

## **Eksperymenty myślowe w służbie esencjalizmu**

Nietrudno zauważyć, że filozofowie, którzy oddają się poszukiwaniom własności istotnych, chętnie sięgają po eksperymenty myślowe, choć nie zawsze posługują się akurat tym określeniem<sup>1</sup>. Powodem tej sympatii jest rzekoma zdolność eksperymentów myślowych do pobudzania w nas intuicji, która pozwala na odróżnianie tego, co konieczne od tego, co jedynie faktyczne, czy możliwe. Celowo nie precyzuję, o jakiej intuicji mowa. Być może jest to intuicja lingwistyczna, być może metafizyczna (czego przykładem mogłaby być intuicja eidetyczna, o której mówią fenomenolodzy). Argumentacja, którą przedstawiam w dalszej części tego artykułu, jest niezależna od rozstrzygnięć w tej sprawie. Związek pomiędzy własnościami istotnymi a koniecznością nietrudno wskazać - własności istotne to takie, które muszą przysługiwać przedmiotowi. Ów mus ma się zaś manifestować w ten sposób, że usunięcie własności istotnej zniszczy przedmiot, z którego ją usunęliśmy. Gdy jest to przedmiot, o którym jedynie myślimy, jego „zniszczenie” polega na tym, że nabieramy w pewnej chwili przekonania, iż w wyniku zaszłej zmiany mamy do czynienia z innym przedmiotem, niż przed tą zmianą. Podstawą tego przekonania ma być zaś wspomniana intuicja.

Mające stymulować naszą intuicję eksperymenty myślowe charakteryzują się dość podobną strukturą. Zazwyczaj jest tak, że autor zarysowuje nam pokrótce jakąś sytuację albo opisuje pewien przedmiot, a następnie przedstawia rozmaite zmiany, którym przedmiot ten podlega, za każdym razem upewniając się, czy słuchacze są skłonni zaakceptować je jako dopuszczalne. Widać to dobrze w klasycznym przypadku opowieści o statku Tezeusza. Mówi się nam, że istnieje statek, który z różnych względów chcielibyśmy zachować, a który musieliśmy z czasem poddawać konserwacji. Dowiadujemy się, że w pewnym momencie wszystkie oryginalne części statku zostały wymienione i naszym zadaniem jest teraz rozsądzenie, czy zmiana, która zaszła, zniszczyła wyjściowy przedmiot, czy też nie.

Warto zauważyć, że pewna dobrze rozpoznana wada eksperymentów myślowych działa w tym przypadku na ich korzyść – mówi się często, że wartość eksperymentów myślowych jest wątpliwa, ponieważ są opisami dziwacznych, mało prawdopodobnych sytuacji. Można mieć wątpliwości, czy ekstrapolowanie wniosków uzyskanych przy okazji rozważań nad tymi dziwnymi przypadkami na przypadki typowe jest zabiegiem rozsądnym. Jest to znany argument, który pojawia się często przy okazji dyskusowania nad rolą eksperymentów myślowych w etyce. W przypadku poszukiwań własności istotnych jest jednak zupełnie inaczej. Nie ma nic zadziwiającego w tym, że odkrywanie tego, co konieczne wymaga od nas wykroczenia daleko poza bezpieczne okolice rzeczywistego świata. Eksperymenty myślowe używane w tym celu wymagają od nas szczególnego „filozoficznego zawieszenia niewiary”. Gdybym zareagował na opowieść o statku Tezeusza, mówiąc, że dobrane

---

<sup>1</sup> Moje uwagi stosują się na przykład do wielu analiz fenomenologicznych, których autorzy mogliby nie być skłonni przystać na określenie „eksperyment myślowy”.

dokładnie takich samych, jak oryginalne, części jest mało prawdopodobne i statek po konserwacji różniłby się nie tylko materiałem, ale i strukturą albo że taka kumulacja cząstkowych zniszczeń wymaga tak długiego czasu, że nie ma się czym przejmować, to zostałbym pouczony, że nie pytamy się tu o to, co jest, a co nie jest prawdopodobne, ponieważ od odpowiedzi na te pytania są nauki przyrodnicze. Czytelnik, proszony o rozważenie eksperymentu myślowego ma protestować tylko wtedy, gdy coś wyda mu się niemożliwe<sup>2</sup>. W szczególności, nie powinniśmy korzystać z naszych zdroworozsądkowych przekonań na temat świata rzeczywistego. Zdrowy rozsądek nie rozróżnia pomiędzy tym, co konieczne a tym, co typowe, lecz akcydentalne. A jednak - zanim uznamy, że odnaleźliśmy idealne zastosowanie dla eksperymentów myślowych, zauważmy, że ich użycie rodzi dwie wątpliwości natury metodologicznej.

Po pierwsze, eksperymenty myślowe testują intuicję wybranej, dość specyficznej grupy ludzi, jakimi są czytelnicy tekstów filozoficznych. Wyniki tego badania są następnie optymistycznie rozciągane na intuicje całej ludzkości. Niefrasobliwość ta podszyta jest niekiedy naiwnością (gdy filozofowie z góry zakładają, że są po prostu modelowymi przedstawicielami ludzkości), a niekiedy arogancją (gdy zakładają, że są do badania własności istotnych w szczególny sposób predestynowani, czy wytrenowani). Druga wada eksperymentów myślowych polega na tym, że niekiedy autorzy stosują je raczej w charakterze zabiegów retorycznych, niż rzeczywistego narzędzia do pobudzania intuicji czytelnika. Być może reakcja czytelników byłaby zupełnie inna, gdyby spotkali się z danym eksperymentem w innym miejscu, niż środek artykułu, który już na nich wpłynął i częściowo ukształtował ich pogląd na dany temat.

Ideał stanowiłby więc taki eksperyment myślowy, za pomocą którego sprawdzilibyśmy intuicje dużej i zróżnicowanej grupy osób (a najlepiej większości ludzkości), bez zdradzania im naszych esencjalistycznych motywów. Sądzę, że eksperymenty takie wykonywane są codziennie i Ty, Czytelniku, byłeś im wielokrotnie poddany. Tym masowym eksperymentem myślowym jest po prostu fikcja literacka.

Gdy filozofowie opowiadają nam o Ziemi Bliźniaczej, stołach zrobionych z zamrożonej wody z Tamizy, czy ołówkach, które okazują się być żywymi organizmami, to choć mogliśmy nie spodziewać się tego typu opowieści w artykule z dziedziny filozofii, nie jesteśmy zadziwieni samą ich treścią. Wielokrotnie spotkaliśmy się z dziwnymi światami, bardzo odległymi od tego, w którym o nich myślimy. Fikcja literacka jest ich pełna i nie chodzi jedynie o współczesne opowieści z dziedziny *science fiction* – baśnie i mity towarzyszą ludzkości od jej początków i równie dobrze nadają się na masowy eksperyment myślowy. Sądzę, że jeśli ktoś wierzy w pobudzającą intuicję moc eksperymentów myślowych i zgadza się na to, że są one obarczone wymienionymi wyżej wadami metodologicznymi, powinien bacznie przyjrzeć się temu, jak intuicja jego czytelników zachowuje się, gdy czytają fikcję. Aby uniknąć nieporozumień – nie chodzi mi o utożsamienie ze sobą światów fikcyjnych i możliwych. Jedyne, co chcę powiedzieć, to że wiele z używanych przy okazji badań nad własnościami istotnymi eksperymentów myślowych można umieścić w dziele literackim i odwrotnie – że wiele fantastycznych

---

2 Podobnie jak w przypadku intuicji, pytanie o charakter możliwości, jaka wchodzi tu w grę, celowo pozostawiam otwarte, ponieważ moja argumentacja jest od tego niezależna. Z pewnością nie chodzi tu o zwykłą możliwość empiryczną, czyli zgodność z prawami przyrody. Co do tego, czy jest to jednak możliwość logiczna, pojęciowa, czy metafizyczna sami esencjaliści nie są zgodni.

opowieści można przenieść do tekstu filozoficznego i wykorzystać w charakterze eksperymentów myślowych. Kłopot w tym, że nasza intuicja zachowuje się odmiennie, w zależności od tego, gdzie umieszczona zostanie dana historia. Zwykle, niefilozoficzne, zawieszenie niewiary okazuje się być znacznie bardziej liberalne, niż filozoficzne – w baśniach, mitach i opowiadaniach *science fiction* dopuszczalne są takie zmiany, które w tekstach filozoficznych uznalibyśmy za niszczące przedmiot. Ludzie mogą ginąć i zmartwychwstawać, zamieniać się w żabę, czy kamień i zostawać w ostatniej chwili znów przemienieni w człowieka, a czytelnik nie uznaje, że właśnie podmieniono mu bohatera, o którym czytał.

Najprostszym wyjaśnieniem tej różnicy w podejściu jest to, że w przypadku fikcji literackiej, w odróżnieniu od tekstów filozoficznych, czytelnicy tolerują sprzeczność. Sądzę, że jest to wyjaśnienie mylące – tolerujemy niezgodności z naszą wiedzą empiryczną i zdrowym rozsądkiem, ale sprzeczność logiczną uważamy już za wysoce niepożądaną wadę fikcji<sup>3</sup> – utrudnia, a często uniemożliwia ona zawieszenie niewiary. Akceptacja niezgodności ze zdrowym rozsądkiem i wiedzą empiryczną jest zaś czymś, co winniśmy czynić również przy okazji rozważania eksperymentów myślowych. Oczywiście może być tak, że fikcja zawiera sprzeczność, której nie dostrzegliśmy, ale to samo można powiedzieć o eksperymentach myślowych – twierdzimy, że opisują sytuacje możliwe, ponieważ nie dostrzegamy w nich sprzeczności.

Sądzę, że powodem, dla którego nasza intuicja zachowuje się w przypadku fikcji inaczej, niż w przypadku eksperymentów myślowych jest coś zupełnie innego – w przypadku fikcji zawieszamy nasze zdroworozsądkowe przekonania na temat tego, co jest, a co nie jest prawdopodobne i co jest, a co nie jest empirycznie możliwe. Ze względów, które postaram się wskazać, nie robimy tego w przypadku eksperymentów myślowych, co w rezultacie czyni je bardziej narzędziami badania granic zdrowego rozsądku, niż tego, co konieczne.

Przyjrzyjmy się następującemu przykładowi: załóżmy, że interesuje mnie ustalenie własności istotnych stołu. Posługuję się w tym celu eksperymentem myślowym, który sprowadza się do opisu trzech scenariuszy.

W **scenariuszu A** jedna ze śrub stołu obluźowała się, przez co stał się on niestabilny.

W **scenariuszu B** stół został rozłożony na części.

W **scenariuszu C** został on porąbany na drzazgi za pomocą topora.

W każdym ze scenariuszy testujemy intuicję naszych rozmówców, pytając, czy opisane zmiany zniszczyły przedmiot (czy przestał on być stołem), czy też nie. Sądzę, że scenariusze A i C są pod tym względem dość niekontrowersyjne. Większość osób powie, że w scenariuszu A nadal mamy do czynienia ze stołem – trzeba tylko dokręcić śrubę. Intuicje większości osób odnośnie do scenariusza C będą zaś takie, że stół przestał w nim istnieć i mamy przed sobą tylko stos porąbanego drewna.

Reakcja na scenariusz B nie będzie jednorodna – gdy prezentowałem ten przykład różnym osobom (studentom i wykładowcom filozofii, a także studentom różnych innych kierunków), moi rozmówcy podzielili się na trzy obozy. Grupa pierwsza uznawała, że scenariusz B jest bardzo podobny do scenariusza A – różnica jest czysto ilościowa,

---

<sup>3</sup> Wystarczy spojrzeć na karkołomne wyjaśnienia, które konstruują wielbiciel powieści, w których znaleziono jakąś sprzeczność.

zamiast jednej śruby i obluzowanych części mamy wiele wykręconych śrub i stół w częściach. Grupa druga uznawała ten przypadek za niepodobny ani do scenariusza A, ani do scenariusza C – stół ma po prostu swoistą „przerwę w istnieniu” (na przykład na czas przeprowadzki). Stół przestał istnieć z chwilą rozłożenia, ale może wrócić do nas, gdy tylko z powrotem go złożymy. Grupa trzecia uznawała zaś, że scenariusz B jest bardzo podobny do scenariusza C - z chwilą rozłożenia stół przestał istnieć i nie da się go już przywrócić. Można jedynie zbudować bardzo podobny stół z części starego stołu (po zakończonej przeprowadzce będziemy więc mieli nowe meble).

Grupę drugą i trzecią możemy tu, dla naszych potrzeb utożsamiać<sup>4</sup>. Obie łączy przekonanie, że natrafiliśmy tu na taką zmianę w przedmiocie, która go niszczy, mamy więc prawo podejrzewać, że wpadliśmy właśnie na trop jakiejś własności istotnej. Jaka to własność? Ludzie zaliczający się do grup 2 i 3 odpowiedzieliby prawdopodobnie, że chodzi po prostu o to, że w wyniku rozłożenia na części, stołu nie da się już, po prostu, używać. Osoby, które uważają, że scenariusz B przypomina scenariusz A, czyli te, które zaliczyliśmy do grupy pierwszej, mogłyby jednak zdziwić się tą odpowiedzią – przecież da się go używać, wystarczy go złożyć. Reakcją na tę uwagę może być co najwyżej grymas i szybka korekta – chodziło oczywiście o to, że nie da się go używać **w tym momencie**. Tak skorygowany pogląd okazuje się jednak prowadzić do mało intuicyjnych konsekwencji. Rozważmy **scenariusz D** – stół został postawiony do góry nogami. Nie ulega wątpliwości, że nie da się go w tym momencie używać. Czy ktokolwiek chciałby jednak twierdzić, że postawiony do góry nogami stół przestaje na moment być stołem albo ginie na zawsze i może być, co najwyżej, zastąpiony niezwykle podobnym stołem, gdy się operację obracania odwróci? Być może powinniśmy zająć się teraz konstruowaniem subtelniejszych eksperymentów myślowych, które pozwoliłyby nam na przyspilenie wymykającej się nam z rąk własności istotnej?

Sądzę, że nie będzie to konieczne. Rzeczywistym powodem, dla którego przedstawiciele grup 2 i 3 uznają, że stół przestaje w scenariuszu B istnieć (choćby tylko czasowo), jest to, że nie można go **z łatwością** używać. Łatwość przywrócenia stanu poprzedniego to główna różnica pomiędzy rozłożeniem stołu na części a jego postawieniem do góry nogami. Kłopot w tym, że zdroworozsądkowe przekonanie na temat średniego czasu składania typowego stołu nie powinno być podstawą ontologicznych rozstrzygnięć i zupełnie nie przypomina pobudzonej przez eksperymenty myślowe intuicji ejdetycznej, której się spodziewaliśmy.

Zostawmy jednak scenariusz B. Być może okazaliśmy się nieco naiwni i wzięliśmy zmianę trudną do szybkiego odwrócenia za zmianę istotną. Na dodatek - od początku istniała co do niego różnica zdań. Mamy przecież jeszcze scenariusz C, co do którego wątpliwości nie było. Skoro wszyscy zgadzają się na to, że stół przestaje w nim istnieć, to może uda się nam dzięki niemu odkryć jakąś własność istotną stołu? Powróćmy teraz do naszych rozważań o możliwości zastosowania fikcji w charakterze eksperymentu myślowego. Załóżmy, że czytamy baśń o pewnym człowieku, którego stół został porąbany

---

4 Rozróżnienie pomiędzy tymi dwiema grupami jest potrzebne głównie wtedy, gdy interesuje nas wykrycie istoty indywidualnej. W artykule tym pomijam rozróżnienie na wykrywanie istot indywidualnych i gatunkowych, ponieważ różnica ta nie jest z mojego punktu widzenia istotna. Jeżeli tylko ktoś wierzy w możliwość wykrywania istot indywidualnych, to do badań zabiera się dokładnie tak samo, jak zabiera się przy okazji poszukiwania istot gatunkowych. Moje uwagi dotyczą zaś jedynie samej techniki badawczej.

na drzazgi. Człowiek ten nie przejął się, ponieważ tak się złożyło, że posiada specjalny magiczny przedmiot – różdżkę do odwracania zmian. Jednym machnięciem różdżki sprawił, że drzazgi złożyły się z powrotem w kompletny stół. Operacja nie trwała dłużej niż postawienie na nogach odwróconego stołu ze scenariusza D. W czym scenariusz C różni się teraz od scenariuszy A i B (przykręcenia obluźwanej śruby i złożenia stołu z części)? Wprawdzie zamiast śrubokręta występuje tu czarodziejska różdżka, ale procedura przywrócenia świetności stołowi jest jeszcze krótsza i prostsza niż w niekontrowersyjnym scenariuszu A. Czy wolno nam przenieść czarodziejską różdżkę ze świata fikcji literackiej do eksperymentów myślowych? Nie widzę powodu, dla którego nie moglibyśmy tego zrobić – nie potrafię w tej opowieści dostrzec żadnej sprzeczności, o ile mi wiadomo, taka różdżka jest nawet empirycznie możliwa.

Początkowo wzięliśmy zmiany trudne do odwrócenia za zmiany istotne. Teraz z góry przyjęliśmy, że zmiana opisana w scenariuszu C jest zmianą **nieodwracalną**. Tymczasem, w świecie, w którym istnieje stosunkowo niezawodny sposób na złożenie stołu z fragmentów, na które został porąbany, kupka drewna nie różni się od zestawu części ze scenariusza B, a gdy tylko naprawa ta jest tak błyskawiczna i łatwa w wykonaniu jak przykręcenie śruby, to nasza opowieść upodabnia się do scenariusza A. Przekonanie o tym, które z zachodzących zmian są nieodwracalne, a które odwracalne (ale z trudnościami), to jedynie zdroworozsądkowe założenia oparte na naszej (niepełnej) wiedzy o istniejących w świecie technologiach, ale nawet gdybyśmy dysponowali technologiczną wszechwiedzą na temat świata rzeczywistego, to niczego by to nie zmieniło, ponieważ i tak byłaby to wiedza, którą mieliśmy na czas rozważania eksperymentów myślowych zawiesić.

Pojawić się może w tym miejscu następująca wątpliwość - być może jest tak tylko w przypadku prób wykrywania własności istotnych artefaktów, które z definicji zależne są od technologii? Czy z istotnymi własnościami przedmiotów należących do gatunków naturalnych jest tak samo?

Nie jestem pewien, czy jest z nimi tak samo, z pewnością jednak nie jest lepiej. Rozważmy następujący przykład - założmy, że pytamy, czy jakiś owad, na przykład karaluch, może mieć 2 metry długości (interesuje nas więc to, czy określona wielkość jest cechą istotną karalucha, czy nie). Na pierwszy rzut oka wydaje się, że jest to dopuszczalne. Kłopot jednak w tym, że istota tej wielkości nie mogłaby oddychać za pomocą systemu tchawek, w który wyposażone są karaluchy. Stajemy więc przed następującym dylematem: czy wymuszona wielkością zmiana układu oddechowego karalucha zmienia jego przynależność gatunkową, czy nie? A może powinniśmy pozostawić szczegóły budowy karalucha w spokoju i zmienić nieco prawa fizyki, tak by dostosować je do anatomii dwumetrowych karaluchów? To prowadzi do kolejnej trudności, ponieważ nie wiadomo, czy świat z innymi prawami fizyki byłby światem z tym samym zestawem gatunków naturalnych. Nie wiadomo też, co począć z krzyżówkami i owocami inżynierii genetycznej. Czy są one raczej artefaktami, czy zmodyfikowanymi gatunkami naturalnymi? Co więcej, nawet jeśli uznamy, że artefaktem jest wszystko, przy czym tylko jakiś człowiek majstrował, to wystarczy wyobrazić sobie taki świat, w którym ta sama modyfikacja jest dziełem natury i znów możemy tym naszą intuicję wyprowadzić w pole.

To, co może się przydarzyć artefaktom, jest też ściśle powiązane z tym, co może

przydarzyć się egzemplarzom gatunków naturalnych i odwrotnie. Gdy zapytacie mnie, czy stół może być zrobiony z dowolnego materiału, na przykład z masła, to jedyną sensowną odpowiedzią, jakiej potrafię udzielić, jest ta, że to zależy od tego, jakie własności może mieć masło, nie przestając być masłem. A to znów może zależeć od tego, jakie własności ma człowiek masła używający i mleko, z którego je uzyskał. Mógłbym tak kontynuować w nieskończoność<sup>5</sup>.

Reasumując, jeśli nawet przedmioty posiadają swe istoty i jeśli nawet istnieją sposoby na ich wykrywanie, to nie należą do nich eksperymenty myślowe. Różnica pomiędzy fikcją a eksperymentami myślowymi polega na tym, że czytając fikcję, zawieszamy naszą niewiarę i dajemy autorowi spory kredyt zaufania – sprowadza się on do tego, że gdy natkniemy się na coś, co uderza nas jako niemożliwe, czekamy, ponieważ autor może wprowadzić do opisywanego świata elementy, które uczynią całość spójną. Tymczasem, eksperymenty myślowe traktujemy jak zagadki – zakładamy, że wszystkie eksperymenty kończą się niewypowiedzianą wprost klauzulą *ceteris paribus*. Iluzją jest jednak sądzić, że da się wykryć własności istotne przedmiotów wziętych w izolacji. To, co może przydarzyć się stołowi bez jego zniszczenia, zależy na przykład od tego, czy w świecie, w którym ten stół stoi, istnieje wspomniana czarodziejska różdżka do odwracania zmian. Jest tak dlatego, że skłonni jesteśmy przymknąć oko nawet na utratę tych własności, o których twierdzimy zazwyczaj, że są własnościami istotnymi. Ważne jest dla nas tylko to, czy jest to utrata szybko i łatwo odwracalna, ale pogląd na to, co jest, a co nie jest odwracalne<sup>6</sup> i jak prosta jest to procedura, czerpiemy już ze zdroworozsądkowego obrazu świata i wiedzy empirycznej, które mieliśmy, podobno, wziąć w nawias.

---

5 Jeśli da się pomyśleć takie technologie, które wydłużają życie w sposób nieograniczony, nie odbierając mi przy tym tożsamości, to zdanie to można nawet potraktować dosłownie.

6 Przykładem takiego postępowania jest nasz stosunek do myślenia, czy rozumności, które niekiedy podajemy jako własność istotną człowieka. Dopiero nieodwracalna utrata tej zdolności sprawia, że niektórzy byliby skłonni odmówić człowieczeństwa tym, którzy ją utracili (na przykład w wyniku wypadku).