

Alvin Plantinga

## EWOLUCYJNY ARGUMENT PRZECIWKO NATURALIZMOWI\*

W ostatnim rozdziale *Warrant and Proper Function* wysunąłem ewolucyjny argument przeciwko naturalizmowi (odtąd skrót: EAPN)<sup>1</sup>. Jako *filozoficzny naturalizm* okreśmy przekonanie, że nie istnieją żadne istoty ponadnaturalne – nie istnieje np. Bóg ani nic w rodzaju Boga, ani żadne inne byty ponadnaturalne<sup>2</sup>. Postawiłem tezę, że naturalizm i współczesna teoria ewolucji pozostają ze sobą w niezgodzie – i to pomimo tego, że tę drugą uważa się potocznie za jeden z głównych filarów, na których opiera się gmach naturalizmu. (Oczywiście nie atakuję teorii ewolucji czy twierdzenia, że istoty ludzkie rozwinęły się na drodze ewolucji z małpich przodków, ani żadnego twierdzenia tego rodzaju. Kwestionuję natomiast koniunkcję *naturalizmu* z poglądem, że istoty ludzkie rozwinęły się w taki sposób. Nie widzę podobnych problemów w przypadku koniunkcji *teizmu* i idei, że istoty ludzkie rozwinęły się na drodze ewolucji w taki sposób, jaki sugeruje współczesna nauka o ewolucji). Mówiąc ściślej, argumentowałem za tym, że koniunkcja naturalizmu z przekonaniem, iż istoty ludzkie rozwinęły się na drodze ewolucji zgodnie ze współczesną doktryną ewolucyjną (w skrócie: ewolucją), jest w pewien interesujący sposób samounieważniająca czy samozwrotnie niespójna. Jeszcze dokładniej: broniłem twierdzenia, że koniunkcja naturalizmu i ewolucji (w skrócie: *N&E*) dostarcza temu, kto ją akceptuje, racji unieważniają-

---

\* *The Evolutionary Argument against Naturalism: An Initial Statement of the Argument*, w: J. Beilby (red.), *Naturalism Defeated? Essays on Plantinga's Evolutionary Argument against Naturalism*, Ithaca: Cornell University Press 2002, s. 1–12. Przekład za zgodą Autora.

<sup>1</sup> A. Plantinga, *Warrant and Proper Function*, New York: OUP 1993.

<sup>2</sup> Gdyby moim celem było podanie analizy filozoficznego naturalizmu, należałoby powiedzieć znacznie więcej (na przykład, jeśli nie wiemy, czym jest naturalizm, to czy pomocne jest wyjaśnianie go w kategoriach ponadnaturalizmu?); w obecnym kontekście możemy pominąć te komplikacje.

cej dla przekonania, że nasze władze poznawcze są rzetelne – i to takiej racji, która sama nie może być unieważniona. Oznacza to jednak, że ta koniunkcja dostarcza również racji unieważniającej dla każdego przekonania wytworzonego przez nasze władze poznawcze, w tym także – w przypadku kogoś, kto akceptuje to przekonanie – *N&E*. Stąd jego samounieważniający charakter.

To tyle tytułem wprowadzenia. Przechodząc do szczegółów: EAPN wychodzi od określonych wątpliwości dotyczących *rzetelności* naszych władz poznawczych. Ogólnie mówiąc, władza poznawcza – pamięć, percepcja, rozum – jest rzetelna, jeśli ogromna część dostarczanych przez nią danych jest prawdziwa<sup>3</sup>. Wątpliwości te są związane z *genezą* naszych władz poznawczych. Zgodnie ze współczesną teorią ewolucji my, istoty ludzkie, podobnie jak inne organizmy, rozwinęliśmy się z pierwotnej jednokomórkowej formy życia przez takie mechanizmy jak selekcja naturalna i zmiany genetyczne, wykorzystujące różnice genetyczne; najpopularniejsza jest przypadkowa mutacja genetyczna. Selekcja naturalna odrzuca większość z tych mutacji (okazują się one szkodliwe dla organizmów, w których występują), jednak niektóre z pozostałych mają wartość adaptacyjną i ułatwiają przystosowanie: rozprzestrzeniają się one w obrębie populacji i w ten sposób trwają. Zgodnie z tą koncepcją cała różnorodność współczesnego życia organicznego powstała dzięki takim lub bardzo podobnym mechanizmom; za ich sprawą powstały również nasze władze poznawcze.

Zgodnie z tradycyjną myślą chrześcijańską (a także żydowską i islamską) my, istoty ludzkie, zostaliśmy stworzeni na obraz Boga. Oznacza to m.in., że Bóg stworzył nas ze zdolnością do osiągnięcia *wiedzy* – o naszym otoczeniu przez percepcję, o innych ludziach przez to, co Thomas Reid nazwał sympatią, o przeszłości przez pamięć i świadectwo, o matematyce i logice przez rozum, o moralności, o naszym własnym życiu mentalnym, o samym Bogu oraz o wielu innych rzeczach<sup>4</sup>. Powyższy ewolucyjny opis

<sup>3</sup> Najogólniej: termometr, który zatrzymał się na 22°C, nie jest rzetelny, nawet jeśli umieści się go w miejscu, gdzie prawie przez cały czas jest 22°C (w San Diego?). Ważne jest również to, co termometr (czy też nasze władze poznawcze) zrobiłby, gdyby sprawy miały się inaczej pod pewnymi (trudnymi do określenia) względami. Ponownie, gdyby naszym celem była analiza *rzetelności*, należałoby powiedzieć znacznie więcej. Zauważmy, że przy takim rozumieniu rzetelności nie wystarczy, aby wytworzone przekonania zwiększały przystosowanie.

<sup>4</sup> Tomasz z Akwinu: „Ponieważ mówi się, że istoty ludzkie są obrazem Boga na mocy posiadania przez nie natury, która zawiera rozum, taka natura jest najbardziej obrazem Boga na mocy tego, iż jest najbardziej zdolna naśladować Boga” (*Summa theologiae*, I, q. 93, a. 4) oraz: „Tylko w istotach racjonalnych odkrywamy podobieństwo do Boga na

naszej genezy jest spójny z teistycznym poglądem, że Bóg stworzył nas na swój obraz<sup>5</sup>. A zatem teoria ewolucji sama w sobie (bez naleciałości towarzyszącego jej często filozoficznego naturalizmu) nie jest niezgodna z ideą, że Bóg stworzył nas i nasze władze poznawcze w taki sposób, że władze te są rzetelne; że (jak lubili mawiać myśliciele średniowieczni) istnieje odpowiedniość między intelektem a rzeczywistością.

Jeśli jednak *naturalizm* jest prawdziwy, to Bóg nie istnieje, a zatem ani on, ani nikt inny nie planuje naszego rozwoju i nie kieruje przebiegiem naszej ewolucji. W tej sytuacji pojawia się natychmiast pytanie, czy jest w ogóle prawdopodobne, że nasze władze poznawcze – zakładając naturalizm i ich ewolucyjną genezę – rozwinęłyby się w taki sposób, by były rzetelne i dostarczały nam w większości prawdziwych przekonań. Wątpliwość tę wyraził sam Darwin: „Prześladuje mnie zawsze – powiedział – nieznośna wątpliwość, czy przeświadczenia umysłu ludzkiego, który powstał w wyniku rozwoju umysłów niższych zwierząt, są w jakimkolwiek stopniu wartościowe lub w ogóle godne zaufania. Czy ktokolwiek ufałby przeświadczeniom w umyśle małpy, o ile w takim umyśle są w ogóle jakieś przeświadczenia?”<sup>6</sup>.

---

sposób obrazu [...]. Jeśli chodzi o podobieństwo do Bożej natury, rozumne stworzenia wydają się niejako dochodzić do przedstawienia [tego] gatunku, o ile imitują Boga nie tylko w tym, że istnieje i żyje, ale zwłaszcza w tym, że myśli” (tamże, I, q. 93, a. 6).

<sup>5</sup> Ktoś mógłby pomyśleć, że jest inaczej: jeśli nasza geneza jest związana z *przypadkową* kombinacją genetyczną, to my oraz nasze władze poznawcze rozwinęliśmy się dzięki *przypadkowi*, a nie dzięki planowi, czego wymagałoby bycie istotami stworzonymi przez Boga na jego obraz. W ten sposób przypisuje się jednak o wiele za mocne znaczenie biologicznemu określeniu „przypadkowa”. Te przypadkowe kombinacje są przypadkowe w tym sensie, że nie wynikają one z planu organizmu i na ogół nie mają znaczenia dla jego zdolności do przeżycia; być może są one również przypadkowe w tym sensie, że nie są przewidywalne. Nie wynika z tego jednak, oczywiście, że są one przypadkowe w znacznie mocniejszym sensie niebycia przyczynowanymi, kierowanymi i organizowanymi przez Boga. Przypuśćmy, że biologowie, czy ktoś inny, *faktycznie* mają na myśli mocniejszy sens „przypadkowości” – wówczas ich teoria (nazwijmy ją T) rzeczywiście pociągałaby to, że istoty ludzkie nie zostały zaprojektowane przez Boga. Jednakże T nie byłaby bardziej prawdopodobna niż nieprawdopodobna ze względu na materiał empiryczny. Istniałaby bowiem empirycznie równoważna teoria (powstała z T przez przyjęcie słabszego sensu „przypadkowej” i dodanie, że Bóg kierował mutacjami), która jest nie-spójna z T, choć w równym stopniu wspierana przez materiał empiryczny; jeśli tak, to T nie jest bardziej prawdopodobna niż nieprawdopodobna ze względu na wchodzący w grę materiał dowodowy.

<sup>6</sup> List do Williama Grahama, Down, 3 lipca 1881 r. według: F. Darwin (red.), *The Life and Letters of Charles Darwin Including an Autobiographical Chapter*, t. 1, London: John Murray, Albermarle Street 1887, s. 315–316. Evan Fales i Omar Mirza zauważyli, że Darwin miał tu prawdopodobnie na myśli nie nasze codzienne przekonania, takie

Tę samą myśl wyraża jeszcze dobitniej Patricia Churchland. Podkreśla ona, że najważniejszą rzeczą dotyczącą mózgu ludzkiego jest to, iż powstał on na drodze ewolucji; oznacza to, mówi ona, że jego główną funkcją jest umożliwienie organizmowi odpowiedniego *poruszania się*:

Sprowadzając rzecz do sedna, układ nerwowy umożliwia organizmowi skuteczną realizację czterech P: pożywiania się, przemieszczania, pokonywania przeciwników oraz prokreacji. Zasadniczym zadaniem układu nerwowego jest kierowanie części ciała tam, gdzie powinny się one znaleźć, aby organizm mógł przetrwać [...] Udoskonalenia w zakresie kontroli sensomotorycznej przynoszą korzyść ewolucyjną: bardziej wyrafinowany sposób reprezentowania jest korzystny *o tyle, o ile jest włączony w sposób życia organizmu i zwiększa szanse organizmu na przetrwanie* [podkreślenie Churchland]. Prawda, czymkolwiek by była, jest najmniej ważna<sup>7</sup>.

Churchland, jak sądzę, ma tu na myśli to, że ewolucja jest bezpośrednio zainteresowana (by się tak wyrazić) wyłącznie *adaptacyjnym zachowaniem* (w szerokim sensie, obejmującym funkcjonowanie fizyczne), a nie prawdziwymi przekonaniem. Naturalnej selekcji nie interesuje to, jakie są wasze *przekonania*, lecz wyłącznie to, jak się *zachowujecie*. Zachodzi ona ze względu na określone rodzaje zachowania: te, które zwiększają przystosowanie. Owo przystosowanie jest bowiem miarą szans na to, że geny danego osobnika będą szeroko reprezentowane w następnych pokoleniach. Nie mają natomiast znaczenia przekonania, chyba że są one odpowiednio powiązane z zachowaniem. To jednak oznacza, że fakt, iż rozwinęliśmy się na drodze ewolucji, gwarantuje co najwyżej, że *zachowujemy się* w określony sposób – przyczyniający się do naszego (naszych przodków?) przetrwania i rozmnażania się w otoczeniu, w którym się rozwinęliśmy. Być może twierdzenie Churchland można zrozumieć jako sugestię, że – jeśli założymy naturalizm i to, że zostaliśmy poskładani w całość przez procesy, na które zwraca naszą uwagę współczesna teoria ewolucji – obiektywne prawdopodobieństwo<sup>8</sup> tego, iż nasze władze są rzetelne, jest niskie. Oczywiście, nie wspomina ona wprost o naturalizmie, trudno jednak oprzeć się wrażeniu, że go zakłada. Gdyby bowiem

---

jak to, że dzbanek jest w kredensie, lecz raczej coś w rodzaju przeświadczeń religijnych i filozoficznych.

<sup>7</sup> P. Churchland, *Epistemology in the Age of Neuroscience*, „Journal of Philosophy”, 84(1987), nr 10, s. 548.

<sup>8</sup> W sprawie charakterystyki obiektywnego prawdopodobieństwa zob. A. Plantinga, *Warrant and Proper Function*, s. 161n.

prawdziwy był teizm, Bóg mógłby kierować i zarządzać zmianami w taki sposób, aby wytworzyć, w dłuższej perspektywie, istoty stworzone na jego obraz i wobec tego zdolne do posiadania wiedzy; wówczas jednak prawda nie byłaby najmniej ważna.

Twierdzenie Churchland możemy przedstawić następująco:

$P(R/N\&E)$  jest niskie

gdzie  $R$  jest sądem, że nasze władze poznawcze są rzetelne,  $N$  jest sądem, że naturalizm jest prawdziwy, a  $E$  – sądem, że rozwinęliśmy się w taki sposób, jak sugeruje to współczesna teoria ewolucji<sup>9</sup>. Sądzę, że właśnie ta myśl – myśl, że  $P(R/N\&E)$  jest niskie – jest również powodem wcześniej przywołanych zmartwień Darwina, dlatego będę ją nazywał „wątpliwością Darwina”.

Czy Darwin i Churchland mają rację? Bez wątpienia słusznie uważają, że naturalna selekcja jest bezpośrednio zainteresowana jedynie zachowaniem, a nie przekonaniem i że jeśli w ogóle interesują ją przekonania, to tylko pośrednio, ze względu na relację między zachowaniem a przekonaniem. Jeśli adaptacyjne zachowanie gwarantuje lub uprawdopodobnia rzetelność władz poznawczych, to być może  $P(R/N\&E)$  będzie dość wysokie: nasze zachowanie – czy raczej zachowanie naszych przodków – jest, przynajmniej w rozsądnych granicach, adaptacyjne, a zatem nasze władze poznawcze muszą być, przynajmniej w rozsądnych granicach, adaptacyjne, wobec czego jest prawdopodobne, że większość naszych przekonań jest prawdziwa. Z drugiej strony, jeśli adaptacyjne zachowanie *nie* gwarantuje czy nawet nie czyni zbyt prawdopodobnym tego, że nasze władze są rzetelne, to przypuszczalnie  $P(R/N\&E)$  będzie raczej niskie. Jeśli, na przykład, zachowanie nie jest przyczynowane czy kierowane przez przekonania, to przekonania byłyby, można powiedzieć, niewidzialne dla naturalnej selekcji; w takim przypadku byłoby nieprawdopodobne, że warunek rzetelności – ogromna przewaga prawdziwych przekonań nad fałszywymi – zostanie spełniony<sup>10</sup>. A zatem odpowiedź na pytanie

<sup>9</sup> W *Warrant and Proper Function* omawiane prawdopodobieństwo było nieco bardziej złożone –  $P(R/N\&E\&C)$ , gdzie  $C$  jest sądem ustalającym niektóre z głównych cech naszego systemu poznawczego (zob. s. 220). Obecnie sądę, że ta dodatkowa komplikacja jest niepotrzebna.

<sup>10</sup> Ewentualnie moglibyśmy powiedzieć, że prawdopodobieństwo jest tu nieokreślalne, tak że nie możemy go oszacować. Zgoda: jest nieprawdopodobne, aby duży zbiór przekonań (porównywalny pod względem wielkości ze zbiorem przekonań posiadanych przez istotę ludzką) zawierał głównie prawdy; to daje nam rację do uznania wchodzącego w grę prawdopodobieństwa za niskie. Z drugiej strony, wiemy coś jeszcze o owym zbiorze są-

o wartość  $P(R/N\&E)$  zależy faktycznie od relacji między przekonaniem a zachowaniem. Fakt, że rozwinęliśmy się i przetrwaliśmy, zwiększa prawdopodobieństwo rzetelności naszych władz poznawczych i prawdziwości większości naszych przekonań tylko wtedy, gdy jest niemożliwe lub nieprawdopodobne, iż istoty mniej lub bardziej nas przypominające zachowywałyby się w sposób ułatwiający przetrwanie, mimo że większość ich przekonań byłaby fałszywa<sup>11</sup>.

Czy *jest* to niemożliwe lub nieprawdopodobne? Zależy to od relacji między przekonaniem a zachowaniem. Jaka byłaby czy mogłaby być ta relacja? Aby uniknąć gatunkowego szowinizmu, proponuję, abyśmy pomyśleli nie o nas samych i o naszym zachowaniu, lecz o populacji istot bardzo do nas podobnych, zamieszkujących planetę bardzo podobną do Ziemi (Darwin zaproponował, abyśmy pomyśleli w tym kontekście o małpach). Istoty te są *racjonalne*, tzn. tworzą przekonania, przeprowadzają wnioski, zmieniają przekonania itp. Wyobraźmy sobie również, że istoty te wraz z ich systemami poznawczymi rozwinęły się zgodnie z mechanizmami, na które zwraca naszą uwagę współczesna teoria ewolucji, niekierowane ręką Boga ani nikogo innego. Jakie jest  $P(R/N\&E)$  określone nie w odniesieniu do nas, lecz do nich? Aby odpowiedzieć na to pytanie, musimy się zastanowić nad relacją między ich przekonaniem

---

dów, mianowicie, że jest to zbiór, którego każdy element jest treścią czyjś przekonania. Jaki ma to wpływ na prawdopodobieństwo? Być może nie wiemy, co powiedzieć, i powinniśmy uznać, że omawiane prawdopodobieństwo jest nieokreślalne. (W tym miejscu jestem wdzięczny Johnowi Haremu).

<sup>11</sup> Czy musimy się zgodzić z Donaldem Davidsonem, który uważa, że „nie można poprawnie utrzymywać, iż każdy mógłby w większości mylić się w kwestii tego, jak mają się rzeczy”? (zob. D. Davidson, *A Coherence Theory of Truth and Knowledge*, w: D. Henrich (red.), *Kant oder Hegel?*, Stuttgart: Klett-Cotta Buchhandlung 1983, s. 535). Nie. Davidson dowodzi jedynie (o ile w ogóle), że nie mogę *zrozumieć* innej istoty, o ile nie założę, że jej przekonania są w większości prawdziwe. Być może tak jest (lub, co bardziej prawdopodobne, tak nie jest); nie wynika z tego jednak, że nie mogłyby istnieć istoty posiadające w większości fałszywe przekonania i *a fortiori* nie wynika, że moje własne przekonania są w większości prawdziwe. Davidson argumentował następnie, że *wszechwiedzący* interpretator byłby zmuszony używać tych samych metod, których my używamy, i dlatego musiałby przyjąć, że jego rozmówca posiada w większości prawdziwe przekonania; zakładając wszechwiedzę wszechwiedzącego interpretatora, doszedł on do wniosku, że jego rozmówca faktycznie posiadałby w większości prawdziwe przekonania. Dochodząc do takiego wniosku, wyraźnie zakłada on, że każdy sąd, który byłby treścią przekonania wszechwiedzącego interpretatora, jest prawdziwy; przesłanka ta bezpośrednio prowadzi do wniosku, że istnieje istota wszechwiedząca (ponieważ każda istota wszechwiedząca warta tej nazwy będzie posiadać przekonanie, że *istnieje* istota wszechwiedząca) – wniosku, do którego Davidson nie chciałby się zobowiązywać w tak bezpośredni sposób. Zob. *Warrant and Proper Function*, s. 80–81.



a zachowaniem. Istnieją tu cztery wzajemnie się wykluczające i łącznie wyczerpujące możliwości<sup>12</sup>.

(1) Pierwszą możliwością jest *epifenomenalizm*<sup>13</sup>: ich zachowanie nie jest przyczynowane przez przekonania. Zgodnie z tą możliwością ich ruchy i zachowania byłyby przyczynowane przez coś – być może impulsy neuronalne – co byłoby przyczynowane przez inne stany organiczne, w tym pobudzenie zmysłowe, samo przekonanie nie należałoby jednak do łańcucha przyczynowego prowadzącego do zachowania. Ten pogląd na relację między zachowaniem a przekonaniem (oraz innymi zjawiskami mentalnymi, takimi jak odczucia, doznania i pragnienia) jest obecnie dość popularny, zwłaszcza wśród tych, którzy ulegają silnemu wpływowi nauk biologicznych. Jak informuje grudniowy numer magazynu „Time” z 1992 r., znany biolog J.M. Smith napisał, że „nigdy nie rozumiał, dlaczego organizmy mają uczucia. Przecież ortodoksyjni ewolucjoniści sądzą, że zachowaniem, bez względu na to jak złożonym, w pełni rządzi biochemia oraz że towarzyszące mu doznania – strach, ból, ciekawość, miłość – są po prostu cieniem rzucanym przez tę biochemię i same w sobie nie mają znaczenia dla zachowania organizmu”. Smith mógłby dodać, że (zgodnie z ortodoksyjną biologią) to samo jest prawdą w odniesieniu do przekonania – przynajmniej jeśli przekonania same nie są jedynie zjawiskami biochemicznymi. O ile ten sposób myślenia jest słuszny w odniesieniu do

<sup>12</sup> W *Warrant and Proper Function* argument obejmuje pięć wzajemnie rozłącznych i łącznie wyczerpujących możliwości; tutaj łączę ze sobą dwie pierwsze.

<sup>13</sup> Po raz pierwszy nazwany tak przez T.H. Huxleya („buldoga Darwina”): „Można przyjąć [...], że zmiany molekularne w mózgu są przyczynami wszystkich stanów świadomości [...] [Czy jednak] istnieje jakikolwiek materiał dowodowy, że te stany świadomości mogą z kolei przyczynować [...] molekularne zmiany [w mózgu], które dają początek ruchom mięśni? Nie dostrzegam żadnego takiego materiału dowodowego [...] [Świadomość wydaje się] [...] całkowicie pozbawiona mocy modyfikowania działań ciała, podobnie jak para [...] w lokomotywie nie ma żadnego wpływu na jej maszynierię” (T.H. Huxley, *On the Hypothesis that Animals are Automata and its History* [1874], w: *Methods and Results*, London: Macmillan 1893, s. 239–240). Nieco dalej w tym samym eseju czytamy: „O ile mi wiadomo, argumentacja, która stosuje się do zwierząt, dotyczy w równej mierze ludzi; a zatem [...] wszystkie stany świadomości w nas, tak jak w nich, są bezpośrednio przyczynowane przez zmiany molekularne w mózgu-substancji. Wydaje mi się, że w przypadku ludzi, tak jak w przypadku zwierząt, nie istnieje żaden dowód na to, że stan świadomości jest przyczyną zmiany w ruchu materii organizmu [...] Jesteśmy świadomymi automatami” (tamże, s. 243–244). (Odnotujmy pojawiającą się tu, szeroko rozpowszechnioną formę argumentacji: „Nie wiem o żadnym dowodzie na to, że *nie-p*; a zatem nie istnieje dowód na to, że *nie-p*; a zatem *p*”). Jednakże tutaj używam terminu „epifenomenalizm” na oznaczenie *dowolnego* poglądu, zgodnie z którym przekonanie nie należy do łańcucha przyczynowego prowadzącego do zachowania, bez względu na to, czy pogląd ten wiąże się z dualizmem, który wydaje się częścią wersji Huxleya.

naszych hipotetycznych istot, to ich przekonania byłyby *niewidzialne* dla ewolucji; a wówczas fakt, że ich mechanizmy tworzenia przekonań powstały w trakcie ich ewolucyjnych dziejów, nadawałyby małe (jeśli w ogóle jakies) prawdopodobieństwo idei, że ich przekonania są w większości prawdziwe lub zbliżone do prawdy. W istocie prawdopodobieństwo tego, że przekonania te byłyby w większości prawdziwe, należałoby uznać za dość niskie (lub nieokreślalne). A zatem przyjmując N&E oraz pierwszą możliwość, prawdopodobieństwo *R* będzie dość niskie.

(2) Drugą możliwością jest *semantyczny epifenomenalizm*: mogłoby być tak, że ich przekonania są faktycznie przyczynowo sprawcze względem zachowania, jednak nie na mocy ich *treści*. Mówiąc współczesnym popularnym żargonem, byłyby to propozycja, zgodnie z którą przekonania są w istocie przyczynowo sprawcze, lecz na mocy ich *syntaktyki*, a nie *semantyki*. Zgodnie z naturalistycznym, a przynajmniej materialistycznym sposobem myślenia, przekonanie mogłoby, być może, stanowić coś w rodzaju długotrwałego układu pobudzenia neuronalnego, być długotrwałym zdarzeniem neuronalnym. Zdarzenie to będzie mieć własności przynajmniej dwojakiego rodzaju. Z jednej strony mamy jego własności neurofizjologiczne czy elektrochemiczne: liczbę neuronów związanych z tym przekonaniem, powiązania między nimi, ich progi pobudzenia, częstotliwość i siła tych pobudzeń, sposób, w jaki zmieniają się one na przestrzeni czasu i w odpowiedzi na inną aktywność neuronalną itd. Nazwijmy te własności przekonania *syntaktycznymi*. Z drugiej jednak strony, jeśli to przekonanie jest naprawdę *przekonaniem*, to będzie przekonaniem, że *p*, dla pewnego sądu *p*. Przypuśćmy, że jest to przekonanie, iż w miejscu, gdzie teraz stoi Metropolitan Opera House, znajdował się kiedyś browar. Sąd ten, jak moglibyśmy powiedzieć, jest *treścią* omawianego przekonania, a zatem oprócz własności syntaktycznych przekonanie będzie również mieć własności *semantyczne*<sup>14</sup> – na przykład: własność bycia przekonaniem, że w miejscu, gdzie teraz stoi Metropolitan Opera House, znajdował się kiedyś browar. (Inne własności semantyczne: bycie prawdziwym lub fałszywym; implikowanie, że istniał przynajmniej jeden browar; bycie spójnym z sądem, że wszyscy ludzie są śmiertelni, itd.). Zgodnie z tą drugą możliwością, przekonanie jest rzeczywiście przyczynowo sprawcze ze względu na zachowanie, jednak dzieje się to na mocy nie jego własności *semantycznych*, lecz syntaktycznych. O ile pierwsza możliwość jest bardzo popularna wśród tych, którzy ulegają wpływowi

<sup>14</sup> Zgoda: analogia między tymi własnościami a syntaktyką i semantyką jest dość odległa i naciągana; podążam tu za przyjmowanym współcześnie zwyczajem.



nauk biologicznych, o tyle druga możliwość jest bardzo popularna wśród współczesnych filozofów umysłu; w istocie, Robert Cummins nie waha się jej nazwać „przyjętym poglądem”<sup>15</sup>.

Zgodnie z tym poglądem, tak jak z poprzednim,  $P(R/N\&E)$  (określone w odniesieniu do tych istot) będzie niskie. Powodem jest to, że prawdziwość lub fałszywość należą, rzecz jasna, do *semantycznych* własności przekonania, a nie do jego własności syntaktycznych. Skoro jednak te pierwsze nie należą do łańcucha przyczynowego prowadzącego do zachowania, to i w tym przypadku przekonania – czy raczej ich własności semantyczne, w tym prawdziwość i fałszywość – będą niewidzialne dla selekcji naturalnej<sup>16</sup>. Wobec tego jednak będzie nieprawdopodobne, że ich przekonania są w większości prawdziwe, a więc i nieprawdopodobne, że ich władze poznawcze są rzetelne. A zatem prawdopodobieństwo  $R$  ze względu na  $N\&E$  połączone z tą możliwością (tak jak z poprzednią) będzie stosunkowo niskie.

(3) Mogłoby być tak, że przekonania są przyczynowo sprawcze – zarówno semantycznie, jak i syntaktycznie – ze względu na zachowanie, ale *nieadaptacyjne*: z punktu widzenia przystosowania istoty te miałyby się lepiej, gdyby nie miały żadnych przekonań. Prawdopodobieństwo  $R$  ze względu na  $N\&E$  w połączeniu z tą możliwością, podobnie jak z dwiema poprzednimi, również wydaje się stosunkowo niskie.

(4) Wreszcie mogłoby być tak, że przekonania naszych hipotetycznych istot są rzeczywiście zarazem przyczynowo powiązane z ich zachowaniem

<sup>15</sup> *Meaning and Mental Representation*, Cambridge: MIT Press 1989, s. 130. W *Explaining Behavior* (Cambridge: MIT Press 1988) Fred Dretske podejmuje odważną (choć – według mnie – nieudaną) próbę wyjaśnienia, w jaki sposób, zakładając materializm w kwestii istot ludzkich, przekonania (i inne reprezentacje) mogłyby – na mocy ich treści czy semantyki – odgrywać przyczynową rolę w wytwarzaniu zachowania. Problem polega częściowo na tym, że koncepcja Dretskego implikuje, iż nie istnieją różne, choć logicznie równoważne przekonania, a w istocie, że nie istnieją różne, choć przyczynowo równoważne przekonania.

<sup>16</sup> Musimy tu również rozważyć możliwość, że syntaktyka i semantyka przekonania są skutkami wspólnej przyczyny: być może istnieje przyczyna posiadania przez przekonanie określonych adaptacyjnych własności syntaktycznych, która jest zarazem przyczyną posiadanych przez przekonanie własności semantycznych (sprawia ona, że wchodzące w grę zdarzenie jest przekonaniem, że  $p$ , dla pewnego sądu  $p$ ); i być może ta przyczyna sprawia, że z wchodzącym w grę przekonaniem (zdarzeniem neuronalnym) zostaje powiązany *prawdziwy* sąd. (W tym miejscu zostałem poinstruowany przez Williama Ramseya i Patricka Kaina). Jakie byłoby prawdopodobieństwo, zakładając  $N\&E$ , że istnieje taka wspólna przyczyna? Przypuszczam, że byłoby ono stosunkowo niskie; dlatego ta wspólna przyczyna miałaby wiązać *prawdziwe* sądy z tymi zdarzeniami neuronalnymi? Być może jednak właściwą odpowiedzią jest nie, że owo prawdopodobieństwo jest niskie, lecz że jest ono nieokreślalne.

i adaptacyjne. (Przypuszczam, że jest to zdroworozsądkowy pogląd na związek między zachowaniem a przekonaniem w naszym własnym przypadku). Jakie jest prawdopodobieństwo (ze względu na to założenie połączone z *N&E*), że nasze władze poznawcze są rzetelne, i jakie jest prawdopodobieństwo, że przekonanie wytworzone przez te władze będzie prawdziwe? Argumentowałem, że nie jest ono tak wysokie, jak mogłoby się początkowo wydawać. Po pierwsze, jeśli zachowanie jest przyczynowane przez *przekonanie*, to jest ono również przyczynowane przez *pragnienie* (a także inne czynniki – takie jak podejrzenie, wątpliwość, aprobatę lub dezaprobatę, strach – które możemy tu pominąć). W przypadku każdego adaptacyjnego działania będzie istnieć wiele kombinacji przekonanie/pragnienie, które mogłyby wytworzyć to działanie; i w przypadku wielu z tych kombinacji odnośne przekonanie jest fałszywe.

Przypuśćmy, że Paweł jest prehistoryczną istotą człekokształtną; zbliża się głodny tygrys. Ucieczka jest być może najodpowiedniejszym zachowaniem. Zwróciłem jednak uwagę, że to zachowanie mogłoby zostać wytworzone przez wiele różnych par przekonanie–pragnienie. Zacytuję tu samego siebie:

Być może Pawłowi bardzo *podoba* się pomysł bycia zjedzonym, jednak widząc tygrysa, zawsze ucieka w poszukiwaniu lepszej okazji, ponieważ sądzi, że jest mało prawdopodobne, iż tygrys, którego widzi, zje go. Mimo że jego przekonanie nie jest prawdziwe, wystarczy ono do umieszczenia części jego ciała tam, gdzie powinny się one znaleźć, aby mógł on przetrwać [...] Albo może myśli on, że tygrys jest dużym, przyjaznym, kudłatym kociakiem i chce się z nim pobawić; sądzi on jednak również, że najlepszym sposobem zabawy jest ucieczka [...] Albo może myśli, że tygrys jest regularnie powracającym złudzeniem, i – pragnąc zrzucić kilka zbędnych kilogramów – postanowił biegać jedną milę najszybciej, jak potrafi, zawsze, gdy napotka takie złudzenie; albo może jest on przekonany, że za chwilę wystartuje w biegu na jedną milę, chce go wygrać i sądzi, że pojawienie się tygrysa jest sygnałem startowym; albo może [...] Nie ulega kwestii, że istnieje wiele systemów przekonanie–pragnienie, które równie dobrze pasują do danego zachowania<sup>17</sup>.

Odpowiednio: istnieje wiele kombinacji przekonanie-pragnienie, które prowadzą do adaptacyjnego działania; w wielu z tych kombinacji przekonania są fałszywe. A zatem bez dodatkowej wiedzy o tych istotach trudno byłoby nam jako wysokie ocenić prawdopodobieństwo *R* ze względu na *N&E* połączone z tą ostatnią możliwością.

<sup>17</sup> A. Plantinga, *Warrant and Proper Function*, s. 225–226.

Przedstawiony dotąd argument napotyka następujący problem. O ile w przypadku tylko *jednego* działania Pawła łatwo zauważyć, że istnieje wiele prowadzących do niego kombinacji przekonanie-pragnienie, o tyle trudniej pojąć, jak większość z jego przekonań (czy nawet wszystkie) mogłaby być fałszywa, a mimo to służyć przystosowaniu i ułatwiać przetrwanie. Czy przekonania Pawła rzeczywiście mogłyby być w większości fałszywe, a mimo to prowadzić do adaptacyjnego działania? W istocie tak; być może najprostszym sposobem unaocznienia tej możliwości jest rozważenie systematycznych sposobów, na jakie jego przekonania mogłyby być fałszywe, a mimo to adaptacyjne. Być może Paweł jest zwolennikiem wczesnego Leibniza i uważa, że wszystko jest świadome (równocześnie przyjmujemy, że jest to nieprawdą); ponadto wszystkie jego sposoby odnoszenia się do rzeczy wykorzystują deskrypcje określone implikujące świadomość, tak że wszystkie jego przekonania przyjmują formę: *Taka-a-taka świadoma istota jest taka-a-taka*. Być może jest on animistą i uważa, że wszystko jest żywe. A być może uważa on, że wszystkie rośliny i zwierzęta w jego otoczeniu są czarownicami i wszystkie jego sposoby odnoszenia się do nich wykorzystują deskrypcje określone implikujące bycie czarownicą. To jednak byłoby w pełni spójne z adaptacyjnością jego przekonania; jest więc, jak sądzę, jasne, że istniałoby wiele sposobów, na które przekonania Pawła mogłyby być w większości fałszywe, będąc mimo to adaptacyjne. Ponadto, z naturalistycznego punktu widzenia, nie musimy się ograniczać do jedynie możliwych przykładów. Większość ludzkości przyjmuje jakiś rodzaj ponadnaturalnych przekonań; według naturalisty takie przekonania są adaptacyjne, choć fałszywe.

Dotychczasowe rozważania pokazują, że istnieją cztery wzajemnie się wykluczające i łącznie wyczerpujące możliwości w odniesieniu do naszej hipotetycznej populacji: epifenomenalizm *simpliciter*, epifenomenalizm semantyczny, możliwość, że przekonania są przyczynowo sprawcze względem zachowania, ale nieadaptacyjne, oraz możliwość, że ich przekonania są zarazem przyczynowo sprawcze i adaptacyjne.  $P(R/N\&E)$  będzie średnią ważoną  $P(R/N\&E\&P_i)$  dla każdej z czterech możliwości  $P_i$  – ważonych przez prawdopodobieństwa tych możliwości za względu na  $N\&E$ . Rachunek prawdopodobieństwa dostarcza nam odpowiedniej formuły:

$$(1) P(R/N\&E) = (P(R/N\&E\&P_1) \times P(P_1/N\&E) + P(R/N\&E\&P_2) \times P(P_2/N\&E) + P(R/N\&E\&P_3) \times P(P_3/N\&E) + P(R/N\&E\&P_4) \times P(P_4/N\&E))$$

Oczywiście sama idea rachunku (sugerująca przypisanie tym różnym prawdopodobieństwom określonych liczb rzeczywistych) jest śmieszna:

w najlepszym razie możemy tu podać niedokładne oszacowania. Niemniej jednak na potrzeby argumentu są one wystarczające. Zgódźmy się, że możliwość – sąd, iż przekonanie jest częścią łańcucha przyczynowego prowadzącego do zachowania zarówno na mocy jego własności neurofizjologicznych, jak też na mocy jego własności semantycznych, będąc jednak nieadaptacyjnym – jest bardzo nieprawdopodobna; powyższa formuła pokazuje jasno, że jej wkład w  $P(R/N\&E)$  można bez szkody pominąć. Zauważmy następnie, że wspólną cechą epifenomenalizmu *simpliciter* oraz epifenomenalizmu semantycznego jest deklarowanie czy implikowanie tego, że treść przekonania nie ma sprawczości przyczynowej ze względu na zachowanie; treść przekonania nie należy do łańcucha przyczynowego prowadzącego do zachowania. Te dwie możliwości możemy więc zredukować do jednej: że treść przekonania nie ma sprawczości przyczynowej. Nazwijmy tę możliwość  $-C$ . Widzieliśmy już wcześniej, że prawdopodobieństwo  $R$  ze względu na  $N\&E\&-C$  jest niskie lub nieokreślalne oraz że prawdopodobieństwo  $R$  ze względu na  $N\&E\&C$  jest również nieokreślalne lub w najlepszym razie umiarkowanie wysokie. Możemy więc uprościć (1) do postaci:

$$(2) P(R/N\&E) = P(R/N\&E\&C) \times P(C/N\&E) + P(R/N\&E\&-C) \times P(-C/N\&E)$$

tj. prawdopodobieństwo  $R$  ze względu na  $N\&E$  jest średnią ważoną prawdopodobieństw  $R$  ze względu na  $N\&E\&C$  i  $N\&E\&-C$  (ważonych przez prawdopodobieństwo  $C$  i  $-C$  ze względu na  $N\&E$ ).

Zauważyliśmy już, że lewy człon pierwszego z dwóch składników prawej strony tej równości jest umiarkowanie wysoki lub nieokreślalny; lewy człon drugiego składnika jest niski lub nieokreślalny. Pozostaje jeszcze oszacować wagi prawej strony obu składników. Jakie jest więc prawdopodobieństwo  $-C$  ze względu na  $N\&E$ : jakie jest prawdopodobieństwo, że któryś z epifenomenalistycznych scenariuszy jest prawdziwy? Przypomnijmy, że według Roberta Cumminsa semantyczny epifenomenalizm jest faktycznie przyjętym poglądem na temat relacji między przekonaniem a zachowaniem. Powodem jest to, że jeśli przyjmiemy materializm, to niezwykle trudno jest sobie wyobrazić, w jaki sposób treść przekonania mogłaby mieć przyczynowy wpływ na zachowanie. Jeśli przekonanie jest po prostu neuronalną strukturą pewnego rodzaju – która w jakiś sposób ma treść – to niezwykle trudno jest zrozumieć, jak treść mogłaby należeć do łańcucha przyczynowego prowadzącego do zachowania. Gdyby bowiem dana struktura tego rodzaju miała inną treść, lecz te same własności neurofizjologiczne, to jej wkład przyczynowy do zachowania byłby, jak się

wyduje, taki sam. Przyczyną związanych z zachowaniem skurczy mięśni są fizjologiczne stany układu nerwowego, w tym fizjologiczne własności tych struktur, które konstytuują przekonania; treść tych przekonań wydaje się przyczynowo nieistotna. A zatem niezwykle trudno zrozumieć, przy założeniu  $N\&E$ , jak treść przekonania może być przyczynowo sprawcza.

Inaczej mówiąc, niezwykle trudno zrozumieć, jak przy założeniu  $N\&E$  można uniknąć epifenomenalizmu – semantycznego lub *simpliciter*. (Pojawiły się wprawdzie pewne śmiałe próby, ale nie są one obiecujące). Wygląda więc na to, że  $P(-C/N\&E)$  należy ocenić jako stosunkowo wysokie; powiedzmy, że przypiszemy mu wartość 0.7. W tej sytuacji  $P(C/N\&E)$  będzie wynosić 0.3. Przyjmijmy ponadto, że  $P(R/N\&E\&-C)$  wynosi – powiedzmy – 0.2. Wówczas  $P(R/N\&E)$  będzie wynosić co najwyżej 0.45, mniej niż 0.5. Oczywiście nietrudno tu o pomyłkę; argument na rzecz przypisania niskiej wartości  $P(R/N\&E)$  nie jest w żadnym razie niepodważalny; przypisanie przez nas rozmaitych wartości prawdopodobieństwu związanym z wartością  $P(R/N\&E)$  ze względu na tę hipotetyczną populację było (naturalnie) zarówno niedokładne, jak i słabo umotywowane. Dlatego ktoś mógłby rozsądnie twierdzić, że właściwym rozwiązaniem jest w tym przypadku po prostu agnostycyzm: nie wiadomo, jakie jest  $P(R/N\&E)$ . Macie wątpliwości, czy jest ono wysokie; nie jesteście jednak również gotowi powiedzieć, że jest ono niskie: nie macie w ogóle określonej opinii na temat tego, ile mogłoby wynosić to prawdopodobieństwo. W takiej sytuacji prawdopodobieństwo to jest dla was *nieokreślalne*. Takie stanowisko również wydaje się racjonalne. Rozsądną rzeczą jest więc przyjęcie, że  $P(R/N\&E)$  jest niskie lub nieokreślalne.

Wracając teraz do wątpliwości Darwina, zauważmy, że jeśli rozsądne jest przyjęcie takiej postawy wobec  $P(R/N\&E)$  w odniesieniu do tej hipotetycznej populacji, to równie rozsądne będzie jej przyjęcie wobec  $P(R/N\&E)$  w odniesieniu do nas. Jeśli  $N\&E$  jest prawdziwe w odniesieniu do nas, to pod istotnymi względami jesteśmy podobni do tych hipotetycznych istot – nasze władze poznawcze mają taką samą genezę, jak ich hipotetyczne władze. Kolejny krok argumentu polegał na zwróceniu uwagi, że każda z tych postaw – pogląd, że  $P(R/N\&E)$  jest niskie, oraz pogląd, że to prawdopodobieństwo jest nieokreślalne – dostarcza naturalście ewolucjoniście racji unieważniającej dla  $R$ . Każda z nich dostarcza racji do wątpienia w  $R$ , racji do powstrzymania się od przyjęcia  $R$ . Argumentowałem za tym przez analogię. Wśród najważniejszych faktów związanych z pytaniem o rzetelność jakiejś grupy władz poznawczych są fakty dotyczące ich genezy. Przypuśćmy, że jestem przekonany, iż stworzył mnie złośliwy demon kartezjański, który czerpie przyjemność z projekto-

wania istot posiadających w większości fałszywe przekonania (a mimo to uważających się za wzór doskonałości poznawczej): mam wówczas rację unieważniającą dla mojego naturalnego przekonania o rzetelności moich władz poznawczych. Przechodząc z kolei do współczesnej wersji tego scenariusza, przypuśćmy, że doszedłem do przekonania, iż zostałem porwany przez supernaukowców z Alfa Centauri, którzy w ramach eksperymentu poznawczego dostarczyli mi w większości fałszywych przekonania: w tym przypadku również mam rację unieważniającą dla *R*. Aby jednak mieć taką rację dla *R*, nie muszę koniecznie być przekonany, że faktycznie zostałem stworzony przez kartezyjańskiego demona lub że zostałem porwany przez supernaukowców z Alfa Centauri. Wystarczy, że rozważyłem te scenariusze i że uznałem, iż prawdopodobieństwo tego, że jeden z nich jest prawdziwy, jest dla mnie nieokreślalne. Wystarczy, że rozważyłem te scenariusze i, w świetle mojej wiedzy czy moich przekonań, jeden z nich jest prawdziwy. W takim przypadku również mam rację do wątpienia, rację do zawieszenia<sup>18</sup> mojego naturalnego przekonania o tym, że moje władze poznawcze są faktycznie rzetelne.

Oczywiście racje unieważniające same mogą zostać z kolei unieważnione. Na przykład: wiem, że jesteś ratownikiem, i na tym opieram moje przekonanie, że jesteś doskonałym pływakim. Następnie dowiaduję się, że 45 procent fryzyjskich ratowników to słabi pływacy, a zarazem wiem, że ty jesteś Fryzyjczykiem: dostarcza mi to racji unieważniającej dla przekonania, że jesteś dobrym pływakim. Nieco później dowiaduję się, że ukończyłeś studia na Wydziale Ratownictwa Uniwersytetu Leeuwarden i że jednym z warunków ukończenia tych studiów jest bycie doskonałym pływakim: dostarcza mi to racji unieważniającej dla racji unieważniającej dla mojego początkowego przekonania – unieważniacza racji unieważniającej, jak moglibyśmy go nazwać<sup>19</sup>. Czy jednak (wracając do naszego argumentu) racja unieważniająca, jaką naturalista ma dla *R*, może zostać z kolei unieważniona? Argumentowałem, że nie może<sup>20</sup>. Mogłaby ona zostać unieważniona tylko przez coś – na przykład przez argument – co obejmuje jakieś inne przekonanie (być może jako przesłankę). Jednakże każde takie przekonanie będzie mieć dokładnie tę samą rację unieważniającą co *R*. A zatem owa racja unieważniająca nie może zostać unieważniona. Jeśli jednak mam nieunieważnialną rację unieważniającą dla *R*,

<sup>18</sup> Będę używał tego terminu na oznaczenie *niebycia przekonanym*, tak że zawieszenie *p* jest albo uznaniem jego zaprzeczenia, albo nieuznaniem ani *p*, ani jego zaprzeczenia.

<sup>19</sup> I jak w istocie nazywa go John Pollock (*Theories of Knowledge*, Totowa: Rowman Littlefield 1986, s. 38–39).

<sup>20</sup> A. Plantinga, *Warrant and Proper Function*, s. 233–234.



to tym samym mam nieunieważnialną rację unieważniającą dla każdego innego przekonania *P* wytworzonego przez moje władze poznawcze, rację dla wątpienia w to przekonanie, rację dla zawieszenia go. Jest tak dlatego, że każde takie przekonanie będzie wytworem władz poznawczych, o których rzetelności nie mogę być racjonalnie przekonany. Jest jasne, że to samo będzie prawdą w odniesieniu do dowolnego przekonania, które jest przez nie wytworzone: jeśli nie mogę być racjonalnie przekonany, że władze, którym zawdzięczam to przekonanie, są rzetelne, to mam rację, by je odrzucić. A zatem zwolennik *N&E* ma rację unieważniającą dla każdego ze swoich przekonań – rację unieważniającą, która jest, jak to określiłem, ostatecznie nieunieważnialna. Oznacza to, że koniunkcja naturalizmu z ewolucją ma charakter samoobalający, tak że nie można jej racjonalnie zaakceptować. Następnie dodałem, że każdy, kto akceptuje naturalizm, powinien zaakceptować ewolucję; dla naturalisty ewolucja jest jedyną dostępną odpowiedzią na pytanie, jak powstało całe bogactwo fauny i flory. I wreszcie, jeśli tak jest w istocie, to naturalizm *simpliciter* ma charakter samoobalający i nie może zostać racjonalnie zaakceptowany – a w każdym razie nie przez kogoś, kto docenia ten argument i dostrzega powiązania między *N&E* a *R*.

*Przełożył Marcin Iwanicki*