

Może błędę, ale jednak myślę. Odpowiedź na uwagi krytyków do książki *Błędę, więc myślę*

Krystyna Bielecka 

Instytut Filozofii

Uniwersytet w Białymstoku

kikab.gazeta.pl @ gmail.com

Przyjęto 29 grudnia 2020; zaakceptowano 30 grudnia 2020; opublikowano 31 grudnia 2020*.

Bardzo dziękuję wszystkim Komentatorom za inspirujące uwagi do mojej książki i koncepcji błędnej reprezentacji (Bielecka, 2019). Odniosę się do poszczególnych zarzutów, które pojawiły się w ich tekstach, a także punktów pomagających rozwinąć mi pewne wątki występujące w książce.

Odniosę się najpierw do uwag profesora Roberta Poczobuta (2020), gdyż te dotyczą w zasadzie całości przedsięwzięcia, jakim jest stworzenie autorskiego szkicu koncepcji błędnej reprezentacji. Następnie zamieszczę odpowiedź na komentarz doktora Michała Piekarskiego (2020), który skupia się wokół wątku normatywności reprezentacji. Do komentarza doktora Pawła Gładziejewskiego (2020) odniosę się niejako przy okazji – z racji tego, że nie widzę w nim zarzutów, a raczej niezwykle ciekawy pomysł na rozwinięcie idei koherencji w filozoficznych koncepcjach reprezentacji umysłowych.

Profesor Poczobut odrzuca moją deklarację, że koherencyjna koncepcja błędnej reprezentacji (KKBR) mogłaby służyć jako podstawa dla naukowych teorii reprezentacji umysłowych. Zwraca przy tym uwagę na to, że KKBR może co najwyżej posłużyć jako jedna z wielu teorii błędnej reprezentacji, której istotowy warunek możliwości wykrywania błędów dla systemu jest dla ludzi trywialny, a dla innych organizmów kontrowersyjny. Jego argumentacja opiera się z jednej strony na innym zgoła założeniu o roli koncepcji filozoficznej w badaniach naukowych, z drugiej zaś na uznaniu dezyderatu o detekcji błędu reprezentacyjnego (błędu dla systemu) za jeden z wielu – ale nie kluczowy – warunek filozoficznej koncepcji reprezentacji umysłowych w duchu realizmu i naturalizmu. Wiąże się z tym dodatkowa, acz ważna kwestia, uznania detekcji błędu za pośrednictwem niespójności za kontrowersyjną.

Profesor Poczobut twierdzi, że filozoficzna koncepcja błędu reprezentacyjnego powinna opierać się na szczegółowej analizie mechanizmów powstawania błędów reprezentacyjnych. Moim zdaniem jednak, jest to rolą nie tyle koncepcji filozoficznej, ile teorii naukowej, co nie znaczy, że koncepcja filozoficzna w duchu realizmu i naturalizmu nie powinna inspirować się przykładami z nauki. W książce starannie odróżniam koncepcje filozoficzne od teorii naukowych; filozofowie nie tworzą zwykle teorii w standardowym znaczeniu tego wyrazu: nie mają one bowiem sprawdzalnych konsekwencji empirycznych, lecz mogą stanowić ramy pojęciowe, które da się treścią empiryczną wypełnić. Takie teorie empiryczne jednak wykraczają poza to, co filozofia ustali w zaciszu gabinetu. KKBR jest koncepcją filozoficzną, która opiera się na analizie naukowych przykładów reprezentacji pojęciowej, obrazowej i strukturalnej, gdzie pojęcie reprezentacji umysłowej jest używane w sposób paradygmatyczny. Konstruję dezyderaty wobec każdej takiej koncepcji filozoficznej, która byłaby zgodna z badaniami naukowymi i praktyką naukową. W szczególności mechanicyzm wydaje się taką koncepcją, która jest zgodna z moimi dezyderatami; KKBR nie jest jednak teorią mechanizmów reprezentacyjnych i dlatego też nie odpowiada na dalsze szczegółowe pytania o mechanizmy powstawania błędów reprezentacyjnych. Nie zgadzam się z Profesorem Poczobutem, że rolą filozofii w naukowej teorii reprezentacji umysłowych jest wyłącznie dostarczenie narzędzi formalno-logicznych, bo z tym doskonale radzą sobie naukowcy tworzący modele (niekiedy z pomocą zawodowych logików); podstawową rolą filozofa nauki jest badanie pojęć oraz struktur argumentacyjnych i ich wykorzystywanie w teoriach naukowych. Szkic mojej koncepcji wykrywania błędów nie jest więc szkicem szeroko zakrojonej teorii wykrywania błędów z podziałem na teorie reprezentacji percepcyjnych czy pojęciowych.

Dezyderat o detekcji błędu przez system jest, moim zdaniem, kluczowy dlatego, że bez możliwości wykrycia poprawności reprezentacji organizm nie może mieć dostępu do treści reprezentacji, czyli *de facto* do własnej reprezentacji. Stanowi to poważny problem dla naturalizacji treści reprezentacji, o czym świadczy chociażby dyskusja wokół tak zwanego trudnego problemu treści (por. Hutto i Myin, 2013). Bez odpowiedzi na ten problem nie można poważnie posługiwać się pojęciem reprezentacji tylko dlatego, że posługują się nim naukowcy, gdyż grozi to możliwością zastąpienia tego pojęcia opisem nieodwołującym się do pojęcia reprezentacji umysłowych; Daniel Hutto i Eric Myin podkreślają, że sama powszechność użycia terminu „reprezentacja” nie legitymizuje go ani w filozofii, ani w kognitywistyce.

Kontrowersyjne jest twierdzenie profesora Poczobuta, że nikt nie zakłada istnienia odwzorowań, do których organizm nie ma dostępu (przypis 13); takimi chociażby są tak zwane reprezentacje śmieciowe (*junk representations*) według terminologii Lawrence’a Shapiro (1997), a więc właśnie reprezentacje bez konsumenta. Zgadzam się z profesorem, że pojęcie konsumenta, a więc odbiorcy reprezentacji, które przyjmuję za Ruth Millikan, jest szeroko zakrojone i niesprecyzowane; uważam to właśnie za zaletę, gdyż pojęcie to można sprecyzować przy zastosowaniu go do opisu poszczególnych mechanizmów wykrywania błędów reprezentacyjnych zarówno na poziomie osobowym (a więc także świadomym), jak i subosobowym.

Dezyderat o detekcji błędu, który spełnia w moim mniemaniu KKBR, jest właśnie dlatego tak ważny, że odpowiada na pytania, co to znaczy mieć dostęp do odwzorowania rzeczywistości w układzie nerwowym i na czym polega korekta błędu (por. s. 10). Odpowiedzi należy szukać nie przez dostęp do źródła błędu, lecz przez dostęp do treści reprezentacji, która nie może się

nią stać, jeśli w organizmie nie ma jej odbiorcy, czyli konsumenta, który jest realizowany w jakiejś części układu nerwowego. Dlatego nie można przypisać zdolności do wykrycia błędu reprezentacyjnego organizmowi, który fizycznie i funkcjonalnie nie posiada takiej części. I tak, żaba z zaburzoną siatkówką nie rozpozna błędu tak, jak mógłby zrobić to naukowiec po szczegółowej inspekcji żaby (Poczobut, 2020, przypis 15). Ponieważ nie jest to błąd rozpoznawalny przez żabę, nie jest to błąd dla systemu, a więc nie jest to reprezentacja żaby. Co najwyżej, można by określić mianem błędu reprezentacyjnego z perspektywy obserwatora, a więc błędu w systemie, coś, co mogłoby być treścią reprezentacji żaby, ale ona nie miałaby do niego dostępu, na przykład zachowanie żaby w labiryncie prowadzące do ominięcia szerokim łukiem muchy, na którą w innych okolicznościach żaba chciałaby zapolować (w takich warunkach Normalnych, jakie są w labiryncie, żabi konsument nie byłby w stanie rozpoznać swojej reprezentacji).

Moja odpowiedź na pytanie, jak możliwa jest korekta błędu, brzmi: możliwa jest dzięki wykrywaniu niespójności. Niespójność wydaje się najbardziej bazowym sposobem wykrywania błędów reprezentacyjnych, a KKBR jest szkicem koncepcji opartej na tej idei. Idea koherencji w zastosowaniu do szerokiej klasy błędów reprezentacyjnych, zarówno percepcyjnych, jak i pojęciowych, została wykorzystana w niezwykle inspirującym komentarzu do *Błędę, więc myślę* przez doktora Gładziejewskiego (2020). Bardzo ciekawie rozwija on ideę koherencji w nawiązaniu do błędnych reprezentacji w oparciu o KKBR oraz koncepcji niespójności opracowanej przez Paula Thagarda (2000). Warto podkreślić, że koncepcja niespójności Thagarda została opracowana w kategoriach sieci neuropodobnych, a więc nie ma charakteru formalnologicznego, lecz cechuje się pewną biologiczną wiarygodnością (choć oczywiście sieci koneksjonistyczne są daleko posuniętą idealizacją działania układów nerwowych). Niespójność jest bazowa właśnie dlatego, że nie jest tym samym co niesprzeczność logiczna wykrywana w zaksjomatyzowanych rachunkach, a więc można ją zaaplikować nie tylko do reprezentacji u ludzi, ale i u zwierząt. Swoją drogą, wspomniane przeze mnie w książce algorytmy uczenia ze wzmacnianiem (zwłaszcza w architekturze aktor–krytyk) są bardzo często odnoszone przez naukowców do badań nad zwierzętami. Podobnie zresztą jest z szacowaniem stopnia rzetelności informacji: chodzi mianowicie o wrodzone bądź nabyte w ramach uczenia zdolności do odróżniania informacji mniej i bardziej istotnych. I to na przykład dzieje się w sytuacji, gdy szczur z mapą poznawczą starego labiryntu przychodzi do nowego labiryntu: jego reprezentacja jest błędna, ale – i tu zgadzam się z profesorem Poczobutem – może się zaadaptować do nowej sytuacji, zastępując stare wzorce nowymi. Odpowiedź, że dzieje się to poprzez wykrytą niespójność, nie tylko wyjaśnia reakcję zdziwienia szczura, ale dostarcza wyjaśnienia, jak to się dzieje, że szczur adaptuje się do nowego otoczenia.

Przypisanie zwierzętom zdolności do wykrywania błędów reprezentacyjnych nie jest antropomorfizacją. Są przecież badania neurobiologiczne wskazujące na ich istnienie, na przykład u zeberek wykrywających błędy występujące w ich własnym śpiewie (Bielecka i Miłkowski, 2020). Nie jest to antropomorfizacja, gdyż mechanizmy wykrywania występują na poziomie

subosobowym (a więc niespójność nie musi być reprezentowana, co wymagałoby metareprezentacji¹), nie jest to też nic automatycznego w takim sensie, w jakim mowa o czynnościach pozbawionych aspektu semantycznego (szacowanie rzetelności ma wpływ na to, jak organizm ustosunkowuje się do tego, co spostrzeża i jak to aktywnie wykorzystuje w działaniu). Zastąpienie zdolności do wykorzystania niespójności tym, że „wykorzystywane są dwa systemy sprawdzające, wzrokowy i dotykowy” (Poczobut, 2020, s. 11), nie pozwalałoby odpowiedzieć na pytanie, co to znaczy „wykorzystywać”.

Główny zarzut doktora Piekarskiego dotyczy normatywności reprezentacji i koncepcji funkcji. Według niego normatywność reprezentacji nie sprowadza się tylko do bycia poprawnym lub niepoprawnym, ale normatywność ma być przede wszystkim wewnętrzną, niezależną od obserwatora, a w szczególności od perspektywy narzuconej w historii adaptacyjnej organizmu, własnością reprezentacji. Doktor Piekarski, posługując się argumentami Paula Daviesa (2001), zachęca, aby przyjąć systemową koncepcję funkcji, gdyż pozwala ona mówić o systemowych własnościach reprezentacji niezależnie od jej historii adaptacyjnej. Moja odpowiedź na ten zarzut jest złożona.

I tak, zgadzam się, że normatywność funkcji powinna występować niezależnie od perspektywy obserwatora, jeśli chcemy realistycznie podejść do problemu reprezentacji umysłowych. Jednak nie zgadzam się z krytyką Daviesa i doktora Piekarskiego, że koncepcje teleosemantyczne, przyjmujące etiologiczne ujęcie funkcji, myślą perspektywę systemu z perspektywą obserwatora. Jest wręcz odwrotnie: to naturalna selekcja pewnych związków między organizmem a środowiskiem w historii adaptacyjnej organizmu pozwala na pojawienie się takich a nie innych warunków nałożonych na poprawność reprezentacji i tylko w ramach tych relacji warunki te ulegają zmianie. To raczej systemowe ujęcie funkcji jest bliższe perspektywie obserwatora (por. Cummins, 1975). Ponieważ jednak nie zawsze historia ewolucyjna organizmu jest kluczowa dla wyznaczenia poprawności reprezentacji, to przyjmuję w swoim ujęciu koncepcję hybrydową, a więc etiologiczno-systemową.

Nie wydaje się, żeby normatywność miała być jakąś specjalną własnością wewnętrzną; sam doktor Piekarski zresztą dystansuje się wobec tego pomysłu na rzecz własności relacyjnej organizmu i środowiska. Bardzo ważny fakt, że środowisko ogranicza swobodę zachowania organizmu, co wpływa na treść jego reprezentacji, świadczyć ma, jak sądzę, o tym, że normatywność reprezentacji nie jest własnością wewnętrzną i niezmienną. O roli środowiska w wyznaczaniu treści mówią zarówno zwolennicy funkcji systemowych (na przykład Mark Bickhard, którego koncepcja opiera się na systemie dalekim od równowagi termodynamicznej), jak i etiologicznych (na przykład Millikan z jej koncepcją warunków Normalnych, które stanowią ograniczenia nie tylko wewnętrzne, wynikające ze struktury organizmu, ale i ze-

¹ Dokładniejsza analiza tego zagadnienia znajduje się w przywołanym artykule. W uproszczeniu: różnica w porównaniu z koncepcjami wykrywania błędów zakładającymi, że wykryty błąd ma charakter metareprezentacji, jest taka, że odrzucam konieczność reprezentowania błędów. W mojej koncepcji wszystkie reprezentacje powinny móc być sprawdzane, a to prowadziło do powstawania kolejnych metareprezentacji, które z kolei powinny być sprawdzane. W istocie jednak sam wykryty błąd nie musi odgrywać takiej samej roli co zwykle reprezentacje. W systemie poznawczym może pozostać tylko informacja semantyczna na temat tego błędów, której adekwatność nie będzie już dalej sprawdzana.

wewnętrzne). Wbrew temu, co twierdzi doktor Piekarski, że „jeśli KKBR ma oferować nietrywialne wyjaśnienie zachowania organizmów w kategoriach reprezentacyjnych, to powinna wziąć pod uwagę nie tylko wewnętrzne komponenty określonych mechanizmów reprezentacyjnych, ale także ograniczenia środowiskowe (...)” (2020, s. 17), KKBR oparte na funkcji hybrydowej już takie wyjaśnienie oferuje i w tym sensie artykuł (Bielecka i Miłkowski, 2020) nie jest rozwinięciem koncepcji z książki (wbrew temu, co twierdzi doktor Piekarski na s. 16). Artykuł ten, podobnie jak (Bielecka i Marcinów, 2017), dostarcza przykładów zastosowania funkcji hybrydowej do wyjaśnienia reprezentowania. Swoją drogą, pierwszy tekst może nieco zaspokoić potrzebę materiału empirycznego, tak podkreśloną przez profesora Poczobuta, gdyż dokładniej opisuje mechanizmy powstawania błędów reprezentacyjnych u zebrek.

Wątek roli czynników motywacyjnych jest bardzo ciekawy i wydaje się mieć duże znaczenie w przypadku wyjaśniania mechanizmów uczenia i pewnych psychopatologii. Zgadzam się z doktorem Piekarskim, że rola motywacji jest bardzo ważna, ponieważ bardzo istotny jest element pragmatyczny w wyznaczaniu treści reprezentacji.

Jeszcze raz dziękuję za otrzymane komentarze i uwagi wszystkim komentatorom, a także uczestnikom sympozjum poświęconego mojej książce, które zorganizował doktor Piekarski w Instytucie Filozofii i Socjologii Polskiej Akademii Nauk. Była to znakomita okazja do pogłębionych i pouczających dla mnie dyskusji.

Finansowanie

Tekst został sfinansowany z grantu NCN kierowanego przez dr Krystynę Bielecką, nr 2016/23/D/HS1/02205.

Bibliografia

- Bielecka, K. (2019). *Błędę, więc myślę: Co to jest błędna reprezentacja?* Warszawa: Wydawnictwa Uniwersytetu Warszawskiego.
- Bielecka, K. i Marcinów, M. (2017). Mental Misrepresentation in Non-human Psychopathology. *Biosemiotics*, 10(2), 195–210. doi: 10.1007/s12304-017-9299-2
- Bielecka, K. i Miłkowski, M. (2020). Error Detection and Representational Mechanisms. W J. Smortchkova, K. Dołęga i T. Schlicht (red.), *What are Mental Representations?* (s. 287–313). New York: Oxford University Press.
- Cummins, R. (1975). Functional Analysis. *The Journal of Philosophy*, 72(20), 741–765.
- Davies, P. S. (2001). *Norms of nature: Naturalism and the nature of functions*. Cambridge: MIT Press.
- Gładziejewski, P. (2020). Koherencja drogowskazem prawdy? Spójność jako źródło błędnych reprezentacji Komentarz do książki Krystyny Bieleckiej *Błędę, więc myślę. Co to jest błędna reprezentacja?* *Avant*, 11(3). doi: 10.26913/avant.2020.03.26
- Hutto, D. D. i Myin, E. (2013). *Radicalizing enactivism: Basic minds without content*. Cambridge Mass.: MIT Press.

- Piekarski, M. (2020). Normatywne mechanizmy reprezentacyjne a koherencyjna koncepcja błędu reprezentacyjnego. *Avant*, 11(3). doi: 10.26913/avant.2020.03.29
- Poczobut, R. (2020). Koherencyjna koncepcja błędu reprezentacyjnego w kontekście badań nad mechanizmami powstawania błędnych reprezentacji pamięciowych. *Avant*, 11(3). doi: 10.26913/avant.2020.03.25
- Shapiro, L. (1997). Junk Representations. *The British journal for the philosophy of science*, 48(3), 345–362. doi: 10.1093/bjps/48.3.345
- Thagard, P. (2000). *Coherence in Thought and Action*. Cambridge, Mass.: MIT Press.