



Koherencyjna koncepcja błędu reprezentacyjnego w kontekście badań nad mechanizmami powstawania błędnych reprezentacji pamięciowych

Robert Poczobut 

Instytut Filozofii

Uniwersytet w Białymstoku

r.poczobut@uwb.edu.pl

Przyjęto 20 listopada 2019; zaakceptowano 25 lutego 2020; opublikowano 28 lipca 2020

Abstrakt

Celem artykułu jest analiza krytyczna stanowiska Krystyny Bieleckiej przedstawionego w książce: *Błądź, więc myślę. Co to jest błędna reprezentacja?* Artykuł zawiera również zarys alternatywnego, naturalistycznego ujęcia błędnych reprezentacji poznawczych opartego na mechanistycznym modelu wyjaśniania naukowego. Według mnie główny problem teorii reprezentacji poznawczych (oraz błędnych reprezentacji) dotyczy mechanizmów ich powstawania: neurobiologicznych, psychologicznych, społecznych i innych. Swoje stanowisko krótko ilustruję przykładami mechanizmów błędnych reprezentacji pamięciowych. Końcowa część artykułu zawiera kilka uwag na temat ogólnej teorii błędnych reprezentacji poznawczych oraz roli, jaką w jej budowie może odegrać filozofia.

Słowa kluczowe: reprezentacja poznawcza; błędna reprezentacja poznawcza; błąd reprezentacyjny; mechanistyczny model wyjaśniania naukowego; błędy pamięci; mechanizmy błędnych reprezentacji pamięciowych

Każdy system, kiedy zawodzi, popełnia pewne charakterystyczne typy błędów. Na szczęście stanowią one cenne źródło informacji – i to nie tylko dla systemu, który dzięki nim się uczy, lecz również dla nas, bo pozwala odkrywać, jak działa system i jakiego jest rodzaju.

Chris Frith (2011, s. 141)

1. Wprowadzenie

Celem artykułu jest analiza krytyczna wybranych twierdzeń koherencyjnej koncepcji błędu reprezentacyjnego (KKBR) przedstawionej w książce Krystyny Bieleckiej (2018): *Błądzą, więc myślę. Co to jest błędna reprezentacja?* oraz zarysowanie projektu badań nad mechanizmami powstawania błędnych reprezentacji poznawczych (BRP). Jednym z głównych zadań realistycznej, a zarazem naturalistycznej koncepcji reprezentacji poznawczych (RP), jest odkrycie wytwarzających je mechanizmów. Z kolei znajomość różnorodnych czynników zaburzających ich działanie umożliwia wyjaśnienie, w jaki sposób w systemach poznawczych powstają BRP¹.

Odkrycie mechanizmów wytwarzających RP wystarcza do odrzucenia stanowiska antyrealistycznego w sporze o status reprezentacji poznawczych, to jest twierdzenia, że RP nie istnieją lub nie pełnią żadnych funkcji eksplanacyjnych. Źródłem problemów z naturalizacją RP jest to, że filozofowie dokonują ich opisu i analizy w oderwaniu od wytwarzających je mechanizmów. Na gruncie nauki problem naturalizacji RP nie pojawia się, ponieważ opis mechanizmów ich powstawania jest częścią naukowej praktyki eksplanacyjnej. Psycholodzy poznawczy czy neurokognitywiści nie zajmują się naturalizowaniem RP, lecz odkrywaniem leżących u ich podstaw mechanizmów i pełniących przez nie funkcji (realizm i naturalizm otrzymujemy „za darmo”). Sformułowanie poprawnego mechanistycznego wyjaśnienia tego, jak powstają, do czego służą i jak działają RP – jakie pełnią funkcje i jakim podlegają dysfunkcjom – rozstrzyga filozoficzny spór na rzecz realizmu i naturalizmu².

W kolejnych sekcjach dokonam analizy krytycznej wybranych twierdzeń KKBR w kontekście badań na temat mechanizmów powstawania błędnych reprezentacji pamięciowych (BRPA). Posłużą one za ilustrację, z konieczności szkieletową, jednej z głównych tez artykułu. W zakończeniu sformułuję zarys interdyscyplinarnego programu badawczego dotyczącego natury i mechanizmów BRP oraz wskażę rolę, jaką może w nim odegrać filozofia.

¹ Mechanistyczne wyjaśnienie RP nie jest tym samym co postulowane przez Pawła Gładziejewskiego (2015) mechanistyczne wyjaśnianie reprezentacyjne. Chociaż w obu wypadkach zakładany jest ten sam model wyjaśniania mechanistycznego, inne jest *eksplanandum*. Gładziejewski skupia się na wyjaśnianiu zachowań systemu poznawczego i jego podsystemów za pomocą mechanizmów reprezentacyjnych, których integralnym składnikiem są reprezentacje strukturalne (ich nośnik pozostaje w relacji podobieństwa strukturalnego do swojego przedmiotu). Natomiast w poniższych analizach chodzi o mechanizmy wytwarzające RP (BRP). Chodzi o problem wyrażony w pytaniu: Jak powstają RP (BRP)?

² Posługuję się terminem „reprezentacje poznawcze”, mimo że wśród filozofów rozpowszechniony jest termin „reprezentacje umysłowe” (traktowany jako zakresowo węższy). Moja decyzja terminologiczna wynika z braku jednoznacznych kryteriów wyróżniających reprezentacje umysłowe w zbiorze reprezentacji poznawczych, a także z tego, że kategoria błędu reprezentacyjnego ma zastosowanie do wszystkich rodzajów RP.

2. Miejsce koherencyjnej koncepcji błędu reprezentacyjnego w ramach ogólnej teorii reprezentacji poznawczych

2.1. Ogólna teoria RP powinna zawierać uzasadnione odpowiedzi na następujące pytania: Czym są RP? Jakie są ich rodzaje i kryteria wyróżniania? Jak powstają RP i jaką rolę odgrywają w działaniu systemów poznawczych (jaki jest ich profil funkcjonalny)? Co konstituuje relacje między nośnikiem, treścią i odniesieniem RP? Czym są BRP i jakie mechanizmy odpowiadają za ich powstawanie? Jakie są rodzaje błędów reprezentacyjnych? Jaką rolę BRP odgrywają w uczeniu się? Na jakie sposoby systemy poznawcze (naturalne i sztuczne) wykrywają błędy reprezentacyjne? Bez wątpienia koncepcja BRP jest integralną częścią ogólnej teorii RP. Analiza mechanizmów powstawania BRP jest źródłem istotnych informacji dla zrozumienia rzeczywistej natury RP.

Przed podjęciem analizy wybranych twierdzeń KKBR, sformułuję główne założenia reprezentacyjnej teorii poznania, których skłonny jestem bronić. Stanowią one również ogólną ramę teoretyczną dla dalszych rozważań:

A) Każdy efektywny proces poznawczy prowadzi do wytworzenia RP: percepcyjnych, pamięciowych, pojęciowych, przekonaniowych, językowych i innych. RP są wytworami procesów poznawczych zawsze realizowanych przez określone mechanizmy. Systemy poznawcze jako systemy reprezentacyjne mogą być systemami naturalnymi (ludzie i inne zwierzęta), sztucznymi (roboty kognitywne) lub hybrydowymi (częściowo naturalnymi, częściowo sztucznymi). Analogiczny podział dotyczy reprezentacji w zależności od tego, czy ich nośnik jest układem naturalnym, sztucznym (artefaktem), czy hybrydowym (mieszanym).

B) Przedmiotami RP mogą być obiekty świata zewnętrznego, wewnętrzne stany systemu poznawczego (autoreprezentacje) lub inne reprezentacje (metareprezentacje). Przedmioty RP mogą istnieć aktualnie (reprezentacje percepcyjne), istnieć w przeszłości (reprezentacje pamięciowe), jeszcze nie istnieć (reprezentacje przyszłych lub tylko możliwych zdarzeń³), być przedmiotami abstrakcyjnymi (reprezentacje matematyczne) lub fikcyjnymi (reprezentacje obiektów nieistniejących, przedstawionych w utworach literackich)⁴.

C) RP jako wytwory procesów poznawczych stanowią formę odwzorowania świata (otoczenia, środowiska), z którym systemy poznawcze wchodzi w różnorodne interakcje i oddziaływania o charakterze przyczynowo-informacyjnym. Wyróżnioną, choć nie jedyną klasę RP tworzą reprezentacje strukturalne. Ponieważ odwzorowania są układami relacyjnymi, RP mają nieredukowalnie relacyjny charakter. Ich adekwatny opis wymaga podania

³ Zdolność do antycypacji (w szczególności antycypacji skutków własnych zachowań) zakłada dysponowanie reprezentacjami przyszłych zdarzeń (nie ma antycypacji bez reprezentacji). Odwrotna zależność nie musi zachodzić, o ile mogą istnieć systemy reprezentujące środowisko wyłącznie aktualnie. Wydaje się również, że nie ma antycypacji bez pamięci (obecność pamięci stwarza możliwość antycypacji). Na temat modelu przewidywania opartego na pamięci por. Hawkins i Blakeslee (2006).

⁴ Reprezentacje przedmiotów istniejących aktualnie, jak reprezentacje percepcyjne, powstają na bazie interakcji przyczynowo-informacyjnych ze środowiskiem, zaś reprezentacje przedmiotów abstrakcyjnych, możliwych lub fikcyjnych są efektem bardziej zaawansowanych form przetwarzania na wyższych poziomach organizacji systemu poznawczego. W każdym wypadku system może generować błędy reprezentacyjne.

nie tylko charakterystyki nośnika i reprezentowanego obiektu, ale także zachodzącej między nimi relacji (z uwzględnieniem użytkownika nazywanego też „konsumentem reprezentacji”).

D) Sposób odwzorowania świata w systemie poznawczym (forma reprezentacji) zależy od specyfiki wchodzącego w grę procesu poznawczego i konstytuujących go mechanizmów. Inaczej odwzorowują otoczenie percepty zmysłowe (wzrokowe, słuchowe, dotykowe), przypomnienia (związane z różnymi rodzajami pamięci), pojęcia (schematy i ramy pojęciowe), przekonania, struktury językowe i teoretyczne czy symulacje komputerowe⁵. Co więcej, przedstawiciele różnych gatunków, wyposażeni w istotnie różne mechanizmy poznawcze, wytwarzają istotnie różne układy RP. Jednak w każdym wypadku jest obecny jakiś rodzaj odwzorowania otoczenia w systemie poznawczym, który umożliwi przedstawicielom różnych gatunków skuteczną nawigację i przetrwanie w zmieniającym się środowisku⁶.

E) Relacja odwzorowania (reprezentowania) zawsze ma charakter przybliżony (uproszczony, schematyczny w stosunku do przedmiotu reprezentacji). Nie ma reprezentacji odwzorowujących obiekty w sposób absolutnie dokładny, w pełni adekwatnie. W szczególności nie czyni tego żaden percept ani żadna reprezentacja pamięciowa. Każda reprezentacja koduje informację zredukowaną i przetworzoną w stosunku do swojego przedmiotu, zaś rodzaj redukcji jest efektem działania mechanizmów przyczynowo-informacyjnych prowadzących do wytworzenia danej reprezentacji. W efekcie poznawczy model świata zewnętrznego i wewnętrznego, doświadczanego dzięki i poprzez RP, nigdy nie pokrywa się w pełni z rzeczywistością. Ta nieuchronna nieadekwatność RP (brak pełnej zgodności ich treści z przedmiotem) nie jest błędem reprezentacyjnym, lecz naturalnym efektem działania mechanizmów poznawczych (to, że trawę widzimy jako zieloną, zaś niebo jako niebieskie, chociaż skądinąd wiemy, że tak nie jest, nie stanowi błędu reprezentacyjnego, lecz jest naturalnym efektem poprawnie działających mechanizmów percepcyjnych)⁷.

⁵ Różnica między reprezentacjami percepcyjnymi i pojęciowymi, choć rzeczywista, wiąże się z innym sposobem odwzorowania obiektów oraz innym sposobem kodowaniem informacji w sieciach neuronowych mózgu. Nie jest to jednak różnica kategoryjalna. Na wyższych poziomach przetwarzania reprezentacje percepcyjne zostają wzbogacone o elementy pojęciowo-pamięciowe (kategoryzacja percepcyjna, percepcja rozumiejąca, widzenie jako). Percepty od pojęć odróżnia poziom redukcji informacji. Ogólność reprezentacji pojęciowych wiąże się z operowaniem znacznie bardziej zredukowaną informacją (wyższym poziomem abstrakcji) niż w wypadku perceptów.

⁶ RP pełnią rolę pośrednika poznawczego: nigdy nie doświadczamy (poznajemy) świata bezpośrednio. Procesy poznawcze mają charakter konstrukcyjny. Rzeczywistość poznajemy za pomocą różnorodnych wytworów poznawczych. Jak trafnie zauważa Frith (2011, s. 141): „Mój mózg odkrywa to, co jest w rzeczywistym świecie, poprzez konstruowanie jego modeli. Nie są to modele arbitralne. Są tak skonstruowane, by umożliwiały jak najlepsze przewidywania moich wrażeń, kiedy działam w rzeczywistym świecie. Nie uzmysławiam sobie jednak tego złożonego mechanizmu. Czego w takim razie jestem świadomy? Treścią percepcji nie jest świat, lecz stworzony przez mózg model świata”.

⁷ W przeciwnym wypadku musielibyśmy uznać wszystkie RP za w pewnym stopniu błędne. Niektórzy mówią w tym kontekście o tzw. wielkiej iluzji. Por. Noë (2002), Clark (2002). Proponuję konwencję terminologiczną, zgodnie z którą o BRP można mówić dopiero wtedy, gdy nastąpi zaburzenie działania jakiegoś mechanizmu reprezentacyjnego. To, że widzimy kij zanurzony w wodzie jako złamany, nie stanowi BRP, ponieważ jest to rezultat poprawnie przebiegającego procesu percepcyjnego, zaś jego efekt iluzoryczny wyjaśniają prawa odbicia i załamania światła w różnych ośrodkach. Jeśli widzimy kij zanurzony w wodzie jako złamany, to widzimy poprawnie. Skądinąd wiadomo, że ludzie z uszkodzonymi mechanizmami percepcyjnymi mogą nie doświad-

F) Możliwość błędu wpisana jest w naturę każdej formy poznawczej reprezentacji świata. Wystąpienie BRP wiąże się z możliwością ingerencji w mechanizmy odpowiadające za wytwarzanie RP, częściowo także z ich ewolucyjnym pochodzeniem. Mechanizmy percepcyjne, pamięciowe czy konceptualne zostały wyselekcjonowane w złożonym procesie ewolucyjnym, umożliwiając wyposażonym w nie organizmom adaptację do środowiska, przetrwanie i reprodukcję. Realizacja poznawczych i pozapoznawczych funkcji RP nie wymaga istnienia doskonałych, bezbłędnych, niepodatnych na zaburzenia i defekty systemów reprezentacyjnych i obsługujących je mechanizmów. Z ewolucyjną genezą mechanizmów reprezentacyjnych nieuchronnie wiąże się ich prowizoryczność, niedoskonałość, niedokładność, podatność na zaburzenia i dysfunkcje (Marcus, 2009; Hoffman, 2012). Możliwość tworzenia BRP jest produktem ubocznym ewolucyjnej genezy mechanizmów leżących u podstaw procesów poznawczych i ich reprezentacyjnych wytworów.

G) W wielu wypadkach, choć nie zawsze, systemy poznawcze potrafią nie tylko rozpoznać fakt błędnego reprezentowania, ale również przyczyny lub mechanizmy odpowiedzialne za wytworzenie BRP. Nie ma jednak koniecznego związku między występowaniem BRP a zdolnością do ich wykrywania. Można błędzić i nie wiedzieć o tym. Można wiedzieć, że się błędzi, ale nie wiedzieć, dlaczego. Można wiedzieć, że się błędzi, i dlaczego się błędzi, ale nie umieć dokonać korekty błędu. Konstatacja błędu reprezentacyjnego bywa niewystarczająca do jego korekty (na przykład w sytuacji, gdy wymaga ona ingerencji w mechanizmy powstawania BRP). Jeśli błąd reprezentacyjny jest wynikiem zaburzenia mechanizmów neuronalnych oraz nośników RP, jego korekta może leżeć poza możliwościami systemu poznawczego. Żaba czy szympanś nie naprawią uszkodzonej siatkówki ani wadliwie działających mechanizmów korowych odpowiadających za wytworzenie w ich systemach wzrokowych błędnych reprezentacji percepcyjnych. Natomiast jeśli rozpoznają ten błąd, będą mogli dokonać korekty swojego zachowania.

Istnieją różne sposoby, za pomocą których organizmy rozpoznają BRP, ich przyczyny i odpowiadające za nie mechanizmy, a także różne sposoby uczenia się na podstawie wykrytych błędów i różne sposoby ich usuwania. Rozpoznanie BRP umożliwia korektę zachowania, podjęcie działań w kierunku wykrycia mechanizmu, który doprowadził do błędu reprezentacyjnego, a następnie korektę (naprawę) mechanizmu wytwarzającego BRP. Mechanizmy wytwarzające RP są tymi samymi mechanizmami, których uszkodzenie lub zaburzenie może doprowadzić do wytworzenia BRP.

2.2. Krystyna Bielecka swoją koncepcję BRP określa mianem „koherencyjnej koncepcji błędu reprezentacyjnego” (KKBR). Przyjrzyjmy się nieco bliżej – w świetle zarysowanej ramy teoretycznej – wybranym twierdzeniom KKBR na temat RP (BRP).

A) Cechą definicyjną RP (autorka używa terminu „reprezentacje umysłowe”, którego zakresu nie precyzuje) jest odwzorowanie środowiska w systemie poznawczym, głównie w stanach jego układu nerwowego:

czać niektórych iluzji. Błąd pojawia się wówczas, gdy cechę perceptu (reprezentacji) przypisujemy jego przedmiotowi w sytuacji, gdy cecha ta mu nie przysługuje. Powyższe zagadnienie jest interesujące samo w sobie i wymaga dalszej dyskusji. Por. Króliczak (1999), Gregory (2009), Hoffman (2012).

Funkcję reprezentowania ujmuję – za Millikan – w kategoriach etiologicznej funkcji właściwej. Pojedyncza reprezentacja ma swój przedmiot (np. muchę) i treść (czyli warunki poprawności), a także można mówić o jej nośniku (stanie układu nerwowego realizującego stan reprezentowania). O poprawności decyduje rzeczywisty stan rzeczy, jednak konsument ma zdolność rozpoznawania adekwatności reprezentacji (nie zakłada się jednak jego nieomyślności) (Bielecka, 2018, s. 208).

W innym zaś miejscu autorka stwierdza:

Koherencyjna koncepcja błędu dla systemu przyjmuje założenia korespondencyjnej teorii prawdy, gdyż o poprawności reprezentacji decydują wyselekcjonowane odwzorowania między reprezentacją a otoczeniem. Element pragmatyczny manifestuje się zaś w tym, że organizmy mogą poznawać odwzorowania w szeroko pojętym działaniu, którego efekt zwykle stanowi dla systemu pewien rodzaj zawodnego wskaźnika tego, czy jego reprezentacja jest poprawna czy też błędna. (...) Reprezentowanie polega (...) na tym, że istnieje reguła odwzorowania między własnościami organizmu a tym, co reprezentowane, i funkcja reprezentowania została wyselekcjonowana właśnie ze względu na cechę reprezentowania (Bielecka, 2018, s. 223, 224).

Zdaniem Bieleckiej, realizacja podstawowych funkcji biologicznych (reprodukcja, adaptacja przetrwanie) wymaga istnienia sprawnego systemu odwzorowań (reprezentacji), które umożliwiają organizmowi poznawanie otoczenia. Można powiedzieć, że właśnie dzięki tym odwzorowaniom mamy do czynienia z systemem poznawczym (system poznawczy jest systemem reprezentacyjnym). Zgodnie z ujęciem teleosemantycznym, którego wariantem jest KKBR, mechanizmy poznawcze zostały wyselekcjonowane w procesie ewolucyjnym, zaś konkretne układy odwzorowań kształtowały się w toku interakcji organizmów ze środowiskiem.

Zakładana przez autorkę koncepcja RP, z gruntu realistyczna, poprzestaje na bardzo ogólnym poziomie analizy. Tym, czego najbardziej brakuje w przedstawionej propozycji, jest szczegółowy opis mechanizmów powstawania odwzorowań środowiska w systemie poznawczym. Realizacja tego zadania wymaga analizy mechanizmów zaangażowanych w powstawanie reprezentacji wzrokowych, słuchowych, pamięciowych, pojęciowych, przekonaniowych i innych. Co prawda, autorka „testuje” dyskutowane w książce koncepcje RP (BRP) na przykładach pojęć, wyobrażeń i map poznawczych (wykazując wyższość KKBR nad innymi koncepcjami w świetle założonych przez siebie dezyderatów), jednak w żadnym wypadku nie wchodzi w szczegóły dotyczące mechanizmów wytwarzających reprezentacje pojęciowe, wyobrażeniowe czy pamięciowe (mapy poznawcze są strukturalnymi reprezentacjami pamięciowymi). Kluczowe pytania, na które chcielibyśmy uzyskać odpowiedzi, są następujące: Jak powstają odwzorowania percepcyjne, pamięciowe, pojęciowe i inne? W jaki sposób, dzięki interakcji ze środowiskiem oraz mechanizmom przetwarzania informacji, plastyczne struktury układu nerwowego wytwarzają różnorodne odwzorowania otoczenia, a także odwzorowania przedmiotów fikcyjnych i abstrakcyjnych⁸?

⁸ Instruktywny przegląd różnych sposobów rozumienia terminu „reprezentacja neuronalna” zawiera Vilarroya (2017). Przez analogię do *psychosemantyki* można mówić o *neurosemantyce* – dyscyplinie zajmującej badaniem, w jaki sposób plastyczne struktury układu nerwowego, ludzi i innych zwierząt, odwzorowują różne typy obiektów. Analizując procesy uczenia się z punktu widzenia aktywności mózgu, Hawkins i Blakeslee (2006,

B) Na gruncie KKBR pojęcie BRP uzyskuje następującą charakterystykę:

Błąd reprezentacyjny jest rodzajem dysfunkcji systemu reprezentacyjnego. Dysfunkcja jest niespełnieniem funkcji właściwej i zachodzi wtedy, gdy nie występują niektóre z szeregu warunków Normalnych danej funkcji właściwej. Organizm (lub jego część) ma dysfunkcję reprezentacyjną, gdy jego zdolności do reprezentowania są zachowane. Błędne reprezentowanie polega na niezgodności treści reprezentacji z reprezentowanym przedmiotem. Błędna reprezentacja pojawia się np. wtedy, gdy reprezentowany przedmiot X nie jest obecny (Bielecka, 2018, s. 208).

Powyższy opis jest na tyle ogólny, że ma zastosowanie do różnych rodzajów RP. Jego główna idea, niezgodność treści RP z przedmiotem, została zaczerpnięta z klasycznej definicji prawdy i podobnie jak ta definicja nie zawiera kryterium odróżniania BRP od reprezentacji poprawnych. Analogia z klasyczną definicją prawdy jest wyraźna: błędne reprezentacje, podobnie jak zdania fałszywe, posiadają treść, ale są niezgodne z przedmiotem. Różnica polega na tym, że klasyczna definicja prawdy dotyczy zdań w sensie logicznym, a więc językowo określonych reprezentacji symbolicznych, tymczasem definicja BRP nadaje temu pojęciu znacznie szerszy zakres.

Skrajny przypadek niezgodności, o jakiej mowa w definicji BRP, zachodzi wówczas, gdy przedmiot reprezentacji nie istnieje, zaś jej konsument, zgodnie z treścią reprezentacji, zakłada jego istnienie. Ma to miejsce na przykład wtedy, gdy konsument reprezentacji ma wygenerowany w korze wzrokowej percept drzewa, zakłada jego istnienie, podejmuje działania będące konsekwencją przyjęcia tego założenia, jednak w rzeczywistości drzewo to nie istnieje (reprezentacja jest omamem wzrokowym, ponieważ nie jest wywołana obecnością drzewa). Podobnie w wypadku reprezentacji pamięciowej: konsument reprezentacji może mieć autobiograficzną reprezentację pamięciową związaną z obecnością w miejscu, w którym nigdy nie był⁹.

Niewątpliwie, halucynacje i fałszywe wspomnienia są jaskrawymi przykładami BRP. Powstaje pytanie: czy źródłem błędu jest wadliwe działanie wchodzących w grę mechanizmów percepcji i pamięci, nieadekwatna reakcja konsumenta na wadliwie wygenerowane reprezentacje, czy jedno i drugie? Chcielibyśmy wiedzieć, jakie dokładnie zaburzenia mechanizmów percepcyjnych (na jakim poziomie funkcjonowania układu wzrokowego) wywarzyły percept drzewa pod nieobecność samego drzewa? Jak wiadomo, czynników

s. 131, 133) piszą: „Możesz myśleć o otaczającym Cię świecie, poruszać się w nim i przewidywać przyszłe zdarzenia dzięki temu, że w korze istnieje model świata. Jedną z najważniejszych idei w tej książce jest to, że w hierarchicznej strukturze kory przechowywany jest model hierarchicznej struktury rzeczywistego świata. Struktura świata jest odzwierciedlana przez strukturę zagnieżdżoną w korze nowej. (...) Budowa kory i sposób, w jaki poznajemy świat, stanowią naturalne odzwierciedlenie hierarchicznych relacji występujących w rzeczywistym świecie”. Różne rodzaje reprezentacji neuronowych (od prostych reprezentacji cech i obiektów po złożone mapy neuronowe) omawia Spitzer (2012, s. 69-96). Autor przyjmuje szerokie rozumienie reprezentacji neuronowej jako odzworowania cechy, obiektu, sytuacji lub reguły działania w aktywności mózgu (pojedynczego neuronu, układu neuronów lub skomplikowanego wzorca dużej populacji zsynchronizowanych neuronów).

⁹ O błędzie można mówić na dwóch poziomach: mechanizmów wytwarzających reprezentacje oraz reakcji konsumenta. Halucynacje i fałszywe wspomnienie mogą być spowodowane zaburzeniami korowych mechanizmów percepcji i pamięci. Natomiast konsument, zwiędziony ich intensywnością, może błędnie uznawać je za zgodne z rzeczywistością.

(zewnątrznych i wewnętrznych) wywołujących halucynacje jest wiele: deprywacja sensoryczna, traumatyczne przeżycia, doświadczenia z pogranicza śmierci, choroby mózgu, substancje halucynogenne i inne. W każdym wypadku są to czynniki zaburzające mechanizmy percepcji na różnych poziomach ich organizacji (nie ma halucynacji bez zaburzenia tychże mechanizmów, mimo że czynniki zaburzające są wielorakie). Mechanistyczne wyjaśnienie BRP polegałoby w tym wypadku na pokazaniu, jakie czynniki i w jaki sposób wpływają na mechanizmy percepcyjne, że prowadzą one do wytworzenia omamów wzrokowych, słuchowych, dotykowych lub innych.

Kluczowe dla teleosemantyki odróżnienie reprezentacji i jej użytkownika (konsumenta) zakłada możliwość dysfunkcji zarówno mechanizmów wytwarzających RP, jak też mechanizmów konsumenckich. Osoby cierpiące na schizofrenię paranoidalną mogą zdawać sobie sprawę, że doświadczają halucynacji (dzięki wiedzy na temat swojej choroby, którą uzyskały od innych), ale nie być w stanie skorygować błędnej reprezentacji wzrokowej, której doświadczają. Jako świadomi konsumenci błędnej reprezentacji mogą nie podejmować rozmowy z osobami, które widzą i słyszą, ale które nie istnieją. Możliwa jest także sytuacja, gdy osoba, która zdaje sobie sprawę z tego, iż doświadcza halucynacji, przeżywa silne emocje oraz wykonuje automatyczne czynności unikowe w reakcji na halucynację. W obu wypadkach inny będzie konsument halucynacji: w pierwszym jest nim świadomy, zdystansowany podmiot zdający sobie sprawę ze swojej choroby, w drugim zaś jego automatycznie uruchamiane podsystemy: emocjonalny i motoryczny¹⁰.

Dalsze problemy wiążą się z możliwością tworzenia reprezentacji zniekształconych, niezgodnych tylko pod pewnymi względami ze swoimi przedmiotami (mogą to być zniekształcenia percepcyjne, pamięciowe, konceptualne i inne). Teoria BRP powinna opisywać mechanizmy oraz przyczyny powstawania reprezentacji całkowicie błędnych, ale także reprezentacji zniekształconych, a więc tylko częściowo (w pewnym stopniu) niezgodnych ze swoimi przedmiotami. Różne rodzaje BRP mogą zakładać różne rodzaje zniekształceń i różne stopnie niezgodności. Błędy konceptualne różnią się pod wieloma względami od błędów percepcyjnych czy pamięciowych. Ogólna teoria BRP powinna zawierać szczegółową klasyfikację BRP w odniesieniu do każdego rodzaju procesów poznawczych¹¹.

¹⁰ Można wyobrazić sobie sytuację, gdy mimo sprawnie działającego producenta, dysfunkcje pojawiają się na poziomie świadomego konsumenta. Nieracjonalny konsument, na skutek wybiórczych zaburzeń świadomości, może błędnie interpretować poprawne reprezentacje, podejmując zgubne dla siebie działania (rzetelne percepcje może uznać za halucynacje). Jak wyżej zauważyłem, pojęcie konsumenta czasami bywa odnoszone do całego systemu poznawczego, czasami zaś do jego wybranych podsystemów. Błędy i dysfunkcje mogą pojawiać się na każdym poziomie sprzężonego układu producent-konsument. Wydaje się, że w układzie poznawczym złożonym z wielu podsystemów każda reprezentacja ma wielu konsumentów. W zasadzie mówienie o całym systemie jako konsumencie ma charakter eliptyczny – jest zgrubną idealizacją. W rzeczywistości w konsumpcję zaangażowane są tylko wyspecjalizowane podsystemy zdolne do zareagowania czy też odczytania danej informacji. Nigdy nie jest to dosłownie rozumiany cały system. Prawdopodobnym powodem mówienia o całym systemie jako konsumencie jest to, że zachowanie kierowane reprezentacją odczytaną przez jakiś podsystem dotyczy najczęściej całego systemu.

¹¹ Nie każda BRP wpływa na funkcjonowanie organizmu w środowisku. Mogą istnieć BRP, z których system poznawczy nie zdaje sobie sprawy, ponieważ nie zakłócają one jego działania w zauważalny dla niego sposób. Bywa, że czyjeś drobne dysfunkcje poznawcze są dostrzegalne dla innych (niezdolność do dyskryminacji niektórych kolorów), ale są niedostępne z wewnętrznej perspektywy systemu poznawczego. Można mieć drobne

C) KKBR nie jest teorią na temat natury, rodzajów i mechanizmów powstawania BRP. Bielecką interesuje głównie to, co nazywa „błędem reprezentacyjnym dla systemu”, czyli taka dysfunkcja reprezentacyjna, którą organizm sam jest zdolny rozpoznać jako błąd. W szczególności autorkę interesuje jeden ze sposobów wykrywania BRP polegający na „reagowaniu na niespójność między reprezentacjami” (Bielecka, 2018, s. 208). Warunkiem wykrycia BRP jest rozpoznanie niespójności między informacjami kodowanymi przez niezgodne ze sobą reprezentacje, co tłumaczy również nazwę zaproponowanej koncepcji.

W tym kontekście Bielecka formułuje kilka kontrowersyjnych, a częściowo także niejasnych twierdzeń: a) rozpoznanie niespójności między reprezentacjami nie wymaga posiadania metareprezentacji, b) nie każdy błąd systemu jest w nim reprezentowany, ale każdy jest w nim rejestrowany (system ma informację o błędzie, której może nie sprawdzać), c) odróżnienie reprezentacji błędnej od poprawnej wymaga dysponowania zdolnością do szacowania poziomu wiarygodności (stopnia rzetelności) poszczególnych reprezentacji. Jak pisze Bielecka (2018, s. 210):

Zdolność do przyjmowania pewnych informacji jako bardziej rzetelnych (...) konstituuje zdolność organizmu do odróżniania reprezentacji błędnych od poprawnych, co oznacza, że warunkuje zdolność do rozpoznawania błędów reprezentacyjnych.

Jeśli chodzi o twierdzenia a-b, autorka nie precyzuje, jakie są minimalne warunki posiadania metareprezentacji ani na czym polega (niereprezentacyjna) rejestracja błędu w systemie. Co dokładnie jest rejestrowane w takim wypadku? W jaki sposób i za pomocą jakich mechanizmów? System może nie tylko nie umieć sprawdzić, czy dana reprezentacja jest błędna; może w ogóle nie podejrzewać, że wytworzył błędną reprezentację. Aby takie podejrzenie powziąć, musi mieć ku temu odpowiednie powody (na przykład niespójne informacje z różnych źródeł, niepowodzenie w działaniu). Czy dysponowanie zdolnością do oceny wartości poznawczej pewnej reprezentacji, zdolnością do porównania kodowanej przez nią informacji z informacją kodowaną przez inną, mniej rzetelną reprezentację, a także zdolnością do korekty błędu reprezentacyjnego, nie świadczy o posiadaniu mechanizmów poznawczych wyższego rzędu umożliwiających tworzenie minimalnych metareprezentacji?

Można zasadnie twierdzić, że system zdolny do ewaluacji, porównywania informacji kodowanych przez poszczególne reprezentacje oraz identyfikacji i korekty błędu reprezentacyjnego powinien być również wyposażony w mechanizmy obliczeniowe (informacyjne) wyższego rzędu umożliwiające tworzenie metareprezentacji. Może to być prosty system wyposażony w mechanizm informacyjnych sprzężeń zwrotnych, dzięki którym potrafi on zdystansować się względem reprezentacji rozpoznanych jako błędne lub obciążone możliwością błędu. Niewykluczone, że żaba jest takim właśnie systemem. Wiele zależy od tego, jakie minimalne warunki nałożymy na posiadanie metareprezentacji. W najprostszym wypadku metareprezentacja byłaby takim stanem układu nerwowego, który koduje informację na temat innych stanów reprezentacyjnych oraz umożliwia podjęcie działań mających na celu

dysfunkcje percepcyjne i nigdy się o tym nie dowiedzieć. O niektórych zaś dysfunkcjach można się dowiedzieć tylko dzięki innym. Istnieją również BRP, których rozpoznanie wymaga zastosowania testów oraz zaawansowanych technologii wykorzystywanych w diagnostyce zaburzeń wzroku, słuchu czy pamięci.

korektę rozpoznanej BRP lub kierowanego nią zachowania. Zdolność do oceny, porównywania i korekty RP, a także uczenia się na podstawie wykrytych błędów, można uznać za warunki minimalne posiadania metareprezentacji¹².

D) Z kolei twierdzenie c) podnosi problem mechanizmów leżących u podstaw zdolności do odróżniania reprezentacji bardziej rzetelnych od mniej rzetelnych. Mówienie o zdolnościach i funkcjach bez znajomości mechanizmów zawsze ma charakter prowizoryczny (roboczy). Ilustrując główne założenie KKBR na przykładzie systemu reprezentacyjnego żaby, Bielecka (2018, s. 212) stwierdza:

Między stanami układu nerwowego żaby a rzeczywistością może zachodzić bardzo wiele różnych odwzorowań, lecz tylko te mają charakter reprezentacyjny, które mogą podlegać weryfikacji, czyli takie, w których sam system poznawczy może rozpoznać błąd.

O ile dobrze rozumiem intencje autorki, skłonna jest ona uznać za reprezentacje tylko takie odwzorowania świata w systemie nerwowym żaby czy innego zwierzęcia, do których ma ono dostęp i może rozpoznać w nich błąd. Powstają pytania: Co to znaczy mieć dostęp do odwzorowania rzeczywistości w układzie nerwowym? Kto może mieć taki dostęp i za pomocą jakich metod może go uzyskać? Jak można rozpoznać błędy takiego odwzorowania? Na czym polega korekta błędu? Czy korekta błędu jest tylko korektą zachowania, czy także korektą dysfunkcyjnego mechanizmu poznawczego, którego wadliwe działanie prowadzi do wytworzenia BRP?

Jak można przypuszczać, zdaniem Bieleckiej obraz siatkówkowy jest odwzorowaniem rzeczywistości, które nie jest reprezentacją, ponieważ żaba nie jest w stanie rozpoznać ani zweryfikować błędu takiego odwzorowania¹³. Uszkodzenie mechanizmów powstawania obrazu siatkówkowego, nawet przy poprawnie działających korowych procesach przetwarzania wzrokowego, może powodować poważne problemy w funkcjonowaniu żaby w jej środowisku naturalnym. Niepowodzenia w działaniu mogą mieć swoje źródło w błędnym odwzorowaniu otoczenia w obrazie siatkówkowym, do którego nie ma ona żadnego dostępu poznawczego¹⁴. Żaba nigdy nie będzie zdolna do rozpoznania ani korekty przyczyn takiego błędu. Pomimo niezdolności do rozpoznania natury błędu, jego przyczyn i mechanizmów, może ona

¹² Autorka odnotowuje ważną pracę (Shea, 2007), w której sformułowano podobny pogląd do zarysowanego powyżej, nie podejmuje jednak tego zagadnienia w sposób systematyczny.

¹³ Powstaje pytanie, czy w układzie nerwowym istnieją jakiegokolwiek odwzorowania rzeczywistości, do których żaden podsystem organizmu nie ma dostępu. Musiałoby to być odwzorowanie bez użytkownika – wytworzone, ale nieskonsumowane. Istnienie takich odwzorowań jest wątpliwe, ponieważ mechanizmy wytwarzające odwzorowania (producenci) nigdy nie są odizolowane od innych mechanizmów. Odwzorowania powstające w systemie nerwowym nie są – i jak się wydaje, nie mogą być – epifenomenami, a więc odwzorowaniami bez konsumenta. Konsumentem odwzorowania może być każdy podsystem organizmu, którego działanie kierowane jest tym odwzorowaniem. W szczególności istnieć może cała sekwencja konsumentów będących przyczynowo-reprezentacyjnymi pośrednikami między odwzorowaniem wyjściowym (obraz siatkówkowy, pamięć sensoryczna) a organizmem jako konsumentem finalnym. Konsumpcję można rozumieć jako wieloetapowy proces angażujący różnorodne podsystemy konsumentów. W literaturze filozoficznej zarówno pojęcie producenta, jak konsumenta reprezentacji, są określane w sposób niezwykle ogólnikowy.

¹⁴ Obrazy siatkówkowe można uznać za reprezentacje strukturalne. Obrazy te są czymś innym niż tzw. reprezentacje receptorowe, jednak im również jestem skłonny przyznać status reprezentacji w odpowiednio rozszerzonym znaczeniu. Interesującą analizę reprezentacji receptorowych przedstawia Gładziejewski (2015, s. 161–172) z konkluzją, która nie jest dla mnie przekonująca. Zagadnienie to wymaga odrębnej dyskusji.

doświadczając jego skutków, a także – dzięki informacjom pochodzącym z innych kanałów zmysłowych – korygować swoje zachowanie w taki sposób, by uzyskać dostęp do pokarmu i przetrwać. Byłoby to jednak rozpoznanie i korekta skutków błędu (dysfunkcyjnego zachowania), a nie samego błędu i leżącego u jego podstaw wadliwego mechanizmu wzrokowego¹⁵.

E) Jednym z głównych założeń KKBR jest twierdzenie, że mechanizm wykrywania BRP obejmuje takie zdolności poznawcze, jak: porównywanie informacji napływających z różnych zmysłów, wykrywanie niespójności między nimi oraz ocena stopnia ich rzetelności. Zdolności tego typu występują już u płazów i gryzoni. Postulowany mechanizm, w barwnym opisie autorki, działa w następujący sposób:

Do żaby dociera wzorzec wzrokowy (ruchomy i podłużny), który wywołuje reprezentację muchy. W rezultacie system motoryczny (konsument tej reprezentacji) powoduje wyrzucenie języka w stronę tego przedmiotu. Bódcze dotykowe z języka wywołują jednak zamiast drugiej reprezentacji muchy, reprezentację trzmieła. Prowadzi to do wykrycia niespójności, a więc błędnej reprezentacji. Gdyby żaba nie reagowała na bodźce z języka mocniej niż na wzrokowe, nie mogłaby od razu odrzucić reprezentacji muchy, a jedynie wykryć, że któraś z dwóch reprezentacji jest błędna. Załóżmy na potrzeby analizy, że rzeczywiście informacje dotykowe są bardziej rzetelne od wzrokowych, przez co błędna jest reprezentacja muchy. Zdolność do reagowania na niespójność (treści) reprezentacji (i informacji) jest warunkiem koniecznym zdolności do rozpoznawania błędu reprezentacyjnego (Bielecka, 2018, s. 213).

Powyższa interpretacja zachowania żaby nasuwa wątpliwość, czy autorka nie przypisuje płazom zbyt dużych kompetencji logiczno-analitycznych. Kategorie wykorzystywane przez nią w opisie zachowania żaby (niespójność, rzetelność informacji, reagowanie na niespójność reprezentacji) z pewnością są adekwatne w odniesieniu do ludzi. Umiejętność wykrywania niespójnych reprezentacji oraz zdolność do szacowania stopnia wiarygodności informacji pochodzących z różnych źródeł stanowią ważne kompetencje, dzięki którym ludzie wykrywają BRP (percepcyjne, pamięciowe, przekonaniowe i inne). Czy stosowanie tych kategorii do opisu zachowania płazów nie jest formą antropomorfizującej projekcji? Czy żaby rzeczywiście mają zdolności, które im przypisuje powyższy opis? Być może, zachowania eksploracyjne żab i innych płazów można opisać i wyjaśnić w języku teorii, która nie przypisuje im tak zaawansowanych kompetencji logicznych. Na przykład, zamiast przypisywać żabie zdolność do wykrywania niespójności między reprezentacjami, wystarczy przyjąć, że zanim dojdzie do konsumpcji owada, wykorzystywane są dwa systemy sprawdzające, wzrokowy i dotykowy. Gdy zawodzi pierwszy, pozostaje jeszcze drugi, który – jeśli działa sprawnie – spowoduje wyplucie owada. W obu wypadkach niezbędne są sprawne systemy pamięciowe: mózg żaby musi porównać informacje receptorowe z odpowiednimi wzorcami przechowywanymi w pamięci wzrokowej i dotykowej. Opisaną sytuację można wyjaśnić z wykorzystaniem znacznie prostszego aparatu pojęciowego.

¹⁵ Naukowcy w laboratorium, po szczegółowej inspekcji żaby z zaburzonym obrazem siatkówkowym, mogliby, jak się wydaje, dokładnie opisać naturę i mechanizm powstania błędu odwzorowania na siatkówce, a także jego konsekwencje (związane z propagacją błędu) dla funkcjonowania żaby w środowisku naturalnym.

Wątpliwości powstają zwłaszcza w odniesieniu do twierdzenia: *gdyby żaba nie reagowała na bodźce z języka mocniej niż na wzrokowe, nie mogłaby od razu odrzucić reprezentacji muchy, a jedynie wykrzyć, że któraś z dwóch reprezentacji jest błędna*. Wiemy, na mocy logicznego prawa niesprzeczności, że jedno z dwóch zdań sprzecznych jest fałszywe (nawet jeśli nie wiemy, które), ale czy taką kompetencję logiczną możemy przypisać żabie w sytuacji żabiego dysonansu poznawczego (posiadania dwóch niespójnych reprezentacji: wzrokowej i dotykowej)? Aby uwiarygodnić swoją interpretację, autorka musiałaby wykazać, że prawa logiki są implementowane w strukturach układu nerwowego żaby (wyznaczają zasady zachodzących w nim procesów przetwarzania informacji oraz określają jej zachowania eksploracyjne). Niewykluczone, że tak właśnie jest. Jednak i wówczas czym innym będzie aktywność kierowana zaimplementowanymi w układzie nerwowym prawami logiki, a czym innym dysponowanie kompetencją logiczną (na przykład wiedzą, że z dwóch sprzecznych reprezentacji jedna jest fałszywa).

Nieco inne wątpliwości budzi trafność podanej przez autorkę interpretacji zachowania szczura przeniesionego do innego labiryntu:

Szczury mają mechanizm reprezentacyjny zdolny do reagowania na niespójność (...) reprezentacji. Niespójność pojawia się wtedy, gdy gryzoń pamiętający położenie pokarmu ze starego labiryntu, reaguje zdziwieniem, gdy nie ma pokarmu w tym samym miejscu, w którym był dotychczas. W tej sytuacji dochodzi do pewnych pobudzeń informujących mózg szczura o nieobecności pokarmu (I). Następnie mechanizm reprezentacyjny szczura rejestruje niespójność poprzedniej mapy poznawczej z informacjami o obecnym położeniu, wobec czego do mózgu szczura przesyłana jest dopamina, a manifestacją rozpoznania niespójności jest reakcja behawioralna zdziwienia (II). W wyniku zbliżenia do przedmiotu reprezentowanego okazuje się (III), że w labiryncie nie ma jedzenia. Szczur rozpoznaje swój błąd tylko wtedy, gdy jest zdolny do reagowania na niespójność, czyli gdy nauczy się, że w tym miejscu labiryntu nie ma jedzenia i nie podejmie tu następnym razem (Bielecka, 2018, s. 221).

Czy w sytuacji przeniesienia szczura do innego labiryntu w ogóle można mówić o BRP? W mózgu szczura powstała mapa poznawcza (przestrzenny obraz pamięciowy), która poprawnie odwzorowuje strukturę labiryntu, w którym doszło do jej wytworzenia. Mapa poznawcza poprawnie kieruje jego zachowaniem. Kolejne pozytywne doświadczenia wzmocniły oczekiwanie, że w określonym miejscu znajduje się pożywienie (antycypacyjna funkcja mapy poznawczej). Czy w sytuacji, gdy szczur zostanie przeniesiony do innego labiryntu, jego mapa poznawcza stanie się BRP? Wydaje się, że nie, ponieważ nie jest to mapa nowego labiryntu. Błąd, z jakim w tym wypadku mamy do czynienia, polega na aplikacji starej mapy poznawczej do nowej sytuacji. Jest to błąd związany z próbą wykorzystania uprzednio wytworzonego obrazu pamięciowego w innym miejscu. W sensie ścisłym nie mamy tu jednak BRP, ponieważ stara mapa poznawcza w dalszym ciągu poprawnie odwzorowuje strukturę starego labiryntu. Szczur w nowej sytuacji, wykorzystując nowe doświadczenia, szybko dokona korekty starej mapy poznawczej, dostosowując jej strukturę do nowego środowiska. Kolejne negatywne doświadczenia osłabiają (wygaszają) oczekiwanie znalezienia pokarmu w miejscu, w którym go obecnie nie ma. Jest to przykład uczenia się polegającego na przebudowie uprzednio wytworzonej mapy poznawczej i związanych z nią oczekiwań.

Czy mamy tutaj do czynienia z reakcją na niespójność? Czy wykorzystanie w tym wypadku logicznej kategorii niespójności jest usprawiedliwione? Niewątpliwie, oczekiwanie związane z wykorzystaniem starej mapy poznawczej w nowej sytuacji nie zostaje spełnione, w związku z czym rozpoczyna się proces przebudowy obrazu pamięciowego i dostosowania go do nowej sytuacji (proces mający swoją neuronową reprezentację w mózgu szczura). Nie mamy tu do czynienia z BRP, nie musimy też wykorzystywać zaawansowanych kategorii logicznych do opisu zachowania szczura w labiryncie. Z pewnością ludzie wykorzystują logiczną kategorię niespójności w wykrywaniu błędów poznawczych, jednak w odniesieniu do ludzi jest to twierdzenie trywialnie prawdziwe. Natomiast zastosowanie tej kategorii do opisu i wyjaśnienia zachowania żaby czy szczura jest co najmniej kontrowersyjne i wymaga przyjęcia założenia, że układy nerwowe płazów i gryzoni implementują podstawowe prawa logiki, które są automatycznie wykorzystywane do kierowania zachowaniem i jego korekty, a także w procesie uczenia się. Niewykluczone, że tak właśnie jest. Należałoby jednak dokładniej zbadać mechanizmy przetwarzania informacji w mózgach płazów i gryzoni, gdyż bez ich szczegółowej znajomości podana interpretacja jest gołosłowna.

W zamierzeniu autorki KKBR nie jest teorią na temat natury, rodzajów i mechanizmów powstawania BRP. Jest raczej szkicem teorii wykrywania błędów reprezentacyjnych akcentującym rolę niespójności jako jednego z kryteriów BRP (z budzącą wątpliwośći metodologiczne aplikacją do zachowań płazów i gryzoni). Co prawda, Bielecka odwołuje się w swojej pracy do wybranych badań empirycznych, jednak czyni w sposób stosunkowo powierzchowny i wybiórczy. Badania empiryczne służą autorce głównie jako narzędzie testowania jej teorii, przy czym sposób, w jaki interpretuje przypadki empiryczne, sprawia wrażenie skrojonego na potrzeby teorii. Opisy zachowania płazów i gryzoni mają ilustrować trafność zakładanej teorii, możliwe są jednak inne opisy i interpretacje (brakuje odniesienia do prac etologów poznawczych).

Bielecka jest przekonana (2018, s. 233), że KKBR może stanowić podstawę szeroko zakrojonych badań empirycznych, których wizję roztacza w przedostatnim akapicie *Zakończenia* swojej książki, skądinąd bardzo interesującym. Moim zdaniem, ta fundacjonistyczna deklaracja, o ile dobrze rozumiem twierdzenie o stanowieniu podstaw, nie znajdzie potwierdzenia w przyszłych badaniach BRP. Punktem wyjścia do budowy ogólnej teorii BRP powinna być wielopoziomowa analiza wyników badań empirycznych dotyczących mechanizmów powstawania i wykrywania BRP. Dostarczają one bogatego zbioru danych na temat natury i rodzajów błędów pojawiających się w procesach poznawczych. Uporządkowanie i analiza wyników tych badań jest dobrym punktem wyjścia dla filozoficznej teorii BRP – teorii zachowującej spójność teoretyczno-metodologiczną z wiedzą naukową.

3. Mechanizmy powstawania błędnych reprezentacji pamięciowych – kilka uwag i komentarzy na marginesie KKBR

Do tej pory najdokładniej opisano mechanizmy powstawania błędnych reprezentacji percepcyjnych (wzrokowych, słuchowych) oraz pamięciowych (głównie w odniesieniu do pamięci epizodycznej i autobiograficznej) z uwagi na ich fundamentalną rolę w działaniu systemów poznawczych. Błędy percepcyjne i pamięciowe mogą powodować olbrzymie koszty, łącznie z utratą życia. Systemy z zaburzonymi mechanizmami percepcji i pamięci, które nie

wytwarzają reprezentacji lub wytwarzają reprezentacje głęboko zniekształcone, giną lub wymagają opieki ze strony innych. W dalszym ciągu artykułu, odwołując się do kilku przykładów badań nad mechanizmami powstawania błędnych reprezentacji pamięciowych (BRPA), sformułuję założenia interdyscyplinarnego programu badawczego, którego celem jest budowa ogólnej teorii BRP.

A) Badania nad mechanizmami powstawania błędnych (fałszywych, zniekształconych) wspomnień dostarczają bogatego materiału empirycznego, który może stanowić punkt wyjścia do budowy ogólnej teorii BRP. Przez BRPA rozumiem wytwory procesów pamięciowych, stanowiące odwzorowanie przeszłych zdarzeń w systemie poznawczym, które w jakimś stopniu i pod jakimś względem są niezgodne ze swoimi przedmiotami¹⁶. Biorąc pod uwagę istnienie wielu rodzajów pamięci, można mówić o różnych rodzajach reprezentacji pamięciowych oraz błędów, które ich dotyczą. Przy czym nie każdy błąd pamięci zaliczymy do klasy BRPA (nie jest oczywiste, czy błędy pamięci proceduralnej mają charakter reprezentacyjny)¹⁷. W dalszym ciągu skupię się na niekontrowersyjnych przykładach BRPA związanych z reprezentacyjnymi formami pamięci, takimi jak reprezentacje autobiograficzne i epizodyczne¹⁸.

Zgodnie ze współczesnymi ujęciami procesy pamięciowe mają charakter dynamiczny, wielofazowy i rekonstrukcyjny. Przypomnienia nie są dosłownie rozumianymi replikami, reprodukcjami czy kopiami zdarzeń doświadczanych w przeszłości, lecz konstruktami utworzonymi z elementów (informacji) kodowanych i przechowywanych w różnych obszarach mózgu (Bahrick, 1996)¹⁹. Do treści wspomnień mogą być włączane również treści zniekształcające o charakterze wyobrażeniowym, związane z wiedzą i przekonaniem jednostki lub zapożyczone od innych osób. Błędy pojawiają się na etapie kodowania, przechowywania i odtwarzania informacji. O reprezentacji pamięciowej najczęściej mówimy, mając na uwadze finalny produkt tych procesów, jednak zaburzenie każdego z nich może prowadzić do wytworzenia BRPA²⁰.

¹⁶ Termin „fałszywe wspomnienia” rezerwuję dla wspomnień mających treść propozycjonalną, które dotyczą przeszłych zdarzeń lub stanów rzeczy oraz dają się wyrazić w formie fałszywych zdań w sensie logicznym. Natomiast terminem „błędy pamięci” proponuję określać wszelkiego typu niezgodności zachodzące między treścią reprezentacji pamięciowych a ich przedmiotami. W literaturze z zakresu psychologii poznawczej często zamiennie używa się terminów: fałszywy, błędny, zniekształcony.

¹⁷ Zdaniem Spitzera (2012, s. 76–79) pamięć proceduralna również ma charakter reprezentacyjny. Umiejętności są reprezentowane w mózgu przez neurony kodujące abstrakcyjne reguły, które określają sposób wykonania danej czynności lub ich sekwencję. Błąd reprezentacyjny spowodowany uszkodzeniem obwodu kodującego takie reguły przejawia się w postaci zaburzeń wykonania wyuczonych czynności. Do wykrycia błędnej reprezentacji nie jest konieczne w tym wypadku rozpoznanie niespójności między reprezentacjami ani ich porównywanie pod względem stopnia rzetelności.

¹⁸ Na temat relacji między tymi dwoma rodzajami reprezentacji pamięciowych por. Niedźwieńska (2004, s. 50–58).

¹⁹ Nie wyklucza to, oczywiście, reprezentacji pamięciowych o fotograficznej dokładności. Jednak również w wypadku niezwykle precyzyjnych reprezentacji pamięciowych, mamy do czynienia z procesem konstruowania obrazu pamięciowego z elementów przetwarzanych i przechowywanych w różnych obszarach mózgu.

²⁰ Różne sposoby rozumienia konstrukcyjnego i rekonstrukcyjnego charakteru pamięci omawia Niedźwieńska (2004, s. 60–61). W pewnym sensie każda reprezentacja jest rekonstrukcją tego fragmentu rzeczywistości, którego dotyczy – rekonstrukcją wytworzoną w systemie poznawczym za pomocą mechanizmów poznawczych, w które jest on wyposażony.

Proces kodowania informacji zawsze wiąże się z jej redukcją oraz podlega wpływowi różnorodnych czynników zewnętrznych, które mogą odpowiadać za specyficzne błędy kodowania. Neuronalna reprezentacja bodźców niosących informację o zdarzeniach ma charakter uproszczony i schematyczny w stosunku do zapamiętywanego zdarzenia. Ślady pamięciowe stopniowo ulegają osłabieniu (zatarciu), przez co treści reprezentacji pamięciowych mogą ulegać zniekształceniu w stosunku reprezentowanego obiektu lub zdarzenia. Badania eksperymentalne pokazują, że przypominanie sobie jest procesem dynamicznym, zależnym od a) mechanizmów neuronalnych (aktywacja obwodów neuronowych, transfer informacji z pamięci długoterminowej do roboczej), b) psychologicznych (emocjonalnych, poznawczych), a także od c) czynników związanych z oddziaływaniem społecznym (kontekst, sugestie, wzorce społeczno-kulturowe). O BRPA można mówić w odniesieniu do podmiotów indywidualnych i kolektywnych – pamięć zbiorowa także podlega zniekształceniom i manipulacjom (Conway, 1997).

B) Zniekształcenia treści wspomnień mogą mieć charakter naturalny (związany z niezaburzonym działaniem systemu pamięci) lub patologiczny (uszkodzenia mózgu, choroby neurodegeneracyjne). W nowszych opracowaniach zakłada się, że naturalne BRPA są skutkiem ubocznym organizacji mechanizmów pamięci i jako takie najczęściej nie pociągają dramatycznych skutków dla organizmu. Adaptacyjny charakter systemów pamięci nie polega na tym, by zapewnić organizmowi pamięć absolutną, doskonale odwzorowującą minione zdarzenia. Biologiczną funkcją pamięci jest zapewnienie przetrwania i odniesienie sukcesu reprodukcyjnego, do tego zaś wystarcza niedoskonały system pamięciowy; system bezbłędny lub dokładny ponad potrzeby organizmu stałby się obciążeniem. Jak zauważa Niedźwieńska (2004, s. 15), „(...) zniekształcenia są kosztem szybkich i ekonomicznych strategii poznawczych, które wprawdzie nie zapewniają absolutnej poprawności, ale umożliwiają szybką orientację, a w konsekwencji bezzwłoczne podejmowanie decyzji i działanie”²¹.

BRPA są skutkiem ubocznym nie tylko naturalnych ograniczeń mechanizmów pamięci, ale również procesów neurodegeneracyjnych związanych ze starzeniem się mózgu i chorobami centralnego układu nerwowego. W przypadku postępującej demencji czy choroby Alzheimera mamy do czynienia z objawami ubytkowymi (wspomnienia niepełne i głęboko zniekształcone, problemy z rozpoznawaniem, narastające trudności z zapamiętywaniem nowych faktów, nakładanie się wspomnień, błędna reprezentacja chronologii zdarzeń), a także z objawami nadmiarowymi, które wiążą się z aktywnością twórczą konsumenta reprezentacji pamięciowych. Klasycznym przykładem efektu nadmiarowego są konfabulacje polegające na dowolnym wypełnianiu luk w pamięci bądź tworzeniu całkowicie wymyślonych narracji na temat przeszłości. Konfabulacje mogą wytworzyć w umyśle chorego rozbudowany system

²¹ W interesującej analizie mechanizmów powstawania zniekształceń pamięciowych Niedźwieńska skupia się na mechanizmach psychologiczno-poznawczych, które uważa za podstawowe. Wiążą się one z istotnymi ograniczeniami (ilościowymi i strukturalnymi) pamięciowego przetwarzania informacji oraz zawodnymi strategiami poznawczymi związanymi z wykorzystywaniem zasobów pamięci. Zniekształcenia o charakterze motywacyjnym (podnoszenie samooceny, wypieranie informacji niezgodnych z żywionymi przekonaniem) oraz społecznym (wpływ grupy na działanie pamięci) Niedźwieńska (2004, s. 16) uznaje za pochodne wobec mechanizmów poznawczych.

fikcyjnych reprezentacji przeszłych zdarzeń obejmujący błędne reprezentacje autobiograficzne. W zniekształconym systemie pamięciowym reprezentacje mające odpowiedniki w realnych zdarzeniach z przeszłości są łączone z treściami fikcyjnymi.

William Hirstein (2005, s. 50–65) zwraca uwagę, że jeśli uszkodzeniom w obszarach skroniowych (wywołujących różnorodne deficyty pamięci) towarzyszą uszkodzenia w obszarach płatów czołowych, chory może nie zdawać sobie sprawy z tego, że jego rzekome wspomnienia są skutkiem działania wyobraźni. Pojawi się wówczas efekt anozognozji: chory nie będzie w stanie rozpoznać błędnej reprezentacji pamięciowej, nie będzie też zdawał sobie sprawy z tego, że konfabuluje (Herzyk, 2002)²². Z punktu widzenia badań nad BRPA przypadek anozognozji jest interesujący, ponieważ wskazuje na istnienie mechanizmów odpowiedzialnych za rozpoznawanie błędnych reprezentacji. Ich uszkodzenie ogranicza lub pozbawia człowieka zdolności do wykrywania określonej klasy błędów reprezentacyjnych, a także zdolności do wykrywania niespójności między różnymi rodzajami reprezentacji (chorzy nie mają poczucia dysonansu między niespójnymi reprezentacjami, brakuje im również zdolności do oceny stopnia wiarygodności niezgodnych danych pochodzących z różnych źródeł)²³. Wszystkie zdolności poznawcze (reprezentacyjne) i metapoznawcze (metareprezentacyjne) opierają się na działaniu określonych mechanizmów, które jeśli ulegają zaburzeniom, prowadzą do charakterystycznych deficytów i dysfunkcji (zależność ta ma charakter systematyczny).

C) Psychologiczne (poznawcze) modele BRPA najczęściej dotyczą funkcjonowania poprawnie działających systemów pamięciowych. Przeprowadzono liczne badania eksperymentalne wykrywające mechanizmy, za pomocą których możliwe jest wytworzenie u ludzi zdrowych wspomnień ze zniekształconą treścią (błędnie odwzorowujących minione zdarzenia). Daniel Schacter w swoich klasycznych badaniach czynnościowych zaburzeń pamięci szczegółowo omawia a) błędy omijania (nietrwałość, roztargnienie i blokowanie) oraz b) błędy przekształcania (błędna atrybucja, podatność na sugestię, tendencyjność i uporczywość)²⁴.

²² Konfabulacje na temat przeszłości, jeśli mają charakter propozycjonalny, są przykładem literalnie rozumianej fałszywej pamięci. Występują one w wielu sytuacjach codziennych, ale towarzyszą również chorobom neurodegeneracyjnym i psychicznym (Prigatano, 2010). U chorych na schizofrenię w wypadku objawów urojeniowych stwierdza się występowanie uporczywie podtrzymywanych przekonań na temat przeszłości, czemu towarzyszą żywe wspomnienia zdarzeń, które nigdy nie miały miejsca lub zostały przekształcone w wyobraźni chorego. Stanowią one szczególny przypadek BRPA. Por. Kępiński (2001, s. 37–38). Jeszcze inna odmiana BRPA występuje w wypadku przeżyć typu *déjà vu*, które także mają postać normalną (występującą u ludzi zdrowych) oraz kliniczną. Mamy tu do czynienia z błędną oceną aktualnej sytuacji jako rzekomo doświadczanej w przeszłości. Oprócz subiektywnej oceny, elementem składowym doświadczenia typu *déjà vu* jest błędna reprezentacja pamięciowa wzbudzona przez aktualne przeżycie o charakterze percepcyjnym. Istnieją różne teorie wyjaśniające powstawanie tej iluzji, która błędnie łączy reprezentacje percepcyjne i pamięciowe. Por. Spalt (2002).

²³ Na temat różnych form anozognozji (m.in. w demencji i zespole Korsakowa) por. Herzyk (1999), Prigatano (2010).

²⁴ Zdaniem Schactera (2003, s. 283–300) wymienione rodzaje błędów nie są efektem wadliwego działania mechanizmów pamięci. Są to raczej efekty uboczne przystosowawczych funkcji pamięci, które autor analizuje w kategoriach egzaptacji i spandrel (w nawiązaniu do koncepcji Stephena J. Goulda i Richarda Lewontina). Badacze pamięci na ogół są zgodni, że doskonały, niezawodny system pamięciowy, który nie popełnia błędów, lecz zawsze odwzorowuje przeszłość w sposób maksymalnie dokładny i poprawny, byłby mniej efektywny i przystosowawczy od systemu niedoskonałego, w który jesteśmy wyposażeni.

Spośród błędów omijania szczególne znaczenie z punktu widzenia mechanizmów BRPA odgrywa nietrwałość, a zwłaszcza jej szczególna odmiana nazywana „interferencją”. Chodzi o sytuacje, gdy nowe zapamiętywane treści nie tylko utrudniają odtworzenie treści starych, ale również powodują ich zniekształcenia (tak samo stare treści mogą utrudniać zapamiętanie treści nowych oraz powodować, że będą one zapamiętywane w sposób zniekształcony). Jednak najważniejsze mechanizmy BRPA należą do drugiej grupy. Z błędną atrybucją mamy do czynienia wówczas, gdy przypominane zdarzenie jest błędnie łączone z niewłaściwą osobą, miejscem lub czasem. Zniekształcenia treści wspomnień, a nawet wspomnienia całkowicie błędne mogą również powstawać pod wpływem silnej sugestii, na przykład sugestywnych pytań. Stwierdzono, że dzieci mogą błędnie zapamiętywać zdarzenia, w których uczestniczyły, pod wpływem sugestii ze strony rodziców lub innych osób, które podały im zmodyfikowany opis zdarzenia (zapamiętywane i przypominane są zniekształcone opisy zdarzeń, a nie zdarzenia oryginalne – zniekształcone przypomnienie może zawierać treści pochodzące z różnych źródeł). Z kolei tendencyjność, jako czynnik związany z powstawaniem BRPA, polega na wpływie wiedzy, przekonań, schematów pojęciowych i uprzedzeń na treści wspomnień. Stwierdzono, że w wielu wypadkach może to być wpływ zniekształcający treści zapamiętywane i odtwarzane (w grę wchodzi tutaj różne mechanizmy szczegółowe, takie jak wypieranie treści, preferencja treści zgodnych z przekonaniami i inne)²⁵.

Odkrycia dotyczące mechanizmów powstawania BRPA są ważne także z punktu widzenia prawa. Możliwość wdrukowywania wspomnień o zniekształconej treści ma znaczenie dla oceny zeznań naocznych świadków zdarzeń podlegających kwalifikacji prawnej (wypadków, morderstw, gwałtów i innych). Dzięki pionierskim badaniom Elizabeth Loftus (1996, 1997, 2000; Loftus i Pickrell, 1995) okazało się, że sposób postawienia pytania naocznym świadkom po zajściu zdarzenia ma istotny wpływ na treści ich przypomnień i składane zeznania. Podobnie, psychoterapeuta może w sposób niezamierzony wywoływać u swoich pacjentów błędne wspomnienia związane z gwałtem lub molestowaniem, co potwierdziły badania eksperymentalne (Ulatowska i Sawicka, 2017; Oronowicz-Jaśkowiak, 2017). Pokazuje to, że procesy pamięciowe nie polegają na kopiowaniu, reprodukowaniu minionych zdarzeń, lecz stanowią ich rekonstrukcję, czasami daleko odbiegającą od pierwowzoru: treści wspomnień mogą ulegać różnorodnym zniekształceniom pod wpływem czynników psychologicznych i społecznych²⁶.

D) Ważną klasę mechanizmów BRPA stanowią mechanizmy neurobiologiczne związane z kodowaniem, przechowywaniem, scalaniem i odtwarzaniem informacji w mózgu. Szczegółne znaczenie ma w tym wypadku fakt, iż różne fragmenty naszych wspomnień są

²⁵ Nie jest moim celem szczegółowa analiza sposobów, na jakie mogą powstawać BRPA, a jedynie ilustracja tezy, że kluczowy problem ogólnej teorii BRP dotyczy mechanizmów ich powstawania. Zestawienie i przegląd różnych klasyfikacji błędów pamięciowych i leżących u ich podstaw mechanizmów (z perspektywy psychologii poznawczej) zawiera praca Niedźwieńskiej (2004, s. 146–157).

²⁶ Interesującym przypadkiem BRPA są tzw. wspomnienia niemożliwe, których doświadczają ludzie rzekomo pamiętający pierwsze dni swojego życia, moment narodzin, epizody związane z przebywaniem w łonie matki, a nawet zdarzenia z poprzednich wcieleń. Określenie „niemożliwe” wskazuje na fakt, że we wczesnym okresie rozwoju mózgu nie są jeszcze ukształtowane struktury umożliwiające długoterminowe zapamiętywanie zdarzeń. Por. Hyman i Pentland (1996), Strange, Sutherland i Garry (2006).

przechowywane w sposób rozproszony w różnych obszarach sieci neuronowej wyspecjalizowanych do przetwarzania określonego typu informacji. Aktywacja nawet prostych reprezentacji pamięciowych o charakterze epizodycznym wymaga zsynchronizowanej aktywności wielu obwodów neuronowych kodujących informacje o różnych aspektach przypominanego zdarzenia. Dzięki scalaniu rozproszonych informacji możliwe jest powstanie jednolitego wspomnienia²⁷. Jeśli jednak jakiś element tego wspomnienia nie zostanie aktywowany (zatarcie śladu pamięciowego, utrata informacji na skutek uszkodzenia obwodu neuronowego, blokada spowodowana działaniem substancji chemicznych) lub też w proces scalania zostaną włączone informacje pochodzące z innych doświadczeń, powstanie wspomnienie o zniekształconej, a nawet całkowicie błędnej treści. Ponieważ są to procesy przebiegające podprogowo, świadomy podmiot nie będzie zdawał sobie sprawy z tego, że doszło do wytworzenia BRPA (będzie miał błędne wspomnienie przy zachowanym subiektywnym poczuciu jego poprawności)²⁸.

Wiedzę na temat tego, jak zaburzenia mechanizmów neuronowych odpowiadających za różne etapy pamięciowego przetwarzania informacji prowadzą do powstawania BRPA, uzyskujemy między innymi dzięki badaniom eksperymentalnym prowadzonym nad systemami pamięciowymi zwierząt. Eksperymenty Kathryn Hunt i Larsa Chittka nad mechanizmami scalania pamięciowego u trzmieła ziemnego (*Bombus terrestris*) wykazały, że na skutek połączenia informacji pochodzących z doświadczeń dwóch różnych rodzajów kwiatów doszło do wytworzenia błędnego wspomnienia będącego mieszanką tych informacji, co ujawniło się w zachowaniu trzmieli. Zdaniem badaczy, scalanie w mózgu informacji pochodzących z różnych doświadczeń jest jednym z neuronowych mechanizmów powstawania BRPA występującym u przedstawicieli wielu gatunków, w tym ludzi (Hunt i Chittka, 2015)²⁹.

Inne badania pokazują, że możliwe jest oddziaływanie na treści wspomnień, a tym samym wywoływanie BRPA, na poziomie reakcji biochemicznych. Podanie szczyrom animozycyny (substancji hamującej syntezę białek) bezpośrednio do ciała migdałowatego blokuje proces powstawania reakcji lękowej w odpowiedzi na bodziec dźwiękowy, któremu wcześniej towarzyszyło podanie elektrowstrząsów. Wykorzystanie stosunkowo prostego mechanizmu biochemicznego umożliwia zakłócenie uczenia się afektywnego, wytworzenie błędnej asocjacji pamięciowej oraz zatarcie wcześniejszych wspomnień (Nader, Schafe i LeDoux, 2000).

²⁷ Neurobiolodzy zwykle posługują się terminem „reprezentacja” w sposób bardzo ogólny. Charakterystyczne jest następujące stwierdzenie (Vetulani, 2011, s. 196): „Engram jest podłożem pamięci i może być definiowany jako trwała zmiana w układzie nerwowym wywołana chwilowym pobudzeniem. Ta zmiana jest reprezentacją pewnego doznania, które stanowiło przedmiot postrzegania, przeżywania lub działania. Pamięć jest procesem tworzenia, przechowywania i przywoływania engramów”. Nietrudno zauważyć, że w wypadku różnych rodzajów pamięci mamy do czynienia z różną trwałością śladów pamięciowych. Mielibyśmy więc reprezentacje ultrakrótkie, krótkoterminowe i długoterminowe, w zależności od czasu przechowywania śladu pamięciowego.

²⁸ Nie jest jasne, na jakim poziomie działania mechanizmów pamięciowych (to samo dotyczy percepcji) dochodzi do wytworzenia reprezentacji pamięciowej. Czy o reprezentacjach można mówić już na poziomie pamięci sensorycznej? Czy rozporozszonym informacjom przechowywanym w różnych obszarach mózgu należy przypisać status reprezentacji cech obiektów? Czy reprezentacjami pamięciowymi są dopiero produkty scalania informacji pojawiające się w pamięci roboczej? Chociaż na każdym poziomie mamy do czynienia z jakąś formą odzwzorowania, najczęściej mówiąc o reprezentacji pamięciowej mamy na uwadze ostateczny produkt pamięciowego przetwarzania informacji, z którego czynimy użytek w działaniu. Zagadnienie to wymaga odrębnej dyskusji.

²⁹ Popularne omówienie wyników eksperymentów Hunt i Chittki zawiera Shaw (2018, s. 75–78).

Z kolei badania nad benzodiazepinami, substancjami spowalniającymi działanie centralnego układu nerwowego, pokazują, że dzięki nim możliwe jest czasowe blokowanie zdolności do tworzenia nowych wspomnień (wywoływanie amnezji następczej). Podanie odpowiedniej dawki tej substancji przed operacją spowoduje, że nie będziemy pamiętać poprzedzających ją wydarzeń lub wspomnienia będą fragmentaryczne, zniekształcone. Wiemy obecnie z dużą dokładnością, które substancje chemiczne i w jaki sposób wpływają na neuronalne mechanizmy pamięci prowadząc do wytworzenia BRPA (Beracochea, 2006).

Niedawno przeprowadzone eksperymenty na komórkach miejsca u myszy (odkrytych przez Johna O'Keefe'a i współpracowników) pozwoliły na wytworzenie u nich BRPA za pomocą bezpośredniej ingerencji w ich mózgi. Jednoczesna stymulacja za pomocą elektrod komórek miejsca (związaną z określoną lokalizacją) oraz ośrodków przyjemności doprowadziła do wytworzenia połączenia między nimi, co spowodowało, że myszy chętniej przebywały w tym miejscu, chociaż nic pozytywnego wcześniej w nim się nie wydarzyło. Poczucie przyjemności zostało powiązane z określonym miejscem (pamięć skojarzeniowa) nie w trybie uczenia się przez doświadczenie, ale dzięki bezpośredniej ingerencji w mózg (Ramirez i in., 2013). Podobne badania prowadzi się obecnie z wykorzystaniem optogenetyki. Sztuczne połączenia w mózgu myszy są wytwarzane za pomocą lasera aktywującego światłoczułe, genetycznie zmodyfikowane neurony. U myszy, które nauczyły się reagować lękiem na miejsca, gdzie doświadczały bólu, można aktywować w ten sposób reakcję lękową na miejsca, gdzie myszy nigdy wcześniej nie przebywały, co interpretuje się jako wytworzenie błędnej asocjacji pamięciowej będącej odmianą BRPA (Robins, 2016)³⁰.

Współczesne badania dotyczące BRPA mają charakter interdyscyplinarny, ponieważ mechanizmy odpowiedzialne za wytwarzanie BRPA obejmują różne poziomy funkcjonowania systemu poznawczego badane na gruncie różnych dyscyplin nauki (Conway, 1997). Pojęcie reprezentacji pamięciowej charakteryzowane jest tutaj w sposób dość swobodny jako odwzorowanie w układzie nerwowym minionych zdarzeń (może to być odwzorowanie strukturalne, jak w wypadku map pamięciowych, lub mające inną formę), które są dostępne systemowi (niekoniecznie jego świadomości) oraz wywierają wpływ na jego funkcjonowanie. Odwzorowania pamięciowe przybierają różną postać na różnych etapach przetwarzania pamięciowego oraz podlegają wielorakim zniekształceniom. Dotychczasowe badania w niewielkim stopniu skupiały się na sposobach, na jakie systemy rozpoznają BRPA, koncentrowały się raczej na mechanizmach ich powstawania: biochemicznych, neurobiologicznych, psychologicznych i społecznych. Jednak również u podstaw zdolności do wykrywania BRPA leżą określone mechanizmy. Ich wielorakie zaburzenia mogą prowadzić do ślepoty na BRPA, czego przykładem są pacjenci z anozognozą towarzyszącą deficytom i zaburzeniom pamięci³¹.

³⁰ Zagadnienie to wymaga dalszej dyskusji. Jak słusznie zwrócił uwagę jeden z anonimowych recenzentów tego artykułu, można mieć wątpliwości, czy w takich sytuacjach (podobnie jak w wypadku omawianego w poprzedniej sekcji przypadku szczura przeniesionego do innego labiryntu) mamy do czynienia z błędną reprezentacją. Wiele zależy od tego, czy uznanie reprezentacji za błędną zależy od sposobu, w jaki została ona wytworzona. Faktem jest natomiast, że w literaturze zagadnienia tego typu przypadki są kwalifikowane jako przykłady błędnego reprezentowania.

³¹ Powyższy przegląd zagadnień dotyczących mechanizmów powstawania błędnych reprezentacji pamięciowych ma charakter szkicowy. Jak trafnie zauważył jeden z recenzentów, nie uwzględnia on wielu badań i teorii,

4. Zakończenie

Ogólna teoria BRP powinna dostarczyć: a) definicji BRP, b) charakterystyki różnych rodzajów BRP (percepcyjnych, pamięciowych, pojęciowych, przekonaniowych, językowych i innych), c) wielopoziomowej analizy mechanizmów oraz przyczyn powstawania BRP (dla każdego rodzaju procesów poznawczych), d) opisu wpływu BRP na funkcjonowanie systemów poznawczych, a także e) analizy sposobów, na jakie systemy poznawcze rozpoznają BRP, dokonują korekty zachowań, uczą się oraz minimalizują negatywne skutki występowania BRP. Mimo że dotychczas takiej teorii nie skonstruowano, istnieją godne uwagi przyczynki (filozoficzne i naukowe) do konstrukcji takiej teorii w przyszłości (Pohl, 2004). Jednym z nich jest również zaproponowana przez Krystynę Bielecką koherencyjna koncepcja błędu reprezentacyjnego. Chociaż w przekonaniu Bieleckiej KKBR może stanowić podstawę szerszej zakrojonych badań naukowych nad naturą, mechanizmami i sposobami wykrywania BRP, to biorąc pod uwagę fakt, iż KKBR jest teorią na temat roli niespójności między reprezentacjami w wykrywaniu BRP (trywialną w odniesieniu do ludzi, kontrowersyjną w odniesieniu do płazów i gryzoni), można mieć uzasadnione wątpliwości, czy teoria ta spełni pokładane w niej nadzieje.

W odróżnieniu od Bieleckiej uważam, że punktem wyjścia do budowy ogólnej teorii BRP powinna być szczegółowa analiza mechanizmów powstawania BRP. Wiele takich mechanizmów odkryto i opisano na gruncie psychologii poznawczej i neurokognitywistyki. Istnieje olbrzymi materiał empiryczny, który może stanowić podstawę oraz punkt wyjścia do konstrukcji ogólnej teorii BRP. Wymaga on analitycznego opracowania, do czego znakomicie nadają się narzędzia logiczno-analityczne wypracowane przez filozofów. Naukowcy posługują się pojęciami, takimi jak reprezentacja, odwzorowanie czy błąd reprezentacyjny, często w sposób niedopracowany teoretycznie (ogólnikowy, roboczy). Tymczasem filozofia może dostarczyć szczegółowych modeli formalnych aplikowalnych do danych empirycznych. Inne zadanie filozofii polega na stawianiu wiedzytwórczych pytań dających początek nowym programom badawczym. Filozofia może wreszcie pełnić rolę platformy badań interdyscyplinarnych integrującej wyniki pochodzące z różnych dyscyplin nauki.

Bibliografia

- Bechtel, W., Abrahamsen, A. (2005). Explanation: A Mechanistic Alternative. *Studies in History and Philosophy of the Biological and Biomedical Sciences*, 36, 421–441.
- Beracochea, D. (2006). Anterograde and retrograde effects of benzodiazepines on memory. *Scientific World Journal*, 6, 1460–1465.
- Bielecka, K. (2018). *Błądzą, więc myślę. Co to jest błędna reprezentacja?* Warszawa: Wydawnictwo Uniwersytetu Warszawskiego.
- Clark, A. (2002). Is Seeing All it Seems? Action, Reason, and the Grand Illusion. *Journal of Consciousness Studies*, 9 (5-6), 181–202.

takich jak teoria rozmytego śladu czy *dual-route model*. Jednak moim celem jest jedynie zwrócenie uwagi, że kluczowym problemem teorii błędnych reprezentacji jest analiza mechanizmów ich powstawania (co zostało zilustrowane przykładami badań nad mechanizmami powstawania błędnych reprezentacji pamięciowych).

- Conway, M.A. (1997) (ed.). *Recovered memories and false memories*. Oxford: Oxford University Press.
- Craver, C., Tabery, J. (2019). Mechanisms in Science. *The Stanford Encyclopedia of Philosophy*, Edward N. Zalta (red.), <<https://plato.stanford.edu/archives/sum2019/entries/science-mechanisms/>>.
- Frith, Ch. (2011). *Od mózgu do umysłu. Jak powstaje nasz wewnętrzny świat?*. Przeł. A. i M. Binder, Warszawa: Wydawnictwo UW.
- Gładziejewski, P. (2015). *Wyjaśnianie za pomocą reprezentacji mentalnych. Perspektywa mechanistyczna*, Warszawa-Toruń: Wydawnictwo UMK.
- Gregory, R.L. (2009). *Seeing Through Illusions*. New York: Oxford University Press.
- Gulla, B. (2017). Ryzyko fałszywych wspomnień w trakcie psychoterapii i innych oddziaływań psychologicznych. *Psychoterapia, 1*, 59–69.
- Hawkins, J., Blakeslee, S. (2006). *Istota inteligencji*. Przeł. T. Walczak. Gliwice: Wydawnictwo Helion.
- Herzyk, A. (1999). Anozognozja: mechanizmy mózgowie i objawy. *Przegląd Psychologiczny, 42(1–2)*, 99–110.
- Hirstein, W. (2005). *Brain Fiction: Self-deception and the Riddle of Confabulation*, Cambridge, Mass.: MIT Press.
- Hoffman, D. D. (2012). The Construction of Visual Reality. W: J.D. Blom, I.E.C. Sommer, E.C. Iris (red.), *Hallucinations. Research and Practice* (s. 7–15). New York: Springer Verlag.
- Hunt, K.I., Chittka, I. (2015). Merging of long-term memories in an insect. *Current Biology, 25(6)*: 741–745.
- Hyman I.E., Pentland, J. (1996). The role of mental imagery in the creation of false childhood memories. *Journal of Memory and Language, 32(5)*, 101–117.
- Kępiński, A. (2001). *Schizofrenia*. Kraków: Wydawnictwo Literackie.
- Króliczak, G. (1999). Błędy percepcyjne czy percepcje normalne? Mechanizmy powstawania iluzji percepcyjnych. W: A. Klawiter, A.L. Nowak, P. Przybysz. (red.). *Umysł a rzeczywistość. Poznańskie Studia z Filozofii Humanistyki, 18*, 303–325.
- Loftus, E.F. (1996). Memory distortion and false memory creation. *Bulletin of the American Academy of Psychiatry and the Law, 24*, 281-295.
- Loftus, E.E. (1997). Creating false memories. *Scientific American, 277*, 70-76.
- Loftus, E.F. (2000). Remembering what never happened. W: E. Tulving (ed.). *Memory, consciousness, and the brain* (s. 106–108). Philadelphia: Psychology Press.
- Loftus, E.F., Pickrell, J.E. (1995). The formation of false memories, *Psychiatric Annals, 25*, 720–725.
- Machamer, P.K., Darden, L., Craver, C.F. (2000). Thinking about Mechanisms. *Philosophy of Science, 67*, 1–25.
- Marcus, G. (2009). *Prowizorka w mózgu*. Przeł. A. Nowak. Sopot: Smak Słowa.

- Nader, K., Schafe, G.E., LeDoux, J.E. (2000). Fear Memories require protein synthesis in the amygdala for reconsolidation after retrieval. *Nature*, 406(6797), 722–726.
- Niedźwieńska, A. (2004). *Poznawcze mechanizmy zniekształceń w pamięci zdarzeń*. Kraków: Wydawnictwo UJ.
- Noë, A. (2002). Is the Visual World a Grand Illusion? *Journal of Consciousness Studies*, 9 (5-6), 1–12.
- Oronowicz-Jaśkowiak, W. (2017). Krytyczna analiza strategii terapeutycznych zorientowanych na odzyskiwanie pamięci. *Przegląd Seksuologiczny*, 3 (51), 34–41.
- Piccinini, G. (2007). Computing Mechanisms. *Philosophy of Science*, 74, 501–526.
- Pohl, R.F. (ed.) (2004). *Cognitive illusions: A handbook on fallacies and biases in thinking, judgement, and memory*. Hove: Psychology Press.
- Prigatano, G.P. (2010) (red.). *Study of anosognosia*. Oxford: Oxford University Press.
- Ramirez, S., Liu, X., Lin, P.A., Suh, J., Pignatelli, M., Redondo, R.I i in. (2013). Creating a false memory in the hippocampus. *Science*, 341(6144), 387–391.
- Robins, S.K. (2016). Optogenetics and the mechanism of false memory. *Synthese*, 5(193), 1561–1583.
- Schacter, D.L. (red.) (1995). *Memory Distortion. How minds, brains, and societies reconstruct the past*. Cambridge, MA: Harvard University Press.
- Schacter, D.L., (2003). *Siedem grzechów pamięci*. Przeł. E. Haman. Warszawa: PIW.
- Shaw, J. (2018). *Oszustwa pamięci. Zapamiętywanie, zapomnianie i fałszywe wspomnienia*. Przeł. A. Cichowicz, Warszawa: Wydawnictwo AMBER.
- Spitzer, M. (2012). *Jak uczy się mózg?* Przeł. M. Guzowska-Dąbrowska. Warszawa: Wydawnictwo Naukowe PWN.
- Shea, N. (2007). Consumers Needs Information: Supplementing Teleosemantics with an Input Condition. *Philosophy and Phenomenological Research*, 75(2), 404–435.
- Spalt J. (2002). Déjà vu: possible parahippocampal mechanisms. *The Journal of Neuropsychiatry and Clinical Neuroscience*, 14(1), 6-10. doi: 10.1176/jnp.14.1.6
- Strange, D., Sutherland, R., Garry, M. (2006). Event plausibility does not determine children's false memories. *Memory*, 14(8), 937–951.
- Thagard, P. (2000). *How scientists explain disease*. Princeton: Princeton University Press.
- Ulatowska, J., Sawicka, M. (2017). Odzyskiwanie wspomnień w praktyce klinicznej – przegląd badań. *Psychiatria Polska*, 51(4), 609–618.
- Vetulani, J. (2011). Neurobiologia pamięci. W: J. Vetulani, *Mózg. Fascynacje, problemy, tajemnice* (s. 195-220). Kraków: Wydawnictwo Homini.
- Vilarroya, V. (2017). Neural Representation. A Survey-Based Analysis of the Notion. *Frontiers in Psychology*, 8, 1458. doi: 10.3389/fpsyg.2017.01458.

The coherence conception of representational error in the context of research into the formation mechanisms of memory misrepresentations

Abstract: The aim of this paper is to provide a critical analysis of Krystyna Bielecka's position presented in her book: *I Err, Therefore I Think. What is Misrepresentation? (Błądź, więc myślę. Co to jest błędna reprezentacja?)*. The paper also contains a sketch of an alternative naturalistic account of cognitive misrepresentations based on the mechanistic model of scientific explanation. In my opinion the main problem of the theory of cognitive (mis)representations concerns formation mechanisms – neurobiological, psychological, social and other. I briefly illustrate my position by examples of mechanism of memory misrepresentations. The final part of the paper contains some comments on the general theory of cognitive (mis)representations and the role that philosophy can play in its construction.

Key words: cognitive representation; cognitive misrepresentation; representational error; mechanistic model of scientific explanation; memory errors; mechanisms of memory misrepresentations

Dr hab. **Robert Poczobut**, prof. UwB – kierownik Zakładu Epistemologii i Kognitywistyki w Instytucie Filozofii UwB. W latach 1993-1999 zatrudniony w Zakładzie Logiki i Metodologii Nauk UMCS, od roku 2000 na Uniwersytecie w Białymstoku. Stypendysta Fundacji na Rzecz Nauki Polskiej (2003), laureat Nagrody im. Tadeusza Kotarbińskiego (2001). Członek zarządu Polskiego Towarzystwa Kognitywistycznego (2014–2020). Opublikował cztery książki oraz kilkadziesiąt artykułów z filozofii logiki, filozofii umysłu i filozofii kognitywistyki. Za pracę *Między redukcją a emergencją. Spór o miejsce umysłu w świecie fizycznymi* (2009) uzyskał wyróżnienie Fundacji na Rzecz Nauki Polskiej. Redaktor i współredaktor licznych prac zbiorowych – m.in. *Analityczna metafizyka umysłu. Najnowsze kontrowersje* (2008), *Przewodnik po filozofii umysłu* (2012), a także dwóch kognitywistycznych numerów (pt. *Cognitive Systems* oraz *Interfiled Integration in Cognitive Science*) czasopisma „Studies in Logic, Grammar and Rhetoric”. Przygotowuje do druku książkę pt. *Umysł i artefakty* na temat natury i wielorakich odmian hybrydowych systemów poznawczych.

Redakcję i publikację tekstu sfinansowano ze środków Ministerstwa Nauki i Szkolnictwa Wyższego na działalność upowszechniającą naukę (DUN), działalność wydawnicza, nr umowy: 711/P-DUN/2019, okres realizacji: 2019–2020.