdzenia empirycznego; żadne rozumowanie takiego sprawdzenia nie zastąpi. Jasne jest też, że dopiero włączenie do horyzontu hipotezy, zgodnie z którą:

(10) Żadne systemy układów scalonych nie są ośrodkami świadomości.  
świadczyłoby o dopuszczalności odrzucenia dualizmu.

#### 4. DIAGNOSTYCZNE EKSPERYMENTY MYŚLOWE

O wiele częściej niż z eksperymentami *quasi-fizycznymi* — spotykamy się w filozofii z eksperymentami myśłowymi, które nazwiemy tu „myśłowymi eksperymentami diagnostycznymi”. Mówiąc swobodnie: eksperyment diagnostyczny służy do ujawnienia czyichś niewyjawionych dotąd przekonań — o różnym stopniu ogólności.<sup>6</sup> Takie nieujawnione (*resp.* niezwerbalizowane) przekonania nazywane bywają „intuicjami” (chodzi oczywiście o jedno ze znaczeń tego słowa).<sup>7</sup>

Otóż chociaż eksperyment diagnostyczny może być zarówno eksperymentem myślowym, jak i rzeczywistym,<sup>8</sup> to jest on szczególnie częsty w formie eksperymentu myślowego w filozofii. Pokazują to zanalizowane przez nas przykłady filozoficznych eksperymentów myślowych,<sup>9</sup> z których kilka przedstawimy poniżej.

<sup>6</sup> Według Nowaczyka i Żołądowskiego — eksperymenty diagnostyczne są „podejmowane w celu uzyskania odpowiedzi na pytania o cechy swoiste badanych zjawisk, pozwalające ustalić miejsce badanego zjawiska w systemie jakiejś klasyfikacji” (1974: 153). W ujęciu Hajduka — eksperyment diagnostyczny jest to eksperyment „służący rozpoznaniu własności obiektu lub jego zaklasyfikowaniu” (2002: 83). Podobnie według Wojtyśiaka — jest to eksperyment „stanowiący podstawę do odpowiedniego zaklasyfikowania danego przedmiotu” (2007: 344). Pozostałe eksperymenty — „podejmowane w celu ustalenia charakteru zależności między zjawiskami” — Nowaczyk i Żołądowski nazywają „eksperymentami wyjaśniającymi” (1974: 153).

<sup>7</sup> Rola, jaką odgrywają intuicje w eksperymentach myślowych, została częściowo zanalizowana w Brendel (2004). Stwierdza się tam, że wszelkie eksperymenty myślowe mogą być zrekonstruowane jako argumentacje, ale ich moc eksplanacyjna płynie właśnie z odwoływania się do intuicji. Zgadza się jedynie po części z tym stwierdzeniem. Prawdą jest, że w takich eksperymentach obecny jest pewien komponent intuicyjny, ale w ostateczności, gdy odpowiadamy na pytanie zawarte w scenariuszu eksperymentu, odwołujemy się do *twierdzeń naukowych*, tj. dobrze uzasadnionych, nie zaś do intuicji.

<sup>8</sup> To, że zamiast eksperymentów rzeczywistych robi się tylko eksperymenty myślowe, ma najczęściej uzasadnienie etyczne (np. dlatego, że eksperymenty rzeczywiste nie są obojętne moralnie).

<sup>9</sup> Diagnostyczne eksperymenty myślowe stosowane są również poza filozofią; zaliczylibyśmy do nich np. Schrödingera eksperyment z kotem i trucizną (1935).

Schemat eksperymentu diagnostycznego wyglądałby następująco. Rozważmy zależność o postaci:

(1)  $\wedge x (Px \Rightarrow Qx)$ .

Zależność (1) jest obiektem eksperymentu. Załóżmy teraz, że mamy sprawdzić hipotezę:

(2) Osoba *O* jest przekonana, że  $\wedge x (Px \Rightarrow Qx)$ .<sup>10</sup>

Aby to sprawdzić — możemy np. osobę *O* zapytać:

(3) Osobo *O*, czy jesteś przekonana, że (1)?

Byłoby to uzasadnienie bezpośrednie. Niekiedy jednak — z różnych powodów — korzystamy ze sprawdzania pośredniego. Rozważmy formułę:

(4)  $P^*a \Rightarrow Q^*a$ .

Predykaty w formule (4) celowo oznaczamy gwiazdkami, wskazując na to, że są w pewnym sensie odpowiednikami predykatów występujących w (1).

Założmy teraz, że:

(5) Formuła (4) jest szczególnym przypadkiem formuły (1).<sup>11</sup>

Pytamy teraz osobę *O*:

(6) Osobo *O*, czy jesteś przekonana, że (4)?

W innej wersji pytanie może brzmieć:

(7) Osobo *O*, co jest pociągane przez to, że  $P^*a$ ?

albo:

(8) Czy ' $P^*a$ ' pociąga ' $Q^*a$ ', czy ' $\sim Q^*a$ '?

Jeżeli osoba *O* odpowie na pytanie (6) pozytywnie (bądź odpowie na pytanie (7) lub (8), stwierdzając, iż ' $P^*a$ ' pociąga ' $Q^*a$ '), to możemy uznać, że hipotezę (b) potwierdziliśmy. Jeżeli natomiast osoba *O* odpowie na pytanie (6) negatywnie (bądź odpowie na pytanie (7) lub (8) stwierdzając, iż ' $P^*a$ ' pociąga ' $\sim Q^*a$ '), to możemy to zinterpretować różnie.

Po pierwsze, może to świadczyć o tym, że osoba *O* odrzuca (1) — i tym samym hipoteza (2) jest obalona.

Po drugie, może to świadczyć o tym, że osoba *O* odrzuca (4) nie dlatego, że odrzuca (1), lecz dlatego, że odrzuca (5), tj. nie uznaje (4) za szczególny przypadek (1).

Przypuśćmy bowiem, że (1) i (4) są schematami odpowiednio zdań:

(9) Oczernianie innych ludzi jest złem.

(10) Osoba *A*, mówiąca coś złego o osobie *B*, jest godna potępienia.

Osoba *O* może deklarować odrzucenie tego, że (10), gdyż uważa, że:

(11) Oczernianie ludzi nie jest złem.

(12) Mówienie o kimś czegoś złego nie (zawsze) jest oczernianiem ludzi.

<sup>10</sup> Sytuację tę można równoważnie opisać mówiąc, że chodzi o ujawnienie przekonań danej osoby co do tego, czy zależność ' $\wedge x (Px \Rightarrow Qx)$ ' ma miejsce, czy nie.

<sup>11</sup> Bycie-szczególnym-przypadkiem w tym wypadku może polegać np. na tym, że predykaty ' $P^*$ ' i ' $Q^*$ ' są po prostu równozakresowe z predykatami ' $P$ ' i ' $Q$ ', bądź na tym, że zakresy predykatów ' $P^*$ ' i ' $Q^*$ ' są podziorami zakresów, odpowiednio, predykatów ' $P$ ' i ' $Q$ '.

Opisany eksperyment ujawnić może więc — w zależności od okoliczności — intuicje etyczne (co do tego, czy oczernianie jest złem) lub intuicje semantyczne (co do znaczenia słów „oczernianie”, „zło” *etc.*) uczestniczącej w nim osoby.

Dodać jeszcze należy, że osobą *O*, «na której» przeprowadza się eksperyment diagnostyczny, może być w szczególności... sam eksperymentator. Chodzi mu wtedy o ujawnienie własnych intuicji. Jeśli z kolei eksperymentator jest różny od osoby *O*, to może mieć różne cele, przede wszystkim perswazyjne. Chodzić może m.in. o:

- (a) uświadomienie osobie *O*, jakie żywi intuicje;
- (b) uświadomienie osobie *O* potrzeby zrewidowania żywionych intuicji;
- (c) uświadomienie osobie *O*, że jej «ukryte» intuicje nie pokrywają się z jej poglądami głoszonymi *explicite*.

## 5. TOR ŚMIERCI

Ograniczymy się tutaj do zreferowania najbardziej znanego eksperymentu myślowego, należącego do eksperymentów diagnostycznych, a mianowicie do eksperymentu z torem śmierci, mającego zresztą rozmaite warianty. Przytoczmy go w wersji Thomson:

Załóżmy, że jesteś maszynistą tramwaju. Tramwaj zakręca i jedzie wprost na pięciu robotników, którzy remontują tory. Tory prowadzą w tym miejscu przez dolinkę o bardzo stromych zboczach, więc żeby nie najechać na robotników, musisz zatrzymać tramwaj. Naciskasz na hamulec, ale okazuje się, że nie działa. Nagle dostrzegasz jednak bocznice prowadzącą w prawo. Możesz w nią skrócić i w ten sposób ocalić pięciu robotników stojących na torze przez tobą. Niestety [...] na bocznym torze pracuje jakiś robotnik. Nie zdąży on zbiec z toru na czas, więc jeśli skierujesz na niego tramwaj, zabijesz go (Thomson 1985: 94).

Opisana wyżej sytuacja — to baza eksperymentu; nazwijmy ją ‘*B*’.

Pytanie diagnostyczne brzmieć może różnie, np.:

- (1) Czy jeżeli znalazłbyś się w sytuacji *B*, to skrzyłbyś w prawo?
- (2) Co zrobiłbyś w sytuacji *B*?
- (3) Czy w sytuacji *B* skrzyłbyś w prawo, czy nie?
- (4) Co powinno się zrobić w sytuacji *B*?
- (5) Czy w sytuacji *B* powinno się / wolno skrzyć w prawo?

Załóżmy, że chodzi o pytanie (5). Zauważmy, że może być ono traktowane jako szczególny wypadek pytań dotyczących wielu różnych norm etycznych, np.:

- (6) Czy jeżeli ktoś może ocalić życie pięciu osób, poświęcając życie jednej osoby, to powinien / może poświęcić życie jednej osoby?
- (7) Czy powinno się / wolno kogoś zamordować, by uratować życie komuś innemu?
- (8) Czy w każdych okolicznościach powinno się / wolno wybierać czyny, które minimalizują ogół cierpienia?

Zauważmy, że czyjeś intuicje językowe mogą być takie, że nie uznałby (5) za szczególny przypadek (6), (7) lub (8) (może np. nie uznać przestawienia zwrotnicy w celu skierowania pociągu na jakąś osobę za zamordowanie tej osoby).

Negatywna odpowiedź na (5) może jednak świadczyć o tym, że osoba, z którą się eksperymentuje, odrzuca (6), (7) lub (8).

Dodajmy, że postawienie kogoś przed koniecznością odpowiedzi na (5) pozwala na uświadomienie sobie, że pewne spontanicznie uznawane normy pozostają względem siebie w konflikcie — np. normy (6)-(8) są w konflikcie z normą „Nie powinno się (nikogo) zabijać”.

## 6. SEMANTYCZNE EKSPERYMENTY MYŚLOWE

Bywają eksperymenty myślowe, których bezpośrednim celem jest ujawnienie intuicji semantycznych — dotyczących denotacji lub konotacji poszczególnych terminów.<sup>12</sup>

Załóżmy, że chodzi o ustalenie konotacji terminu  $T$ , przy czym są dwie «kandydatury» na tę konotację —  $K_A$  i  $K_B$  — gdyż:

(1) Każdy znany desygnat terminu  $T$  posiada zarówno własności składające się na  $K_A$ , jak i własności składające się na  $K_B$ .

(2) Żaden znany nie-desygnat terminu  $T$  nie posiada łącznie własności składających się na  $K_A$  ani własności składających się na  $K_B$ .

Inaczej mówiąc, zarówno  $K_A$ , jak i  $K_B$ , dobrze wyznaczają zbiór znanych desygnatów terminu  $T$ .

Pytanie, która z proponowanych konotacji jest istotnie konotacją terminu  $T$ , można niekiedy rozstrzygnąć przez odwołanie się do eksperymentu myślowego. Załóżmy, że chcemy potwierdzić hipotezę semantyczną, zgodnie z którą  $K_A$  jest konotacją  $T$ . Wyobrażona baza takiego eksperymentu wskazywałaby sytuację, w której:

(3) Pewien desygnat terminu  $T$  nie posiada jakiegoś elementu  $K_B$ , bądź:

(4) Pewien przedmiot posiada własność  $K_B$ , ale nie jest desygnatem terminu  $T$ .

Baza byłaby więc pewnym «granicznym» przypadkiem użycia terminu  $T$ .<sup>13</sup>

<sup>12</sup> Poza omówionymi niżej — do eksperymentów semantycznych należą m.in. takie, jak Locke'a z księciem i szewcem (1690: 479-280), Zenona z Elei z ciągłością czasoprzestrzeni — a w szczególności tzw. paradoksów dychotomii, Achillesa i żółwia oraz strzały i stadionu (Ajdukiewicz 1931), Quine'a z uprzywilejowaną logiką języka (1960: 58-59) i Fregego z prawdziwością (1893: 14-15).

<sup>13</sup> Dodajmy od razu, że ktoś, kto chciałby kontrargumentować, mógłby kontrargumentować — i utrzymywać, że wskazany w (a) przedmiot wcale nie byłby desygnatem terminu  $T$ , a w wypadku (b) — że wskazany przedmiot do denotacji  $T$  należy. Swoje stanowisko uzasadniać mógłby jedynie, podobnie zresztą jak  $A$ , swoimi i cudzymi intuicjami pojęciowymi. Pamiętajmy, że reguły semantyczne — wiążące wyrażenia z ich konotacjami — są zawsze konwencjonalne i mogłyby być dowolne, gdyby nie to, że niekiedy jest pożądane, by odpowiadały intuicjom użytkowników języka. Do żadnej wyższej instancji — poza intuicją tych użytkowników — w opisanym sporze odwołać się nie można.

Eksperymenty myślowe służą w takich wypadkach do uzasadniania pewnych twierdzeń analitycznych i przyjmowanych reguł semantycznych.

## 7. STATEK TEZEUSZA

Zacznijmy od eksperymentu ze statkiem Tezeusza, znanym w starożytności, a tak opisanym i uzupełnionym przez Hobbesa:

Jakieś ciało można porównywać z nim samym w coraz to innym czasie. Stąd między filozofami zrodził się wielki spór o *zasadę indywidualizacji* [...], a mianowicie o to, w jakim znaczeniu jakieś ciało raz może być uważane za tożsame, a kiedy indziej za różne od tego, jakim było uprzednio. [...] Wedle pierwszego zdania nie byłby tym samym człowiek, który grzeszy i który podlega karze, a to z tej racji, że ciało ludzkie nieustannie podlega zmianom, i nie byłoby tożsame państwo, które w jednym wieku pewne prawo ustanowiło, a w innym je odwołuje [...]. Wedle drugiego zdania dwa ciała istniejące jednocześnie byłyby jednym i tożsamym co do liczby; oto bowiem na przykład ów okręt Tezeusza byłby tożsamy co do liczby z tym, jakim był od początku, po zmianie wszystkich desek (kiedyś sofisci ateńscy dysputowali o tym, czy będzie on tożsamy, czy inny, gdy stopniowo będzie się wyjmowało z niego stare deski i na ich miejsce będzie się wstawiało wciąż nowe). Otóż gdyby ktoś zachowywał owe stare deski w mare, jak się je wyjmuje, i z zachowanych, zestawiając je w tym samym porządku, później na nowo zbudował okręt, to jest rzeczą niewątpliwą, iż ten okręt byłby tożsamy co do liczby z tym, jaki był na początku; i mielibyśmy dwa okręty tożsame co do liczby, co jest rzeczą jak najbardziej niedorzeczną. Wedle trzeciego zdania żadna w ogóle rzecz nie jest tożsama z tym, czym była; tak że nawet człowiek, który stoi nie byłby tożsamy z człowiekiem, które poprzednio siedział, i nie byłaby tożsama woda, która znajduje się w naczyniu i która niebawem zostaje wylana. Tak więc należy uważać, że zasada indywidualizacji nie zawsze spoczywa w samej materii ani też nie zawsze w samej formie. [...] Okręt, o ile tym mianem oznacza się materię w pewien sposób ukształtowaną, będzie ten sam, jeśli cała materia jest ta sama; jeśli zaś żadna część materii nie jest ta sama, to całkiem różny co do liczby będzie też i okręt; jeśli zaś pozostała część materii, część zaś się zmieniła, to okręt będzie w części ten sam, w części zaś różny (Hobbes 1642: 156-158).

Rozważmy teraz następujące (nie wszystkie!) znaczenia „tożsamości”:

- (1)  $\forall x \forall y [x \text{ jest tożsamy}_M z y \Leftrightarrow \wedge C (C \text{ jest częścią } x \wedge C \text{ jest częścią } y)]$ .
- (2)  $\forall x \forall y [x \text{ jest tożsamy}_F z y \Leftrightarrow \forall f (x \text{ ma formę } f \wedge y \text{ ma formę } f)]$ .
- (3)  $\forall x [x \text{ zachowuje tożsamość}_G \text{ w okresie } T \Leftrightarrow \wedge t_1 \wedge t_2 (t_1, t_2 \in T \Rightarrow x \text{ w chwili } t_1 \text{ jest tożsamy}_M \text{ i tożsamy}_F z x \text{ w chwili } t_2)]$ .

Jeżeli teraz założymy, że tożsamość<sub>M</sub> (co do części) i tożsamość<sub>F</sub> (co do formy) są wzajemnie niezależne, to np. zarazem:

- (4) Jeżeli *A* i *B* mają różne części, ale tę samą formę (kształt), to *A* i *B* są tożsame<sub>F</sub>, ale nietożsame<sub>M</sub>.
- (5) Jeżeli *A* i *B* mają te same części i tę samą formę (kształt), to *A* i *B* są zarówno tożsame<sub>F</sub>, jak i tożsame<sub>M</sub>.
- (6) Jeżeli *A* w chwili *t*<sub>1</sub> jest tożsamy<sub>F</sub> i tożsamy<sub>M</sub> z *A* w chwili *t*<sub>3</sub>, ale nie jest taki w chwili *t*<sub>2</sub> (późniejszej od *t*<sub>1</sub> i wcześniejszej od *t*<sub>3</sub>), to *A* nie zachowuje tożsamości<sub>G</sub> w okresie *t*<sub>1</sub>-*t*<sub>3</sub>.



Uproszczone przykłady tych sytuacji są różne przypadki «modyfikacji» statku Tezeusza. Przyjmijmy za bazę eksperymentu stan rzeczy, w którym nastąpiła wymiana wszystkich desek — z zachowaniem kształtu całości. W tych okolicznościach statek utraciłby tożsamość<sub>M</sub> (ale zachowałby tożsamość<sub>F</sub>). Statek zestawiony z desek usuniętych przy tamtej wymianie byłby tożsamy<sub>M</sub> (i ewentualnie tożsamy<sub>F</sub>) ze statkiem przed «modyfikacją», ale utraciłby tożsamość<sub>G</sub>.

Zauważmy, że wszelkie znane nam z codziennego doświadczenia przedmioty, które są tożsame<sub>M</sub>, są także tożsame<sub>F</sub>. Dopiero przypadek graniczny, jakim jest «rozegrany» statek Tezeusza uwiadamia potrzebę rozróżnienia dwóch pojęć: tożsamości<sub>M</sub> i tożsamości<sub>F</sub>. Potraktowany jako eksperyment diagnostyczny pozwala natomiast danej osobie uświadomić sobie, które z tych pojęć odpowiada jej intuicjom związanym z terminem „tożsamość”.

## 8. UZASADNIONE PRZEKONANIE

Opis eksperymentu z uzasadnionym przekonaniem pochodzi od Gettier'a. Przyjmiemy się następującemu opisowi tego eksperymentu, który — jak się wydaje — jest skądinąd przykładem eksperymentu nieudanego:

W ostatnich latach podejmowano różne próby ustalenia koniecznych i wystarczających warunków tego, że ktoś wie, że jest tak jak mówi dane zdanie. Dotychczasowe próby można przedstawić w następującej formie [...]:

(a) *S* wie, że *p*

wtedy i tylko wtedy, gdy

(i) '*p*' jest prawdziwe

(ii) *S* jest przekonany, że *p*

oraz

(iii) *S*-a przekonanie, że *p*, jest uzasadnione. [...]

Będę argumentował, że (a) jest fałszywe, ponieważ warunki tam ustalone nie są WYSTARCZAJĄCE dla prawdziwości zdania „*S* wie, że *p*”. Argumentując w ten sam sposób pokażę, że (b) i (c) nie zachodzą, jeśli wyrażenia: „jest oczywiste dla” lub „ma prawo być pewnym, że” będą zastąpione przez wyrażenie: „przekonanie (że) jest uzasadnione”. Zauważmy najpierw dwie rzeczy. Po pierwsze, w tym sensie „uzasadnienia”, w którym bycie *S*-a w sposób uzasadniony przekonany, że *p*, jest koniecznym warunkiem wiedzy *S*-a, że *p*, istnieje możliwość, że ktoś będzie przekonany, że uzasadnione jest zdanie, które faktycznie jest fałszywe. Po drugie, dla dowolnego zdania '*p*', jeśli przekonanie *S*-a, że *p*, jest uzasadnione i '*p*' pociąga za sobą '*q*' i *S* wyprowadza '*q*' z '*p*' i przyjmuje '*q*' jako wynik tej dedukcji, to przekonanie *S*-a, że *q*, jest również uzasadnione. Pamiętając o tych dwóch rzeczach, przedstawię teraz [...] [przykłady sytuacji], w których warunki ustalone w (a) są spełnione dla pewnego zdania, chociaż jednocześnie fałszem jest, że dana osoba w danej sprawie posiada wiedzę. [...]

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

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

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

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

Przypuśćmy teraz, że Smith dostrzega wynikanie (e) z (d) i uznaje (e) na podstawie (d), które jest dla niego zupełnie oczywiste. W tym przypadku przekonanie Smitha, że (e) jest prawdziwe, istotnie jest uzasadnione.

Lecz wyobraźmy sobie dalej, że to właśnie Smith (a nie Jones) dostanie posadę, choć nie będzie o tym wiedział, oraz że to on będzie miał dziesięć monet w kieszeni i tego również nie będzie wiedział. Wówczas zdanie (e) jest prawdziwe, chociaż zdanie (d), z którego Smith wyprowadził (e), jest fałszywe. W naszym przykładzie wobec tego, to wszystko jest prawdziwe: (i) (e) jest prawdziwe, (ii) Smith jest przekonany, że (e) jest prawdziwe oraz (iii) przekonanie Smitha, że (e) jest uzasadnione. Jednakże jasne jest też, że Smith *nie wie*, że (e) jest prawdziwe; (e) jest bowiem prawdziwe z racji ilości monet w kieszeni Smitha, gdy tymczasem Smith nie wie, ile monet jest w kieszeni Smitha, i opiera swoje przekonanie, że (e), na obliczeniu monet w kieszeni Jonesa, co do którego jest fałszywie przekonany, że jest tym człowiekiem, który otrzyma posadę (Gettier 1963: 94-95).

A oto nasza rekonstrukcja.

«Klasyczna» definicja „wiedzy” jest następująca:

(1) Osoba *O* wie, że *p*  $\Leftrightarrow$  [zdanie '*p*' jest prawdziwe  $\wedge$  (osoba *O* jest przekonana, że *p*  $\wedge$  przekonanie osoby *O*, że *p*, jest uzasadnione)].

Eksperyment Gettier'a to taka odmiana eksperymentu semantycznego, w której «obala» się (w intencji) pewną hipotezę semantyczną, nie proponując żadnej w zamian. Bazą eksperymentu jest sytuacja opisywana przez zdania (2)-(3):

(2) Osoba *O* jest przekonana, że *q*  $\wedge$  przekonanie osoby *O*, że *q*, jest uzasadnione.

(3) Osoba *O* wywnioskowuje '*p*' z '*q*'.

Gettier zdaje się ponadto zakładać, że:

(4)  $q \Rightarrow p$ .

(5) [(Osoba *O* jest przekonana, że *q*  $\wedge$  przekonanie osoby *O*, że *q*, jest uzasadnione)  $\wedge$  (osoba *O* wywnioskowuje '*p*' z '*q*'  $\wedge$   $q \Rightarrow p$ )]  $\Rightarrow$  (osoba *O* jest przekonana, że *p*  $\wedge$  przekonanie osoby *O*, że *p*, jest uzasadnione).

Czyli: jeżeli coś wywnioskujemy w sposób niezawodny z zasadnie uznanej racji, to nasz wniosek jest uznany zasadnie.

Z (2)-(5) wynika, że:

(6) Osoba *O* jest przekonana, że *p*  $\wedge$  przekonanie osoby *O*, że *p*, jest uzasadnione.

Założmy teraz, że:

(7) Zdanie '*p*' jest prawdziwe.

Z (1), (6) i (7) wynika, że:

(8) Osoba *O* wie, że *p*.

Rozumowanie Gettier'a miało — w jego intencji — pokazać, że to, że *p*, może zachodzić i można mieć uzasadnione przekonanie, że *p*, a mimo to nie wiedzieć, że *p*. Otóż Gettier uważa, że zdanie (d) pociąga zdanie (e). Zdanie (d) ma formę:

(9)  $Pa \wedge Qa$ ,

gdzie ' $P$ ' to tyle, co „jest tym, który dostanie posadę”, ' $Q$ ' — „ma dziesięć monet w kieszeni”, a ' $a$ ' oznacza Jonesa.

Zdanie (e) można zaś interpretować na kilka sposobów:

(10)  $\wedge x (Px \Rightarrow Qx)$ .

(11)  $\wedge x (Qx \Rightarrow Px)$ .

(12)  $\wedge x (Px \Leftrightarrow Qx)$ .

Żadna z formuł (10)-(12) nie wynika jednak z formuły (9) — bez dodatkowych założeń, a zatem o osobie  $O$  (i o nikim w ogóle) nie można powiedzieć, że ma uzasadnione przekonanie co do którejkolwiek z formuł (10)-(12), a więc nie dotyczy jej formuła (5), czyli nie żywi ona uzasadnionego przekonania, że  $p$ .

Założmy więc, że dokładnie jedna osoba dostanie posadę:

(13)  $\wedge x \wedge y [(Px \wedge Py) \Rightarrow x = y]$ .

Odpadają dzięki temu interpretacje (11) i (12) (gdyż co najmniej dwie osoby mają w kieszeni dziesięć monet: Jones i Smith) — pozostaje tylko interpretacja (10). Aby jednak uznać (10) na podstawie uznania formuły (9), trzeba jeszcze uznać, że:

(14)  $x = a$ .

Tymczasem osoba  $O$  nie ma dostatecznych podstaw do uznania formuły (14),<sup>14</sup> która zresztą jest fałszem, gdyż:

(15)  $x = b$ ,

gdzie ' $b$ ' oznacza Smitha.

W tej sytuacji osoba  $O$  myliłaby się sądząc, że jej przekonanie, że (10) jest uzasadnione. Opisana sytuacja — wbrew intencji Gettier'a — nie obala więc formuły (1), tj. krytykowanej definicji „wiedzy”.<sup>15</sup>

## 9. ZIEMIA BLIŻNIACZA

Eksperyment z Ziemią Bliźniaczą opisany jest przez swego autora, Putnama, w sposób następujący:

W celu wykazania, że stan psychiczny nie wyznacza ekstensji, posłużę się odrobiną science fiction. Na użytek przykładów z dziedziny science fiction, które za chwilę wystąpią, przypuśćmy, że gdzieś w naszej galaktyce znajduje się planeta zwana Ziemią Bliźniaczą. Ziemia Bliźniacza jest bardzo podobna do Ziemi; ludzie na Ziemi Bliźniaczej nawet mówią *po polsku*. Pomijając różnice, które zostaną wymienione w naszych przykładach z dziedziny science fiction, czytelnik może założyć, że Ziemia Bliźniacza jest *dokładnie* taka sama, jak Ziemia. Może nawet założyć, jeżeli zechce, że ma tam swojego *Doppelgänger* — identyczną kopię siebie samego — aczkolwiek dla moich bajeczek jest to obojętne.

<sup>14</sup> Smith nie może za taką podstawę uważać tego, że „prezes spółki zapewnił go, że w końcu zostanie wybrany Jones”.

<sup>15</sup> Warto dodać, że inne wersje tego eksperymentu, w których występują przykłady zdań z alternatywą logiczną, są raczej argumentacją przeciw zwyczajowo przyjętej w logice interpretacji naturalnojęzykowej alternatywy.

Wprawdzie niektórzy ludzie na Ziemi Bliźniaczej mówią po polsku, istnieją jednak, co nie jest niespodzianką, drobne różnice między polskim dialektem występującym na Ziemi Bliźniaczej a standardowym językiem polskim. Owe różnice wynikają z pewnych osobliwości Ziemi Bliźniaczej.

Jedną z nich jest okoliczność, że płyn zwany na Ziemi Bliźniaczej „wodą” nie jest  $H_2O$ , ale jakimś innym płynem o składzie chemicznym wyrażającym się bardzo długim i skomplikowanym wzorem. Będziemy w skrócie zapisywać go jako XYZ. Założmy, że XYZ jest nieodróżnialny od wody w zwykłym zakresie temperatury i ciśnienia. W szczególności, smakuje jak woda i gasi pragnienie jak woda. Założmy również, że jeziora, morza i oceany Ziemi Bliźniaczej są wypełnione XYZ, a nie wodą, że tamtejsze deszcze składają się z kropeł XYZ, a nie wody, itd.

Kiedy statek kosmiczny z Ziemi odwiedzi Ziemię Bliźniaczą, przybysze z początku będą mniemać, że wyraz „woda” ma takie samo znaczenie na Ziemi i na Ziemi Bliźniaczej. Owo przypuszczenie zostanie skorygowane z chwilą, gdy odkryją, że „woda” na Ziemi Bliźniaczej jest XYZ, co Ziemianie stwierdzą mniej więcej za pomocą słów:

„Na Ziemi Bliźniaczej wyraz «woda» znaczy XYZ”.

(Na marginesie: powyższy sposób użycia słowa „znaczy” jest zgodny z doktryną, w myśl której ekstensja jest jednym z sensów wyrazu „znaczenie”. Zauważ jednak, że choć „znaczy” znaczy w przytoczonym przykładzie coś w rodzaju *ma ekstensję*, nikt nie twierdziłby:

„Na Ziemi Bliźniaczej znaczeniem wyrazu «woda» jest XYZ”.

chyba że — ewentualnie — fakt, iż „woda jest XYZ”, byłby znany wszystkim mówiącym po polsku dorosłym mieszkańcom Ziemi Bliźniaczej. Rzecz wyjaśni się na gruncie teorii znaczenia rozwijanej poniżej; na razie zaznaczymy tylko, że choć czasownik „znaczy” niekiedy znaczy: „ma ekstensję”, rzeczownik odsłowny „znaczenie” *nigdy* nie znaczy: „ekstensja”).

Na odwrót, kiedy statek kosmiczny z Ziemi Bliźniaczej odwiedzi Ziemię, przybysze z początku będą mniemać, że wyraz „woda” ma takie samo znaczenie na Ziemi Bliźniaczej i na Ziemi. Owo przypuszczenie zostanie skorygowane z chwilą, gdy odkryją, że „woda” na Ziemi jest  $H_2O$ , co Bliźniaczanie stwierdzą mniej więcej takimi słowami:

„Na Ziemi<sup>16</sup> wyraz «woda» znaczy  $H_2O$ ”.

Zauważ, że nie ma żadnego problemu z ekstensją nazwy „woda”. Wyraz ten ma po prostu dwa różne znaczenia (jak twierdzę): w sensie używanym na Ziemi Bliźniaczej, w sensie wody<sub>ZB</sub>, to co *my* nazywamy „wodą” po prostu nie jest wodą; podczas gdy w sensie używanym na Ziemi, w sensie wody<sub>Z</sub>, to co Bliźniaczanie nazywają „wodą” po prostu nie jest wodą. Ekstensją „wody” w sensie wody<sub>Z</sub> jest zbiór wszystkich całości złożonych z cząsteczek  $H_2O$ , lub coś w tym rodzaju; ekstensją wody w sensie wody<sub>ZB</sub> jest zbiór wszystkich całości złożonych z cząsteczek XYZ, lub coś w tym rodzaju.

Cofnijmy teraz czas mniej więcej do roku 1750. Chemia nie była wówczas jeszcze rozwinięta ani na Ziemi, ani na Ziemi Bliźniaczej. Typowy ziemski Polak nie wiedział, że woda składa się z wodoru i tlenu, a typowy Polak bliźniaczański nie wiedział, że „woda” składa się z XYZ. Niech Oskar<sub>1</sub> będzie takim typowym ziemskim Polakiem, a Oskar<sub>2</sub> jego odpowiednikiem na Ziemi Bliźniaczej. Można założyć, że wszystkie przekonania Oskara<sub>1</sub> na temat wody Oskar<sub>2</sub> żywi w stosunku do „wody”. Można nawet założyć, jeżeli sobie życzysz, że Oskar<sub>1</sub> i Oskar<sub>2</sub> są dokładnie jednakowi pod względem wyglądu zewnętrznego, uczuć, myśli, monologów wewnętrznych itd. Niemniej ekstensją nazwy „woda” na Ziemi w 1750 roku była, tak sa-

<sup>16</sup> Lub stwierdzą: „Na Ziemi Bliźniaczej (*Bliźniaczańska nazwa planety Terra — H.P.*) wyraz „woda” znaczy  $H_2O$ ”.

mo jak w roku 1950,  $H_2O$ ; a ekstensją nazwy „woda” na Ziemi Bliźniaczej w 1750 roku była, tak samo jak w roku 1950, XYZ. Oskar<sub>1</sub> i Oskar<sub>2</sub> różnie rozumieli nazwę „woda” w 1750 roku, mimo że znajdowali się w jednakowym stanie psychicznym, i mimo że, biorąc pod uwagę ówczesny stan nauki, ich społeczności uczonych dopiero pięćdziesiąt lat później odkryją, że rozumieli oni nazwę „woda” różnie. Zatem ekstensja nazwy „woda” (i jej „znaczenie”, w intuicyjnym, przedanalizycznym użyciu tego słowa) *nie* jest funkcją stanu psychicznego.

Można wszakże zapytać, dlaczego mielibyśmy zgodzić się z twierdzeniem, że nazwa „woda” ma jednakową ekstensję w roku 1750 i 1950 (na obu Ziemiach). Logika nazw oznaczających rodzaje naturalne, jak „woda”, stanowi skomplikowane zagadnienie, niemniej poniżej przedstawię zarys odpowiedzi. Przypuśćmy, że pokazuję palcem na szklankę wody i mówię: „ten płyn nazywa się «woda»” (lub „to się nazywa «woda»”, jeżeli charakteryzator „płyn” jasno wynika z kontekstu). Moja „definicja ostensywna” wody zakłada następującą presupozycję empiryczną: porcja płynu, na którą pokazuję, pozostaje w pewnej relacji jednakowości (*sameness relation*, *x jest tym samym płynem, co y*, lub *x i y są jednakowe*) z większością tworów, które ja i inni członkowie mojej społeczności językowej przy innych okazjach nazywali „wodą”. Jeżeli ta presupozycja jest fałszywa, dajmy na to, jeżeli szklanka, o czym nie wiem, jest napełniona ginem, a nie wodą, wówczas nie oczekuję akceptacji mojej definicji ostensywnej. Zatem definicja ostensywna przynosi coś, co można by nazwać odwoływalnym warunkiem koniecznym i wystarczającym: warunkiem koniecznym i wystarczającym tego, aby być wodą, jest znajdowanie się w relacji jednakowości z płynem w tej oto szklance; co wszakże jest warunkiem koniecznym i wystarczającym tylko wtedy, gdy wspomniana presupozycja empiryczna jest spełniona. W przeciwnym razie zostaje uruchomiony któryś z szeregu, że tak powiem, warunków „zapasowych”.

Kluczowa sprawa polega na tym, że relacja jednakowości<sub>L</sub> jest relacją *teoretyczną*: rozstrzygnięcie, czy dwa płyny są jednakowe, czy nie, może być kwestią nie kończących się badań naukowych. Co więcej, nawet jeśli, dzięki badaniu naukowemu lub na skutek zastosowania jakiegoś „zdroworozsądkowego” testu, uzyska się „ostateczną” odpowiedź, będzie ona *odwoływalna*: dalsze badanie może wywrócić na nice najbardziej „pewny” przykład. Toteż fakt, że Polacy w 1750 roku mogliby nazywać XYZ „wodą”, a w 1800 czy 1850 już tak by jej nie nazywali, nie dowodzi, że przez ten czas dla przeciętnego Polaka zmieniło się „znaczenie” słowa „woda”. W 1750 roku, lub w 1850 czy w 1950, można było wskazać na, powiedzmy, płyn w dolnym biegu Wisły jako na przykład „wody”. Zmieniłoby się tylko tyle, że w 1750 roku fałszywie sądziłoby się, że XYZ znajduje się w relacji jednakowości<sub>L</sub> z płynem w dolnym biegu Wisły, podczas gdy w 1800 czy w 1850 roku wiedzielibyśmy, że tak nie jest (pomijam fakt, że w drugiej połowie XX wieku płyn w dolnym biegu Wisły jest tylko mniej więcej wodą<sup>17</sup>) (Putnam 1975: 105-109).

W naszym przekonaniu za całą tą długą historyjką kryje się rozumowanie następujące.

Założmy, że:

(1) *A* i *B* używają wyrażenia *W* w odniesieniu do przedmiotu *x* po to, aby przedmiotowi *x* przypisywać konotację *K*.

<sup>17</sup> W oryginale zamiast o mówiących po polsku jest mowa o „English speakers” (mówiących po angielsku), a zamiast o dolnym biegu Wisły, o Jeziorze Michigan. Dlatego nie należy tego fragmentu taktować jako wyraz poglądu Putnama na temat stanu czystości wód w polskich rzekach. — Przyp. tłum.



(2) *A* używa wyrażenia *W* po to, aby wskazywać przedmioty należące do denotacji *D<sub>A</sub>*.

(3) *B* używa wyrażenia *W* po to, aby wskazywać przedmioty należące do denotacji *D<sub>B</sub>*.

(4)  $D_A \neq D_B$ .<sup>18</sup>

Założmy jeszcze, że:

(5) Znaczenie wyrażenia *W* = konotacja wyrażenia *W*.

Przy tym założeniu z (1) wynika, że:

(6) *A* i *B* używają wyrażenia *W* w tym samym znaczeniu.

Założmy teraz, że:

(7) Znaczenie wyrażenia *W* = denotacja wyrażenia *W*.

Przy tym założeniu z (2)-(4) wynika, że:

(8) *A* i *B* używają wyrażenia *W* w różnych znaczeniach.

Po pierwsze więc, okazuje się, że wyrażenie *W* — przy różnych znaczeniach słowa „znaczenie” — może być używane «zarazem» w tym samym znaczeniu (przy jednym znaczeniu „znaczenia”) i w różnych znaczeniach (przy różnych znaczeniach „znaczenia”).

Po drugie, jest jasne, że rozstrzygnięcie, czy rzeczywiście jest tak, jak głosi założenie (4) (nawiasem mówiąc to samo dotyczy pozostałych założeń), co najmniej niekiedy nie jest proste, i wobec tego trzeba się liczyć z pomyłką.

Tutaj — znowu — do samego przeprowadzenia tego (zgódźmy się dość banalnego) rozumowania nie potrzebne są żadne „bajeczki” o Ziemi Bliźniaczej, które pełnią wyłącznie funkcje ilustracyjną.

## 10. CHIŃSKI POKÓJ

Eksperyment z chińskim pokojem jest autorstwa Searle’a, który pisze następująco:

Mieć umysł, to coś więcej niż realizować formalne czy syntaktyczne operacje. Nasze stany umysłowe, na mocy definicji, mają zawsze jakąś treść. [...] Jeżeli myśl jest zawsze myślą o CZYMŚ, przeto dany ciąg symboli musi mieć jakieś ZNACZENIE, by stać się myślą. [...] Powód, dla którego komputerowy program nie może być umysłem, jest prosty; komputerowy program ma cechy syntaktyczne, umysły mają coś więcej niż syntaktyka. [...]

By zilustrować tę tezę, proponuję rozważenie następującego obmyślonego przeze mnie eksperymentu myślowego. Wyobraźmy sobie, że zespół programistów napisał program, który jest w stanie symulować rozumienie języka chińskiego. Zatem, na przykład, jeśli komputerowi zadaje się pytanie w tym języku, może porównać je ze swoją pamięcią lub bazą danych i wyprodukować w języku chińskim odpowiedź na zadane pytanie. Założmy, dla celów dyskusji, że odpowiedzi te są tak dobre jak odpowiedzi osoby, której ojczystym językiem jest język chiński. Czy w takiej sytuacji komputer rozumie język chiński, czy rozumie go dokładnie tak, jak użyt-

<sup>18</sup> Podkreślimy, że jeśli na konotację danego wyrażenia nakłada się tylko warunek, aby była to własność istotna przedmiotów należących do denotacji tego wyrażenia, to wyrażenie o danej denotacji może mieć więcej niż jedną konotację.

kownicy języka chińskiego rozumieją swój ojczysty język? Wyobraźmy sobie teraz, że ktoś z nas jest zamknięty w pokoju, i że w pokoju tym jest szereg koszy wypełnionych znakami z języka chińskiego. Załóżmy, że osoba ta, podobnie jak autor myślowego eksperymentu, nie zna języka chińskiego, otrzymała jednak napisaną w jej ojczystym języku książkę reguł manipulowania znakami języka chińskiego. Reguły te opisują używanie symboli w sposób czysto formalny, opisują manipulowanie nimi w sposób syntaktyczny, nie semantyczny. Mogą mieć postać: „Wybierz ten znak z podwójnym zakreślasem z kosza numer jeden i połóż go za znakiem z dwoma zawijasami z kosza numer dwa”. Przyjmijmy teraz, że w pokoju pojawiają się jakieś nowe symbole, a osoba w nim siedząca otrzymuje instrukcje, jakie chińskie symbole ma wysłać z pokoju w odpowiedzi na te, które się pojawiły. Załóżmy, że siedzący w pokoju nie wie, iż wysyłane przez ludzi z zewnątrz do pokoju symbole nazywane są przez nich pytaniami, zaś symbole, które siedzący w pokoju wysyła na zewnątrz, nazywane są odpowiedziami na pytania. Przyjmijmy poza tym, że programiści napisali na tyle dobry program, zaś siedząca w pokoju osoba jest do tego stopnia dobra w manipulowaniu symbolami, że jej odpowiedzi są nieodróżnialne od odpowiedzi osoby faktycznie znającej język chiński. Zatem jakaś osoba zamknięta jest w pokoju, w którym wybiera symbole chińskie i wysyła je w odpowiedzi na inne pojawiające się w pokoju chińskie symbole. W sytuacji, jaką tu opisałem, nie ma możliwości, by w wyniku takiej manipulacji formalnie zdefiniowanymi symbolami nauczyć się języka chińskiego.

Meritum mojego myślowego eksperymentu jest następujące: realizując taki formalny komputerowy program, z punktu widzenia obserwatora z zewnątrz, zachowujemy się dokładnie tak, jak byśmy rozumieli język chiński, jednocześnie jednak nie znamy ani jednego słowa z tego języka naturalnego. Jeśli wykonanie komputerowego programu symulującego rozumienie języka chińskiego nie jest wystarczające dla nas, byśmy ten język rozumieli, NIE MOŻE BYĆ WYSTARCZAJĄCE TAKŻE DLA MASZYNY liczącej. Wyjaśnienie tego stanu rzeczy jest całkiem proste. Jeśli w takiej sytuacji człowiek nie rozumie chińskiego, to inne „komputery” go nie rozumieją, gdyż żaden komputer cyfrowy nie może, dzięki temu, że zrealizuje program, uzyskać zdolności, której nie uzyskuje człowiek. To wszystko, co ma komputer, i co my mamy siedząc w chińskim pokoju, jest jedynie formalnym programem, pozwalającym na manipulowanie niezinterpretowanymi chińskimi symbolami. Powtórzmy to — komputer opanowuje syntaktykę, nie zaś semantykę. Cały morał naszej historii o chińskim pokoju przypomina nam to, co już wiemy. Rozumienie języka — a z pewnością i posiadanie wszelkich innych stanów umysłowych — wymaga czegoś więcej niż garści symboli formalnych. Konieczna tu jest jeszcze interpretacja lub znaczenie związane z tymi symbolami. Komputer cyfrowy zaś, na mocy definicji, nie może mieć niczego więcej poza formalnymi symbolami, gdyż operacje komputerowe, jak to już powiedziano, definiuje się w terminach maszynowej zdolności do wykonywania programów. Programy natomiast mają czysto formalną charakterystykę, to znaczy nie mają semantycznych treści.

Możemy się przekonać o sile tego argumentu, jeśli porównamy sytuację, w której otrzymujemy pytania i udzielamy odpowiedzi w naszym ojczystym języku, z sytuacją, w której pytają nas i odpowiadamy w języku, w którym nie znamy znaczenia ani jednego słowa. Wyobraźmy sobie, że jesteśmy w chińskim pokoju, gdzie zadają nam w naszym ojczystym języku pytania dotyczące takich spraw, jak nasz wiek czy fakty z naszej biografii, i że odpowiadamy na te pytania. Jaka jest różnica pomiędzy posługiwaniem się językiem chińskim a posługiwaniem się językiem polskim? Otóż, jeśli ktoś nie zna języka chińskiego, a zna język polski, różnica jest oczywista. Rozumiemy pytania zadane po polsku, bo wyrażone są one symbolami, których znaczenie jest nam znane. Analogicznie, odpowiadając po polsku, produkujemy symbole, które mają dla nas znaczenie. W przypadku języka chińskiego zaś nic takiego nie występuje. W przy-

kownicy języka chińskiego rozumieją swój ojczysty język? Wyobraźmy sobie teraz, że ktoś z nas jest zamknięty w pokoju, i że w pokoju tym jest szereg koszy wypełnionych znakami z języka chińskiego. Założmy, że osoba ta, podobnie jak autor myślowego eksperymentu, nie zna języka chińskiego, otrzymała jednak napisaną w jej ojczystym języku książkę reguł manipulowania znakami języka chińskiego. Reguły te opisują używanie symboli w sposób czysto formalny, opisują manipulowanie nimi w sposób syntaktyczny, nie semantyczny. Mogą mieć postać: „Wybierz ten znak z podwójnym zakrętasem z kosza numer jeden i połóż go za znakiem z dwoma zawijasami z kosza numer dwa”. Przyjmijmy teraz, że w pokoju pojawiają się jakieś nowe symbole, a osoba w nim siedząca otrzymuje instrukcje, jakie chińskie symbole ma wysłać z pokoju w odpowiedzi na te, które się pojawiły. Założmy, że siedzący w pokoju nie wie, iż wysyłane przez ludzi z zewnątrz do pokoju symbole nazywane są przez nich pytaniami, zaś symbole, które siedzący w pokoju wysyła na zewnątrz, nazywane są odpowiedziami na pytania. Przyjmijmy poza tym, że programiści napisali na tyle dobry program, zaś siedząca w pokoju osoba jest do tego stopnia dobra w manipulowaniu symbolami, że jej odpowiedzi są nieodróżnialne od odpowiedzi osoby faktycznie znającej język chiński. Zatem jakaś osoba zamknięta jest w pokoju, w którym wybiera symbole chińskie i wysyła je w odpowiedzi na inne pojawiające się w pokoju chińskie symbole. W sytuacji, jaką tu opisałem, nie ma możliwości, by w wyniku takiej manipulacji formalnie zdefiniowanymi symbolami nauczyć się języka chińskiego.

Meritum mojego myślowego eksperymentu jest następujące: realizując taki formalny komputerowy program, z punktu widzenia obserwatora z zewnątrz, zachowujemy się dokładnie tak, jak byśmy rozumieli język chiński, jednocześnie jednak nie znamy ani jednego słowa z tego języka naturalnego. Jeśli wykonanie komputerowego programu symulującego rozumienie języka chińskiego nie jest wystarczające dla nas, byśmy ten język rozumieli, NIE MOŻE BYĆ WYSTARCZAJĄCE TAKŻE DLA MASZINY liczącej. Wyjaśnienie tego stanu rzeczy jest całkiem proste. Jeśli w takiej sytuacji człowiek nie rozumie chińskiego, to inne „komputery” go nie rozumieją, gdyż żaden komputer cyfrowy nie może, dzięki temu, że zrealizuje program, uzyskać zdolności, której nie uzyskuje człowiek. To wszystko, co ma komputer, i co my mamy siedząc w chińskim pokoju, jest jedynie formalnym programem, pozwalającym na manipulowanie nieinterpretowanymi chińskimi symbolami. Powtórzmy to — komputer opanowuje syntaktykę, nie zaś semantykę. Cały morał naszej historii o chińskim pokoju przypomina nam to, co już wiemy. Rozumienie języka — a z pewnością i posiadanie wszelkich innych stanów umysłowych — wymaga czegoś więcej niż garści symboli formalnych. Konieczna tu jest jeszcze interpretacja lub znaczenie związane z tymi symbolami. Komputer cyfrowy zaś, na mocy definicji, nie może mieć niczego więcej poza formalnymi symbolami, gdyż operacje komputerowe, jak to już powiedziano, definiuje się w terminach maszynowej zdolności do wykonywania programów. Programy natomiast mają czysto formalną charakterystykę, to znaczy nie mają semantycznych treści.

Możemy się przekonać o sile tego argumentu, jeśli porównamy sytuację, w której otrzymujemy pytania i udzielamy odpowiedzi w naszym ojczystym języku, z sytuacją, w której pytają nas i odpowiadamy w języku, w którym nie znamy znaczenia ani jednego słowa. Wyobraźmy sobie, że jesteśmy w chińskim pokoju, gdzie zadają nam w naszym ojczystym języku pytania dotyczące takich spraw, jak nasz wiek czy fakty z naszej biografii, i że odpowiadamy na te pytania. Jaka jest różnica pomiędzy posługiwaniem się językiem chińskim a posługiwaniem się językiem polskim? Otóż, jeśli ktoś nie zna języka chińskiego, a zna język polski, różnica jest oczywista. Rozumiemy pytania zadane po polsku, bo wyrażone są one symbolami, których znaczenie jest nam znane. Analogicznie, odpowiadając po polsku, produkujemy symbole, które mają dla nas znaczenie. W przypadku języka chińskiego zaś nic takiego nie występuje. W przy-

padku chińskiego posługujemy się po prostu formalnymi symbolami, zgodnie z komputerowym programem, jednak nie potrafimy przypisać tym elementom żadnego znaczenia.

Zarówno specjaliści od sztucznej inteligencji, jak i psychologowie oraz filozofowie rozmaicie odpowiadali na moje argumenty. Wszystkie te odpowiedzi miały wspólną cechę: wszystkie były chybione. Stanie się oczywiste, dlaczego były one takie, jeśli oprzemy nasze rozumowanie na prostej logicznej prawdzie, stwierdzającej, że syntaktyka jako taka nie wystarcza dla semantyki, i że komputery w związku z tym, że są komputerami, z definicji, mają tylko syntaktykę (Searle 1984: 28-30).

Mamy tu do czynienia z ilustracją następującego rozumowania.

Założmy, że:

(1) *A* rozumie język *J*, gdy zna znaczenie wyrażen języka *J*.

(2) Jeśli *A* dysponuje zestawem wyrażen języka *J* i zna reguły syntaktyczne budowania odpowiedzi właściwych na pytania zadawane w języku *J*, to *A* umie zbudować odpowiedzi właściwe na pytania zadawane w języku *J*.<sup>19</sup>

Zapytajmy:

(3) Czy jeśli *A* umie zbudować odpowiedź właściwą na pytania zadawane w języku *J*, to *A* rozumie język *J*?

Otóż z (3) i (1) wynika, że:

(4) Jeśli *A* umie zbudować odpowiedzi właściwe na pytania zadawane w języku *J*, to *A* zna znaczenie wyrażen języka *J*.

Formuła (4) jest wersją krytykowanej w eksperymencie kryterialnej definicji „rozumienia”.

Z (2) i (4) wynika, że:

(5) Jeśli *A* dysponuje zestawem wyrażen języka *J* i zna reguły syntaktyczne budowania odpowiedzi właściwych na pytania zadawane w języku *J*, to *A* zna znaczenie wyrażen języka *J*.

Założmy, że zbudowaliśmy *A*, które dysponuje zestawem wyrażen języka *J* i zna reguły syntaktyczne budowania odpowiedzi właściwych na pytania zadawane w języku *J*, a nie zna znaczenia wyrażen języka *J*.

Zatem:

(6) Istnieje *A*, które dysponuje zestawem wyrażen języka *J* i zna reguły syntaktyczne budowania odpowiedzi właściwych na pytania zadawane w języku *J*, a nie zna znaczenia wyrażen języka *J*.

Formuła (6) to — ujęta w zacytowanym opisie w atrakcyjną formę — baza eksperymentu myślowego.

Z (6), (2) i (1) zaś wynika, że:

(7) Istnieje *A*, które umie zbudować odpowiedzi właściwe na pytania zadawane w języku *J* i nie rozumie języka *J*.

Zatem odpowiedź na pytanie (3) jest negatywna.

<sup>19</sup> Odpowiedź właściwa na dane pytanie to dowolne zdanie będące podstawieniem danej pytania, przy czym dana pytania o postaci „Dla którego *x* jest tak, że *Px*?” to funkcja „*Px*”. Nie każda odpowiedź właściwa jest, rzecz jasna, prawdziwa.



Oczywiste jest jednak przy tym, że nie samo założenie, że (6), tylko zbudowanie *A*, o którym mowa w (6) daje gwarancję tego, że odpowiedź na pytanie (3) jest negatywna.

## 11. MÓZG W NACZYNIU

Szczególną odmianą eksperymentu semantycznego jest eksperyment Putnama z mózgiem w naczyniu, który miałby stanowić argument przeciwko epistemologicznemu pan-iluzjonizmowi — przez wykazanie, że teza ta nie da się w żaden sposób uzasadnić.

Przytoczmy istotne fragmenty opisu tego eksperymentu:

Oto dyskutowana przez filozofów możliwość rodem z *science fiction*: wyobraźmy sobie, że pewien człowiek (Czytelnik może wyobrazić sobie siebie w tej roli) poddał się operacji wykonanej przez niegodziwego uczonego. Jego (twój) mózg został usunięty z ciała i umieszczony w naczyniu wypełnionym pożywką, która podtrzymuje mózg przy życiu. Zakończenia nerwowe zostały podłączone do superkomputera, który powoduje, że osoba, której mózg wyjęto, doświadcza iluzji, iż wszystko jest w najlepszym porządku. Ma ona złudzenie istnienia osób, przedmiotów, nieboskonu itd.; natomiast w rzeczywistości wszystko, czego ów człowiek doznaje (ty doznajesz), jest następstwem impulsów elektronicznych, płynących od komputera do zakończeń nerwowych. Komputer jest tak sprytnie zaprogramowany, że kiedy ofiara eksperymentu usiłuje podnieść rękę w górę, dzięki sprzężeniu zwrotnemu „widzi” i „czuje”, że ręka podnosi się w górę. Odpowiednio modyfikując program, niegodziwy uczony może spowodować, że jego ofiara „doświadczy” (lub dozna złudzenia) uczestnictwa w dowolnie zaaranżowanej przez niegodziwca sytuacji lub obecności w dowolnie zaprojektowanym otoczeniu. Może on również wymazywać pamięć mózgu zamkniętego w naczyniu tak, że nieszczęsnej ofierze będzie się wydawało, że zawsze przebywała w otoczeniu przeznaczonym jej przez eksperymentatora. Ofierze może nawet wydawać się, że siedzi i czyta te słowa o zabawnym, lecz zgoła absurdalnym domniemaniu, iż pewien niegodziwy uczony usuwa ludziom mózgi i umieszcza je w naczyniu z pożywką, która podtrzymuje je przy życiu. Zakończenia nerwowe są jakoby podłączone do superkomputera, który powoduje, że osoba, której mózg wyjęto, doświadcza iluzji, iż...

Celem takich opowieści w toku wykładu z teorii poznania jest oczywiście postawienie w nowoczesnej formie klasycznego problemu sceptycyzmu wobec świata zewnętrznego. (*Skąd wiesz, że nie znajdujesz się w takim żałosnym położeniu?*) Niemniej opisane położenie nadaje się również doskonale do zilustrowania kwestii stosunku umysłu do rzeczywistości.

Zamiast rozważać tylko jeden mózg w naczyniu, moglibyśmy wyobrazić sobie, że wszyscy ludzie (a może wszystkie istoty zmysłowe) są mózgami w naczyniu (lub układami nerwowymi w naczyniu, w przypadku istot o najprostszym układzie nerwowym, jaki pozwala zaliczyć je do istot „zmysłowych”). Niegodziwy uczony musi oczywiście znajdować się na zewnątrz — ale czy aby na pewno? Może nie ma żadnego niegodziwego uczonego, może (choć to absurd) kosmos po prostu przypadkiem składa się z urządzeń automatycznych służących zaopatrzeniu naczynia wypełnionego mózgami i układami nerwowymi.

Tym razem przypuśćmy, że owe automatyczne urządzenia są zaprogramowane tak, aby wywoływać u nas pewną *zbiorową* halucynację, a nie odrębne, pozbawione wzajemnych związków, rozmaite halucynacje. Kiedy więc wydaje mi się, że rozmawiam z tobą, tobie wy-



daje się, że słyszysz moje słowa. Nie znaczy to, rzecz jasna, że moje słowa faktycznie docierają do twoich uszu – ponieważ (w rzeczywistości) nie masz uszu, a ja w rzeczywistości nie mam ust ani języka. Zamiast tego, kiedy wypowiadam me słowa, bodźce wychodzące z mojego mózgu płyną do komputera, który powoduje, że „słyszę”, jak mój własny głos je wymawia, i „czuję”, jak porusza się mój język, itd., oraz powoduje, że ty „słyszysz” moje słowa, „widzisz” mnie, jak mówię itd. Skoro tak, to w pewnym sensie autentycznie ze sobą rozmawiamy. Nie myślę się co do twojego rzeczywistego istnienia (tylko co do istnienia twojego ciała i „świata zewnętrznego” poza mózgami). Z pewnego punktu widzenia fakt, że „cały świat” jest zbiorową halucynacją, nie jest nawet szczególnie istotny; w końcu słyszysz przecież moje słowa, kiedy mówię do ciebie, mimo że mechanizm tego zjawiska jest inny, niż przypuszczamy. (Gdybyśmy byli parą kochanków oddających się miłosnym uściskom, a nie tylko dwójgiem ludzi zajętych rozmową, wówczas, rzecz jasna, sugestia, że jesteśmy tylko mózgami w naczyniu, mogłaby być nieco kłopotliwa.)

Pragnę teraz zadać pytanie, na pozór głupkowate i oczywiste (przynajmniej dla niektórych, w tym również paru nader wyrafinowanych filozofów), które jednak dość szybko doprowadzi nas do autentycznych głębi myśli filozoficznej. Przypuśćmy, że cała nasza historyjka jest rzeczywiście prawdziwa. Czy moglibyśmy, będąc takimi mózgami w naczyniu, *powiedzieć* lub *pomyśleć*, że nimi jesteśmy?

Zamierzam udowodnić, że odpowiedź brzmi: „Nie, nie moglibyśmy”. A właściwie zamierzam udowodnić, że domniemanie, iż naprawdę jesteśmy mózgami w naczyniu, choć nie gwałci żadnego prawa fizyki i w żaden sposób nie kłóci się z całym naszym doświadczeniem, nie może być prawdziwe. *Nie może być prawdziwe*, ponieważ w pewnym sensie obala samo siebie. [...]

Domniemanie samo siebie obala, jeżeli z założenia o jego prawdziwości wynika, że jest fałszywe. [...]

W dalszym ciągu wywodu wykazę, iż domniemanie, że jesteśmy mózgami w naczyniu, ma tę właśnie właściwość. Jeżeli mamy możliwość zastanawiać się nad tym, czy jest prawdziwe, czy fałszywe, wówczas nie jest prawdziwe (co jest do wykazania). A zatem nie jest prawdziwe. [...]

Mieszkańcy możliwego świata, którzy naprawdę są mózgami w naczyniu, [...] mogą wprowadzić pomyśleć i „wypowiedzieć” każde słowo, które my potrafimy pomyśleć lub wypowiedzieć, lecz ich słowa nie mogą (jak twierdzę) *odnosić się* do tego samego, co nasze słowa. W szczególności nie mogą oni pomyśleć ani powiedzieć, że są mózgami w naczyniu (*nawet myśląc: „jesteśmy mózgami w naczyniu”*) (Putnam 1983: 302-306).

Oto nasza rekonstrukcja tego wywodu.

Bazę eksperymentu (w wersji mocnej: „Wszyscy jesteśmy mózgami w naczyniach”) sformułujemy w sposób następujący:

(1)  $\neg x \wedge y$  ( $x$  ma dostęp poznawczy do  $y \Rightarrow y$  jest przedmiotem pozornym).

Czyli: przedmioty jawiące się w świadomościach «nabudowanych» na mózgach w naczyniach są jedynie pozorne.<sup>20</sup>

A oto nasza rekonstrukcja horyzontu tego eksperymentu.

Po pierwsze, zgódźmy się, że jeśli twierdzimy coś zasadnie o jakimś przedmiocie, to mamy do tego przedmiotu dostęp poznawczy:

<sup>20</sup> Pomijamy tu trudność dotyczącą statusu ontycznego, który przyznaje się w analizowanym wywodzie „niegodziwemu uczonemu”, którego doznania nie są chyba pozorne.

daje się, że słyszysz moje słowa. Nie znaczy to, rzecz jasna, że moje słowa faktycznie docierają do twoich uszu – ponieważ (w rzeczywistości) nie masz uszu, a ja w rzeczywistości nie mam ust ani języka. Zamiast tego, kiedy wypowiadam me słowa, bodźce wychodzące z mojego mózgu płyną do komputera, który powoduje, że „słyszę”, jak mój własny głos je wymawia, i „czuję”, jak porusza się mój język, itd., oraz powoduje, że ty „słyszysz” moje słowa, „widzisz” mnie, jak mówię itd. Skoro tak, to w pewnym sensie autentycznie ze sobą rozmawiamy. Nie myślę się co do twojego rzeczywistego istnienia (tylko co do istnienia twojego ciała i „świata zewnętrznego” poza mózgami). Z pewnego punktu widzenia fakt, że „cały świat” jest zbiorową halucynacją, nie jest nawet szczególnie istotny; w końcu słyszysz przecież moje słowa, kiedy mówię do ciebie, mimo że mechanizm tego zjawiska jest inny, niż przypuszczamy. (Gdybyśmy byli parą kochanków oddających się miłosnym uściskom, a nie tylko dwójkiem ludzi zajętych rozmową, wówczas, rzecz jasna, sugestia, że jesteśmy tylko mózgami w naczyniu, mogłaby być nieco kłopotliwa.)

Pragnę teraz zadać pytanie, na pozór głupkowate i oczywiste (przynajmniej dla niektórych, w tym również paru nader wyrafinowanych filozofów), które jednak dość szybko doprowadzi nas do autentycznych głębi myśli filozoficznej. Przypuśćmy, że cała nasza historyjka jest rzeczywiście prawdziwa. Czy moglibyśmy, będąc takimi mózgami w naczyniu, *powiedzieć* lub *pomyśleć*, że nimi jesteśmy?

Zamierzam udowodnić, że odpowiedź brzmi: „Nie, nie moglibyśmy”. A właściwie zamierzam udowodnić, że domniemanie, iż naprawdę jesteśmy mózgami w naczyniu, choć nie gwałci żadnego prawa fizyki i w żaden sposób nie kłóci się z całym naszym doświadczeniem, nie może być prawdziwe. *Nie może być prawdziwe*, ponieważ w pewnym sensie obala samo siebie. [...]

Domniemanie samo siebie obala, jeżeli z założenia o jego prawdziwości wynika, że jest fałszywe. [...]

W dalszym ciągu wywodu wykażę, iż domniemanie, że jesteśmy mózgami w naczyniu, ma tę właśnie właściwość. Jeżeli mamy możliwość zastanawiać się nad tym, czy jest prawdziwe, czy fałszywe, wówczas nie jest prawdziwe (co jest do wykazania). A zatem nie jest prawdziwe. [...]

Mieszkańcy możliwego świata, którzy naprawdę są mózgami w naczyniu, [...] mogą wprowadzić pomyśleć i „wypowiedzieć” każde słowo, które my potrafimy pomyśleć lub wypowiedzieć, lecz ich słowa nie mogą (jak twierdzę) *odnosić się* do tego samego, co nasze słowa. W szczególności nie mogą oni pomyśleć ani powiedzieć, że są mózgami w naczyniu (*nawet myśląc: „jesteśmy mózgami w naczyniu”*) (Putnam 1983: 302-306).

Oto nasza rekonstrukcja tego wywodu.

Bazę eksperymentu (w wersji mocnej: „Wszyscy jesteśmy mózgami w naczyniach”) sformułujemy w sposób następujący:

(1)  $\forall x \forall y (x \text{ ma dostęp poznawczy do } y \Rightarrow y \text{ jest przedmiotem pozornym})$ .

Czyli: przedmioty jawiące się w świadomościach «nabudowanych» na mózgach w naczyniach są jedynie pozorne.<sup>20</sup>

A oto nasza rekonstrukcja horyzontu tego eksperymentu.

Po pierwsze, zgódźmy się, że jeśli twierdzimy coś zasadnie o jakimś przedmiocie, to mamy do tego przedmiotu dostęp poznawczy:

<sup>20</sup> Pomijamy tu trudność dotyczącą statusu ontycznego, który przyznaje się w analizowanym wywodzie „niegodziwemu uczonemu”, którego doznania nie są chyba pozorne.

(2)  $\forall x \forall y$  ( $x$  twierdzi zasadnie, że  $Py \Rightarrow x$  ma dostęp poznawczy do  $y$ ).  
 Po drugie, przedmiot pozorny to przedmiot nierzeczywisty:  
 (3)  $\forall y$  ( $y$  jest przedmiotem pozornym  $\Leftrightarrow y$  nie jest przedmiotem rzeczywistym).  
 Po trzecie, tylko rzeczywiste byty posiadają zasadne przekonania:  
 (4)  $\forall x \forall y$  ( $x$  twierdzi zasadnie, że  $Py \Rightarrow x$  jest przedmiotem rzeczywistym).  
 Zgódźmy się, że w (4) twierdzi się coś o  $x$ -ach, tj. o przedmiotach posiadających za-

sadne przekonania.  
 Zapytajmy teraz: czy jest ktoś, kto twierdzi zasadnie, że (4):

(5)  $\forall z$  ( $z$  twierdzi zasadnie, że (4))?

Zauważmy, że z (1)-(3) wynika, że:

(6)  $\forall x \forall y$  ( $x$  twierdzi zasadnie, że  $Py \Rightarrow y$  nie jest przedmiotem rzeczywistym).

Załóżmy teraz, że odpowiedź na pytanie (5) jest pozytywna, czyli:

(7)  $\forall z$  ( $z$  twierdzi zasadnie, że (4)).

Przypuśćmy, że  $A$  byłoby takim  $z$ -em. Na mocy (7) i (6) byłoby, że:

(8)  $A$  nie jest przedmiotem rzeczywistym.

Natomiast na mocy (7) i (4) byłoby z kolei, że:

(9)  $A$  jest przedmiotem rzeczywistym.

Okazuje się, że założenie (7) prowadzi na gruncie poprzednich założeń do sprzeczności. Wynika z tego, że:

(10)  $\sim \forall z$  ( $z$  twierdzi zasadnie, że (4)).

Inaczej mówiąc — na gruncie przyjętych założeń — nikt nie może twierdzić zasadnie, że (4). W konsekwencji — nie może też twierdzić zasadnie, że (6). Jeśli zaś uznamy (6) za adekwatną interpretację tezy, że świat jest iluzją, to trzeba się zgodzić, że nikt tej tezy — nazwijmy ją dla uniknięcia historycznych obciążeń innych terminów „pan-iluzjonizmem” — nie może jej zasadnie stwierdzić (oczywiście przy wyszczególnionych założeniach).

„Opowiadka” o mózgach w naczyniu jest — w naszym przekonaniu — niczym innym jak sugestywną ilustracją tego typu rozumowania.

## 12. KONKLUZJA

Nie dziwi nas, że eksperymenty myślowe są tak bardzo popularne wśród filozofów, a zwłaszcza wśród czytelników niektórych pism filozoficznych.

Jak pokazaliśmy — są to bądź «niedoszlę» eksperymenty rzeczywiste, bądź stimulatory do ujawniania poglądów osób tym eksperymentom «poddanych», w szczególności do ujawniania intuicji semantycznych takich osób. Poza funkcją ilustracyjną i *a fortiori* perswazyjną — filozoficzne eksperymenty myślowe nie pełnią żadnej doniosłej funkcji teoretycznej. Stanowią jednak egemplifikacje — nieraz pomysłowe i dowcipne — odpowiednich «czystych» rozumowań, a że już od najmłodszych lat lubimy czytać książki z... obrazkami — stąd ich wspomniane wzięcie. Nie byłoby skądinąd niczego złego w posługiwaniu się nimi, gdyby teksty, w których są opi-

sywane, zawierały ponadto jasne wysłownienie bronionych twierdzeń i przejrzyste przedstawioną argumentację. Niestety — znanym nam tekstom filozoficznym, ubarwionym opisami eksperymentów myślowych, jasności i przejrzystości — nie dostaje.

Być może zresztą stąd właśnie bierze się wielość prac, w których się te teksty komentuje. Jak wiadomo, im tekst mniej jasny i przejrzysty, tym więcej ma możliwych interpretacji. W owych komentarzach przy tym zamiast (a nie — obok!) jasno sformułowanych twierdzeń i przejrzystych argumentacji pojawiają się na ogół — nowe eksperymenty myślowe...

## BIBLIOGRAFIA

- Ajdukiewicz, Kazimierz  
(1931) Paradoksy starożytnych, [w:] *Język i poznanie*. Tom I, Warszawa 1960, PWN, s. 137-144.
- Brendel, Elke  
(2004) Intuition Pumps and the Proper Use of Thought Experiments, *Dialectica* vol. 58, No. 1, s. 89-108.
- Brożek, Anna, Jacek Jadacki  
(2012) Eksperymenty myślowe w nauce, *Filozofia Nauki*, nr 1, s. 5-32.
- Frege, Gottlob  
(1893) *Grundgesetze der Arithmetik, begriffsschriftlich abgeleitet*. Band I, Jena, Verlag Hermann Pohle.
- Gettier, Edmund L.  
(1963) Is Justified True Belief Knowledge?, *Analysis* vol. 23, No. 6, s. 121-123.
- Hajduk, Zygmunt  
(2002) Eksperyment, [w:] Andrzej Maryniarczyk *et al.* (red.), *Powszechna encyklopedia filozofii*. Tom 3, Lublin, Polskie Towarzystwo Tomasza z Akwinu, s. 83-84.
- Hobbes, Thomas  
(1642) *Elementy filozofii*, Warszawa 1956, PWN.
- Locke, John  
(1690) O tożsamości i różnicy, [w:] *Rozważania dotyczące rozumu ludzkiego*. Tom I, Warszawa 1955, PWN, s. 460-493.
- Nowaczyk, Adam & Żołądowski, Zenobiusz  
(1974) Eksperyment, [w:] *Logika i metodologia badań naukowych dla lekarzy*, Warszawa, PZWL, s. 152-153.
- Putnam, Hilary  
(1975) Znaczenie wyrazu „znaczenie”, [w:] Putnam 1998, s. 93-184.  
(1983) Mózgi w naczyniu, [w:] Putnam 1998, s. 295-324.  
(1998) „Wiele twarzy realizmu” i inne eseje, Warszawa 1998, Wydawnictwo Naukowe PWN
- Quine, Willard van Orman  
(1960) *Word and Object*, Cambridge (Massachusetts), The MIT Press.

Schrödinger, Erwin

(1935) Die gegenwärtige Situation in der Quantenmechanik, *Naturwissenschaften* B. XXIII, s. 807-812, 823-828, 844-849.

Searle, John S.

(1984) Czy komputery mogą myśleć?, [w:] *Umysł, mózg i nauka*, Warszawa 1995, Wydawnictwo Naukowe PWN, s. 25-37.

(1992) *The Rediscovery of Mind*, Cambridge (Massachusetts) 1994, The MIT Press.

Thomson, Judith Jarvis

(1985) The Trolley Problem, [w:] *Right, Restitution, and Risk. Essays in Moral Theory*, Cambridge (Massachusetts) & London (England) 1986, Harvard University Press, s. 94-116.

Wierzbicki, Piotr

(2011) *Boski umysł. Esej o tropieniu świata*, Warszawa, Wydawnictwo Sic!

Wojtyśiak, Jacek

(2007) Słownik wybranych terminów filozoficznych, [w:] Antoni B. Stępień, *Wstęp do filozofii*, Lublin, Towarzystwo Naukowe KUL, s. 327-438.