

Uniwersytet Szczeciński
Instytut Filozofii i Kognitywistyki

Janina Mękarska

**Wpływ uprzedzeń na interpretacje badań procesów poznawczych
zwierząt innych niż ludzie na tle debaty nad teorią umysłu**

Praca doktorska napisana pod kierunkiem:
dr hab. Adrian Schetz, prof. US

Szczecin, 2025

Spis treści

Wstęp:	5
Rozdział I: Czym są uprzedzenia i jakie mechanizmy leżą u ich podstaw	9
1.1. Rozważania dotyczące wpływu uprzedzeń na interpretacje badań naukowych.....	9
1.2. Podstawowe kwestie terminologiczne	16
1.2.1. Zdolności z zakresu poznania społecznego	17
1.2.2. Uprzedzenia, stereotypy, dyskryminacja	22
1.2.3. Klasyfikacja uprzedzeń	25
1.3. Wpływ uprzedzeń międzygrupowych na interpretacje badań nad poznaniem ZINL	29
1.4. Podsumowanie	33
Rozdział II: Wybrane mechanizmy poznawcze leżące u podstaw tendencyjności związanej z uprzedzeniami	35
2.1. Mechanizm kategoryzacji społecznej.....	35
2.2. Mechanizm samo-licencjonowania	41
2.3. Myślenie analogiczne.....	44
2.4. Efekty oczekiwań.....	50
2.5. Teoria dysonansu poznawczego	56
2.6. Słodka reakcja.....	69
2.7. Efekt halo	77
2.8. Podsumowanie	79
Rozdział III: Narracje naukowe konstytuujące porządek społeczny w kontekście relacji człowiek—inne zwierzęta.....	82
3.1. Historia badania umysłów zwierząt pozaludzkich	82
3.2. Dwie perspektywy interpretowania eksperymentów z udziałem ZINL	97
3.3. Komponent poznawczy	101
3.3.1. Mity i inne narracje	102
3.3.2. Stereotypy.....	110
3.4. Behawioralny komponent uprzedzeń	115
3.4.1. Tendencyjność	116
3.4.2. Heurystyki i związane z nimi tendencje w zachowaniu i podejmowaniu decyzji istotne w kontekście uprzedzeń międzygatunkowych	127
3.5. Podsumowanie	171
Rozdział IV: Kluczowe koncepcje teorii umysłu.....	173

4.1. Współczesne filozoficzne koncepcje mindreadingu.....	173
4.2. Model zdolności czytania umysłów innych Sinoma Baron-Cohena	178
4.3. Test fałszywych przekonań	182
4.4. Stopniowalność teorii umysłu w ujęciu Robina Dunbara i José Luisa Bermúdeza.....	186
4.5. Podsumowanie	195
Rozdział V: Tendencyjność w badaniu teorii umysłu	197
5.1. Teoria umysłu poza gatunkiem ludzkim: argumenty i dowody	197
5.1.1. Przełomowe badania Premacka i Woodruffa.....	199
5.1.2. Badania Daniela Povinelliego i Kurta Nelsona	207
5.1.3. Badania Calla i Tomasello.....	211
5.1.4. Najnowsze dane empiryczne	214
5.2. Perspektywa krytyczna w kwestii teorii umysłu u zwierząt innych niż ludzie: wybrane badania	231
5.2.1. Daniel Povinelli i Timothy Eddy	234
5.2.2. Stanowisko Sary Shettleworth.....	247
5.3. Podsumowanie	260
Rozdział VI: niesprawiedliwość wynikająca z uprzedzeń względem zwierząt pozaludzkich	262
6.1. niesprawiedliwość związana z uprzedzeniami względem zwierząt pozaludzkich	262
6.2. Podstawowe pojęcia w koncepcji niesprawiedliwości epistemicznej	263
6.3. niesprawiedliwość testymonialna.....	267
6.4. Uprzedzenia tożsamościowe a niesprawiedliwość testymonialna względem ZINL	272
6.5. niesprawiedliwość hermeneutyczna	291
6.6. Uprzedzenia tożsamościowe a niesprawiedliwość hermeneutyczna względem ZINL.....	296
6.7. Podsumowanie	300
Zakończenie:.....	302
Bibliografia:	304
Netografia:	346
Summary:.....	357

Wstęp:

Możliwości poznawcze zwierząt pozaludzkich od dawna stanowiły obiekt zainteresowania filozofii, nauk przyrodniczych oraz psychologii. Jednak mimo rosnącego korpusu badań empirycznych, pewne pytania nadal pozostają bez odpowiedzi. W literaturze z różnych okresów obserwujemy próby wyznaczenia jasnej linii demarkacyjnej między ludźmi a innymi zwierzętami, wskazując na określone zdolności, które mają być wyłącznie ludzkie. Cechy te stają się dystynktywnymi atrybutami człowieka, które nie tylko podkreślają jego wyjątkowość, ale często także wyższość nad innymi gatunkami. Od lat jedną ze zdolności wyznaczającą linię demarkacyjną jest *mindreading*, dzięki której człowiek potrafi przewidywać działania innych, rozumieć intencje i podejmować złożoną współpracę (Tomasello, 2015). Chociaż współczesne eksperymenty procesów poznawczych zwierząt innych niż ludzie (ZINL) sugerują, że ich możliwości, w wielu aspektach wskazanej zdolności nie odbiegają od tych ludzkich. Dlatego coraz bardziej zasadne staje się pytanie, czy granice wyznaczone międzygatunkowo są rzeczywistym odzwierciedleniem różnic poznawczych, czy raczej efektem utrwalonych uprzedzeń i ram interpretacyjnych?

W rozważaniach z zakresu poznania społecznego funkcjonują dwa nurty interpretacyjne dotyczące zdolności przypisywania stanów mentalnych innym — perspektywa mentalistyczna oraz behawioralna¹. Pierwsza z wymienionych zaleca wyjaśnianie zachowania zwierząt pozaludzkich poprzez odwołanie do stanów mentalnych. Natomiast zwolennicy drugiego z wymienionych nurtów prezentują sceptyczne postawy wobec odwołań tego rodzaju. Powołują się na regułę parsymonii — zasadę Morgana² — konstruując wyjaśnienia oparte na argumencie *ad ignorantiam*³. Moim celem jest porównanie wskazanych nurtów interpretacyjnych funkcjonujących w obszarze *mindreadingu*, ze szczególnym uwzględnieniem teorii umysłu⁴. Wskazana zdolność odpowiada za przewidywanie zachowania na podstawie fałszywych

¹ Szczegółowy opis założeń każdej ze wskazanych perspektyw znajduje się w części 3.2. *Dwie perspektywy interpretowania eksperymentów z udziałem zwierząt pozaludzkich*.

² Szeroki opis zasady Morgana znajduje się w części 3.1. *Historia badania umysłów zwierząt pozaludzkich*.

³ Szeroko opisany w części 2.5. *Teoria dysonansu poznawczego*.

⁴ Rozumiana w niniejszej dysertacji jako zdolności do przewidywania zachowania innego na podstawie jego fałszywych przenonań, co wymaga zrozumienia, że inny podmiot może mieć wyobrażenia o świecie niezgodne z rzeczywistością, i że na ich podstawie może podejmować działania.

przekonań, podzielalnych przez agenta. Wymaga to zrozumienia, że ten ma swoje własne wyobrażenia o świecie, które nie muszą zgadzać się z faktami. Dlatego coraz bardziej zasadne staje się pytanie, czy granice wyznaczone międzygatunkowo rzeczywiście odzwierciedlają faktyczne różnice poznawcze, czy raczej są produktem utrwalonych uprzedzeń i ram interpretacyjnych.

Choć przez dekady przeprowadzono już wiele badań, a same narracje wyjaśniające zdolności poznawcze zwierząt stały się znacznie bardziej złożone, trudno oprzeć się wrażeniu, że nie zbliżyły nas one znacząco do jednoznacznego rozstrzygnięcia. Nadal toczy się debata zapoczątkowana niemal pół wieku temu przez klasyczny już artykuł Premacka i Woodruffa (1978) zatytułowany *Does the chimpanzee have a theory of mind?* Pomimo 47 lat intensywnych badań i powszechnego założenia, że ludzkie zdolności poznawcze muszą mieć podstawę ewolucyjną, przypisywanie zwierzętom pozaludzkim nawet najbardziej podstawowych mechanizmów *mindreadingu* – takich jak detekcja intencjonalności czy rozumienie kierunku patrzenia – wciąż budzi kontrowersje.

Uważam, że współczesne spory wokół teorii umysłu w odniesieniu do ZINL ujawniają głębsze problemy epistemologiczne, metodologiczne i kulturowe. Niniejsza dysertacja stanowi próbę połączenia krytyki epistemicznej, analizy poznawczej oraz refleksji etycznej. Jej celem nie jest rozstrzygnięcie sporu o teorię umysłu, lecz uwrażliwienie na sposoby, w jakie struktury poznawcze i kulturowe kształtują nasze rozumienie (i nierozumienie) innych umysłów – zarówno ludzkich, jak i nie-ludzkich. W centrum rozważań znajduje się problematyka uprzedzeń poznawczych i kulturowych, które wpływają na interpretację zdolności umysłowych zwierząt, a także konsekwencje epistemiczne, metodologiczne i etyczne wynikające z przyjęcia określonego stanowiska interpretacyjnego. Szczególnie istotne są tu: napięcie między podejściem mentalistycznym a behawioralnym, asymetria w analizie wyników eksperymentów z udziałem ludzi i zwierząt innych niż ludzie, a także rola kulturowych narracji i mitów w utrwalaniu wykluczających założeń. Praca nie opiera się na jednej hipotezie badawczej, lecz przyjmuje charakter analityczno-krytyczny, dążąc do ujawnienia mechanizmów epistemicznych stojących za dominującymi interpretacjami poznania zwierząt.

Rozdział pierwszy stanowi wprowadzenie do problematyki pracy i przedstawia główne kierunki rozumienia teorii umysłu oraz ich implikacje dla badań z udziałem zwierząt. Omawiam w nim podstawowe założenia poznawcze i filozoficzne, które

stanowiąc będą ramy dla dalszej analizy. Rozdział drugi poświęcony jest wybranym mechanizmom poznawczym, które leżą u podstaw tendencyjności i uprzedzeń, w tym kategoryzacji społecznej, infrahumanizacji, efektowi jednorodności grupy obcej czy mechanizmowi samo-licencjonowania. Ich analiza ukazuje, jak głęboko zakorzenione są procesy automatyczne, które mogą wpływać na ocenę zdolności umysłowych innych. Rozdział trzeci rekonstruuje konteksty historyczne, epistemiczne i kulturowe, które wpływają na interpretację poznania zwierząt. Szczególnie istotne miejsce zajmuje tu analiza narracji antropomorficznych i antroponegacjonistycznych oraz wskazanie na funkcje mitów w strukturze poznawczej. Rozdział czwarty skupia się na współczesnych modelach teorii umysłu, w szczególności metareprezentacyjnych, analizując ich przydatność w badaniach nad poznaniem społecznym ZINL. Omawiam również model Baron-Cohena i problematykę testów fałszywych przekonań, a także złożoność teorii umysłu, powołując się na teorie jej stopniowości. W rozdziale piątym przedstawiona została analiza asymetrii interpretacyjnej w badaniach nad teorią umysłu u ludzi i zwierząt oraz krytyka redukcjonistycznych ujęć behawioralnych. Zwracam uwagę na to, w jaki sposób przyjęcie określonych założeń teoretycznych wpływa na ocenę wyników empirycznych i możliwość przypisywania stanów mentalnych. Rozdział szósty stanowi próbę rozszerzenia kategorii niesprawiedliwości epistemicznej (w duchu Mirandy Fricker) na relację międzygatunkową. Omawiam w nim nie tylko systematyczne wykluczanie zwierząt jako podmiotów epistemicznych, ale także marginalizację osób reprezentujących podejście mentalistyczne.

W tym kontekście warto zaznaczyć, że chociaż w dysertacji przyjmuję, iż uprzedzenia mają charakter postawy, nie zdecydowałam się na wydzielenie osobnej sekcji poświęconej komponentowi afektywnemu. Nie wynika to z przekonania o jego marginalnym znaczeniu, lecz raczej z dążenia do utrzymania spójności obecnych, pogłębionych analiz. Uważam, że emocjonalny wymiar uprzedzeń zasługuje na osobne opracowanie – i mam nadzieję, że przedstawione tu rozważania będą stanowić impuls do dalszych badań i refleksji właśnie w tym kierunku. Szersze uasadnienie wskazanego wyboru znajdzie się w dalszych rozważaniach oraz podkreślone zostanie w kilku rozdziałach dysertacji. Chcę jednak zaznaczyć, że taki zabieg strukturalny nie wyklucza wskazanego komponentu z zawartych niżej rozważań. Zagadnienie afektu zostanie szeroko omówione szczególnie w rozdziałach *II: Wybrane mechanizmy poznawcze leżące u podstaw tendencyjności związanej z uprzedzeniami* oraz *III: Narracje naukowe konstytuujące porządek społeczny w kontekście relacji człowiek — inne zwierzęta*.

Struktura pracy może zostać odczytana jako odbiegająca od konwencjonalnego, linearnego układu rozprawy naukowej. Świadomie przyjąłam taką konstrukcję, gdyż pozwala ona na wielowarstwową i asocjacyjną analizę treści, co wynika z charakteru analizowanego problemu. Zagadnienie zdolności poznawczych ZINL oraz interpretacji ich zachowań nie daje się zamknąć w jednoznacznych ramach pojęciowych czy metodologicznych, dlatego kolejne rozdziały podejmują temat z różnych perspektyw, niekiedy powracając do wcześniejszych wątków i uzupełniając je o nowe konteksty. Taka struktura odzwierciedla metodologiczną postawę, zgodnie z którą pogłębione zrozumienie złożonych zjawisk poznawczych wymaga łączenia różnych obszarów analizy.

Rozdział I: Czym są uprzedzenia i jakie mechanizmy leżą u ich podstaw

1.1. Rozważania dotyczące wpływu uprzedzeń na interpretacje badań naukowych

Sposób w jaki ludzie interpretują informacje dotyczące otaczającego ich świata, zależy nie tylko od bodźców zmysłowych, lecz także od ich kontekstu w perspektywie przekonań czy oczekiwań. Uprzedzenia prezentowane wobec zwierząt innych niż ludzie, kształtują interpretacje wyników badań dotyczących procesów poznawczych, takich jak teoria umysłu. Jest to naturalna konsekwencja wynikająca z własności systemu poznawczego ludzi, której elementem są struktury poznawcze utrzymujące porządek społeczny. Wskazane interpretacje stanowią formę narracji naukowych konstytuujących dobór narzędzi badawczych, takich jak np. paradygmaty stosowane w ramach procedury eksperymentalnej. Uważam to za interesujące, gdyż automatyczny charakter wielu mechanizmów związanych z uprzedzeniami przyczynia się do łamania dobrych praktyk epistemicznych, osłabiając zdolność jednostek do rzetelnej oceny rzeczywistości. Percepcja i interpretacja danych są kształtowane przez intencje, oczekiwania czy afekty jednostki, co prowadzi do systematycznych błędów poznawczych, a w konsekwencji – do utrwalania nierówności społecznych i błędnych założeń naukowych. W tym kontekście refleksja dotycząca wpływu uprzedzeń na procesy poznawcze nie jest jedynie zagadnieniem teoretycznym, lecz kluczowym wyzwaniem nie tylko dla epistemologii, ale też praktyki badawczej. Co więcej, kulturowe uprzedzenia ukryte kształtują sposób interpretacji danych, co przekłada się na hipotezy badawcze formułowane na gruncie psychologii porównawczej. Wspomniany proces wiąże się m.in. z tendencjami do przyjmowania lub ignorowania informacji, co zniekształca wyciągane z nich wnioski. Ludzki system poznawczy operuje na uproszczonych regułach wnioskowania (heurystykach), co prowadzi do tendencyjności w ocenie zdolności poznawczych ZINL. Uważam, że omawiane postawy widoczne są w debacie dotyczącej teorii umysłu zwierząt⁵. Ludzie często intuicyjnie przyjmują, że zdolność do przypisywania stanów mentalnych innym jest unikalna dla *homo sapiens*, co może prowadzić do braku precyzji w ocenie wiarygodności interpretacji naukowych. Analiza wpływu uprzedzeń na interesujące mnie narracje, jest niezbędna

⁵ Chociaż dla niektórych informacja ta może być oczywista, to chcę podkreślić, że w niniejszym opracowaniu termin „zwierzęta” odnosi się do przedstawicieli królestwa zwierząt, czyli bardzo szerokiego zbioru gatunków, do którego wlicza się także człowieka.

dla refleksyjności nauki i rozważań nad indywidualnym unikaniem błędów poznawczych. Uznanie, że każdy komponent uprzedzenia wynika z określonych tendencji, pozwala na bardziej krytyczne podejście do metodologii badań oraz interpretacji ich wyników. Dla przykładu, jeśli badacze nieświadomie operują na antropocentrycznych albo antroponegacjonistycznych schematach poznawczych, mogą nadinterpretować dane zgodnie z oczekiwaniami.

W swoim interesującym artykule *Bias and Perception* opublikowanym w książce *Knowledge, Justice, and the Social Mind*, Susanna Siegel, (2020) argumentuje za poglądem głoszącym, że percepcja jest kształtowana przez uprzedzenia, co rodzi pytania o wiarygodność ludzkich spostrzeżeń i sądów. Wskazuje przy tym szereg procesów istotnych w interesującym mnie obszarze, dlatego niektóre z nich omówię w niniejszym opracowaniu. Uważam, że uprzedzenia międzygatunkowe zniekształcają sposób w jaki ludzie postrzegają intencje, wiedze lub działania przedstawicieli innych gatunków zwierząt, co konstytuuje zachowanie wobec nich. Mechanizmy leżące u podstaw wykazywania tego rodzaju dyskryminacji, są analogiczne do tych będących źródłem postaw rasowych (Gilbert, 2013; Sikkens, Bosman, Olcese, 2019). Kluczowym mechanizmem leżącym u podstaw zniekształceń percepcyjnych wynikających z uprzedzeń, jest penetracja poznawcza (Siegel, 2012; 2020; Main, 2023; Rousay, 2023). Zjawisko to dobrze prezentuje klasyczny już eksperyment, w którym prezentowano jak postawy rasowe prowadzą do błędnej interpretacji niejednoznacznych bodźców (Payne, 2001). Uczestnikom badania przedstawiano serię zdjęć różnych narzędzi, a ich zadaniem było, jak najszybsze naciśnięcie przycisków, gdzie jeden odpowiadał za broń palną, a drugi za inne narzędzie (np. łopaty). Czas ekspozycji na bodziec był na tyle krótki, aby umożliwić ich dostrzeżenie, lecz uniemożliwić pełną identyfikację obiektu. Przed bodźcem właściwym, podprogowo wyświetlano prymę, którą były zdjęcia czarnoskórego i białego człowieka. Wyniki wykazały, że gdy uczestnicy widzieli pierwszego z wymienionych mężczyzn, częściej klasyfikowali obiekt jako broń palną. Co więcej oczekiwania mogą również powodować, że uwaga jednostki będzie wybiórczo rejestrować wybrane cechy otoczenia, ignorując te uznane za mniej istotne.

Umysł wypełnia luki w informacji, wykorzystując doświadczenia z przeszłości np. wnioskując o brakujących słowach w zdaniu na podstawie kontekstu (Rousay, 2023). Zjawisko to, prowadzi do systematycznych błędów w interpretacji rzeczywistości (Rousay, 2023). Czyli postawy takie, jak antropomorfizm czy

antroponegacjonizm kształtują cały proces poznawczy, już na etapie percepcji (Gilbert 2013; Sikkens, Bosman, i Olcese, 2019; Main, 2023). Dla przykładu, stosunek jednostki do przepisów prawa, konstytuuje jej zachowanie podczas jazdy samochodem. Tak samo, uprzedzenia międzygatunkowe skutkują tendencyjnym odbiorem wyników badań i ich interpretacją zgodnie ze stereotypami i istniejącymi narracjami. Na przykład, zakładając, że osoba popierająca określone interpretacje nieświadomie przyjmuje, że tylko ludzie posiadają zdolność do intencjonalnego działania, to nawet w obliczu dowodów wskazujących na ową intencjonalność w zachowaniach zwierząt, może on interpretować je jako automatyczne lub warunkowe, odrzucając bardziej złożone wyjaśnienia. W ten sposób uprzedzenia międzygatunkowe mogą wpływać na konstruowanie i utrzymywanie określonych interpretacji naukowych, podobnie jak wcześniejsze schematy poznawcze mogą utrudniać adaptację do nowych warunków w innych obszarach. Przytoczony przypadek stanowi rażące złamanie dobrej praktyki epistemicznej, jaką jest wrażliwość. Głosi ona, że śledzenia prawdy jest możliwe, wtedy i tylko wtedy gdy jest wrażliwe. Przekonanie (P) jest wrażliwe, w sytuacji, w której gdyby było fałszywe to podmiot (O) by je porzucił (Nozik, 1981; Holroyd, Puddifoot, 2019). Natomiast status epistemiczny (P) (wiarygodny/niewiarygodny), określa się przez ilość i jakość przesłanek świadczących za jego słuszością. Inną dobrą praktyką związaną z uprzedzeniami jest bezpieczeństwo, które stanowi, że (P) nie może być prosto zweryfikowane jako fałszywe (Williamson, 2000; Pritchard, 2005; Holroyd, Puddifoot, 2019).

Procesy przetwarzania odgórnego integrują wcześniejszą wiedzę z napływającymi danymi, wpływają na sposób, w jaki jednostki interpretują rzeczywistość oraz odgrywają kluczową rolę w budowaniu jej naukowego wizerunku. W przeciwieństwie do przetwarzania oddolnego⁶, które operuje na surowych danych sensorycznych, odgórna analiza sytuacji, zaczyna się od procesów wyższego rzędu, takich jak pamięć czy oczekiwanie (Main, 2023; Rousay, 2023). Pozwala to skuteczne

⁶ Dla lepszego zrozumienia różnicy między przetwarzaniem odgórnym i oddolnym, warto przytoczyć przykład drugiego z wymienionych. Wyobraźmy sobie, że jednostka po raz pierwszy spaceruje po parku, kiedy zauważa na ziemi kolorowy obiekt. W miarę zbliżania się rejestruje jego kształt, kolor i fakturę, co pozwala na rozpoznanie, że to frisbee. W takim przypadku procesy percepcyjne zaczynają się od surowych danych sensorycznych, które są stopniowo przetwarzane, aż do momentu rozpoznania obiektu. Interpretacja nie jest kształtowana przez wcześniejsze oczekiwania czy wiedzę, a wyłącznie przez informacje zmysłowe prowadzące do rozpoznania bodźca, taki rodzaj analizy nazywamy oddolnym.

nadawanie sensu niejednoznacznym lub niekompletnym bodźcom, korzysta jednak z informacji, które mogą zniekształcać percepcję (Main, 2023; Rousay, 2023). Dla przykładu podczas powrotu do domu, gdy jednostka przechodzi nocą przez ciemną alejkę, wiedza i oczekiwania, mogą wpłynąć na to w jaki sposób interpretuje ona informacje ze świata (ciemne alejki są często kojarzone z niebezpieczeństwem). W takiej sytuacji, każdy dźwięk lub cień postrzegany jest jako potencjalne zagrożenie: szelest, automatycznie może wywołać przyjęcie fałszywych założeń, np. że ktoś ukrywa się w pobliskim krzaku, choć w rzeczywistości jest to jedynie wiatr. Ten sam mechanizm umożliwia interpretację tekstu, podczas czytania, gdzie wcześniejsza wiedza pozwala na szybkie dekodowanie niejednoznacznych słów (Main, 2023; Rousay, 2023). Chociaż przetwarzanie odgórne znacząco ułatwia funkcjonowanie w złożonym środowisku, niesie ze sobą również pewne zagrożenia. Szybsze przetwarzanie informacji zwiększa skuteczność rozpoznawania wzorców, kosztem zmniejszenia elastyczności poznawczej. Dzięki redukcji obciążenia w znanych sytuacjach umożliwia sprawniejsze działanie, lecz ogranicza zdolność do adaptacji w nowych warunkach (znaczące wady oraz zalety przetwarzania odgórnego przedstawione zostały w tabeli 1).

Tabela 1. Wady i zalety przetwarzania odgórnego

Korzyści	Wady
Przyspiesza percepcję w złożonych środowiskach (Schendan, Ganis, 2015).	Grozi błędną interpretacją bodźców (np. pomylenie narzędzi z bronią); (Payne, 2001; Correll, Park, Judd, Wittenbrink, 2002).
Zwiększa rozpoznawanie wzorców (Nigam, Schwiedrzik, 2024).	Utrwala stereotypy lub uprzedzenia (Bodenhausen, 1988).
Redukuje obciążenie poznawcze w znanych zadaniach (Simon, Tusch, Holcomb, Daffner, 2016).	Ogranicza zdolność adaptacji do nowych sytuacji (Allen, Sherman, Conrey, Stroessner, 2009).

[Opracowanie własne]

To jaki stosunek jednostka wykazuje w relacji do odbieranego świata, zależy m. in. od kształtowanych kulturowo narracji pozwalających wyjaśnić często niezrozumiałą rzeczywistość, za pomocą relatywnie prostych struktur poznawczych. Mity naukowe, którym poświęcam w niniejszej dysertacji szczególną uwagę, stanowią przykład, powszechnie obserwowanych w nauce narracji (Czeremski, 2021). Adaptacyjną funkcją kultury jest umożliwienie jednostkom i społecznościom skuteczne dostosowanie do środowiska poprzez rozwój technologii, norm społecznych oraz innych mechanizmów, które zwiększają szanse na przetrwanie i sukces społeczny (Smolla, Akçay, 2023). Zmienność kulturowa z powodzeniem rozpatrywana jest w perspektywie zasad ewolucji, co pozwala lepiej zrozumieć, jak kultura rozwija się i dostosowuje do zmieniających się warunków. Ewolucja kulturowa jest dynamicznym procesem, który umożliwia szybkie adaptacje i innowacje, co jest kluczowe dla ludzkiego sukcesu w różnych środowiskach. Podlega usprawnieniom oraz selekcji, następuje poprzez akceptację lub odrzucenie nowych praktyk i idei przez społeczność (Daszkiewicz, 2015). Badania pokazują, że kultura i geny ewoluują współzależnie co kształtuje rozwój ludzkości (Gintis, 2011; Whitehead, Laland, Rendell, Thorogood, Whiten, 2019). Sama kultura może wpływać na genetyczne dostosowania, a geny mogą determinować zdolność do nabywania i przekazywania kultury (Waring, Wood, 2021). Zatem, na to jak jednostka ustosunkowuje się do świata, który odbiera, wpływają zarówno narracje kulturowe, jak i doświadczenia osobiste kształtujące, wiedzę czy umiejętności, dające się szybko dostosowywać do zmieniających się warunków środowiskowych.

Człowiek rozpatrywany jako gatunek jest zwierzęciem stadnym, które osiągnęło ogromny sukces ewolucyjny oparty na współpracy. Życie w grupie było kluczowe dla przetrwania – samotny osobnik miał minimalne szanse w środowisku naturalnym ze względu na ograniczone możliwości zdobywania pożywienia, obrony przed drapieżnikami czy wychowania potomstwa. Ewolucja społeczna *homo sapiens* obejmowała rozwój złożonych struktur współpracy, w tym komunikacji (np. języka) oraz kultury, które umożliwiły tworzenie trwałych społeczności (Foley, 1996; Foley, Gamble, 2009). Doniesienia empiryczne dotyczące zachowań społecznych (nie tylko u naczelnych⁷), wskazują, że mechanizmy premiujące np. agresję czy altruizm podobne

⁷ Naczelne to grupa ssaków obejmująca zarówno małpy szerokonose, jak i małpy wąskonose, w tym małpy człekokształtne, do których zaliczają się goryle, szympansy i orangutany, a także ludzie. Ludzie,

są u ludzi i innych zwierząt, co sugeruje ich wspólne pochodzenie ewolucyjne (Foley, 1996). Strategia rozrodcza człowieka, oparta na nielicznym potomstwie wymagającym długotrwałej opieki, wymusiła życie w grupie. Wsparcie społeczne zwiększało szanse przeżycia dzieci, a podział obowiązków (np. zbieractwo, polowania) optymalizował wykorzystanie zasobów (Foley, Gamble, 2009). Ewolucja zachowań altruistycznych, zarówno krewniaczych czy odwzajemnionych, ugruntowała mechanizmy współpracy, tj. jednostki pomagające innym zwiększały szansę na przetrwanie całej grupy (Śliwak, 1999; Tomasello, Melis, Tennie, Wyman, Herrmann, 2012; Adamus, 2018). Przykładowo, opieka nad starszymi przedstawicielami plemienia, pozwalała na przekazywanie wiedzy między pokoleniami (Foley, Gamble, 2009).

Przetrwanie w grupie wymagało szybkiego oceniania stanów mentalnych takich jak: intencje, motywacje, emocje czy wiedza, a zatem cechami preferowanymi ewolucyjnie, były te, umożliwiające przyswajanie norm społecznych, co ułatwiało koordynację działań w dużych społecznościach (Centola, Baronchelli, 2015). Osoby przejawiające altruizm⁸ wobec krewnych lub sojuszników, zwiększały szanse na przekazanie genów (Boyd, 1985). Najbardziej podstawowym mechanizmem leżącym u podstaw wykazywania zaufania lub nieufności do obiektów dostępnych poznawczo, jest kategoryzacja poznawcza. Umożliwia ona szybkie i efektywne przetwarzanie informacji w środowisku. Procesy przyporządkowywania elementów środowiska do kategorii umożliwia redukcję złożoności świata społecznego, pozwalając na rozpoznanie i sklasyfikowanie innych według istotnych cech, takich jak przynależność grupowa, status czy rola społeczna (Tajfel, 1982). Tym samym, pozwala na przydzielanie różnych zwierząt (w tym oczywiście ludzi) do określonych kategorii, co ułatwia podejmowanie decyzji i reagowanie na sytuacje społeczne. Najbardziej podstawową kategorią społeczną jest *ja — inny*, natomiast bardziej złożone formy przyjmują postać *my — oni*. Zatem, mechanizm kategoryzacji społecznej kreuje postawy wobec dostępnych poznawczo aktorów, w zależności od strategii przetrwania gatunku. U zwierząt stadnych, takich jak naczelné, ssaki czy niektóre owady, kategoryzacja *my — oni* jest powszechna, a jej funkcją adaptacyjną jest utrzymanie spójności grupy i koordynacji działań w celu osiągnięcia wspólnych korzyści.

jako gatunek *homo sapiens* są częścią tej samej linii ewolucyjnej, co inne naczelné, z którymi dzielimy wspólnego przodka (Lamża, 2017).

⁸ Przy tym należy pamiętać, że w perspektywie etologii altruizm rozumiany jest, jako zachowanie przynoszące korzyść innemu, kosztem aktora (De Waal, 2008, str. 280).

W przypadku zwierząt samotniczych, takich jak pająki, mechanizmy kategoryzacji społecznej są mniej złożone niż u zwierząt stadnych, ale nadal istnieją. Zwierzęta te zdolne są do odróżnienia siebie od innych osobników, zwłaszcza w kontekście rywalizacji o zasoby takie jak pokarm, terytorium lub partnerzy. Wykorzystują sygnały chemiczne (feromony) do rozpoznawania innych osobników i określania ich statusu (np. płeć, stan rozrodczy) oraz do oznaczania swojego terytorium (Napiórkowska, 2017).

W pewnych przypadkach, opisywany mechanizm, pozwala na unikanie konfliktów i optymalizację dostępu do zasobów, takich jak pokarm czy terytorium. Zarówno w sytuacji zorientowanej na indywidualne przetrwanie oraz rywalizację o zasoby, jak i współpracę grupową. W obydwu przypadkach kategoryzacja społeczna stanowi źródło dla przetrwania gatunku, gdyż umożliwia szybkie podejmowanie decyzji w środowisku, co jest kluczowe w kontekście unikania zagrożeń lub wykorzystania okazji. Zatem, postawy oparte na strachu przed obcymi (np. ksenofobia), chronią jednostkę przed konfliktami z przedstawicielami konkurencyjnych grup (Apicella, Norenzayan, Henrich, 2020). Można więc powiedzieć, że prezentowanie postaw takich, jak uprzedzenia jest wynikiem ewolucyjnej tendencji do kategoryzowania świata na swoich (*my*) i obcych (*oni*), co minimalizuje ryzyko współpracy z wrogami (Centola, Baronchelli, 2015; Apicella, Norenzayan, Henrich, 2020). Jak wszystkie postawy, uprzedzenia składają się z trzech komponentów. Poznawczy, obejmuje stereotypy⁹, mity (konstytuujące porządek społeczny, o czym więcej w 3.3.1. *Mity i inne narracje*) oraz zestawy przekonań. Drugi określa afekty, które jednostka odczuwa względem obiektu, jak np. niechęć lub lęk wobec przedstawicieli określonej grupy społecznej. Ostatnim istotnym elementem uprzedzeń jest behavior, określany mianem dyskryminacji, może przyjmować formę unikania przedstawicieli innych grup społecznych, a w skrajnych przypadkach wykazywaniem wobec nich zachowań agresywnych. Przetwarzanie informacji w ramach każdego aspektu postawy cechuje się charakterystycznymi tendencjami, czyli preferencjami określonych schematów myślowych¹⁰.

⁹ Więcej o tym, znajduje się w częściach 1.2.2. *Uprzedzenia, stereotypy, dyskryminacja* oraz 3.3. *Komponent poznawczy*.

¹⁰ Szczegółowe rozwinięcie tematu, dotyczącego uprzedzeń wpływających na interpretacje badań procesów poznawczych zwierząt innych niż ludzie, znajduje się w rozdziale III: *Narracje aukowe konstytuujące porządek społeczny w kontekście relacji człowiek—inne zwierzęta*.

Ludzka natura społeczna jest produktem milionów lat ewolucji, która nagradzała współpracę, ale także utrzymywała uproszczone schematy myślowe. Zrozumienie tych mechanizmów pozwala, zarówno docenić korzyści życia w grupie, jak i krytycznie analizować źródła uprzedzeń czy błędów decyzyjnych. Przyporządkowywanie obiektów do określonych kategorii zachodzi w ramach kryteriów nazywanych stereotypami (Ashmore, Del Boca, 1979; Mackie, 1999). Pozwala to na zmniejszenie warunków niepewności, poprzez wypełnienie deficytów wiedzy, reprezentacjami własności wynikających z przynależności do określonej grupy społecznej. Jak powszechnie wiadomo, taki system podatny jest na błędy poznawcze, wynikające z tendencji ukształtowanych przez mechanizmy ewolucyjne. Utrzymanie wielu licznych przekonań, często sprzecznych ze sobą, możliwe jest dzięki narracją konstytuującym porządek społeczny (Olszak, 2006; Czeremski, 2021). Zatem stereotypy oraz narracje ułatwiają identyfikację z grupą własną, co w konsekwencji prowadzi do wzmocnienia jej spójności. Tendencje przyspieszają podejmowanie decyzji w warunkach zagrożenia, co pozwala na zaoszczędzenie czasu i zasobów umysłowych, lecz w złożonym, współczesnym świecie często prowadzą do irracjonalnych wyborów.

Zarysowane w tej części zagadnienia stanowią punkt wyjścia do dalszej refleksji nad epistemologicznymi i metodologicznymi konsekwencjami przyjmowanych założeń dotyczących umysłowości zwierząt innych niż ludzie. Celem niniejszej dysertacji jest wstępne przeanalizowanie, w jaki sposób prezentowanie uprzedzeń antropomorficznych i antroponegacjonistycznych prowadzi do systematycznych błędów w interpretacji wyników badań nad teorią umysłu u zwierząt pozaludzkich. Jak pokazano, obie postawy – choć pozornie przeciwstawne – mogą w równym stopniu zaciemniać obraz badanych zjawisk, wpływając na projektowanie eksperymentów, ich interpretację oraz zakres formułowanych wniosków. W tym kontekście szczególnego znaczenia nabiera krytyczna analiza samych kategorii używanych do opisu poznania zwierząt, a także uważność na ukryte założenia obecne w języku i perspektywach teoretycznych. Rozpoznanie tych mechanizmów stanowi niezbędny krok w kierunku bardziej adekwatnego ujęcia złożoności zwierzęcego poznania i jego badawczej reprezentacji.

1.2. Podstawowe kwestie terminologiczne

Poznanie społeczne definiowane jest jako zbiór procesów obejmujących rozpoznawanie stanów poznawczych i emocjonalnych innych jednostek, rozumienie ich

działań oraz refleksję nad nimi. Czyli można powiedzieć, że omawiany zbiór zajmuje się kodowaniem, przechowywaniem, wyszukiwaniem i przetwarzaniem informacji o życiu społecznym jednostki i innych ją otaczających osób (Nowogrodzka, 2014). Obejmują one percepcję społeczną, empatię, teorię umysłu, przyjmowanie perspektywy innego i wiele innych (Eisenberg i in., 1989; Hoffman, 1996; Eisenberg Wentzel, Harris, 2001; Nowogrodzka, 2014; Baron-Cohen, 2009; Benslah, Caillies, Anduze, 2016). Choć szczegółowe omówienie wszystkich jej aspektów wykracza poza ramy tematyczne niniejszego opracowania, uważam za konieczne, krótkie opisanie wybranych zdolności.

1.2.1. Zdolności z zakresu poznania społecznego

Poznanie społeczne definiowane jest jako zbiór procesów obejmujących rozpoznawanie stanów poznawczych i emocjonalnych innych jednostek, rozumienie ich działań oraz refleksję nad nimi. Czyli można powiedzieć, że omawiany zbiór zajmuje się kodowaniem, przechowywaniem, wyszukiwaniem i przetwarzaniem informacji o życiu społecznym jednostki i innych ją otaczających osób (Floridi, 2010, 2011; Nowogrodzka, 2014). Obejmuje ono m.in. percepcję społeczną, empatię, teorię umysłu, przyjmowanie perspektywy innego (Eisenberg i in., 1989; Hoffman, 1996; Eisenberg Wentzel, Harris, 2001; Nowogrodzka, 2014; Baron-Cohen, 2009; Benslah, Caillies, Anduze, 2016). Choć szczegółowe omówienie wszystkich jej aspektów wykracza poza ramy tematyczne niniejszego opracowania, uważam za konieczne, krótkie opisanie wybranych zdolności z tego zakresu, nazywane w psychologii rozwojowej *mindreadingiem*, czyli czytaniem umysłów innych (Baron-Cohen, 2009). W dalszej części pracy będę posługiwać się przede wszystkim terminem anglojęzycznym, ponieważ jest on szeroko rozpowszechniony w literaturze przedmiotu, a jego stosowanie sprzyja precyzji i spójności terminologicznej. W niektórych miejscach pojawi się jednak również polskie sformułowanie — czytanie umysłów innych — traktowane synonimicznie, czyli używane zamiennie, głównie ze względów stylistycznych.

W rozważaniach dotyczących moralności i sprawczości moralnej (*moral agency*) można dostrzec dużą różnorodność definicji kluczowych pojęć (empatia, teoria umysłu, przyjmowanie perspektywy innego), co prowadzi do nieścisłości terminologicznych (Aaltola, 2013). Niektóre definicje funkcjonujące w obszarze jednego terminu, wzajemnie się wykluczają, co wskazuje na konieczność precyzyjnego

i interdyscyplinarnego podejścia do analizowanych zjawisk. W bieżącej części szczególną uwagę poświęcam rozróżnieniu zdolności do przyjmowania perspektywy innego, teorii umysłu oraz empatii, gdyż zrozumienie różnic we wskazanych zjawiskach jest kluczowe dla poniższych rozważań. Warto jednak podkreślić, że wszystkie wskazane procesy tworzą zintegrowany system, funkcjonujący jako całość, a nie jako zbiór oddzielnych mechanizmów. Historia rozważań z zakresu poznania społecznego może być rozpatrywana w aż trzech perspektywach badawczych (Aaltola, 2013). Jednak definicje zawarte w niniejszym opracowaniu wywodzą się z tradycji, nazywanej Hume’owską (Eisenberg i współpracownicy, 1989; Hoffman, 1996; Wentzel, Harris, 2001; Baron-Cohen, 2009; Benslah, Caillies, Anduze, 2016). David Hume (2007) uważał, że to emocje, a nie rozum, stanowią fundament moralności, a kluczowym pojęciem w jego filozofii jest sympatia, która pełni rolę wczesnego odpowiednika współczesnej koncepcji empatii. Hume definiował ją jako naturalną zdolność człowieka do „odzwierciedlania” emocji innych osób. Według autora *Traktat o naturze ludzkiej* to właśnie sympatia umożliwia tworzenie więzi społecznych i stanowi podstawę naszych moralnych ocen (Hume, 2007, Księga II). Zachowania oceniane jako dobre lub złe wzbudzają emocje, takie jak współczucie, podziw czy oburzenie, a te kształtują moralne wybory. W swojej filozofii określa dwa komponenty składające się na sympatię: idee (*ideas*), które odpowiadają za myśli i przekonania, oraz wrażenia (*impressions*), które określają afektywny wymiar omawianego zjawiska (Hume, 2007, Księga I). To właśnie ten podział stał się inspiracją dla późniejszych teorii empatii (Prinz 2007).

Kluczową teorią zdolności z zakresu poznania społecznego, dla prowadzonych w niniejszej pracy rozważań jest koncepcja rozwijana przez Nancy Eisenberg i współpracowników (1989; 1998, 2001a, 2001b). Stanowi ona kompleksowe ujęcie empatii, które oddziela ją od innych zdolności z zakresu poznania społecznego. Jej analiza pozwala na określenie roli wskazanego mechanizmu w zachowaniach prospołecznych i rozwoju moralnym. W tym modelu empatia definiowana jest jako reakcja afektywna na sytuację innej osoby. Jest ona bardzo podobna do tego, co druga osoba czuje lub tego, co jednostka oczekuje, że obserwowana osoba będzie czuła. W swojej teorii Eisenberg wyróżnia dwa kluczowe stany afektywne związane z omawianym zjawiskiem: sympatię oraz dystres osobisty (*personal distress*). Pierwsza z wymienionych ma charakter pozytywny, dlatego afekt ten sprzyja zachowaniom prospołecznym. Natomiast druga rozumiana jest jako negatywne napięcie, dające się załagodzić tylko na dwa sposoby. Jedną ze strategii jest reakcja odpowiednią emocją na

wykryty stan innego (zarażenie emocjonalne¹¹) oraz wykonanie działania łagodzącego jego cierpienie, niwelując tym samym własny dystres. Druga polega na uniknięciu bodźca, wywołującego napięcie za pośrednictwem tendencji do unikania interakcji z osobą, która jest źródłem takiego pobudzenia. Co ważne, im wyższy poziom sympatii, tym większa szansa na to, że jednostka podejmie się zachowania prospołecznego, oraz analogicznie, im silniejszy dystres osobisty, tym większa szansa, że jednostka wykona zachowanie egoistyczne (unikające).

Przyjmowanie perspektywy innego definiowane jest jako specyficzna¹² zdolność do świadomego wyobrażenia sobie, jak wygląda świat z punktu widzenia innej osoby, co obejmuje zarówno jej emocje, jak i myśli (Eisenberg, Zhou, Koller, 2001a; Eisenberg, Morris, 2001b; Bensalah, Caillies, Anduze 2016). Jest to proces poznawczy, który pozwala jednostce na lepsze zrozumienie sytuacji, w której znajduje się ktoś inny i przewidywanie jego reakcji. Omawiane zjawisko wymaga zaangażowania wielu zdolności poznawczych, co odróżnia ją od automatycznych, afektywnych reakcji empatycznych. Dlatego trudniejsza do uchwycenia jest różnica między przyjmowaniem perspektywy innego, a teorią umysłu. Szczegółowy opis koncepcji istotnych w kontekście najbardziej interesującej mnie zdolności znajduje się w rozdziale *IV: Kluczowe koncepcje teorii umysłu*. Na ten moment zaznaczę jedynie, że w niniejszym opracowaniu przyjmuję, dwie definicje tego zjawiska. Pierwsza opisuje teorię umysłu w sensie szerokim, głosząc, że jest to zdolność do atrybucji stanów mentalnych innym (Premarck, Woodruff, 1978). Druga wprowadzona przez Michela Tomasello i Josepha Calla (2011) zawęża jej zakres do wnioskowania na podstawie fałszywych przekonań innego. Żadna z wybranych definicji, nie stoi w sprzeczności z opisaną wcześniej koncepcją Eisenberg, gdzie teoria umysłu rozumiana jest jako bardziej ogólna¹³

¹¹ Zarażenie emocjonalne (*emotional contagion*) to podstawowa zdolność z zakresu poznania społecznego, polegająca na automatycznym „przechwytywaniu” emocji innych osób. Proces ten ma charakter afektywny, w którym jednostka mimowolnie przejmując emocje obserwowane u innych. Mechanizm ten umożliwia natychmiastową reakcję na stany emocjonalne innych, tworząc podstawę do bardziej złożonych zdolności empatycznych (Eisenberg, Wentzel, Harris, 1998). Jednym z podstawowych mechanizmów omawianego zjawiska jest *rapid mimicry*, czyli szybka (do 1s) reakcja mimiczna taka jak odpowiedź uśmiechem na uśmiech innej osoby (Palagi, Nicotra, Cordoni, 2015).

¹² Specyficzna, ponieważ odnosi się do konkretnego procesu aktywnego wyobrażania sobie, jak świat jest odczuwany, czy inaczej przeżywany poznawczo i afektywnie przez daną jednostkę.

¹³ Obejmuje szeroki zakres procesów lub umiejętności, przez co jest mniej specyficzna w swoim zastosowaniu.

zdolność umożliwiająca jednostce rozumienie, że inni ludzie mają własne myśli, przekonania, intencje, pragnienia i emocje, które mogą różnić się od jej własnych (Hoffman, 1996; Eisenberg, Wentzel, Harris, 1998; Eisenberg, Zhou, Koller, 2001a; Eisenberg, Morris, 2001b). Pozwala na przewidywanie i wyjaśnianie zachowania innych na podstawie ich stanów umysłowych. W odróżnieniu od przyjmowania perspektywy innego, teoria umysłu dotyczy reprezentowania stanów mentalnych innych osób w sposób bardziej uniwersalny, obejmując zarówno intencje, motywacje, przekonania jak i emocje. Eisenberg traktuje teorię umysłu jako kluczowy element zrozumienia relacji społecznych, który wspiera rozwój empatii i zachowań prospołecznych, ale zwraca uwagę, że sama zdolność do odczytywania stanów umysłowych nie zawsze prowadzi do działań empatycznych — w przeciwieństwie do przyjmowania perspektywy innego, które jest ściśle powiązane z motywacją prospołeczną (Eisenberg, Wentzel, Harris, 1998; Eisenberg, Zhou, Koller, 2001a; Eisenberg, Morris, 2001b).

Wskazane podejścia do zdolności z zakresu poznania społecznego, zgodne jest z fundamentalną zasadą Tinbergena (1963), głoszącą, że należy oddzielać konsekwencje zachowania od jego motywacji (Mayr, 1961; Tinbergen, 1963; De Waal, 2008). Z ewolucyjnego punktu widzenia, doborowi naturalnemu podlegają konsekwencje zachowania, nie zaś leżące u jego podstaw motywacje¹⁴. Ideę tę dobrze obrazuje przykład zachowania altruistycznego u pszczoły miodnej. Dotychczasowe ustalenia podsumowuje tabela 2.

¹⁴ Jest to także zgodne, z opisanym w części 3.3. *Komponent poznawczy* mechanizmem ewolucji mitów (Czeremski, 2021).

Tabela 2. Porównanie teorii umysłu i przyjmowania perspektywy innego

Teoria umysłu (ToM)	Przyjmowanie perspektywy innego
Jest ogólną zdolnością poznawczą.	Jest specyficznym procesem poznawczym.
Dotyczy rozumienia stanów umysłowych w szerokim zakresie (myśli, przekonania, intencje, emocje).	Koncentruje się na aktywnym wyobrażaniu sobie perspektywy innej osoby. Obejmuje zarówno rozpoznanie stanów mentalnych innego, jak i współodczuwanie z nim.
Może występować bez empatii (np. zrozumienie, że ktoś ma inne przekonanie, bez współodczuwania).	Angażuje empatię, teorię umysłu oraz wiele innych.

[Opracowanie własne (Hoffman, 1996; Eisenberg, Wentzel, Harris, 1998; Eisenberg, Zhou, Koller, 2001a; Eisenberg, Morris, 2001b; Bensalah, Caillies, Anduze 2016)]

Zachowanie altruistyczne, przynosi korzyść innemu, kosztem aktora, jego podręcznikowym przykładem jest pszczoła miodna, która żądli intruza, tracąc przy tym własne życie (de Waal, 2008). Chociaż jego funkcją adaptacyjną jest obrona ula, motywacją do takiego zachowania jest uzyskanie korzyści osobistych (przetrwanie własnych genów — przede wszystkim królowej), motywacja zwierzęcia ma charakter agresywny. Takie zachowanie altruistyczne nazywane jest samolubnym. (Mysterud, 1999). W swojej przełomowej pracy *On aims and methods of ethology* (1963), Tinbergen wyznaczył ramy teoretyczne dla analizy z zakresu biologii ewolucyjnej, wyróżniając poziomy wyjaśniania zachowania zwierząt. Sformułował także cztery pytania, które pogrupować można według dwóch różnych rodzajów wyjaśnień: ostatecznych (*ultimate*)¹⁵ oraz bezpośrednich (*proximate*)¹⁶. Zatem, w etologii wyróżnia

¹⁵ Rozumiana jako korzyść, którą organizm lub jego bliscy czerpią z zachowania. Stanowi ona prawdopodobny powód, przez który było ono faworyzowane na drodze doboru naturalnego. (de Waal, 2008, s 278).

¹⁶ Rozumiana jako sytuacja, która wyzwała zachowanie i mechanizm fizjologiczny, neuronalny czy poznawczy, który je umożliwia (de Waal, 2008, s 278).

się pojęcie przyczyny ostatecznej (*ultimate cause or goal*), które odnosi się do adaptacyjnych i ewolucyjnych źródeł danej cechy, w tym wypadku zachowania, uwzględniając korzyści, jakie przynosi, oraz bezpośredniej (*proximate cause*) które dotyczy mechanizmów na drodze, których wywoływane oraz regulowane jest konkretne zachowanie. Mówi się zatem, że wyjaśnienia *ultimate* odpowiadają na pytania: *dlaczego?*, natomiast *proximate*: *jak?*. Rozróżnienie to podkreśla konieczność niezależnego rozpatrywania przyczyn i efektów zachowania (Mayr, 1961; Tinbergen, 1963; Triversen, 2002). Tinbergen podkreśla, że motywacja powinna być badana na poziomie mechanizmów bezpośrednich (mechanizm i rozwój), a nie mylona z jej funkcją ewolucyjną. Ustalenia te obrazuje tabela 3.

Tabela 3. Cztery poziomy wyjaśniania zachowania zwierząt

Ultimate (dlaczego?)	Proximate (jak?)
<u>Ontogeneza:</u> Jak przebiega rozwój zachowania w ciągu życia?	<u>Mechanizm:</u> Jakie mechanizmy fizjologiczne, neurobiologiczne i psychiczne leżą u podstaw danego zachowania?
<u>Ewolucja:</u> Jak to zachowanie wyewoluowało w linii danego gatunku?	<u>Funkcja:</u> Jakie korzyści adaptacyjne przynosi to zachowanie?

[Opracowanie własne (Tinbergen, 1963; de Waal, 2008)]

1.2.2. Uprzedzenia, stereotypy, dyskryminacja

Łączenie się w grupy jest fundamentalną cechą ludzkiej natury, stanowiącą konsekwencje strategii społecznej charakterystycznej dla zwierząt stadnych. Z ewolucyjnego punktu widzenia przynależność do społeczności ułatwia przetrwanie (np. ochronę przed zagrożeniami oraz wychowanie potomstwa). Konsekwencją takiej formy organizacji społecznej, jest powstanie kategorii poznawczych takich jak np. *grupa własna* (my) — *grupa obca* (oni). Natomiast skłonność do identyfikowania się z różnymi grupami wpływa na sposób, w jaki ludzie postrzegają siebie i otaczający ich świat (Tajfel, 1982). Co ważne, badania pokazują, że przynależność do różnych

społeczności określa tożsamość jednostki oraz stanowi istotny czynnik kształtujący samoocenę (Nelson, 2003, str. 22). Jak zauważa Gordon Allport (1954), uprzedzenia wynikają z potrzeby organizowania świata w kategorie oraz z potrzeby potwierdzania własnej wartości i pozycji w grupie (Allport, 1954). W swojej książce *Błądzą wszyscy (ale nie ja). Dlaczego usprawiedliwiamy głupie poglądy, złe decyzje i szkodliwe działania?* Elliot Aronson i Carol Travis (2007) zaznaczają, że omawiane postawy pozwalają nam zachować status, władze i przywileje. Todd Nelson (2003) w klasycznym już podręczniku dotyczącym uprzedzeń wskazuje, nawet, że stanowiły one przyczynę wszystkich wojen i bitw w historii ludzkości. Uważam, że takie twierdzenie może być zbyt generalizujące, dlatego proponuję przyjąć bardziej ostrożne podejście, a mianowicie, że omawiany rodzaj postaw stanowi podstawę aktów grupowej przemocy w przynajmniej większości wypadków.

W literaturze znajdziemy kilka nieścisłości terminologicznych związanych z pojęciami uprzedzeń, stereotypów, afektów i dyskryminacji. Najwcześniejsze definicje interesujących mnie postaw, kładły nacisk na podłoże afektywne (przede wszystkim negatywne) tego zjawiska, co często prowadziło do jego utożsamienia z negatywnymi emocjami (Nelson, 2003). Nawet Allport w książce *The Nature of Prejudice* definiuje je jako „antypatie wynikającą z błędnego lub nieelastycznego uogólnienia” (1954, str. 9). Podejście to skutkowało powstaniem licznych niejednoznacznych definicji, które w efekcie utrudniły prowadzenie badań nad opisywanym zjawiskiem (Nelson, 2003). Jak uważa Nelson (2003, str. 33) to właśnie problemy z pomiarem, doprowadziły do przyjęcia, że struktura uprzedzeń zgodna jest z tradycyjnym ujęciem postaw, zakładającym istnienie trzech składowych komponentów:

- a. Komponent poznawczy: opisuje przekonania, narracje i stereotypy jednostki, które biorą udział w kształtowaniu postaw wobec konkretnej grupy czy osoby.
- b. Komponent afektywny: odnosi się do typu emocji związanych z określoną postawą (wobec danego obiektu) oraz poziomowi ich intensywności.
- c. Komponent behawioralny: opisuje działania podejmowane przez osobę uprzedzoną. Są one reakcją na przeżywane emocje oraz stanowi widoczny efekt prezentowania konkretnych postaw (dyskryminacja).

Pomimo porzucenia praktyki utożsamiania uprzedzeń z negatywnym afektem, współczesne opracowania nadal traktują przede wszystkim o negatywnych

uprzedzeniach, definiowanych jako „tendencja do wykazywania wrogości lub negatywnych opinii wobec osób należących do pewnych grup społecznych, wyłącznie ze względu na fakt, że należą do tych grup” (Zakrzewska 2015, str. 1). Takie podejście jest zrozumiałe, warto jednak zauważyć, że pozytywne uprzedzenia, choć mniej dostrzegalne, również mogą prowadzić do niesprawiedliwości. Przekonania takie jak *Azjaci są świetni z matematyki* czy *kobiety są naturalnie empatyczne* wydają się nieszkodliwe, jednak w rzeczywistości, także mogą być krzywdzące dla jednostek, których cechy nie są zgodne ze stereotypem. Oczekiwania wynikające z takich przekonań mogą ograniczać możliwości rozwoju tych osób lub prowadzić do nieuzasadnionych wymagań względem nich. Co więcej, przekonanie o cudzych atrybutach wiąże się z oczekiwaniami, jakie możemy mieć wobec dostępnego poznawczo obiektu, co może skutkować uprzedzeniami utrudniającymi obiektywną ocenę jednostki i jej rzeczywistych kompetencji. W tym miejscu warto podkreślić, że uprzedzenia są związane z przynależnością do danej kategorii społecznej, a nie z jakąkolwiek cechą obiektu. Najczęściej oparte są na wyobrazeniowych własnościach osób przynależących do danej grupy. Co więcej, w przypadku uprzedzeń obserwuje się unifikację przekonań o całej grupie, przy jednoczesnym pominięciu indywidualnych cech jednostki (Zakrzewska, 2015).

Kolejnym pojęciem, które wymaga omówienia na wstępie, jest stereotyp. Jego użycie w literaturze naukowej bywa problematyczne ze względu na liczne różnice w definicjach przyjmowanych przez badaczy. Z tego względu zagadnienie to zostanie szczegółowo rozwinięte w części 3.3.2. *Stereotypy*, jednak w tym miejscu, dla uporządkowania dalszych rozważań, przyjmuję, że stereotypy stanowią kryteria przyporządkowywania obiektów dostępnych poznawczo do kategorii mentalnych (Mackie, 1999). Uznaję także, że stanowisko w ramach którego postrzeganie stereotypów, jako zespołów przekonań na temat atrybutów pewnej grupy społecznej, nie jest konkurencyjne, a kompatybilne z przyjętą przeze mnie definicją (Ashmore, Del Boca, 1979). W obydwu stanowiskach kluczowy jest proces przypisywania jednostek do określonych kategorii na podstawie cech, postrzeganych jako typowe dla jej przedstawicieli. Proces ten nazywany jest stereotypizacją. Mechanizmy leżące u ich podstaw wskażę szerzej w części 2.1. *Mechanizm kategoryzacji* oraz 3.3.2. *Stereotypy*.

1.2.3. Klasyfikacja uprzedzeń

Istnieje wiele różnych klasyfikacji uprzedzeń, jednak w kontekście niniejszej dysertacji istotne będą dwie: uprzedzenia indywidualne — kulturowe oraz jawne — ukryte (Simpsona, Yingera, 1985, Nelson, 2003, Zakrzewska, 2015). Ze względu na zakres tematyczny dysertacji mniej uwagi poświęcę uprzedzeniom jawnym oraz indywidualnym. Należy jednak zaznaczyć, że nie znaczy to, iż ich wpływ na interpretację badań z udziałem zwierząt innych niż ludzie nie jest znaczący. Należy pamiętać, że w perspektywie dymensjonalnej¹⁷ trudno rozróżnić pewne elementy poznawcze, gdyż stanowią one zintegrowany system i nigdy nie działają autonomicznie. Tak samo, w przypadku dwóch wybranych klasyfikacji, zarówno uprzedzenia jawne, jak i ukryte mogą mieć charakter indywidualny lub kulturowy. W tym sensie koncepcje te są kompatybilne, nie zaś konkurencyjne. Co więcej, chociaż opisane dalej rodzaje uprzedzeń ukrytych, subtelnych oraz automatycznych różnią się od siebie, ich synteza pozwala na lepsze zrozumienie złożoności omawianych postaw. Uważam, że taka perspektywa pozwoli na szczegółową analizę tendencyjności, prowadzącą do określenia, preferencji występujących w procesie formułowania interpretacji eksperymentów z udziałem zwierząt innych niż ludzie ZINL. Dlatego w niniejszej dysertacji używając każdego ze wskazanych określeń, będę miała na myśli syntezę trzech wybranych koncepcji (teorie uprzedzeń ukrytych, subtelnych oraz automatycznych).

¹⁷ Carl Hempel w swojej pracy z 1991 roku zauważa, że większość typologii naukowych rozpoczyna się od klasyfikacji kategoryjnej, ale każda z nich dąży do dymensjonalnej, pozwalającej na bardziej złożone opisywanie zjawisk. Pierwsza z wymienionych dąży do stworzenia klas zdefiniowanych rozłącznie. To znaczy, że np. mechanizmy poznawcze, zdolności poznawcze, zaburzenia psychiczne, tworzą odrębne klasy zjawisk, które nie nakładają się na siebie. Warto tutaj zaznaczyć, że ze względu na swoje praktyczne i teoretyczne zalety typologia ta jest dobrze ugruntowana w naukach medycznych – także w psychiatrii (Modrzejewska, 2011). Jednak podejście to obecne jest przede wszystkim w starszych kryteriach diagnostycznych stosowanych np. w psychologii tzn. w DSM (1-4) oraz ICD (1-10), gdzie zaburzenia psychiczne traktowane były jako kategorie objawów. Natomiast, w typologii dymensjonalnej tworzy się pewien model określonego fragmentu życia psychicznego, do którego realne przypadki jedynie przystają. Typy nie występują rozłącznie, ale współwystępują ze sobą wykazując cechy podobne, dlatego trudno je czasem rozróżnić. Perspektywa ta przyjmowana jest w najnowszych kryteriach diagnostycznych takich jak DSM – 5 oraz ICD – 11, gdzie obserwujemy wyraźną zmianę, w ramach której typy osobowości traktowane są raczej jako kontinuum (spektrum). Dlatego w niniejszej dysertacji wszystkie zdolności, mechanizmy i struktury poznawcze traktowane są w perspektywie dymensjonalnej, przyjmowanej w psychologii od lat 80 XX wieku (Hempel, 1991).

Klasyfikacja uprzedzeń międzygrupowych określana jest w literaturze na kilka różnych sposobów: jawne — ukryte (*explicit and implicit*), jawne — subtelne (*blatant and subtle*) oraz automatyczne — kontrolowane (*automatic processes and controlled processes*); (Pettigrew, Meertens, 1995, Devine, 1989, Nelson, 2003). Choć wskazane podejścia różnią się od siebie, w kontekście niniejszej dysertacji istotne jest, że pewne ich twierdzenia uzupełniają się nawzajem, dlatego ich synteza stanowi ważny punkt wyjścia dla poniższych rozważań. Zaskoczeniem może być zaliczenie do tego grona teorii komponentów automatycznych i kontrolowanych. Patricia Devine (1989) podchodzi krytycznie do analizy uprzedzeń jako postaw (Fiske, 1994). Wydaje się, że z łatwością można zaimplementować jej teorię do przyjętej wcześniej definicji uprzedzeń, uznając, że proponowane przez nią komponenty: automatyczny — kontrolowany są analogiczne do stworzonej przez Anthony'ego Greenwalda i Mahzaira Banajiego (1995), klasyfikacji jawne — ukryte (*explicit—implicit*); (Nelson, 2003).

W artykule *Implicit social cognition: Attitudes, self-esteem, and stereotypes*, opublikowanym w czasopiśmie *Psychological Review* w 1995 roku Greenwald i Banaji opisują nieświadome procesy leżące u podstaw zachowań społecznych, tym samym wprowadzając jedną z kluczowych koncepcji uprzedzeń w psychologii społecznej. Autorzy przedstawiają dwupoziomowy model poznania społecznego, który wyróżnia postawy jawne i ukryte. U podstaw pierwszego z wymienionych, leżą procesy świadome, a co za tym idzie, refleksyjne i możliwe do wyrażenia wprost. Natomiast za przetwarzanie na poziomie drugim odpowiadają procesy automatyczne, co oznacza, że zachodzą bez świadomej kontroli, dzięki czemu mogą pozostawać w sprzeczności z deklarowanymi przekonaniami. Ich aktywacja zachodzi w reakcji na bodziec w postaci np. osoby z grupy obcej, co może wpłynąć na decyzje i działania jednostki, nawet jeśli świadomie wyznaje ona inne wartości. Greenwald i Banaji (1995) omawiają również koncepcję ukrytej samooceny, czyli nieświadomych ocen siebie, które mogą różnić się od tej deklarowanej. Zwracają uwagę na znaczenie tych procesów w kontekście relacji społecznych, decyzji podejmowanych przez jednostkę, szczególnie w sytuacjach wymagających szybkiego działania, oceny przy ograniczonych informacjach lub stanie obniżonej kondycji poznawczej. Koncepcja ta wspierana jest także przez stanowisko Thomasa Pettigrew i Roela Meertensa (1995). Chociaż w swoim artykule *Subtle and blatant prejudice in Western Europe* Pettigrew i Meertens nie odnosili się bezpośrednio do koncepcji Greenwalda i Banaji'ego, ich prace można uznać za uzupełniające. Obie eksplorują różne formy uprzedzeń, przy czym Greenwald

i Banaji koncentrują się na procesach poznawczych o charakterze automatycznym, i kontrolowanym, podczas gdy Pettigrew i Meertens badają różnice w sposobie manifestowania się uprzedzeń w zachowaniach i postawach społecznych. Można by zatem uznać, że subtelne uprzedzenia opisane przez Pettigrew i Meertensa mają pewne cechy wspólne z ukrytymi Greenwalda i Banaji'ego, choć nie są to tożsame koncepcje.

Jak już wspomniałam, autorzy *Subtle and blatant prejudice in Western Europe* (1995) koncentrują się na sposobach, w jaki omawiany rodzaj postawy manifestuje się w zachowaniu i myśleniu. Uprzedzenia jawne charakteryzują się silnym poczuciem zagrożenia wobec przedstawicieli obcych społeczności, a co za tym idzie mocnym oporem przed kontaktami z przedstawicielami tych grup. Natomiast te ukryte charakteryzują się zaprzeczaniem przeżywanym emocjom, a ich konsekwencją jest silna tendencja do obrony tradycyjnych wartości oraz podkreślanie różnic międzygrupowych (Pettigrew, Meertens, 1995; Zakrzewska, 2015, str. 79). Jednym z kluczowych mechanizmów napędzających ukryte uprzedzenia jest silna identyfikacja z grupą własną – realną lub wyobrażeniową (Zakrzewska, 2015, str. 79). Taka identyfikacja prowadzi do tworzenia wyraźnych granic między *nami* a *nimi*, co wzmacnia podziały i utrwała różnice międzygrupowe. Obrona tradycyjnych wartości, często podkreślana jest w retoryce osób przejawiających ukryte uprzedzenia, stanowiąc narzędzie legitymizujące te postawy oraz nadając im pozór neutralności lub wręcz moralnego uzasadnienia. Warto podkreślić, że badania Pettigrew i Meertensa miały istotny wpływ na rozwój narzędzi badawczych, takich jak kwestionariusze mierzące subtelne uprzedzenia, które pozwalają na dokładniejszą diagnozę tych postaw w różnych kontekstach społecznych. Zrozumienie tego rozróżnienia ma kluczowe znaczenie dla projektowania interwencji mających na celu redukcję dyskryminacji oraz promowanie tolerancji i integracji międzygrupowej.

Ukryte uprzedzenia wobec ZINL mogą przejawiać się w postawie antropocentrycznej charakteryzującej się stereotypem, w ramach którego zdolności poznawcze takie jak np. teoria umysłu, są unikalne dla człowieka i stanowią punkt odniesienia do oceny innych gatunków. Podejście to może prowadzić do tendencyjności w projektowaniu eksperymentów, w postaci np. nieuwzględniania sposobu poznawania świata zwierząt biorących udział w badaniu. W efekcie ignorowane mogą być przejawy inteligencji zwierząt pozaludzkich, które różnią się od tych występujących u ludzi, ale są dostosowane do specyficznych wymagań ekologicznych danego gatunku. Co więcej, subtelne uprzedzenia mogą wpływać na wybór języka, a co za tym idzie, na dobór ram

interpretacyjnych używanych w publikacjach naukowych. Opisywanie zachowań zwierząt jako „prymitywnych” czy „prostszych” w porównaniu do ludzi może sugerować ich niższość, co utrwała mity o ograniczonych zdolnościach poznawczych innych gatunków. Procesy automatyczne mogą prowadzić do ignorowania danych, które wskazują na większą złożoność poznawczą zwierząt pozaludzkich, takich jak ich zdolność do komunikacji, empatii czy planowania. W kontekście badań nad teorią umysłu opisywane postawy mogą dodatkowo wzmacniać struktury poznawcze świadczące o tym, że tylko ludzie są zdolni do przyjęcia perspektywy innej istoty, uwzględniając jej fałszywe przekonania. Badania pokazują jednak, że wiele gatunków, w tym naczelne, delfiny czy ptaki, wykazuje zachowania – o czym będzie mowa w dalszych częściach rozważań – sugerujące istnienie podstawowych mechanizmów *mindreadingu*, choć mogą one różnić się od ludzkich pod względem formy, a także zakresu. Ignorowanie takich różnic może prowadzić do tendencyjnej interpretacji danych i niedoszacowania potencjału poznawczego ZINL¹⁸.

Klasyfikacja indywidualne — kulturowe zapoczątkowana została przez George'a Simpsona i Milтона Yingera w książce *Racial and Cultural Minorities: An Analysis of Prejudice and Discrimination* (1985). Pierwsze z wymienionych są specyficzne dla jednostki i wynikają z osobistych doświadczeń, które kształtują jej przekonania, emocje oraz zachowanie wobec przedstawicieli innych społeczności. Mogą być efektem bezpośrednich interakcji z przedstawicielami określonych grup (zarówno pozytywnych, jak i negatywnych) lub bazować na subiektywnych interpretacjach takich kontaktów. Postawy te formowane są także przez informacje i przekonania podzielane przez przedstawicieli grup własnych – czyli tych, z którymi jednostka się identyfikuje. Uprzedzenia indywidualne mogą przybierać formę specyficznych przekonań, emocji oraz zachowań, które niekoniecznie są reprezentatywne dla wszystkich innych z danej wspólnoty, do których należy jednostka. Co istotne, są one podatne na zmianę w wyniku nowych doświadczeń lub świadomego wysiłku, w celu zmiany zachowania. Natomiast uprzedzenia kulturowe mają charakter społeczny i kolektywny. Odzwierciedlają szeroko rozpowszechnione normy, wartości, przekonania oraz tradycje, które są przyjmowane i podtrzymywane w ramach danej kultury, społeczności lub narodu. Uprzedzenia te są zakorzenione w systemach społecznych i instytucjonalnych, takich jak system edukacji, media, religia czy prawo, co sprawia, że ich oddziaływanie

¹⁸ Teza zostanie szerzej omówiona w części 3.4. *Behawioralny komponent uprzedzeń*.

jest trwalsze i trudniejsze do zmiany. Przekazywane są z pokolenia na pokolenie, co prowadzi do ich utrwalania w świadomości społecznej. Upředzenia kulturowe mogą wpływać na kształtowanie stereotypów dotyczących grup mniejszościowych, prowadząc do ich marginalizacji, wykluczenia społecznego lub upředzeń instytucjonalnych. Przykładami takich mechanizmów są m.in. dyskryminacja w dostępie do edukacji, pracy czy opieki zdrowotnej, a także obecność reprezentacji przedstawicieli mniejszości w kulturze masowej. Ważnym aspektem analizy upředzeń indywidualnych i kulturowych jest ich wzajemna zależność. Drugie z wymienionych mogą wzmacniać i legitymizować pierwsze, dostarczając jednostce społecznego uzasadnienia dla jej postaw. Jednocześnie upředzenia indywidualne, szczególnie w przypadku osób zajmujących wpływowe pozycje w społeczeństwie, mogą przyczyniać się do wzmacniania tych kulturowych, np. poprzez wpływ na kształtowanie polityki, mediów czy systemów edukacji.

1.3. Wpływ upředzeń międzygrupowych na interpretacje badań nad poznaniem ZINL

Upředzenia to zintegrowany system trzech komponentów, których nie należy rozpatrywać odrębnie. W niniejszej dysertacji przedmiotem zainteresowania są interpretacje badań dotyczących *mindreadingu* u zwierząt pozaludzkich. Stanowią one formę narracji naukowych, które odzwierciedlają zarówno indywidualne, jak i kulturowo utrwalone stereotypy. Konstytuują tym samym porządek społeczny, a co za tym idzie, determinują sposób, formułowania hipotez, konstruowanie procedur eksperymentalnych oraz interpretowanie danych empirycznych. Należy zauważyć, że wiele klasycznych teorii postaw społecznych – jak teoria dominacji społecznej, teoria tożsamości społecznej czy model upředzeń ukrytych – może być z powodzeniem zaadaptowany do analizy relacji człowiek – inne zwierzęta (Tajfel, 1982; 1984; Greenwald, Banaji, 1995, Pratto, Sidanius, Stallworth, Malle, 1994). Choć pierwotnie koncentrowały się one na relacjach międzyludzkich, ich struktura pozwala na analizę procesów kategoryalnych, tożsamościowych i afektywnych, które równie dobrze funkcjonują w odniesieniu do granicy gatunkowej.

W świetle powyższych założeń, w dalszej części pracy skoncentruję się przede wszystkim na poznawczych i behawioralnych komponentach postaw, których szczegółowa analiza zostanie przedstawiona w rozdziale *III: Narracje naukowe*

konstituujące porządek społeczny w kontekście relacji człowiek — inne zwierzęta. Choć afektywny wymiar uprzedzeń – obejmujący emocje i reakcje uczuciowe wobec zwierząt innych niż ludzie – stanowi integralną część systemu postaw i odgrywa istotną rolę w ich kształtowaniu, nie zostanie on omówiony w odrębnej sekcji. Nie oznacza to jednak jego pominięcia: wskazany komponent będzie obecny w dalszej analizie, zwłaszcza tam, gdzie emocje okazują się kluczowe dla omawianych procesów leżących u podstaw *mindreadingu*. Taka decyzja strukturalna podyktowana jest potrzebą zachowania spójności wywodu oraz koncentracją na tych aspektach, które najpełniej ilustrują wpływ uprzedzeń na procesy badawcze w obrębie relacji międzygatunkowych.

Przykładem istotnego, w kontekście przetwarzania postaw i przekonań, aspektu komponentu afektywnego jest mechanizm autowaloryzacji¹⁹. Jak zauważa Wojciszke (2009), jest on jednym z najpotężniejszych pozabiologicznych motywatorów, głęboko zakorzenionym w mechanizmach samooceny i przynależności grupowej. W obszarze relacji międzygrupowych autowaloryzacja przejawia się m.in. poprzez faworyzację grupy własnej, realizowaną kosztem deprecjonowania grup obcych. W przypadku zwierząt innych niż ludzie, postrzeganych jako grupa obca, skutkuje to systematycznym umniejszaniem ich możliwością poznawczym — nie ze względu na brak empirycznych dowodów, ale dlatego, że ich uznanie mogłoby zagrozić symbolicznemu statusowi człowieka. Tendencja ta znajduje wyraz w efekcie potwierdzenia²⁰, który – jak wskazuje Nickerson (1998, str. 205) – może być interpretowany nie tylko, jako ograniczenie systemów poznawczych, ale także, przejaw motywacyjnych potrzeb jednostki. Aspekt ten nabiera szczególnego znaczenia w badaniach nad teorią umysłu u zwierząt pozaludzkich, gdzie sukces poznawczy bywa warunkowany zgodnością zachowań zwierząt z oczekiwanym, ludzkim wzorcem²¹. W nurcie behawioralnym²² często obserwuje się tendencję do interpretowania zachowań zwierząt przez pryzmat *ludzkiej wyjątkowości* – mit ten pełni funkcję motywacyjną, wzmacniając pozytywny obraz człowieka jako istoty nieporównywalnej pod względem zdolności poznawczych,

¹⁹ Więcej na jego temat znajduje się w częściach 2.1. *Mechanizm kategoryzacji społecznej* oraz 2.2. *Mechanizm samo-licencjonowania*.

²⁰ Szeroki opis tego mechanizmu znajduje się w części 2.5. *Teoria dysonansu poznawczego*.

²¹ Szerokie omówienie tego tematu znaleźć można w sekcji 2.4. *Efekty oczekiwań*.

²² Szczegółowe omówienie dwóch perspektyw obecnych w interpretacji badań dotyczących procesów poznawczych ZINL znajduje się w części: 3.2. *Dwie perspektywy interpretowania eksperymentów z udziałem zwierząt pozaludzkich*.

emocjonalnych czy społecznych. W ten sposób, przedstawiciele innych gatunków, testowani są przez pryzmat systemu poznawczego ludzi. W konsekwencji wyniki eksperymentów często interpretowane są, jako dowód na brak wskazanej zdolności u ZINL. To skutkuje zawężeniem pola interpretacyjnego i potencjalnym zniekształceniem wyników. Tego rodzaju narracje mają nie tylko charakter poznawczy, ale również aksjologiczny – odzwierciedlają wartości przypisywane granicy gatunkowej oraz hierarchii istot żyjących. Tym samym, choć w niniejszej pracy nie zostanie wydzielona odrębna sekcja poświęcona komponentowi afektywnemu, to jego obecność – w szczególności poprzez motywację autowaloryzacyjną, efekt potwierdzenia i mechanizmy kategoryjne – będzie zauważalna w dalszych rozważaniach.

W niniejszej dysertacji szczególną uwagę poświęcam analizie wybranych wzorców postaw społecznych, które — choć nie odnoszą się bezpośrednio do relacji człowiek — inne zwierzęta — ilustrują mechanizmy uprzedzeń oraz ich kulturową trwałość i wpływ na narracje społeczne czy naukowe. Analizy te służą dwóm celom: po pierwsze, mają ukazać, w jaki sposób systemy postaw mogą być zakorzenione w długoterminowych, emocjonalnie nacechowanych opowieściach, nazywanych przez Hochschild (2016) „głębokimi historiami”; po drugie – stanowią punkt wyjścia do analogicznego spojrzenia na uprzedzenia wobec ZINL, w których również ujawnia się rola tożsamości, norm społecznych oraz afektywnego komponentu postaw. Uważam, że zarysowanie zagadnienia głębokich historii i ich wpływu na przyjmowanie postaw jest kluczowe, szczególnie dla późniejszych analiz zawartych w częściach 3.3. *Komponent poznawczy* i 3.4. *Behawioralny komponent uprzedzeń*.

Głęboka historia jest silnie związana z afektywnym komponentem uprzedzeń, a jej zrozumienie możliwe jest tylko za pośrednictwem narzędzia, nazywanego przez Arlie Hochschild *mostem empatii* (Hochschild, 2016). Pojęcie to ma charakter metaforyczny, nie zaś ściśle naukowy, a jego szczegółowe omówienie wykracza poza ramy teoretyczne niniejszej dysertacji. Most empatii stanowi jednak istotny element metodologii badawczej przyjętej przez Hochschild. Jest to określenie odnoszące się do pewnego poziomu zrozumienia, którym musi wykazać się eksperymentator, aby umożliwić swobodny sposób porozumiewania się z osobami pochodzącymi z „innych światów”. Jest to możliwe nie tylko dzięki poznaniu autobiograficznych informacji o uczestniku badania, ale także poprzez ujęcie opowieści, konstytuujących jego porządek społeczny. Dla zobrazowania, w badaniu przedstawionym w książce *Obcy we własnym kraju. Gniew i żal amerykańskiej prawicy* (2016), Hochschild zastosowała

metodę obserwacji uczestniczącej, czyli jakościowe narzędzie wymagające długotrwałego życia w obserwowanej społeczności. Autorka przez pięć lat prowadziła badania terenowe w miejscowości w stanie Luizjana, dążąc do zrozumienia postaw społecznych, a w szczególności ich komponentu afektywnego, mającego wpływ na decyzje polityczne uczestników. Jej głównym celem było wyjaśnienie tzw. wielkiego paradoksu: dlaczego osoby najbardziej dotknięte skutkami zanieczyszczeń środowiska jednocześnie opowiadają się za deregulacją prawa w tym obszarze (Hochschild, 2016; Lakoff, 2008). Analizując postawy wyborców prawicy w południowych Stanach Zjednoczonych, Hochschild pokazuje, że ich przekonania opierają się na głęboko zakorzenionej, emocjonalnej narracji, która nadaje sens rzeczywistości społecznej.

Zatem głęboka historia, w rozumieniu Hochschild, to szeroki zbiór mitów i przekonań obecnych w danym społeczeństwie. Choć ich szczegółowy opis wykracza poza ramy tematyczne niniejszej dysertacji, wprowadzenie do tego zagadnienia uznałam za konieczne, z punktu widzenia uzasadnienia decyzji o niewyodrębnianiu komponentu afektywnego w osobnej sekcji. Jego pogłębiona analiza wymagałaby bowiem zastosowania odmiennych metod badawczych niż te przyjęte w niniejszym opracowaniu. Analogicznie, narracje o zwierzętach innych niż ludzie również bywają podtrzymywane przez symboliczne opowieści o *naturalnym porządku rzeczy*, *hierarchii istot żywych* i *wyjątkowości człowieka*, które są odporne na doniesienia empiryczne. Co więcej, biorąc pod uwagę ogromne zróżnicowanie gatunkowe w obrębie zwierząt innych niż ludzie, niemożliwe byłoby precyzyjne uchwycenie i przedstawienie tak złożonego zagadnienia w ramach jednej pracy teoretycznej. Dlatego w niniejszym opracowaniu nie ma możliwości szczegółowego omówienia głębokich historii przywoływanych badaczy. Tematyka, na której koncentruje się rozprawa: interpretacje naukowe dotyczące przypisywania zwierzętom teorii umysłu — jest na tyle złożona, że jej pełne ujęcie wymagałoby znacznie szerszych badań, zarówno socjologicznych, jak i filozoficznych. Analiza tego typu musiałaby uwzględniać nie tylko zróżnicowanie kulturowe i obowiązujące w tych społeczeństwach konwencje, ale również biograficzne uwarunkowania każdego z badaczy formułujących interesujące mnie interpretacje — co, z oczywistych względów, przekracza ramy tej pracy. W niniejszej dysertacji ograniczam się więc do wskazania szerokiego, historycznie uwarunkowanego kontekstu relacji człowiek — ZINL, koncentrując się przede wszystkim na odmiennych ujęciach charakterystycznych dla tradycji amerykańskiej i europejskiej.

1.4. Podsumowanie

Rozdział pierwszy poświęcony jest wprowadzeniu do tematu uprzedzeń, ze szczególnym uwzględnieniem ich wpływu na interpretację badań naukowych oraz mechanizmów poznawczych i społecznych, które je podtrzymują. Uprzedzenia wobec zwierząt innych niż ludzie kształtują sposób, w jaki interpretujemy dane dotyczące procesów poznawczych, takich jak teoria umysłu, prowadząc do systematycznych błędów poznawczych i utrwalania nierówności w nauce i społeczeństwie. Mechanizmy te mają charakter automatyczny i są głęboko zakorzenione w strukturach poznawczych, które służą utrzymaniu porządku społecznego, co sprzyja naruszeniom dobrych praktyk epistemicznych, takich jak wrażliwość i bezpieczeństwo weryfikacji przekonań. Należy także nadmienić, że każdy człowiek prezentuje pewne postawy, nawet gdy nie zdaje sobie tego sprawy.

Kluczowym mechanizmem wskazanym w omawianej części, jest penetracja poznawcza, która kształtuje procesy kognitywne już na poziomie percepcji. Na drodze przetwarzania odgórnego, uprzedzenia wpływają na interpretację bodźców, co obrazuje klasyczny eksperyment Payne'a (2001) pokazujący, jak stereotypy rasowe prowadzą do błędnej klasyfikacji obiektów. Uważam, że analogiczne mechanizmy działają także w przypadku uprzedzeń międzygatunkowych, powodując tendencyjny odbiór zachowań zwierząt i interpretację ich działań zgodnie z antropocentrycznymi lub antroponegacjonistycznymi schematami poznawczymi. Procesy przetwarzania odgórnego, integrują wcześniejszą wiedzę i oczekiwania z danymi sensorycznymi, choć ułatwiają funkcjonowanie w złożonym środowisku, zwiększają ryzyko błędów poznawczych oraz przyczyniają się do utrwalania stereotypów, ograniczając elastyczność poznawczą i adaptację do nowych warunków.

Istotnym wnioskiem płynącym z rozdziału jest fakt, że uprzedzenia są nie tylko psychologicznym zjawiskiem jednostkowym, ale mają również charakter systemowy i kulturowy — to znaczy, są utrzymywane przez instytucje, narracje i praktyki społeczne. Ich zrozumienie wymaga nie tylko refleksji nad procesami poznawczymi, ale także krytycznej analizy struktur epistemicznych, w których te procesy są osadzone. Ewolucyjnie i kulturowo ukształtowane mechanizmy kategoryzacji społecznej, takie jak podział na grupy my — oni służą adaptacji i przetrwaniu, umożliwiając szybkie podejmowanie decyzji w środowisku społecznym i naturalnym, ale jednocześnie stanowią źródło uprzedzeń i dyskryminacji. Jak wszystkie postawy składają się

z komponentów: poznawczego (stereotypy, mity, przekonania), afektywnego (emocje takie jak lęk czy niechęć) oraz behawioralnego (dyskryminacja), dlatego kompletna analiza uprzedzeń, nie może ograniczać się tylko do jednego wymiaru — konieczne jest podejście zintegrowane, które uwzględnia złożoność wszystkich wymienionych obszarów.

W rozdziale podkreślam, że zrozumienie wpływu uprzedzeń na interpretację badań procesów poznawczych u zwierząt innych niż ludzie, jest kluczowe dla refleksyjności nauki i unikania błędów poznawczych, szczególnie w procesie konstruowania badań nad zdolnościami poznawczymi oraz w interpretacji ich wyników. Co więcej, dostrzegam potrzebę krytycznego namysłu nad definicjami badanych zdolności, czy paradygmatami badawczymi, które mogą wzmacniać istniejące uprzedzenia i nierówności epistemiczne.

Rozdział II: Wybrane mechanizmy poznawcze leżące u podstaw tendencyjności związanej z uprzedzeniami

2.1. Mechanizm kategoryzacji społecznej

System poznawczy odbiera i analizuje dane ze świata zewnętrznego i wewnętrznego względem jednostki. Jednak kodowanie i przetwarzanie wszystkich dostępnych informacji nie miałyby sensu z punktu widzenia wymagań sytuacji, w jakiej znajduje się organizm. Co więcej, umysł nie jest w stanie przetworzyć ich wszystkich, ze względu na ograniczenia związane z pamięcią czy zasobami. Dlatego, aby uniknąć przeciążenia informacyjnego, systemy poznawcze selekcionują dane poprzez mechanizmy uwagowe oraz angażując kontrolę poznawczą hamują, zbędne informacje. Zamiast szczegółowej osobnej analizy dla wszystkich odbieranych elementów, obiekty łączone są w klasy na podstawie pewnych kryteriów, np. cech fizycznych lub wspólnych własności. Proces ten nazywany jest kategoryzacją poznawczą, a jego podstawową funkcją adaptacyjną jest redukcja różnorodności w ramach danej reprezentacji pojęciowej (kategorii). Proces ten jest szeroko opisywany w literaturze, gdzie wskazuje się, że kluczowym dla niego mechanizmem jest efekt grupowania kategoryjnego, polegający na łączeniu elementów dostępnych poznawczo w grupy²³(Bousfield, 1953; Nelson, 2003; Nęcka, 2013). Mechanizm ten umożliwia powiększenie pojemności pamięci roboczej (Bousfield, 1953). Jeden z pionierów badań nad grupowaniem kategoryjnym, William Bousfield (1953), zdefiniował wskazane zjawisko jako, mechanizm kodowania informacji przez kojarzenie ich ze strukturami wyższego rzędu, reprezentującymi nazwę kategorii. Przyjmuje on czasem formę nadawania znaczenia poszczególnym elementom zapamiętywanego materiału (Bousfield, 1953). Badania Bousfielda (1953) wykorzystują paradygmat rozprzestrzeniającej się aktywacji, wywodzący się z teorii sieci semantycznych. Zakłada on, że pamięć zorganizowana jest w sieć semantyczną, składającą się z węzłów (*unity*) połączonych ze sobą relacjami. W takiej architekturze pobudzenie jednego pojęcia wywołuje aktywację rozprzestrzeniającą się jednocześnie w wielu kierunkach, co

²³ Np. 123987 można zapamiętać jako 123 i 987. Mechanizm ten przyjmuje czasem formę nadawania znaczenia poszczególnym elementom zapamiętywanego materiału (Necka, 2013).

proceedi do jednoczesnego pobudzenia wszystkich elementów sieci, powiązanych z początkowym węzłem (Samani, Sharifian, 1997).

W wyniku wskazanego mechanizmu, przypominanie jednego pojęcia, aktywuje całą kategorię, co ułatwia odtwarzanie pozostałych elementów (Bousfield, Cohen, 1953). James Deese (1959) wykazał także, że związki skojarzeniowe pomiędzy słowami są predykatorem liczby od pamiętanych słów oraz liczby intruzji, czyli słów należących do określonej kategorii, ale nie będących elementem listy. Wyniki jego badań wspierają teorię rozprzestrzeniającej się aktywacji, szczególnie biorąc pod uwagę, że intruzje występują częściej w materiale łatwym do grupowania.

Szczególnie interesującym w kontekście niniejszej dysertacji jest mechanizm kategoryzacji społecznej. Polega on na przyporządkowywaniu agensów do specyficznych kategorii poznawczych. Procesy realizujące to zadanie wykonywane są w ramach kryteriów, nazywanych stereotypami. (Mackie, 1999). Oczywiście niektóre kategorie mają większe znaczenie dla tożsamości niż inne. Do tych ważniejszych należą płeć, orientacja seksualna, religia, poglądy polityczne, przynależność etniczna czy narodowość (Travis, Aronson, 2007, str. 75). Chociaż przynależność gatunkowa nie jest wymieniana w literaturze *explicite*, to rozważania dotyczące szowinizmu gatunkowego czy karnizmu bardzo często przywołują argumenty odnoszące się do koncepcji tożsamości społecznej (Dhont, Hodson, Loughnan, Amiot, 2019). Teoria tożsamości społecznej zakłada, że ludzie mają skłonność do dzielenia osób na kategorie, gdzie jedna z nich zawsze identyfikuje grupę, do której należy jednostka – grupa własna, *my*, natomiast pozostałe identyfikują grupy obce – *oni* (Tajfel, Turner, Austin, Worchel, 1979). Dodatkowo wskazuje, że przynależność do cenionych grup społecznych wspiera samoocenę. Czyli na drodze kategoryzacji społecznej przynależność do określonej społeczności może być wykorzystywana jako mechanizm autowaloryzacji. Dlatego przekonanie, że *grupa własna jest lepsza niż obce* stanowi sposób obrony ego (mechanizm obronny²⁴). Psychologia ewolucyjna wskazuje, że etnocentryzm, czyli

²⁴ Mechanizmy obronne to nieświadome procesy psychiczne, wykorzystywane przez *ego* do zmniejszenia wewnętrznego napięcia i ochrony przed silnymi emocjami oraz stresem (Cullari, Kusché, 2025). Ich podstawową funkcją jest ochrona integralności psychicznej jednostki poprzez zarządzanie konfliktami wewnętrznymi i zewnętrznymi zagrożeniami emocjonalnymi. Mechanizmy te rozwijają się od wczesnego dzieciństwa i z wiekiem stają się coraz bardziej złożone i adaptacyjne, co sprzyja lepszemu radzeniu sobie z trudnościami życiowymi oraz utrzymaniu zdrowia psychicznego. W kontekście poprawy ego, mechanizmy obronne pełnią rolę nieświadomych zasobów, które pozwalają jednostce zachować

przekonanie, że *nasza grupa etniczna jest lepsza od innych*, pomaga w przetrwaniu oraz wzmacnia więzi społeczne (Wojciszke, 2009, str. 58). To wpływa na gotowość do działania na rzecz własnej społeczności oraz w czasach kryzysu, pozwala na podjęcie decyzji o oddaniu za nią życia. W rozdziale czwartym swojej książki Wojciszke (2009) wskazuje, że opisana autowaloryzacja jest jednym z najsilniejszych ludzkich pozabiologicznych motywatorów.

Tajfel i współpracownicy (1971) przeprowadzili badania wskazujące, że powstawanie kategorii *my* — oni jest automatyczne. W ramach eksperymentu prezentowano chłopcom kropki i proszono o oszacowanie ich liczby. Następnie uczestnicy losowo podzieleni zostali na 2 grupy, gdzie członkom jednej powiedziano, że przeszacowywali ilość kropek, natomiast drugiej, że zaniżyli ich liczbę. Później wszyscy uczestnicy byli kolejno zapraszani do kabiny, w której mieli oceniać wyniki wszystkich chłopców biorących udział w eksperymencie. Okazało się, że uczestnicy preferowali chłopców, którzy tak jak oni przeszacowali lub niedoszacowali liczby kropek. Co więcej, gdy chłopcy wychodzili z kabiny pytani byli przez innych uczestników, do której grupy należą. Odpowiedzi wzbudzały aplauz u osób z tej samej grupy i „buczenie” chłopców z kategorii przeciwnej. W innym badaniu przedstawiono uczestnikom sylaby bez znaczenia, które zostały losowo powiązane ze słowami odnoszącymi się do grup: własnej (*my, nasz, nas*) oraz obcej (*oni, ich*), natomiast w grupie kontrolnej prezentowano sylaby bez znaczenia powiązane z neutralnymi słowami (*ona, on, jej*); (Perdue, Dovidio, Gurtman, Tyler, 1990). Następnie poproszono ich o ocenę, które z bezsensownych sylab wydawały się przyjemne, a które nieprzyjemne. Okazało się, że sylaby bezsensowne połączone ze słowami odnoszącymi się do grupy własnej wydają się uczestnikom przyjemniejsze, niż te połączone ze słowami odnoszącymi się do grupy obcej. Co ciekawe uczestnicy nie zdawali sobie sprawy z tego, dlaczego w ten sposób odbierali bezsensowny materiał. Ciekawych informacji dotyczących automatyczności kategoryzacji społecznej, dostarczyło badanie pokazujące, że jeżeli niedzielny mecz ligi uniwersyteckiej zakończył się sukcesem drużyny szkolnej, to w poniedziałek więcej osób ubierało ciuchy z emblematem uczelni,

stabilność psychiczną, chroniąc ją przed przytłaczającymi uczuciami i umożliwiając adaptacyjne funkcjonowanie w społeczeństwie. Przykłady takich mechanizmów to wyparcie, racjonalizacja, projekcja czy sublimacja, z których niektóre są uważane za bardziej dojrzałe i sprzyjające rozwojowi osobistemu (Cullari, Kusché, 2025).

pokazuje to że sukces w rywalizacji międzykategorialnej nasilił identyfikację z grupą własną (Crisp, Rubin i Hewstone, 1976).

Eksperymenty z zakresu psychologii społecznej wskazują, że u podstaw kategoryzacji społecznej leżą mechanizmy akcentuacji, faworyzacji grupy własnej przy jednoczesnej deprecjacji oraz infrahumanizacji grupy obcej (Wojciszke, 2009; Tajfel i in. 1971; Crisp, Rubin, Hewstone, 1976). Zjawiska te obserwowane są nie tylko w kontekście dobrze znanych kategorii (np. płciowych, narodowych, etnicznych), ale także tych, które powstają sytuacyjnie. Tajfel (1981) przeprowadził eksperyment, w ramach którego wykazał, że gdy człowiek ocenia dwie kategorie społeczne pod względem zmienności określonej cechy, po pierwsze, ludzie spostrzegają więcej różnic międzygrupowych, a po drugie spada liczba dostrzeganych różnic wewnątrzgrupowych. Dla przykładu stereotyp mówiący, że *mężczyźni są bardziej agresywni od kobiet*, powoduje akcentowanie różnic międzygrupowych pod względem agresji oraz redukuje je wewnątrzgrupowo. Taki sposób spostrzegania sprawia, że podział między dwiema grupami wydaje się znacznie większy, niż jest w rzeczywistości. Mechanizm akcentuacji jest skutkiem społecznych stereotypów dotyczących każdej płci. Dlatego kobiety są postrzegane jako łagodne, często łagodniejsze niż są w rzeczywistości oraz analogicznie ludzie nie widzą różnic wewnątrzgrupowych w kategorii *mężczyzna*, więc uważają, że wszyscy mężczyźni są jednakowo agresywni. Czyli kiedy dwie osoby zostaną zidentyfikowane jako przedstawiciele tej samej kategorii, spostrzegani są jako bardziej podobni do siebie, a gdy reprezentują dwie odrębne grupy, to wydają się bardziej od siebie różnić. Wielu autorów uważa, że redukcja różnorodności w ramach kategorii dotyczy przede wszystkim grup obcych (Wojciszke, 2009). Zjawisko to nazywane jest efektem jednorodności lub homogenizacją grupy obcej. Wyraża się nie tylko w przekonaniu, że *oni* [np. Chińczycy] *są tacy sami*, ale też w faktycznym częstszym myleniu przedstawicieli grup obcych (np. w wyniku międzygrupowych problemów z rozpoznawalnością twarzy). Dodatkowo okazuje się, że ludzie wykazują tendencję do kodowania informacji o przedstawicielach grupy obcej jako, danych o grupie, natomiast informacji o przedstawicielach grupy własnej jako informacji o jednostce. Zjawisko faworyzacji grupy własnej, czyli skłonność do przypisywania pozytywnych cech przedstawicielom kategorii *my* (np. polscy licealiści przypisują więcej zalet Polakom, niż przedstawicielom innych narodowości); (Stone, Łukaszewski, Krueger, Canedo, 2020). Deprecjacja, czyli umniejszenie przedstawicielom kategorii *oni*, zwiększa się, gdy jednostka znajduje się w sytuacji zagrożenia samooceny. Bogdan

Wojciszke (2009) opisuje eksperyment w ramach, którego kobiety i mężczyźni przechodzili rzekomy test na inteligencję emocjonalną, a następnie po upływie tygodnia przyjeżdżali w konkretne miejsce, aby odebrać wyniki i wykonać dodatkowe zadania. Poproszeni zostali o ocenę osoby kierowniczej, której opis zawierał się na jednej stronie kartki. Okazało się, że mężczyźni, którzy usłyszeli wynik plasujący się poniżej oczekiwań (np. średni), byli bardziej skłonni do przydzielania wyższych ocen kierowników, deprecjonując przy tym kierowniczkę. Analogiczne zachowywały się kobiety, faworyzując kierowniczkę oraz deprecjonując kierowników. Co ciekawe, zależność ta zanikała u osób, które otrzymały wysoki wynik w teście.

Innym mechanizmem związanym z kategoryzacją społeczną, omawiany jest na gruncie teorii infrahumanizacji grupy obcej, która zakłada, że systemy poznawcze mogą reprezentować kategorie na dwa sposoby: za (1) pomocą cech przetwarzanych obiektów lub (2) porównując poznawany element świata społecznego z innym obiektem pokrewnej kategorii. Poznanie (np. uczenie się nowych obiektów), odbywa się poprzez (2), czyli wyjaśnianie ich za pomocą starych kategorii, co możliwe jest dzięki myśleniu analogicznemu (opisanemu bardziej szczegółowo w części 2.3. *Myślenie analogiczne*). Literatura wskazuje, że odpowiadając na pytanie *jaka jest natura człowieka?* ludzie częściej wybierają strategię (2), tworząc dwie kategorie (Kagan 2004, str. 77). Jedną określaną jest przez cechy wyjątkowo ludzkie, drugie stanowią tylko jakościowe ulepszenia zdolności ewolucyjnie starszych. W literaturze zaznacza się także, że kiedy człowieczeństwo definiowane jest w kategoriach *ludzkiej wyjątkowości* („większego człowieczeństwa”), jest ono rozumiane w sensie wyraźnie porównawczym (Kagan, 2004). Co potwierdza wcześniej opisane twierdzenia, w ramach których kategoryzacja społeczna jest narzędziem do autowaloryzacji, stanowiącym mechanizm obronny, który pomaga w utrzymaniu stereotypów i mitów o wyjątkowości kategorii *my*. Tym samym ludzie wykazują tendencję do odmawiania przypisywania przedstawicielom grup obcych cech uznawanych za wyjątkowo ludzkie. Odrębność infrahumanizacji od faworyzacji i deprecjacji wykazana została w licznych publikacjach (Leyens, Paladino, Rodriguez-Torres, Vaes, Demoulin i in. 2000; Leyens, Cortes, Demoulin, Dovidio, Fiske, i in., 2003; Brewer, Brown, 1998; Haslam, Bain, Douge, Lee, Bastian, 2005). Pierwszy z wymienionych mechanizmów różni się od drugiego w przynajmniej dwóch aspektach:

- a. Po pierwsze nie można go wyjaśnić przez mniejszą znajomość grupy obcej, a co za tym idzie, mniejszą złożonością kategorii *oni*,
- b. Po drugie nie można go sprowadzić do tendencji polegającej na przypisywaniu przedstawicielom grupy własnej bardziej pożądanym cech, ponieważ zjawisko przypisywania większego człowieczeństwa przedstawicielom grupy własnej dotyczy zarówno cech pozytywnych, jak i negatywnych.

Haslam, Bain, Douge, Lee, Bastian (2005) przeprowadzili szereg badań wykazujących, że ludzie tworzą dwie różne kategorie reprezentujące dwa odrębne zbiory przekonań dotyczących człowieczeństwa. Pierwsza zawiera w sobie (a.) stereotypy określające zasady atrybucji albo odmówienia cech unikalnych dla człowieka, nazywanych *esencjonalnymi dla natury ludzkiej (human essence)*. Posiadanie tych cech świadczy o „większym człowieczeństwie”. Druga kategoria określana jest przez (b.) stereotypy reprezentujące cechy niespecyficzne dla człowieka, czyli takie, które dzielone są z innymi zwierzętami. Teoria infrahumanizacji przyjmuje trzy założenia:

- a. Możliwe jest istnienie więcej niż dwóch koncepcji człowieczeństwa, a co za tym idzie, ludzka wyjątkowość może nie być jedyną perspektywą wynikającą z infrahumanizacji.
- b. Infrahumanizacja występuje zarówno w porównaniach interpersonalnych oraz międzygrupowych.
- c. Nie jest pewne, czy cechy *esencjonalne dla człowieka* są rozumiane w sposób esencjalistyczny, jak twierdzą Leyens i in. (2003).

Drugi z pięciu przeprowadzonych eksperymentów polegał na ocenie siebie i innych w perspektywie (a) i (b); (Haslam, Bain, Douge, Lee, Bastian 2005). Wyniki pokazały, że przedstawicielom kategorii *my* przypisywano więcej cech uznawanych za esencjonalne, natomiast kategoria *oni* związana jest ze stereotypami odnoszącymi się do cech niespecyficznie ludzkich. Trzeci i czwarty eksperyment miały charakter replikacyjny oraz powieliły wyniki wcześniejszych badań. Wyniki te zdają się potwierdzać inne doniesienia empiryczne, które pokazały, że emocje pierwotne są tendencyjne kojarzone z grupami obcymi oraz ludzie preferują unikane atrybucji emocji wtórnych przedstawicielom kategorii *oni* (poprzez ignorowanie przesłanek za ich

doświadczeniem); (Gaunt, Leyens, Sindic, 2002; Castano, Yzerbyt, Paladino, Sacchi, 2002; Gaunt, Leyens, Sindic, 2004). Uczestnicy eksperymentu chętniej przypisywali emocje wtórne uważane za wyjątkowo ludzki przedstawicielom grupy własnej (Gaunt, Leyens, Sindic, 2004). Stereotypy charakteryzują grupę własną jako kategorię, której przysługują przymioty takie, jak inteligencja czy głębokość emocjonalna. Natomiast kategoria *oni* charakteryzuje się głupotą, nie zna miłości czy wyrzutów sumienia (Haslam, Bain, Douge, Lee, Bastian, 2005). Nie zaobserwowano podobnej korelacji odnośnie do cech pierwotnych, powszechnie uznawanych za dzielone z innymi zwierzętami. Autorzy innego badania sugerują, że emocje wtórne wyrażane przez przedstawicieli grupy obcej, częściej spotykają się z negatywnym odbiorem (Vaes, Paladino, Castelli, Leyens, Giovanazzi, 2003). Co więcej, gdy przedstawiciele *oni* wyrażają swoje potrzeby w kategoriach emocji, to jednostka będzie mniej skłonna, te potrzeby zaspokoić.

2.2. Mechanizm samo-licencjonowania

W literaturze psychologicznej znajduje się wiele odniesień do mechanizmu własnej wyższości (*grandiosity*), charakteryzującego się przekonaniem o *byciu lepszym od innych*. W psychologii społecznej uznaje się, że ludzie mają naturalną potrzebę posiadania pozytywnego obrazu siebie. Jednym ze sposobów jego budowania, jest wnioskowanie oparte na przekonaniu o *własnej wyższości*. Zjawisko to jest bardzo powszechne, pełni funkcję mechanizmu autowaloryzacji – np. gdy ktoś uważa, że jest bardziej etyczny, sprawiedliwy lub świadomy społecznie niż inni (niektórzy wegetarianie, weganie, przedstawiciele grup religijnych czy osoby zaangażowane społecznie); (Haidt, 2012). Jonathan Haidt w swojej książce *The Righteous Mind: Why Good People are Divided by Politics and Religion* (2012) zauważa, że w kontekście podziałów ideologicznych i politycznych, ludzie mają tendencję do przekonania, że ich poglądy są bardziej moralne, co prowadzi do polaryzacji i eskalacji konfliktów. Poczucie własnej wyższości może przybierać różne formy, na przykład dostrzegane jest w efekcie Dunninga-Krugera, który polega na tym, że osoby o niskich kompetencjach w danej dziedzinie mają skłonność do przeceniania swoich umiejętności (mechanizm własnej wyższości), podczas gdy osoby wysoko kompetentne często nie doceniają swojego poziomu wiedzy (Kruger, Dunning, 1999). W literaturze zjawisko to jest elementem szerszego zbioru procesów poznawczych, nazywanych samo-

licencjonowaniem (*self-licensing*), polega ono na nieświadomym zwiększeniu zaufania do swoich wyborów (np. moralnych), co wywołuje skłonność do dokonania wyboru niemoralnego (Merritt, Effron, Monin, 2010). Jest to rodzaj samousprawiedliwienia się, dzięki któremu jednostka może wykonać niemoralne zachowanie, przy jednoczesnym nieodczuwaniu konsekwencji takiego zachowania (wyrzuty sumienia, dysonans poznawczy). Benoît Monin i Dale Miller (2001) opublikowali artykuł, w ramach którego przedstawili trzy eksperymenty obrazujące wskazany mechanizm. Ze względu na podobieństwo wszystkich trzech badań poświęcę szczególną uwagę tylko pierwszemu. Dotyczyło ono seksizmu oraz obejmowało trzy grupy: kontrolną²⁵ i dwie badawcze. W pierwszym etapie zadaniem uczestników było oszacowanie poprawności pięciu następujących twierdzeń²⁶:

- „(a) Niektóre/większość kobiet są lepsze w wykonywaniu prac domowych i opiece nad dziećmi.
- (b) Emocjonalność mężczyzn sprawia, że są oni bardziej odpowiedni do pracy w polityce niż przynajmniej większość kobiet.
- (c) Najlepszymi zawodami dla niektórych/większości kobiet są prace takie jak kucharka, pielęgniarka, nauczycielka.
- (d) Niektóre/większość kobiet potrzebują mężczyzny, żeby się nimi opiekował.
- (e) Niektóre/większość kobiet nie jest tak naprawdę inteligenta” [tłum. własne z tekstu Monin, Miller, 2001, str. 35].

Podzielenie próby eksperymentalnej na dwie grupy miało na celu wprowadzenie manipulacji „niektóre/większość”, część respondentów szacowało poprawność zdań zawierających słowo „niektóre”, a pozostali „większość”. Autorzy eksperymentu zakładali, że uczestnicy oceniający twierdzenia zawierające drugie wymienione słowo, będą mocniej utwierdzać się w przekonaniu, że *nie prezentuje uprzedzeń*, niż grupa, w której zdaniach pojawiło się „niektóre”. W ramach kolejnego etapu wszyscy (grupa

²⁵ Grupa kontrolna nie brała udziału I etapie eksperymentu.

²⁶ Oryginalne twierdzenia:

„Some/Most women are better off at home taking care of the children.

Men are more emotionally suited for politics than are some/most women.

The best job for some/most women is something like cook, nurse, or teacher.

Some/Most women need a man to protect them”.

Some/Most women are not really smart” (Monin, Miller, 2001, str. 35)

kontrolna zaczynała od niżej opisanego zadania) uczestnicy zapoznawali się z następującym tekstem²⁷:

„Imagine that you are the manager of a small (45-person) cement manufacturing company based in New Jersey. Last year was a particularly good one, and after you invested in increasing the output capacity of your plant, you decide that it would be very fruitful if you could find clients in other states to increase your business. Because you cannot spend too much time away from the plant, you decide to appoint someone to go around to prospective clients and negotiate contracts. This is a highly specialized market, and the job will mostly consist in going from one building site to another, establishing contacts with foremen and building contractors. It is also a highly competitive market, so bargaining may at some points be harsh. Finally, it's a very technical market, and a representative that did not exude confidence in their technical skills would not be taken seriously by potential clients. Realizing how useful such a help would be for you, you decide to give the person chosen one of the top-five salaries in your company. Do you feel that this job is better suited for one gender rather than the other?” (Monin, Miller, 2001, str. 35).

Zadaniem respondentów było udzielenie odpowiedzi na pytanie zawarte w ostatnim zdaniu tekstu. Należało tego dokonać poprzez użycie umieszczonej pod tekstem 7-stopniowej skali, gdzie -3 oznaczało „tak, zdecydowanie lepsze dla mężczyzn”, natomiast 3: „ani trochę się z tym nie zgadzam”. Wyniki potwierdziły stawianą przez autorów hipotezę, tj. osoby, które nie zgodziły się ze zdaniem prezentowanymi w pierwszym etapie eksperymentu, były bardziej skłonne do oznaczania punktów od -3 do -1, niż ci, którzy zgadzali się z prezentowanymi wcześniej zdaniem. Jednak najciekawszym rezultatem był ten uzyskany odnośnie do manipulacji „niektóre/większość”. Efekt zgodny z założeniem obserwowany był tylko u mężczyzn, co zdziwiło autorów, gdyż zakładano, że różnice płciowe nie zostaną zaobserwowane. Literatura nie dostarcza najmniejszej sugestii świadczącej o tym, że ludzie różnią się płciowo w kontekście potrzeby bycia postrzeganym jako *pozbawionym uprzedzeń*. Zaproponowano dwa wyjaśnienia takiej sytuacji, w ramach pierwszego, uznano, że treści zawarte w eksperymencie stanowiły dla uczestniczek personalne zagrożenie. Możliwość zgodzenia się z twierdzeniem (e) pojawiającym się w pierwszym etapie, wywoływało w kobietach złość i poczucie upokorzenia, co doprowadziło do

²⁷ Fragment został pozostawiony w oryginalnym języku ze względu na potencjalne ryzyko utraty precyzji i niuansów językowych w przypadku tłumaczenia, co mogłoby wpłynąć na zatrącenie jego stereotypowego charakteru.

przeciwstawienia się zawartym w nim twierdzeniom. Za słuszością wskazanego wyjaśnienia przemawia fakt, że niektóre uczestniczki wyrażały jawną irytację po ocenie poprawności pięciu wyżej wymienionych pozycji. Monin i Miller (2001) zaznaczają także, że warunek ze słowem „większość” może implikować przyjmowanie bardziej seksistowskich postaw, tak jak słuchanie nietolerancyjnych wypowiedzi, skłania odbiorców do przyjmowania uprzedzeń zgodnych z ich treścią (Blanchard, Crandall, Brigham, Vaughn, 1994).

Rozważania z zakresu motywacji, wskazują, że samo-licencjonowanie jest mechanizmem redukcji napięcia, jakie powstaje na skutek konkurencji między mechanizmami motywującymi do działania i hamującym zachowanie (Lewin, 1958). Czyli opisywany proces pomaga w przekonaniu samego siebie, że *ja* mam prawo do mniej etycznych działań w określonym obszarze, ponieważ wykazuję zachowania moralne w innym. Przykładem może być jednostka, usprawiedliwiająca swoje moralnie wątpliwe działania np. drobne kłamstwa, faktem, że przekazuje darowiznę na cele charytatywne (Merritt i inn. 2010, str. 346). Mechanizm ten jest zastanawiający w kontekście teorii dysonansu poznawczego, opisanego niżej, jednak literatura w tym zakresie pozostaje uboga.

Przytoczone badania sugerują, że omawiany mechanizm pozwala jednostkom na zachowanie spójnego obrazu siebie, jako osoby moralnej, mimo podejmowania działań nieetycznych w określonych sytuacjach. Wprawdzie ubytek w literaturze wymaga uzupełnienia, jednak szczegółowe zgłębianie tej problematyki wykracza poza ramy niniejszej rozprawy doktorskiej. Inną istotną luką jest brak analiz samo-licencjonowania w kontekście uprzedzeń międzygatunkowych, wydaje się jednak, że mechanizm ten powinien występować także w przypadku szowinizmu gatunkowego czy karnizmu, a także, że ludzie uzasadniają własne wybory konsumpcyjne (jedzenie mięsa) i konsumenckie (noszenie futer/skór) równoważąc je działalnością prozwierzęcą w postaci podpisywania petycji, czy wpłat na schroniska. Takie samousprawiedliwienie redukuje napięcie poznawcze i pozwala na utrzymanie pozytywnego obrazu siebie mimo niespójności.

2.3. Myślenie analogiczne

Pojęcie analogii wywodzi się ze starożytnej Grecji, gdzie Arystoteles definiował ją jako równość stosunków (Aubenque, 2010; Kaproń-Charzyńska, 2012). Czyli

analogia to reguła wnioskowania, w ramach której A i B ma się do siebie jak C i D. Ciekawy wkład w rozwijanie koncepcji analogii mieli Tomasz z Akwinu i jego komentatorzy (Szymanek, 2008, str. 71). Uznawali oni, że omawiana reguła leży u podstaw jednego z trzech sposobów orzekania o świecie: synonimicznie, wieloznacznie i analogicznie. Pokazuje to, jak dużą rolę odgrywa omawiana reguła w systemach poznawczych człowieka. Współcześnie o analogii mówimy, gdy istnieją struktury powiązane ze sobą relacją podobieństwa (Szymanek, 2008). Pojęcie to ma bardzo szerokie zastosowanie m.in. w biologii, gdzie struktury analogiczne przeciwstawia się często strukturom homologicznym. Pierwsze wymienione to przykład struktur nieposiadających wspólnego pochodzenia ewolucyjnego, ale pełniących te same funkcje np. skrzydła ptaka i skrzydła ważki. Natomiast struktury homologiczne posiadają wspólne pochodzenie ewolucyjne, ale pełnią zupełnie różne funkcje, przykładem są skrzydła nietoperza i ręce człowieka.

Podobieństwo jest najpierwotniejszym zjawiskiem prowadzącym do spontanicznego kreowania kategorii (Szymanek, 2008). Ewolucyjne wyposażenie systemu poznawczego w zdolność do rozpoznawania cech ma swoje ograniczenia m.in., jak już wspomniałam, takie „naturalne”/fizyczne podobieństwo, często nie gwarantuje trafności podejmowanych na jego podstawie decyzji. Dodatkowo są takie podobieństwa, których nie sposób uchwycić zmysłowo np. między jajkiem i nasionem. W związku z tym kierowanie się wskazaną relacją prowadzi często do błędów. Antropomorfizm polega na przypisywaniu ludzkich cech nie tylko ZINL, ale także obiektom nieożywionym, na podstawie podobieństwa fizycznego. Natomiast antroponegacjonizm, jest uporczywym negowaniem występowania cech podobnych do ludzkich (*human-like*) u zwierząt innych niż człowiek.

Historia nauki dostarcza wielu przykładów wpływu błędów płynących z niewłaściwego zastosowania analogii. Błędy tego typu mogą manifestować się poprzez ignorowanie *Umweltu*²⁸ określonego gatunku zwierząt. Książka estońskiego biologa, która po raz pierwszy ukazała się w Niemczech w 1934 roku, była wyrazem jego niezadowolenia. Krytykował w niej ówczesne podejście badaczy zwierząt pozaludzkich, zaznaczając, że ich zobiektywizowane metody badawcze, sprawiają, że uczestniczące w badaniu zwierzęta jawią się jako obiekty podobne do maszyn. Sam autor twierdził, że wszystkie zwierzęta – bez względu na swoją złożoność – muszą być

²⁸ *Umwelt* rozumiany jako subiektywny świat organizmu (Uexküll, 2013).

rozumiane jako ukonstituowane podmioty światów konstruowanych przez ich specyficzne sposoby postrzegania i funkcjonowania w środowisku. Jednym z badaczy zainspirowanych koncepcją indywidualnego świata kształtowanego przez zmysły, doświadczenia i sposób percepcji, otaczającego zwierzę świata jest Donald Griffin (1944). Był on jednym z pierwszych naukowców, którzy zauważyli i zbadali zdolność zwierząt do używania echolokacji. W swoich badaniach wykazał, że nietoperze, używają tej zdolności, aby zbierać informacje o otoczeniu. Jego prace nad nietoperzami nie tylko przyczyniły się do rozwoju wiedzy o tym gatunku, czy do rozwoju technologicznego (późniejsze urządzenia wykorzystujące echolokację), ale również wpłynęły na rozwój koncepcji *Umweltu*, pokazując, w jaki sposób zdolności percepcyjne organizmów kształtują ich doświadczenie otoczenia.

Figura nietoperza wykorzystana została także przez Tomasa Nagela (1997), aby wyjaśnić prywatność świadomych doświadczeń. W odpowiedzi na tytułowe pytanie otrzymujemy odpowiedź, że nigdy się nie dowiemy. Autor eseju „Jak to jest być nietoperzem?”, odnosi się do doznań świadomych, które nadal zostają zagadką dla nauki i filozofii. Chociaż szczegółowe rozważania dotyczące świadomości są interesujące, nie mieszczą się w ramach tematycznych niniejszego opracowania. Uważam jednak za konieczne podkreślenie myśli Nagela, ponieważ przyrównywana jest ona do idei podniesionej w *Dociekaniach filozoficznych* (2019), w kontekście koncepcji języka Wittgensteina (de Waal, 2016). Żeby wiedzieć, jak to jest być nietoperzem, jednostka musiałaby stać się nim. Wówczas nie dowiedziałaby się, jak to jest być tym zwierzęciem, bo ona sama przestałaby istnieć jako podmiot poznawczy. Gdyby zaś dowiedziała się, jak to jest być nietoperzem dla niej, to wciąż nie wiedziałaby, jak to jest nim być, bo *de facto* nigdy nie była. W tym sensie Nagel uważa, że odpowiedź na pytanie zawarte w tytule jego tekstu, nie jest możliwe do rozstrzygnięcia. Twierdzi on, że nie dowiemy się literalnie, pierwszosobowo, czy proceduralnie, jak to jest być inną osobą. Wittgenstein (2019) sugeruje, że różnice w doświadczeniu, percepcji świata i sposobie pojmowania rzeczywistości między różnymi gatunkami są tak fundamentalne, że nie jest możliwe zrozumienie punktu widzenia czy perspektywy innego organizmu. W ramach takiego stanowiska podkreśla się, że systemy poznawcze innych zwierząt konstruują świat w sposób diametralnie różny od człowieka, w związku z czym, człowiek nigdy nie będzie w stanie zrozumieć innych zwierząt. Stanowisko to można zilustrować przez aforyzm „Gdyby lew nauczył się mówić, i tak byśmy go nie zrozumieli” (2019, §327). Wittgenstein postulował, że

język jest nierozzerwalnie związany z praktyką społeczną, w której funkcjonuje. Tym samym, nawet gdyby lew potrafił mówić, to porozumiewałby się w języku właściwym dla jego gatunku, osadzonym w odmiennym porządku sensów i form życia, co czyniłoby jego wypowiedzi niezrozumiałymi dla człowieka. Z kolei opanowanie przez lwa języka ludzkiego, rozumiane jako wejście w ludzką formę życia, oznaczałoby przekształcenie go w istotę ludzką, wówczas możliwe byłoby wzajemne zrozumienie.

Istotne w kontekście rozważań nad analogią zdaje się pytanie, czy biorąc pod uwagę wyżej przedstawioną dyskusję, korzystanie z analogii porównującej elementy systemu poznawczego ludzi z elementami systemu poznawczego innych zwierząt jest zasadne? Faktem jest, że zwierzęta pozaludzkie mają zupełnie inny zestaw umiejętności, które są dostosowane do ich środowiska, zachowania i potrzeb. Wiemy, że ich *Umwelt*, kształtowany jest przez ich zmysły, doświadczenia i sposoby percepcji świata, a co za tym idzie, biorąc pod uwagę te aspekty, można spróbować zrozumieć funkcjonowanie systemów poznawczych innych niż ludzki. Co więcej, takie porównania mogą być użyteczne do zrozumienia pewnych aspektów zachowania i funkcjonowania zwierząt. Dla przykładu długo przed odkryciem echolokacji zauważono, że głuchy nietoperz, jest jak ślepy człowiek (de Waal, 2016, str. 19). Przykład ten pokazuje, że odpowiednie zrozumienie *Umweltu* ZINL, jest kluczem do tworzenia skutecznych modeli poznawczych. Klasycznym już przykładem paradygmatów, które nie uwzględniały *Umweltu* uczestniczących w badaniu zwierząt, są te z zakresu rozwiązywania problemów, a konkretnie korzystania z narzędzi. Przez wiele lat uważano, że gibony i słonie nie są do tego zdolne. W przypadku gibbonów było to interesujące, ponieważ były to jedyne naczelne, u których takiej zdolności nie wykryto. Eksperyment przeprowadzany był w paradygmacie korzystania z narzędzi, gdzie uczestnicy mieli za zadanie, wybrać narzędzie w postaci długiego kija i przyciągnąć do siebie, za jego pomocą przynętę umieszczoną poza zasięgiem ręki zwierzęcia. Dopiero w latach '60 XX wieku Benjamin Beck (1967), wykazał, że prezentowanie gibbonom przedmiotów położonych na płaskiej powierzchni, znacząco utrudnia im podniesienie obiektu, gdyż budowa dłoni tych zwierząt dostosowana jest do poruszania się po drzewach. Cechuje się długimi palcami rąk i nóg, jednak kciuk jest bardzo mały. Czyli ręce gibbonów przypominają raczej haki niż elastyczne organy chwytne. Rozmieszczenie zestawu potencjalnych narzędzi na pionowej powierzchni, znajdującej się na wysokości ramion gibona, umożliwiło uczestnikom chwycenie pożądanego obiektu i przyciągnięcie przynęty. Podobnie było w przypadku słoni, gdzie

okazało się, że zwierzęta te nie podnosiły narzędzi z powierzchni płaskiej, ponieważ zgięcie trąby zaburzyłoby ich zmysł węchu, a to znacząco utrudniłoby zlokalizowanie przynęty, którą powinny przyciągnąć za pomocą narzędzia (Foerder, Galloway, i in., 2011). Gdy zastąpiono kije dużymi skrzyniami, które słoń mógł kopnąć i dosięgnąć przynęty powieszanej ponad jego zasięgiem, zwierzę wykorzystywało pudło, aby dosięgnąć celu.

De Waal uważa, że u podstaw stanowczych twierdzeń dotyczących braku określonych zdolności u ZINL, leży błąd polegający na ignorowaniu zasady, że brak dowodu, nie jest dowodem na brak (de Waal, 2018, str. 24). Wydaje mi się jednak, że ignorowanie przytoczonej zasady jest raczej efektem pewnych uproszczeń w rozumowaniu, jakich dokonały systemy poznawcze osób tworzących wskazane modele. Zignorowano *Umwelt* gibbonów, ponieważ dokonano prostej analogii – ręce gibbonów są jak ręce ludzi. Skoro ręce człowieka nie mają problemu z podniesieniem przedmiotów z powierzchni płaskiej, nie było powodu, aby sądzić, że ręce gibona będą miały z tym problem. Natomiast w przypadku słoni, gdzie zastosowano analogię: ręce człowieka – trąba słonia, przy czym zignorowano cechy trąby, które są odmienne od cech rąk. Przede wszystkim fakt, że trąba stanowi dla słonia także narząd węchu, a jej zginanie powoduje duże zaburzenie tego zmysłu. Przez lata, aż do czasu pierwszych badań sugerujących błędy w procedurze eksperymentalnej, badacze nie brali pod uwagę *Umweltu* gibbonów i słoni, co skutkowało efektem podłogi. Wynik ten był błędnie odczytany jako prawomocny, czyli świadczący za niezdolnością zwierząt uczestniczących w badaniu do korzystania z narzędzi.

Czyli, zjawisko analogii jest silnie związane z mechanizmami kreowania kategorii poznawczych, porównywanych ze sobą zjawisk i obiektów m.in. pod względem podobieństwa²⁹. Zatem mechanizm ten pozwala traktować rzeczy nieidentyczne jako tożsame, a jego podstawą adaptacyjną jest radykalna redukcja informacji, konieczna do przetworzenia sytuacji i podejmowania decyzji. Dzięki procesowi kategoryzacji podczas podejmowania decyzji system poznawczy nie bierze pod uwagę setek nieistotnych informacji np. właściwości obiektów, ponieważ przynależność do kategorii określa cechy istotne egzemplarzom obecnym w ich

²⁹ Jak zaznacza Szymanek (2008, str. 61), sama relacja „a przypomina b” jest raczej związana z percepcją, niż z rozumowaniem. Gdy mówimy o podobieństwie obiektów, mamy na myśli, że przysługują im pewne wspólne charakterystyki, które wykorzystywane są w procesach kategoryzacji.

obszarze. Wiele kategorii funkcjonujących potocznie jest nieadekwatnym odwzorowaniem rzeczywistości, ponieważ powstają one spontanicznie, bez wysiłku, a co za tym idzie z niewielkim udziałem mechanizmów kontrolnych. Co więcej, fundamentalnym mechanizmem dla procesów kategoryzacji jest generalizacja (Szymanek, 2008). Jej forma ma charakter indukcyjny, czyli jeśli grupa obiektów należących do kategorii A posiada cechę x, to wszystkie obiekty tej kategorii posiadają tę cechę. Niepoprawne automatyczne generalizacje wykorzystywane w omawianym procesie, mogą utrudniać racjonalne wnioskowanie i działanie.

Chcę także zaznaczyć, że termin „szowinizm gatunkowy” pierwszy raz użyty został przez Richarda Rydera (1980), definiującego ten rodzaj postaw jako „formę uprzedzeń, dotyczących wyglądu – jeśli osobnik wygląda inaczej, to jest oceniany jako gorszy, pod względem moralnym” (Ryder, 1980, str. 39-47). Wskazuje to na silny związek szowinizmu gatunkowego z postrzeganiem przedstawicieli innych gatunków przez pryzmat podobieństwa do człowieka. Takie kolokwialne postrzeganie analogii stoi w sprzeczności z jej naukowym znaczeniem, ponieważ, jak zaznaczyłam wcześniej, struktury analogiczne są często do siebie niepodobne pod względem fizycznym. Zatem poleganie na moralnie nieistotnych różnicach fizycznych, opartych na cechach odbieranych zmysłowo, podczas dokonywania osądów moralnych i jak wykażę później — atrybucji cech, może być wynikiem tendencyjności u podstaw której nie rzadko leżą preferencje dotyczące wyglądu (Kahneman, Tversky, 2019). Co więcej, w przypadku szowinizmu gatunkowego lub karnizmu człowiek jest bardziej skłonny do przyjmowania pozytywnych postaw w stosunku do zwierząt, które wydają mu się w jakimś stopniu bardziej atrakcyjne fizycznie, np. gdy ZINL jest bardziej podobne do człowieka — w przypadku szowinizmu gatunkowego, natomiast w ramach karnizmu — czy jest ładniejsze pod względem fizycznym. Ciekawym uprzedzeniem jest bambizm, ponieważ u jego podstaw leży mechanizm, który początkowo pełnił funkcję obronną dla ludzkiego potomstwa. Słodka reakcja to mechanizm biologiczny manifestujący się przez uczucie przyjemności, któremu często towarzyszy wdawanie dźwięków zachwyty, jak np. westchnienia (Sherman, Haidt, Coan, 2009). Zatem pełni on funkcję motywacyjną, do ochrony ludzkiego potomstwa. Uznaje się, że słodka reakcja wyzwalana jest przez dostrzeżenie cech określanych w języku polskim jako „cechy dziecka” (*baby face*). Jest to okrągła twarz, duże oczy i mały nos. Z czasem mechanizm

ten uległ generalizacji, w wyniku czego słodka reakcja wyzwalana jest także przez przedstawicieli innych gatunków zwierząt, cechujących się wskazanymi cechami³⁰.

Pomimo tego, że analogia rozumiana jest jako podobieństwo stosunków, w systemie poznawczym pojawia się zawsze tam, gdzie istnieją systemy powiązanych ze sobą elementów, które łączy relacja podobieństwa. Dlatego jej rola jest znacznie szersza i wiąże się z występowaniem określonych błędów. W literaturze wyróżniamy pięć najważniejszych zastosowań analogii w systemie poznawczym. Po pierwsze, pełni ona istotną funkcję w procesach rozumowania i uczenia się. Jej rolą jest ułatwienie rozumienia nowych obiektów, poprzez wyjaśnienie ich za pomocą starych kategorii. Jedną z zasad kierujących rozumowaniem z zakresu tworzenia teorii umysłu jest założenie, że „osoba, którą obserwuje, jest taka jak ja” (Szymanek, 2008, str. 59). Po drugie, analogia zapewnia spójność ocen i decyzji, gdzie podstawową zasadą kierującą wnioskowaniem jest „podobne traktować podobnie”, jest to jedna z najważniejszych zasad ferowania osądów moralnych (Khaneman, 2012). Po trzecie, naukowcy często posługują się modelami i paradygmatami, które stanowią pewien przypadek idealny oraz wyznaczają sposób wypełniania określonej funkcji. W każdym z wymienionych przypadków mamy do czynienia z tworzeniem nowych form i struktur na podobieństwo wybranego wzorca. Po czwarte analogia ma bardzo istotne zastosowanie w języku, gdyż pewne treści są bardzo trudne do przekazania, bez sjej zastosowania. Ostatnią, piątą rolą omawianej relacji, jest ta heurystyczna, wyjątkowo ważna dla niniejszego opracowania. Szrokie omówienie wskazanej funkcji znajduje się w części 3.4.2. *Heurystyki i wynikające z nich tendencje w zachowaniu i podejmowaniu decyzji istotne w kontekście relacji międzygatunkowej*.

2.4. Efekty oczekiwań

Kolejnym istotnym w kontekście niniejszej pracy zjawiskiem są efekty oczekiwań. Można powiedzieć, że są one skutkiem wpływu oczekiwań, przekonań i sugestii na zachowanie, a co za tym idzie przebieg zdarzeń, czy wydajność zwierząt w różnych sytuacjach. Próbę uchwycenia tego zjawiska znajdziemy na gruncie psychologii i socjologii. Istnieje wiele efektów oczekiwań, które w literaturze występują pod licznymi nazwami m.in.: efekt Pygmaliona, mądrego Hansa, Hawthorne, placebo

³⁰ Szersze omówienie mechanizmu słodkiej reakcji znajduje się w późniejszej części tego rozdziału (2.6. *Słodka reakcja*).

czy samospełniające się proroctwo, to wszystkie te terminy odnoszą się do sytuacji, w której obserwator nieświadomie i nieintencjonalnie wpływa na zachowanie obserwowanego podmiotu. Różnicuje się je, w zależności od sytuacji, w jakiej zachodzą oraz relacji, jaka łączy obserwatora i podmiot. W celu zobrazowania szerokiego spektrum omawianego zjawiska przedstawię pokrótce najważniejsze rodzaje efektów oczekiwań, wydaje się jednak, że najważniejsze w kontekście interpretacji wyników eksperymentów dotyczących teorii umysłu u zwierząt pozaludzkich, są oczekiwania interpersonalne oraz eksperymentatora. Mniej istotnym w określonym temacie, jednak nie mniej interesującym, jest także samospełniająca się przepowiednia, która obecna jest w dyskusji dotyczącej uprzedzeń.

Efekt Hawthorne zauważony w badaniach dotyczących tego, jakie warunki pracy lub otoczenia skutkują poprawą wydajności i motywacji do pracy osób zatrudnionych w fabrykach przemysłowych (Mayo, 1945). Badanie prowadzone było w latach 1924-1932 w Western Electric znajdującej się niedaleko Chicago (Mayo, 1945). Okazało się, że samo zainteresowanie pracodawcy warunkami pracy wpływało na poprawę wydajności pracowników, niezależnie od faktycznych zmian. Aktualnie wiedza o tego rodzaju efekcie oczekiwań wykorzystywana jest w naukach dotyczących zarządzaniem zasobami ludzkimi (Berny, Leśniewski, Górski, 2012).

Efekt placebo oraz nocebo to wpływ oczekiwań odnośnie do skuteczności leczenia (fizycznego/psychicznego), które mogą wpływać na reakcję pacjenta (Stolarska, Kaciński, 2011). Jeśli pacjent oczekuje poprawy zdrowia po przyjęciu substancji niemającej działania o określonym skutku, to pomimo braku fizycznego działania substancji człowiek ją przyjmujący może faktycznie odczuć ulgę, czy znaczącą poprawę samopoczucia (efekt placebo). W przeciwnym razie, gdy pacjent oczekuje negatywnych skutków ubocznych, może ich doświadczyć, nawet jeśli substancja przez niego przyjmowana jest nieszkodliwa (efekt nocebo). Badania we wskazanym zakresie obrazują, jak oczekiwania wpływają na fizyczne parametry organizmu.

Efekt samospełniającej się przepowiedni polega na tym, że oczekiwania jednej osoby wobec sytuacji wpływają na zachowanie osób zaangażowanych w daną sytuację, w taki sposób, że oczekiwania stają się samospełniającą się przepowiednią (Merton, 2005). Na przykład, jeśli ktoś wierzy, że dana osoba jest agresywna, może zachowywać się wobec niej defensywnie, co w rezultacie sprawi, że pierwsza rzeczywiście stanie się bardziej agresywna. Trudno określić kto jako pierwszy opisał omawiany efekt, jednak

w socjologii za pioniera uznaje Thomasa Mertona (2005), który definiuje samospełniające się proroctwo jako „fałszywą definicję sytuacji, wywołującą nowe zachowanie, które powoduje, że początkowo nieprawdziwa koncepcja staje się koncepcją prawdziwą” (Merton, 2005, str. 363). Pomimo tego, że powszechnie uznaje się, że to Merton jako pierwszy szczegółowo opisał wskazane zjawisko, w literaturze socjologicznej znaleźć można wiele wcześniejszych odniesień do opisywanego mechanizmu. W eseju *Samospełniające się proroctwo* (2005) autor wskazuje także, że to właśnie ten efekt był przyczyną m.in. upadku Last National Bank w 1932. To wtedy szybko rozprzestrzeniająca się plotka o niewypłacalności banku spowodowała, że klienci zaczęli masowo wypłacać pieniądze, co w konsekwencji faktycznie doprowadziło do upadku instytucji. Co więcej, Merton (2005) zauważa, że samospełniające się proroctwo leży u podstaw wielu uprzedzeń na tle etnicznym i rasowym³¹. Za przykład wybiera postawy związane z edukacją czarnoskórych obywateli, pokazując, że przekonania o ich mniejszych zdolnościach intelektualnych, prowadzą do domagania się np. przekazywania mniejszych środków na edukację dla tej grupy społecznej, co w rezultacie przekłada się na gorsze wyniki intelektualne.

Jedną z najszerzej opisywanych grup omawianego efektu, są efekty oczekiwań interpersonalnych (Sternberg, 2013). Odnoszą się one do wpływu oczekiwań, opinii lub postaw jednej osoby (nadawcy) na zachowanie, osiągnięcia lub reakcje innej osoby (odbiorcy). Nadawca może determinować także sposoby, zachowania odbiorcy np. metody osiągania celów czy rozwiązywania problemów. Zjawisko to może działać w wielu kontekstach, na przykład w relacjach międzyludzkich, edukacji, pracy czy życiu społecznym. Jego wpływ może mieć charakter pozytywny, jak i negatywny, w zależności od tego, jakie są oczekiwania nadawcy i jak wpływają one na zachowanie czy osiągnięcia odbiorcy. Teorie opisujące omawiany efekt, podkreślają istotną rolę przekonań, oczekiwań i interakcji między ludźmi w kształtowaniu zachowań i osiągnięć członków społeczeństwa. Jego najbardziej znanym przykładem jest efekt Rosenthala (nazywany również Pygmaliona), który dotyczy wpływu oczekiwań nauczyciela na osiągnięcia ucznia. Wiara nauczyciela, że dany uczeń jest bardzo zdolny, może go zmotywować i faktycznie przyczynić się do poprawy jego wyników. Klasyczną już

³¹ Chcę także podkreślić, że określenie „rasy” odnośnie ludzi ma charakter jedynie kulturowy i historyczny, gdyż z biologicznego punktu widzenia w obrębie gatunku ludzkiego nie wyróżnia się ras, a jedynie klasty (Templeton, 2013).

pracą w tym temacie jest artykuł Roberta Rosenthala i Lenore Jacobson (1968), w którym autorzy przedstawiają wyniki swoich badań dotyczących wpływu oczekiwań nauczyciela na ucznia. W szkole podstawowej (*Oak School*) przeprowadzono testy z udziałem uczniów sześciu klas. Test ten przedstawiony był jako sprawdzający potencjał rozwoju intelektualnego dzieci. Następnie nauczycielom przedstawiono, którzy z podopiecznych uzyskali najwyższe wyniki, a co za tym idzie, wskazano uczniów „specjalnie uzdolnionych”. Na koniec roku szkolnego sprawdzono wyniki uczniów klas biorących udział w badaniu i okazało się, że dzieci uznawane przez nauczycieli za wyróżniające się uzyskały wyższy przyrost inteligencji, niż dzieci w grupie kontrolnej. Analizując wyniki 348 badań potwierdzających klasyczny eksperyment w Oak School, Rosenthal i Rubin (1978) wyróżnili cztery czynniki komunikacyjne wpływające na spełnienie oczekiwań nauczycieli:

- a. Atmosfera (*climate*): nauczyciele bardziej życzliwie (*warmer*) podchodzili do uczniów uważanych za wyróżniających się. Zauważono, że życzliwość ta była przynajmniej częściowo komunikowana w sposób pozawerbalny.
- b. Informacja zwrotna (*feedback*): nauczyciele udzielali uczniom uznawanym przez siebie za specjalne więcej zróżnicowanych informacji zwrotnych. Dotyczyły one tego, jak uczeń radzi sobie z wykonywanym zadaniem. Wskazówki te miały charakter zarówno werbalny, jak i niewerbalny.
- c. Materiał (*input*): nauczyciele przekazywali specjalnym uczniom znacznie bardziej bogaty materiał podczas zajęć. Bogactwo materiału polegało zarówno na ilości przekazywanych informacji, jak i poziomie jego trudności.
- d. Możliwości (*output*): nauczyciele dawali specjalnym uczniom zdecydowanie więcej okazji do wypowiedzi podczas zajęć, a także dawali im więcej czasu na odpowiedź.

Kolejnym istotnym efektem oczekiwań w kontekście interpretacji badań z udziałem zwierząt pozaludzkich dotyczących teorii umysłu, jest wpływ eksperymentatora, na uczestników. Efekt ten został wielokrotnie udowodniony w licznych badaniach (Fode, 1960; Rosenthal, Fode, Friedman, Vikan, 1960; Sebeok, T. A., Rosenthal, R. E. 1981). Jednym z najpopularniejszych przykładów tego fenomenu jest opisana niżej (3.1. *Historia badania umysłów zwierząt pozaludzkich*) historia Mądrego Konia Hansa, który czytając pozawerbalne wskazówki ludzi, potrafił odpowiedzieć na każde zadawane mu pytanie. Dla przykładu ludzie zadający pytanie

wymagające liczenia pochylali głowę w dół, aby policzyć uderzenia kopytem, a gdy odpowiednia ilość stuknięć została naliczona, nadawca podnosił wzrok oraz głowę do góry, co było jasnym sygnałem dla konia, że ma przestać wykonywać ruch nogą (Trojan, 2013, min. 8.20-9.15).

Omawiany efekt odnosi się do badań z udziałem wszystkich zwierząt (także ludzi). Chociaż niniejsze opracowanie dotyczy interpretacji badań z udziałem zwierząt pozaludzkich, uważam, że badania z człowiekiem dostarczają wielu istotnych wskazówek także dla interesującego mnie tematu. Dla przykładu dwuetapowy eksperyment, który w I etapie wykorzystuje klasyczne już badanie dotyczące efektu oczekiwań eksperymentatora (Adair, Epstein, 1968). Polega on na poproszeniu uczestników badania o ocenę *czy osoby znajdujące się na prezentowanych im fotografiach odniosły sukces, czy nie*. Ludzie biorący udział w badaniu widzieli na zdjęciach tylko twarze, więc nie mogli wyciągnąć wniosków np. na podstawie ubrań. W etapie II uczestnicy otrzymywali instrukcje przez audiodeskrypcję (bez fizycznej obecności eksperymentatora). Wyniki badań ujawniają, że w obydwu przypadkach uczestnicy odpowiadają zgodnie z oczekiwaniami eksperymentatorów, którzy wcześniej poinformowani zostali o osiągnięciach osób widniejących na zdjęciach. Wyniki tych badań są bardzo istotne, ponieważ pokazują, że do wystąpienia efektu oczekiwań wystarczą wskazówki audialne. To doprowadziło do próby weryfikacji znaczenia wskazówek pozawerbalnych, co ujawniło, że zarówno wskazówki wizualne, jak i audialne są wystarczające do pojawienia się badanego efektu (Zoble, Lehman, 1969).

Pierwszym badaczem, który zauważył wpływ omawianego zjawiska w badaniach z udziałem zwierząt pozaludzkich, jest Maier (1956; Rosenthal, Fode, 1963). Szczegółowy opis celów eksperymentu nie jest konieczny w kontekście niniejszego opracowania, warto jednak powiedzieć³², że dotyczył on zjawiska fiksacji³³ u szczurów. Więc w założeniu, procedura eksperymentalna, miała prowadzić do wykonania określonej reakcji stresowej u uczestników. Autora artykułu *Frustration theory: Restatement and extension* zastanowiły wyniki grupy szczurów, biorące udział

³² W swoim bardzo ciekawym artykule Maier opisał wyniki eksperymentu, w ramach którego wykorzystał teorię frustracji do określenia problemów związanych ze zmianą punktu widzenia w nauce. Jego eksperyment skupiał się na zjawisku fiksacji, czyli nieumiejętności odejścia od konwencjonalnych funkcji obiektów znajdujących się w przestrzeni problemowej (Nęcka, 2013, str. 514-515).

³³ Polega na powtarzaniu przez szczura specyficznej reakcji, gdy zostanie postawiony przed problemem niemożliwym do rozwiązania (odczuwanie frustracji).

w badaniach przeprowadzonych przez studentów Meiera w ramach seminarium. Uczucie porażki i lęk przed karą są czynnikami wywołującymi wysoki poziom fiksacji. Natomiast pomimo to, szczury będące pod opieką jednego z seminarzystów nie wykazywały pożądanej reakcji. Różnica w częstotliwości pojawiania się zachowania stresowego u różnych szczurów, wyjaśniona została przez odniesienie do traktowania zwierząt uczestniczących w badaniu. Maier (1963, str. 375-376) przywołuje anegdotę o dwóch asystentach badawczych, biorących udział w eksperymencie sprzed dwóch lat. Problemy były podobne, jednak każdy pracował nad innym eksperymentem. Jednak tylko jeden z nich uzyskał typową liczbę fiksacji w każdej z czterech grup, z którymi pracował. Drugi asystent, nie był w stanie uzyskać ani jednej fiksacji w żadnej grupie. Od początku sprawdzania przyczyn takich wyników wykluczono aspekt motywacji eksperymentatora, gdyż ten w ramach badań do swojej dysertacji doktorskiej potrzebował zaobserwowania fiksacji. Okazało się jednak, że odczuwał współczucie do szczurów biorących udział w badaniu, co prawdopodobnie powodowało, że głaskał je częściej między próbami w eksperymencie. To natomiast przyczyniło się do obniżenia stresu, a co za tym idzie, zmniejszyło szansę na pojawienie się fiksacji.

Literatura dostarcza jednak wcześniejszych sugestii wpływu eksperymentatora na wyniki badań z udziałem ZINL. Dla przykładu w wymianie listów między Conway'em Zirke (1958) a Gregory'm Razanem (1959) w czasopiśmie *Science*. Dyskusja wywołana została przez artykuł opublikowany w 1957 roku, przez drugiego z wymienionych badaczy. Zawierał on informację o werbalnej deklaracji Pawłowa, który poddał w wątpliwość wyniki eksperymentu z udziałem myszy, którego wyniki opublikował lata wcześniej. Pawłow był żywo zainteresowany teorią Lamarcka, która opiera się na założeniu, że cechy nabyte podczas życia osobnika mogą być przekazywane potomstwu³⁴ (Lamarck, 1994). Nie uważam, że szczegółowy opis lamarckizmu jest konieczny, gdyż aktualnie dominującą teorią ewolucji jest koncepcja

³⁴ Lamarckizm zakłada, że organizmy zmieniają się w odpowiedzi na środowisko poprzez swoje działania i że te zmiany są dziedziczone przez następne pokolenia. Mimo że teoria Lamarcka miała znaczący wpływ na rozwój myśli biologicznej, obecnie jest uznawana za przestarzałą. Odkrycia z dziedziny genetyki i biologii ewolucyjnej wykazały, że większość cech nie jest dziedziczona poprzez nabyte zmiany, ale przez dziedziczenie genów. Należy tu jednak zaznaczyć, że aktualne badania nad epigenomem potwierdzają, że pewne cechy środowiskowe mogą zostać odziedziczone. Dla przykładu aktualnie istnieje wiele doniesień o epigenetycznym dziedziczeniu kulturowej traumy (Lehrner, Yehuda, 2018).

oparta na pojęciu doboru naturalnego, zaproponowana przez Charlesa Darwina. Początkowe badania opublikowane w *Science* zdawały się potwierdzać, założenia Lamarcka (Pawłow, 1923). Myszy uczone były reakcji poszukiwania pożywienia w określonym przez eksperyment miejscu, pod wpływem dźwięku dzwonka. Pawłow deklarował, że zwierzęta w pierwszym pokoleniu potrzebowały 300 lekcji, aby przyzwyczaić je do podążania na miejsce karmienia po usłyszeniu dzwonka, natomiast pokolenie drugie wymagało już tylko 100 prób. Pokolenie trzecie miało prezentować wyuczoną reakcję już po 30 powtórzeniach a ostatnie, które Pawłow widział, wykonywały pożądane zachowanie już po 5 próbach. Dopiero w 1929 roku miał zauważyć, że u podstaw uzyskanych wyników nie leżało lamarckowskie uczenie się, a oczekiwania jego asystenta Studentsova, który zajmował się szkoleniem myszy (Rosenthal, 1981, Broadhurst, 1966). Według doniesień Pawłow miał zwrócić uwagę na to, że uzyskane wyniki nie były efektem szybszego uczenia się zwierząt, a ulepszanego uczenia (*improved learning*) przez eksperymentatora.

2.5. Teoria dysonansu poznawczego

Istotnych informacji dotyczących mechanizmów leżących u podstaw uprzedzeń, dostarcza teoria dysonansu poznawczego stworzona w 1957 roku przez Leona Festingera (2007). Sam dysonans poznawczy definiowany jest jako stan napięcia, jakie powstaje na skutek współistnienia przynajmniej dwóch sprzecznych elementów poznawczych (mogą to być myśli, opinie czy przekonania) lub gdy postawa człowieka jest niespójna (np. występuje sprzeczność komponentu behawioralnego z poznawczym). Jednostka doświadczająca takiego napięcia dąży do jego redukcji (Festinger, 2007, str. 16-18). Uznaje się, że funkcją dysonansu poznawczego jest utrzymanie pozytywnego obrazu samego siebie, a co za tym idzie utrzymaniu samooceny na wysokim poziomie.

Festinger skonstruował swoją teorię na podstawie badań terenowych prowadzonych w latach 1951-1956 (Festinger, 2007). Badacze dołączyli do sekty New Age, której członkowie wierzyli, że przed świętem 21.12.1954 roku nadejdzie koniec świata, jednak osoby wierzące zostaną uratowane przez kosmitów, którzy zjawią się o północy i zabiorą ich na obcą planetę. Celem prowadzonych obserwacji było sporządzenie szczegółowego opisu i analizy zachowania osób należących do kultu, gdy przepowiednia się nie spełni. Szczegółowe zapiski wydarzeń nocy 20.12-21.12.1954 wskazują, że najtrudniejszy czas przypadał na godziny 00:05-4.45 rano, czyli od

momentu, gdy stało się jasne, że przybysze z innej planety nie pojawią się w ramach misji ratunkowej. O godzinie 4.45 doszło do ogłoszenia przez guru sekty — Dorothy Martin — iż wielka wiara i gorliwe modlitwy sprawiły, że Bóg zmienił zdanie i postanowił uchronić świat przed zagładą (Krawczyk, 2022). Hipoteza Festingera mówi, że wyjaśnienie nienadejścia końca świata polegające na przyjęciu nowego przekonania, które pasowało do istniejących struktur poznawczych, pozwala na zachowanie pozytywnego obrazu siebie. Odrzucenie całego systemu przekonań wymagałoby przyjęcia, że *ja* zachowałem się w sposób nieracjonalny, np. rozdając cały swój dobytek w tym m.in. dom. Wybranie wyjaśnienia korzystniejszego dla własnego *ego*, prowadziło do wzmożonej aktywności członków sekty polegającym na nawracaniu innych ludzi.

Chociaż pierwsze wydanie książki Festingera zawierająca szczegółowy opis teorii dysonansu poznawczego, pojawiło się w 1957 roku, już w 1955 podczas Symposium on Cognition można było usłyszeć wystąpienie zawierające kilka informacji dotyczących toczących się prac autora. Jack William Brehm zainspirowany wysłuchanym wystąpieniem przeprowadził własny eksperyment, który wykazał, że dysonans poznawczy sprawia, że ludzie postrzegają wybraną przez siebie alternatywę jako lepszą niż przed jej dokonaniem (Brehm, 1956). Brehm poprosił kobiety o ocenę 8 produktów, a następnie dał im możliwość wyboru jednego z dwóch ocenianych urządzeń. Po dokonaniu decyzji niektóre z uczestniczek usłyszały informacje dotyczące zalet i wad produktów, między którymi dokonywały wyboru. Następnie poproszono o ponowną ocenę ośmiu urządzeń. Wyniki pokazały, że u uczestniczek, którym prezentowano informacje dodatkowe, wytworzony został dysonans poznawczy, który był redukowany przez uczynienie wybranej alternatywy bardziej atrakcyjną. Wskazuje to na fakt, iż omawiane napięcie oraz towarzyszące mu mechanizmy pomagają w godzeniu się z podjętymi decyzjami, oraz potwierdzaniu ich słuszności.

Teoria dysonansu poznawczego głosi także, że im bardziej decyzja człowieka jest obciążająca, w kategoriach czasu, pieniędzy, wysiłku czy niedogodności (np. dyskomfort czy ból) i im bardziej nieodwracalne są jej konsekwencje, tym silniejsze jest odczuwane napięcie. Zjawisko to dobrze prezentuje klasyczny już eksperyment przeprowadzony przez Aronsona i Millsa (1959), który pokazał, że im większy trud (odczuwalne poczucie wstydu) zostanie włożony w jakieś działanie, tym wyższy poziom zadowolenia z osiągnięcia celu. W takiej sytuacji zadowolenie będzie wyższe niż gdyby jednostka uzyskała taką samą korzyść mniejszym kosztem. Grupą badawczą

były studentki Uniwersytetu Stanforda, które otrzymały zaproszenie dołączenia do grupy zajmującej się psychologią seksu, osoby zainteresowane proszone były o przybycie na spotkanie. Uczestniczki podzielone zostały na 2 grupy badawcze oraz jedna kontrolną. Po wyjaśnieniu zasad funkcjonowania dyskusji, wszystkim uczestniczką zadano pytanie: *czy uważają, że będą w stanie bez skrępowania rozmawiać o seksie w przedstawionych im warunkach?* Po udzieleniu odpowiedzi twierdzącej kobiety z grupy kontrolnej usłyszały nagranie z rzekomo toczącej się w danym momencie dyskusji. Tutaj należy zaznaczyć, że przygotowana rozmowa miała być w założeniu nudna i bezwartościowa. Natomiast kobiety dwóch pozostałych grup poproszone zostały o odczytanie tekstów o zróżnicowanym poziomie trudności pod względem potencjalnego wstydu, jaki miały odczuwać. Następnie odsłuchiwały tego samego nagrania, co osoby z grupy kontrolnej. Po zakończeniu nagrania proszono uczestniczki o ocenę odsłuchanego materiału. Wyniki pokazały, że kobiety, które przechodziły inicjację o najwyższym poziomie trudności, oceniały usłyszaną rozmowę najwyżej. Autorzy uważają, że u osób, które dobrowolnie godzą się na dyskomfortowe doświadczenia, aby być częścią określonej grupy społecznej, odczuwają dysonans poznawczy, który redukuje przez wyższy poziom zadowolenia z osiągnięcia określonego celu. Napięcie powstaje, gdyż udzielnie zgody na poniżającą aktywność, aby dostać się do grupy dyskusyjnej, wymaga przyznania, że decyzja własna była nie do końca racjonalna (Aronson, Travis, 2007).

Innym ciekawym badaniem w kontekście wpływu nieodwracalności decyzji na poziom dysonansu poznawczego, badano decyzje dokonywane podczas zakładów wyścigowych (Knox i Inkster, 1968). W ramach eksperymentu uczestnicy zobowiązali się obstawić \$2 na konie wyścigowe w naturalnych warunkach (podczas prawdziwego wyścigu). Zadaniem było oszacowanie szansy na wygraną, obstawionego przez siebie konia, jedna grupa uczestników wykonywała je przed podejściem do okienka bookmacherskiego, natomiast drugiej zadawano pytanie natychmiast po dokonaniu zakładu. Okazało się, że osoby z grupy drugiej szacowały wyższą szansę wygranej, niż osoby z grupy pierwszej. W kolejnym eksperymencie osoby obstawiające w wyścigach oceniały pewność trafności swojego wyboru konia przed i po dokonaniu zakładu (w siedmio-stopniowej skali). Także wyniki tego eksperymentu potwierdziły założenia teorii dysonansu poznawczego, czyli wykazały, że pewność rośnie wraz z podjęciem decyzji.

Wielu autorów wskazuje, że większość ludzi uważa siebie za osoby moralne oraz posiadające wysoką samoocenę w tym aspekcie (Sykes, Mtza, 1957; Siemaszko, 1993; Aronson, Aronson, 2007; Wypler, 2013). Co za tym idzie, starają się przestrzegać obowiązujących norm społecznych, a zachowanie z nimi niezgodne powoduje pojawienie się dysonansu poznawczego, wyrzutów sumienia i samopotępienia. System przekonań dotyczących siebie wyposażony w samoocenę, określany jest mianem *ja* lub koncepcji *ja*, a poznanie i zachowanie jednostki ukierunkowane są na utrzymanie jego spójności. Wykazuje on także tendencję do zniekształcania informacji, interpretowania oraz reinterpretowania zdarzeń w sposób pozwalający na utrzymanie pozytywnego obrazu siebie (Pervin, John, 2002). Zdarza się także, że jednostki aktywnie zbierają informacje spójne z koncepcją *ja* (Wypler, 2013). Jest to przyczyna, przez którą ludzie unikają przekonania o tym, że są uprzedzeni, gdyż zdanie sobie sprawy z występowania u siebie takich postaw wywołuje dysonans poznawczy (Travis, Aronson, 2007, str.76). Jest tak, ponieważ uprzedzenia uznawane są za moralnie lub społecznie niewłaściwe, co koliduje z przekonaniem, że są ludźmi moralnym, a ich antypatie opierają się na racjonalnych i uzasadnionych przesłankach. Nieprzyjemne napięcie zredukować można na trzy sposoby: poprzez zmianę zachowania, przyjęcie nowego przekonania (usprawiedliwienia) oraz odrzucenie przekonania (Aronson, Aronson, 2007, str. 171). Okazuje się jednak, że rzadko odbywa się to w ramach pierwszej z wymienionych strategii. Ludzie są bardziej skłonni przyjąć nowe przekonanie (pośredniczące), pozwalające utrzymać istniejące już struktury poznawcze w postaci stereotypów i mitów lub odrzucić często całe zespoły przekonań stojące w sprzeczności z istniejącymi już strukturami poznawczymi, lub przyznać, że ich zachowanie było niewłaściwe (Festinger, 1962). Redukcja dysonansu poznawczego może mieć charakter natychmiastowy, jak i rozciągnięty w czasie. W pierwszej sytuacji człowiek doświadczający dysonansu, nawet go nie zauważa, natomiast w drugim przypadku jednostka początkowo zdaje sobie sprawę z niewłaściwego zachowania, jednak wraz z upływem czasu zniekształca wspomnienie, przedstawiając sytuację w zupełnie innym świetle niż początkowo. Co więcej, jednostka dokonująca takiego zniekształcenia jest pewna, że od początku myślała o zdarzeniu w taki sam sposób. Zatem uważa się, że mechanizmy te mają charakter obronny, automatyczny i często nieuświadomiony (Wypler, 2013). Człowiek dysponuje licznymi mechanizmami pozwalającymi na realizację wyżej opisanych strategii neutralizacji dysonansu poznawczego. W niniejszej części pragnę skupić się na trzech z nich: samousprawiedliwieniu, efekcie

potwierdzenia i egotyzmie atrybucyjnym. Należy zaznaczyć, że strategie te nie funkcjonują jako autonomiczne, a raczej tworzą pewien zintegrowany system.

W pierwszej kolejności skoncentruję uwagę na egotyzmie atrybucyjnym, który jako cecha pamięci bierze udział kształtowaniu wspomnień. Szczególnie wymieniany jest jako jedna z przyczyn zapominania lub zniekształcania zapamiętywanych wydarzeń. Polega na łatwiejszym zapominaniu informacji niezgodnych z własnymi przekonaniem. Jej konsekwencją jest tendencja do wyjaśniania zachowań własnych w zależności od ich efektów, tj. sukcesy tłumaczone są czynnikami wewnętrznymi (zdolności czy elementy charakteru), a porażki czynnikami sytuacyjnymi. Wiąże się z wyolbrzymianiem własnych zasług, przy jednoczesnym zaniżaniu odpowiedzialności za własne niepowodzenia (Mezulis, Abramson, Hyde, Hankin, 2004). Jones i Kohler, (1958) przeprowadzili dwa eksperymenty, w ramach których wykazali łatwość zapamiętywania twierdzeń spójnych z przekonaniem, jakie prezentowali uczestnicy badania oraz trudność (mierzoną przez szybkość zapominania) w zapamiętywaniu twierdzeń niezgodnych z postawami przez nich prezentowanymi. Autorzy potwierdzili trzy hipotezy:

- a. Uczestnicy badania o poglądach prosegregacyjnych ucą się twierdzeń wiarygodnych na temat prosegregacji i nieprzekonujących dotyczących antysegregacji, znacznie skuteczniej, niż nieprzekonujących twierdzeń prosegregacyjnych i wiarygodnych twierdzeń antysegregacji.
- b. Uczestnicy o poglądach antysegregacyjnych ucą się twierdzeń wiarygodnych na temat antysegregacji i nieprzekonujących dotyczących prosegregacji, znacznie skuteczniej, niż nieprzekonujących twierdzeń antysegregacyjnych i wiarygodnych twierdzeń prosegregacyjnych.
- c. Uczestnicy o poglądach neutralnych lub pośrednich ucą się równie dobrze wszystkich typów stwierdzeń.

Gdy dojdzie do naruszenia osobistych standardów jednostki, ma ona skłonność do wykonania reakcji samousprawiedliwienia (Bandura, 2007). Mechanizm ten pozwala ludziom na zachowanie pozytywnego obrazu siebie (przynajmniej utrzymanie samooceny na tym samym poziomie, a niekiedy nawet do jej podniesienia, przy jednoczesnym przekraczaniu własnych granic moralnych). Crendal i Eshleman (2003) wskazują, że usprawiedliwianie pozwala na wyrażanie swoich negatywnych przekonań

o określonej grupie, zapewniając bezpieczeństwo i chroniąc przekonanie o egalitarnym charakterze własnych działań. Uczynienie z pewnej grupy niewolników oraz odbieranie im podstawowych praw, powoduje przyjmowanie przekonań usprawiedliwiających takie traktowanie (Travis, Aronson, 2007, str. 76). Unikamy w ten sposób wyrzutów sumienia wynikających z własnego złego zachowania. Badania empiryczne dotyczące dyskryminacji i stereotypów pokazują, że niekiedy przyjmowanie nowych przekonań usprawiedliwia dyskryminację, a innym razem dyskryminacja usprawiedliwia przyjmowanie nowych negatywnych przekonań (Travis, Aronson, 2007). W literaturze znajdziemy opisy wielu technik neutralizujących negatywne skutki niespójności koncepcji *ja*, są to (Tsang, 2002; Kosslyn, Rosenberg, 2006; Bandura, 2007, Sykes, Matza, 1957):

- a. Przerzucanie winy i dehumanizację ofiar,
- b. Uchylenie się od odpowiedzialności za wykonywanie zachowanie,
- c. Ignorowanie negatywnych konsekwencji zachowania np. przez unikanie informacji dotyczących jego negatywnych skutków lub łagodzenie szkód z niego wynikających,
- d. Zmianę w interpretacji zachowania w sposób przedstawiający je jako umożliwiającą osiągnięcie ważniejszego celu moralnego,
- e. Porównanie zachowania do zachowań innych ludzi oraz argumentowanie, że w takiej perspektywie nasze działania nie są takie złe, jak mogłoby się wydawać,
- f. Popieranie zachowania zasadami religijnymi, ideologicznymi lub porządkiem społecznym,
- g. Używanie eufemizmów podczas opisów zachowania,
- h. Potępienie potępiających, czyli zaatakowanie osoby krytykującej zachowanie jednostki. Realizowane jest przez skupienie się na motywacjach i zachowaniach osoby wysuwającej zarzutu, co służy zdyskredytowaniu jej opinii.

Sykes i Matza (1957) uważają, że techniki redukcji nieprzyjemnych skutków wynikających z braku spójności *ja*, angażowane są przed wykonaniem zachowania, co umożliwia behawioralne naruszenie normy moralnej. Co ważne, usprawiedliwienie jednego złego zachowania określonego typu, przygotowuje grunt pod kolejne akty dyskryminacji, w tym agresji (Travis, Aronson, 2007). Gdyż wykonanie agresywnego zachowania wywołuje dysonans poznawczy, w wyniku czego pojawia się potrzeba jego

redukcji, co może prowadzić do potwierdzenia słuszności własnych działań. Jak w przypadku badanej przez Festingera sekty, ludzie próbują to osiągnąć poprzez powielanie zachowania, np. namawiając innych do jego wykonywania.

Inną kluczową strategią pomagającą w zredukowaniu dysonansu, jest potwierdzanie słuszności swoich zachowań i struktur poznawczych, co możliwe jest dzięki efektowi potwierdzenia/konfirmacji. Definiowany jest on jako tendencja do utrzymania lub niechęci do porzucenia aktualnie preferowanej hipotezy/przekonań/opinii/poglądów (Klayman, 1995, str. 386). Omawiany efekt może przyjmować co najmniej pięć form (Klayman, 1995):

- a. Po pierwsze może przejawiać się nadmierną pewnością odnośnie do swoich przekonań, nawet po otrzymaniu nowych danych podważających przyjmowaną hipotezę/teorię/model/opinię.
- b. Po drugie może polegać na selektywnym wyszukiwaniu danych, co zniekształci badany temat na korzyść własnego stanowiska.
- c. Po trzecie interpretacja informacji może być stronnicza, czyli dane potwierdzające mogą być bezpodstawnie uznane za wiarygodne, a dane przeczące za niewiarygodne.
- d. Po czwarte może przejawiać się przez niewłaściwą aktualizację hipotezy, przyjmując przekonanie na temat niskiej siły zebranych dowodów.
- e. Po piąte nawet gdy jednostka postanowi o konieczności odrzucenia swoich twierdzeń, może mieć problem z tworzeniem alternatywnych wyjaśnień.

Omawiany efekt postrzegany jest jako ograniczenie ludzkich systemów poznawczych lub jako mechanizm motywacyjny³⁵ (Nickerson, 1998, str. 205).

Mocno ugruntowane w kulturze (także tej popularnej) uprzedzenia w stosunku do ZINL, składają się z komponentu poznawczego, tj. stereotypów i mitów charakterystycznych dla danej społeczności oraz komponentu afektywnego, który

³⁵ Efekt potwierdzenia jest mechanizmem, motywującym do podjęcia pewnych działań np. aktywowany jest w sytuacji, gdy jednostka odczuwa potrzebę do posiadania racji – podczas dyskusji. Dla przykładu w sytuacji społecznej, jaką jest przekonywanie innych ludzi do swojej racji (Nickerson, 1998, str. 205), jednostka może namawiać innych do przyjęcia postawy przynajmniej podobnej do własnej. W przypadku członków społeczności badanej przez zespół Festingera, członkowie kultu wykazywali tendencję do zachęcania innych ludzi do przyjęcia swojej religii, a co za tym idzie, werbowania do swojej sekty nowych członków (Festinger, 2007). Gdyż nawrócenie innego wcześniej nieprzekonanego człowieka traktowane jest jako potwierdzenie słuszności własnego wyboru.

zawiera w sobie wszystkie afekty określonej jednostki zarówno te bez świadomej refleksji, jak i te świadome. Dla przykładu dysonans poznawczy może być zredukowany bez użycia świadomej refleksji, między innymi dzięki efektowi potwierdzenia. Pomaga w zmniejszeniu zużycia zasobów poznawczych kosztem dokładności wykonywanych procesów. Z kolei trzeci komponent, behawioralny, odnosi się do dyskryminacji manifestującej się, poprzez stronnicze analizowanie uzyskanych danych, co w konsekwencji prowadzi do potwierdzania swoich przekonań i spełniania własnych oczekiwań.

Silnie zakorzenione opinie mogą przetrwać całkowitą dyskredytację dowodów, które je zapoczątkowały (Lord, Ross, Lepper, 1979). Osoby posiadające zdecydowane opinie na temat złożonych problemów społecznych, analizują doniesienia empiryczne w sposób stronniczy. Chętniej akceptują informacje potwierdzające ich opinie, jednocześnie poddając surowej krytyce dowody im przeciwnie. W wyniku czego wyciągają wnioski na korzyść własnych struktur poznawczych. To doprowadziło do postawienia hipotezy, że gdyby podczas debaty społecznej, wystawić dwie przeciwnie strony na ten sam zbiór informacji, to zamiast doprowadzenia przynajmniej do przekonania, że sprawa jest zdecydowanie bardziej złożona, niż wcześniej się wydawało, sprawi to, że ludzie będą utrzymywać swoje struktury poznawcze, odrzucając argumenty sprzeczne z ich poglądami, co doprowadzi do silniejszej polaryzacji. Dzieje się tak, ponieważ sama dostępność sprzecznych dowodów rzadko jest wystarczająca, aby skłonić człowieka do zmiany swoich postaw i przekonań. Aby sprawdzić hipotezę o zwiększeniu polaryzacji na skutek przedstawiania różnych dowodów dwóm grupom sporu, poproszono zwolenników i przeciwników kary śmierci o przeczytanie dwóch artykułów naukowych. Pierwszy potwierdzał skuteczność kary śmierci w odstraszaniu od popełniania brutalnych przestępstw, drugi natomiast krytykował stanowisko, że kara śmierci jest skutecznym narzędziem w odstraszaniu potencjalnych brutalnych przestępców. Czyli jeden artykuł był zgodny drugi niezgodny z poglądami uczestnika. Zarówno zwolennicy, jak i przeciwnicy kary śmierci oceniali argumenty, które były spójne z ich przekonaniami za bardziej wiarygodne i niezwykle wartościowe. Natomiast ostro krytykowali te niezgodne ze swoimi strukturami poznawczymi. Czyli w konsekwencji wszyscy utwierdzali się w swoich opiniach, co prowadziło do zwiększenia poziomu pewności wobec własnych stanowisk. To oraz inne badania wskazują, że silnie zakorzenione przekonania mogą nie tylko przetrwać ekspozycję na sprzeczne informacje, ale także pozostają niezmienione nawet w sytuacji całkowitej

destrukcji dowodów, które je zapoczątkowały (Ross, Lepper, Hubbard, 1975; Ross, 1977). Potwierdza to stanowisko głoszące, że sama dostępność sprzecznych dowodów rzadko zdaje się wystarczająca, aby skłonić jednostkę do zmiany postaw i przekonań.

W innym badaniu — Westerna i współpracowników (2006) — prowadzonym w zakresie neuropsychologii, wykonywano pomiar za pomocą funkcjonalnego obrazowania metodą rezonansu magnetycznego (fMRI). Jest to pierwszy eksperyment prezentujący neuronalne korelaty rozumowania motywacyjnego, a konkretnie konstrukcję ukrytej regulacji afektów obrony psychologicznej – neutralizacji dysonansu poznawczego, za pomocą efektu potwierdzenia. Jest to także pierwsza praca prezentująca neuronalne korelaty tworzenie sądów (*political judgment*) i podejmowania decyzji politycznych. W jego trakcie mierzono aktywność neuronalną podczas doświadczenia dysonansu poznawczego. Osobom biorącym udział w badaniu prezentowano zdania dotyczące określonych polityków. Ich treść była zgodna albo niezgodna z poglądami politycznymi osób biorących udział w badaniu. Czyli jeżeli ktoś żywił antypatię do konkretnych kandydatów, prezentowano mu twierdzenia pozytywne o tym polityku, a do polityków obdarzanych sympatią, przyporządkowane były twierdzenia negatywne. Takie przedstawienie bodźców prowadziło do wzbudzenia dysonansu poznawczego, natomiast chęć jego neutralizacji motywowała system do usunięcia sprzeczności z koncepcji *ja*. Wykonane badanie potwierdza odrębność struktur neuronalnych odpowiedzialnych za rozumowanie bez silnych afektywnych motywatorów (np. za pomocą dysonansu). Okazało się, że twierdzenia sprzeczne ze strukturami poznawczymi jednostki (czyli zagrażającymi koncepcji *ja*), przetwarzane są obwodami neuronalnymi odpowiedzialnymi za obieg danych afektywnych. Natomiast obwody neuronalne odpowiedzialne za racjonalne rozumowanie, przetwarzają głównie informacje zgodne z własnymi przekonaniami. Wyniki przedstawiają także podstawy neuronalne dla ukrytej i ponownej regulacji oceny informacji. Grzbietowo-przyśrodkowa kora mózgowa powiązana z odniesieniem do siebie i odczuwaniem sympatii, wykazuje inny wzorzec aktywności podczas ukrytej regulacji afektu oraz regulacji z udziałem świadomości. Gdy uczestnik rozumował w sposób motywowany, tylna część zakrętu obręczy była aktywna (czyli podczas dokonywania oceny działania zarówno własnego, jak i innych co możliwe jest dzięki neuronom lustrzeanym). Gdy informacje przetwarzane są drogą motywowaną jednostka jest bardziej skłonna wybaczyć określone działanie, także w przypadkach dylematów moralnych. Gdy uczestnicy konfrontowani byli z informacjami angażującymi koncepcję *ja*, w pewnym

momencie widoczna była duża aktywność w obszarach odpowiedzialnych za przetwarzanie emocjonalne, co odzwierciedlało ulgę lub nawet nagrodę wywołaną zwróceniem stabilności poprzez wytworzenie przekonania pozwalającego na jej osiągnięcie.

Wiele rozważań wskazuje na to, że efekt potwierdzenia jest równie powszechny w nauce (Holton, 1973; Cohen, 1985; Nickerson, 1998). Wniosek ten jest szczególnie interesujący m.in. ze względu na charakterystykę demarkacji nauki, a konkretnie na regułę weryfikacji nie wprost stosowaną w jej obszarze. Zasada falsyfikacji głosi, że teoria/hipoteza/model nie mogą zostać uznane za naukowe, jeśli są niepodatne na empiryczną falsyfikację (Popper, 2015). Oznacza to, że weryfikacja ta nie powinna być nakierowana na potwierdzenie, a na znalezienie dowodów ich błędności. Zakłada się, że w systemie nastawionym na falsyfikację możliwe jest znacznie szybsze wykrycie fałszywości proponowanych twierdzeń. Przyjmuje się także kryterium replikacji, co oznacza, że wszystkie eksperymenty powinny być powtarzane w dużej ilości, aby sprawdzić, czy wyniki są replikowalne, a także określić zróżnicowanie wyników i stworzyć kolejne możliwe interpretacje rezultatów badań, gdyż dowody naukowe opierają się nie na ostateczności dowodów, a ich przewadze pod względem ilościowym (Nickerson 1998, str. 207). Pomimo założeń zasady falsyfikacji w systemie nastawionym na weryfikację pozytywną, łatwo ulegamy efektowi konfirmacji. Jest to spowodowane wieloma czynnikami, np. trudnościami związanymi z publikacją negatywnych wyników badań oraz replikacji eksperymentów (Chambers, 2017). Należy przyjąć, że bardzo wielu eksperymentatorów ulega mechanizmom ukrytej regulacji afektów, czyli opisanemu wyżej samousprawiedliwianiu, poruszanemu teraz efektowi potwierdzenia czy omówionemu niżej egotyzmowi atrybucyjnemu. Zjawisko to jest tak powszechne, gdyż są to naturalne i silnie zakorzenione elementy systemu poznawczego obecne u każdego człowieka. Natomiast zasady demarkacji nauki sugerują obiektywność wyników uzyskanych na drodze empirycznej. To skutkuje ignorowaniem własnych uproszczeń rozumowania. Ten brak uważności prowadzi do przetwarzania informacji za pomocą rozumowania motywowanego zamiast racjonalnego. Stronnicze analizowanie i interpretowanie danych wspierane jest także przez system publikowania danych eksperymentalnych (Travis i Aronson, 2007; Chambers 2017).

Travis i Aronson (2007) czy Chambers (2017) wskazują, że mechanizm konfirmacji odpowiedzialny jest za uleganie *argumentum ad ignorantiam*, czyli argumentowi odwołującemu się do braku dowodu świadczącego za określoną tezę. Gdy

są one skuteczne, wyciągany jest wniosek, o fałszywości bronionej przez oponenta tezy. Problem *ad ignorantiam* przez lata powszechnie opisywany był w literaturze, przykładem są wspomniane wcześniej rozważania de Waala (2018, str. 24), który zwraca uwagę na to, że naukowcy zajmujący się prowadzeniem badań z udziałem zwierząt pozaludzkich, wykazują tendencję do uznawania braku dowodu, za dowód na brak. Ignorują przy tym fakt, że argument *ad ignorantiam* bywa zawodny. Schopenhauer (2007) przedstawia omawiany zabieg w Erystyce jako pozycję 37 (Schopenhauer, 2014). W opisie zaznacza, że wystarczy wybrać jeden słaby argument oponenta, odeprzeć go i podkreślić, że skoro przeciwnik w dyskusji nie znalazł lepszej obrony własnej tezy, podważono jej kluczowe twierdzenia. Dlatego argument *ad ignorantiam* często wykorzystywany jest w debatach dotyczących problemów społecznych wzbudzających duże emocje. Wielu autorów i niektóre podręczniki do logiki uznają omawiany argument za błąd, istnieją także autorzy przyjmujący stanowisko, że istnieją argumenty *ad ignorantiam* mające charakter merytoryczny (Walton, 1999a, 1999b; Wagemans, 2003, Wiecek, 2017; Van Eemeren, Grootendorst, i in., 2002, str. 23-30; Wagemans, 2003; Chambers, 2017; de Waal, 2018).

Wiedząc, że traktowanie braku dowodu, jako dowodu na brak może być wynikiem ulegania efektowi potwierdzenia, należy zastanowić się, które eksperymenty z udziałem zwierząt pozaludzkich opierają się na racjonalnych argumentach z niewiedzy, a które na motywowanych. Opisane wyżej badania wskazują, że doniesienia empiryczne odmawiające gibbonom i słoniom zdolności do korzystania z narzędzi są wynikiem błędów w procedurach eksperymentalnych (Beck, 1967; Foerder, Galloway, i in., 2011). Wniosek ten wynika z wieloletniej historii relacji człowiek — inne zwierzęta, charakterystycznej dla kultury eksperymentatora i osób interpretujących dane oraz z założenia o powszechności opisywanych w niniejszym podrozdziale mechanizmów. Ugruntowane kulturowo postawy są mocno zakorzenione w ludzkich systemach poznawczych. Stronnicze konstruowanie i interpretowanie eksperymentów empirycznych (dyskryminacja), wynika ze stereotypów i mitów (komponent poznawczy) związanych z przekonaniami, że człowiek różni się od innych zwierząt pod względem posiadania określonych zdolności poznawczych. Czyli w koncepcji *ja* istnieją przekonania dotyczące *wyjatkowości umysłu człowieka*. W takiej sytuacji wiele informacji wskazujących, że zwierzęta pozaludzkie posiadają zdolności uznawane za dystynktywne dla człowieka, będą skutkować pojawieniem się dysonansu poznawczego (komponent afektywny). Potrzeba jego redukcji będzie wywoływała

tendencję do odmawiania ZINL złożonych zdolności poznawczych (komponent behawioralny).

W celu przybliżenia zagadnienia argumentów z niewiedzy, porównam dwie koncepcje — Waltona i van Eemeren i Grootendorsa. Pierwszy wymieniony autor stoi na stanowisku głoszącym, że w niektórych przypadkach użycie argumentu *ad ignorantiam* jest racjonalne (*reasonable*); (Walton, 1999a, 1999b). Natomiast Eemeren i Grootendors (2002) uważają, że sam argument z niewiedzy zawsze jest błędny (*fallacious*). Zaznaczają natomiast, że istnieją argumenty, które udają *ad ignorantiam*, ale w rzeczywistości są do nich tylko powierzchwniowo podobne. Porównanie dwóch wyżej wskazanych koncepcji prowadzi do konkluzji, że różnica między analizą poszczególnych przypadków argumentów ma charakter terminologiczny (Wagemans, 2002). Mianowicie, gdy Walton mówi o rozsądnym użyciu *ad ignorantiam*, von Eemeren i Grootendors twierdzą, że są to argumenty powierzchwniowo przypominające te z niewiedzy, ale zasada domknięcia nie została naruszona.

Walton dokonuje obserwacji, że pewnym argumentom formalnie odpowiadającym tym z niewiedzy, nie można odmówić wiarygodności, natomiast inne są nieważne (*invalidity*), ale nie błędne (*fallacious*). Powołuje się na przykład Robinsona (1971), który porusza temat argumentów *ad ignorantiam* pojawiających się w nauce.

(1) Nigdy nie udowodniono, że quintczana jest szkodliwa dla człowieka.

Walton uważa, że powyższa konkluzja jest racjonalna, ponieważ opiera się nie tylko na niewiedzy, ale też na dowodach empirycznych z użyciem quintczana. Czyli *ad ignorantiam* nie jest błędne, gdy istnieje możliwość eksplikacji przesłanki „gdyby zaszło x, znaleźlibyśmy y”, która jest wskaźnikiem zasadności omawianego argumentu. To związane jest z przeniesieniem ciężaru dowodu (konkretnie odwróceniem)³⁶, w kontekście jego użycia (Wagemans, 2003). Racjonalny wariant *ad ignorantiam* opiera się na kryterium złych konsekwencji, gdzie ważna jest niemożliwość gromadzenia dowodów, a raczej powaga jego skutków w przypadku braku określonej wiedzy. Dla

³⁶ Ciężar dowodu rozumiany jako jeden z parametrów dialogu argumentacyjnego. Obowiązek jego spełnienia obciąża uczestnika dialogu broniącego (A), gdy ten wysuwa określoną tezę (T). Gdy zaprezentowane są kolejne twierdzenia (R) i (S), nakładany jest ciężar dowodu (Walton, 2004).

przykładu, obchodzenie się z bronią, o której nie wiemy, czy jest naładowana i z jakiegoś powodu nie możemy tego sprawdzić, powinno skutkować zachowaniem, tak jakby broń była załadowana. Natomiast błędność omawianego argumentu, wynika z trudności z wywiązaniem się z ciężaru dowodu, zwłaszcza gdy trudno znaleźć alternatywę (co jak opisałam wyżej, może być powodowane przez efekt potwierdzenia).

W ramach stanowiska pragma-dialektycznego podzielanego przez von Eemeren i Grootendorsa, błędy rozumiane są jako naruszenie zasad dotyczących tego, jakie konsekwencje powinny wynikać z niepowodzenia obrony stanowiska w końcowej fazie dyskusji (2002, str. 23-30). Jedną z nich jest reguła domknięcia, która głosi, że: nieudana obrona stanowiska musi skutkować wycofaniem swoich twierdzeń, natomiast, gdy obrona jest skuteczna, to oponent musi wycofać swoje wątpliwości. Zasada domknięcia może zostać złamana na wiele sposobów, np. poprzez uznanie prawdziwości stanowiska, na podstawie jego słabej obrony – co określa się mianem „absolutyzacji niepowodzenia obrony”, zjawisko to odpowiada *argumentum ad ignorantiam*. Von Eemeren i Grootendors wyjaśniają, że o argumencie z niewiedzy mówić można w przypadku popełnienia dwóch błędów utrudniających rozstrzygnięcie sporu. Po pierwsze osoba, której argumenty zostały uznane za nieważne lub błędne ignoruje ciężar dowodu związany z własną rolą w sytuacji komunikacyjnej, po drugie ignoruje możliwość przyjęcia neutralnego stanowiska. Ulega w ten sposób złudzeniu, że dyskusja musi się zakończyć skutkiem pozytywnym (przyjęcie określonego stanowiska) albo negatywnym (odrzućcie określonego stanowiska). Pragma-dialektyczna perspektywa wskazuje, że istnieją argumenty odpowiadające temu, co Walton nazywa racjonalnym zastosowaniem *ad ignorantiam*, jednak są one do nich tylko powierzchwnie podobne. Dla przykładu zasada domniemania niewinności nie spełnia kryteriów bycia argumentem z niewiedzy, ponieważ jej konkluzja polega na tym, że oskarżonemu nie udowodniono winy, nie zaś, na tym, że oskarżony jest niewinny. W wyniku czego nie dochodzi do absolutyzacji niepowodzenia obrony. Inną bardziej istotną dla niniejszej pracy formą łamania zasady domknięcia jest sytuacja, w której przyjmujemy za tezę twierdzenie, które w rzeczywistości jest jedną z możliwych hipotez. Może się tak zdarzyć, gdy próba sfalszowania tezy/hipotezy nie powiedzie się. Dochodzi wtedy do absolutyzacji niepowodzenia obrony, co sprawia, że argument powinien zostać uznany za nie ważny.

Ciekawym zjawiskiem społecznym pokazującym jak wykorzystuje się *ad ignorantiam* do potwierdzania własnych przekonań, jest panika moralna³⁷, która prezentuje fakt, że człowiek ulegający efektowi potwierdzenia jest odporny na logiczne argumenty³⁸ (Travis, Aronson, 2007). Świetnym przykładem jest panika satanistyczna (*satanic panic*), która opanowała społeczeństwo amerykańskie w latach 80-tych (Huges, 2016). Zjawisko to dotyczyło szybko rozprzestrzeniającego się lęku, przed rzekomo działającymi na terenie USA sektami satanistycznymi. Organizacje te miały porywać i mordować dzieci w celach rytualnych. Gdy FBI ogłosiło wyniki dochodzenia w tej sprawie, głosząc, że nie znaleziono żadnych, dowodów na istnienie tak złożonej szajki przestępczej, osoby ulegające panice uznały to za dowód istnienia takiego przedsięwzięcia. Tłumacząc wynik śledztwa rzekomymi wpływami sekt oraz uznali, że ich członkowie są bardzo przebiegli i mają swoich ludzi wszędzie.

2.6. Słodka reakcja

Kolejnym mechanizmem, któremu warto się bliżej przyjrzeć, jest słodka reakcja (*cute response*), która stanowi silny motywator pozabiologiczny, popychający ludzi do działań na rzecz ZINL. Tym samym stanowi podstawę motywacyjną do tworzenia kampanii edukacyjnych, czy inicjatyw terapeutycznych (z udziałem zwierząt pozaludzkich); (Borgi, Cogliati-Dezza, Brelsford, Meints, Cirulli, 2014). Jak wspomniałam wcześniej, słodka reakcja pojawia się w odpowiedzi na bodziec w postaci twarzy charakteryzującej się określonym zestawem cech nazywanych schematem

³⁷ Panika moralna stanowi rodzaj niepokoju społecznego (Zielińska, 2004), rozumianego jako szybko rozprzestrzeniający się lęk w określonej społeczności. Pojawia się najczęściej w sytuacji kryzysu lub transformacji systemowej oraz wiąże się z trudnościami w przystosowaniu się do zachodzących w zmian. Panika moralna ograniczona jest czasowo i przestrzennie, a jej źródłem jest naruszenie porządku społecznego lub podzielanych przez daną społeczność wartości.

³⁸ Technika perswazyjna, która polega na przedstawieniu interlokutorowi logicznych argumentów opiera się na zmianie relacji między osobą lub grupą, której dotyczą uprzedzenia (obiektem) a atrybutem im przypisywanym (Orosz i inn, 2016). Do wywierania wpływu poprzez jej realizację, stosuje się narzędzia takie jak np. kampanie społeczne czy filmy dokumentalne. Badania z zakresu zmiany postaw i przekonań wskazują, że technika ta osiąga w najniższy współczynnik skuteczności (Orosz i inn, 2016). Co ciekawe, połączenie jej z inną, której celem jest operowanie poziomem identyfikacji między osobą, która posiada pewne określone przekonania a obiektem, którego te przekonania dotyczą, uzyskuje najwyższy wskaźnik skuteczności.

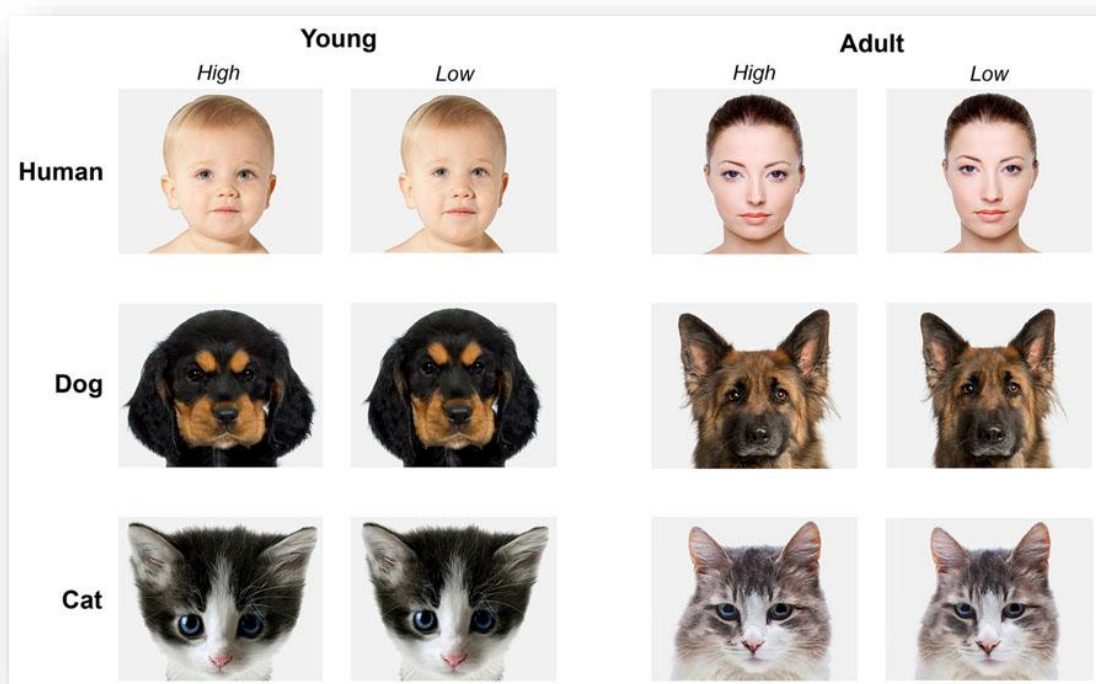
dziecka (*baby schema*, niem. *Kindchenschema*): dużą głowę, okrągłą twarz, wysokim i wystającym czołem, dużymi oczami, oraz małym nosem i ustami (Lorenz, 1943). Omawiana reakcja wiąże się ze zwiększeniem skupienia uwagi na wyznaczonych przez *baby schema* obszarach twarzy. Towarzyszy temu pozytywny afekt, pojawienie się zachowań opiekuńczych (*protective behavior*), a nawet ze zmniejszonym prawdopodobieństwem agresji wobec zwierzęcia cechującego się twarzą dziecka (czyli charakteryzującą się wysokimi parametrami *baby schema*); (Lorenz, 1943; Alley, 1983; Brosch, Sander, Scherer, 2007; Glocker, Langleben, Ruparel, Loughhead, Gur, Sachser, 2009a; Sherman Haidt, Coan, 2009; Nittono Fukushima, Yano, Moriya, 2012; Little, 2012).

Koncepcja schematu dziecka została zaprojektowana jako zbiór cech dziecięcych o wysokiej atrakcyjności dla ludzi (Borgi, Cogliati-Dezza, Brelsford, Meints, Cirulli, 2014). Uważa się, że stanowi ona kluczowy czynnik wyzwalający potrzebę opiekuńczości i wyznaczający orientację afektywną wobec niemowląt (Lorenz, 1943). Badania z 2009 roku przeprowadzone za pomocą neuroobrazowania, pokazują na poziomie neurofizjologicznym, że słodka reakcja pełni istotną rolę w promowaniu zachowań wychowawczych (czyli wspiera np. rozwój zdrowia moralnego u dzieci (Grochmal-Bach, 1997; Glocker, i in. 2009b)). Należy pamiętać, że *baby schema* jest filogenetycznie młodszy od człowieka, co oznacza, że występuje także u innych zwierząt (domowych i dzikich). Jego funkcją adaptacyjną jest motywacja do obrony potomstwa własnego gatunku. Jednak, istnieje szeroki zakres literatury wskazujący, że słodka reakcja wyzwalana pod wpływem percepcji schematu dziecka uległa generalizacji, w wyniku czego możliwa jest projekcja afektu na przedstawicieli innych gatunków (Wilson, 1984; Deloache Pickard, Lobue, 2011; Archer, Monton, 2011; Lobue, Bloom-Pickard, Sherman, Axford, Deloache, 2012; Little, 2012; Borgi, Cirulli, 2013). Jak wskazują Borgi i współpracownicy (2014), posiadanie cech twarzy dziecka minimalizuje szanse pogryzienia przez psa.

Borgi i współpracownicy (2014) przeprowadzili eksperymenty z udziałem dzieci w wieku od 3 do 6 lat. W ich ramach uczestnicy oglądali zdjęcia zwierząt (ludzi, psów, kotów), a ich zadaniem była ocena prezentowanych postaci np. podając opinię dotyczącą urody. Dodatkowo dzieci badane były za pomocą metod *eye-tracking*'owych, co pozwoliło na przeanalizowanie wzorców fiksacji uczestników w warunkach słodkości (*cute task*). Sprawdzono, które twarze przyciągają najwięcej uwagi, pomimo że aktorzy

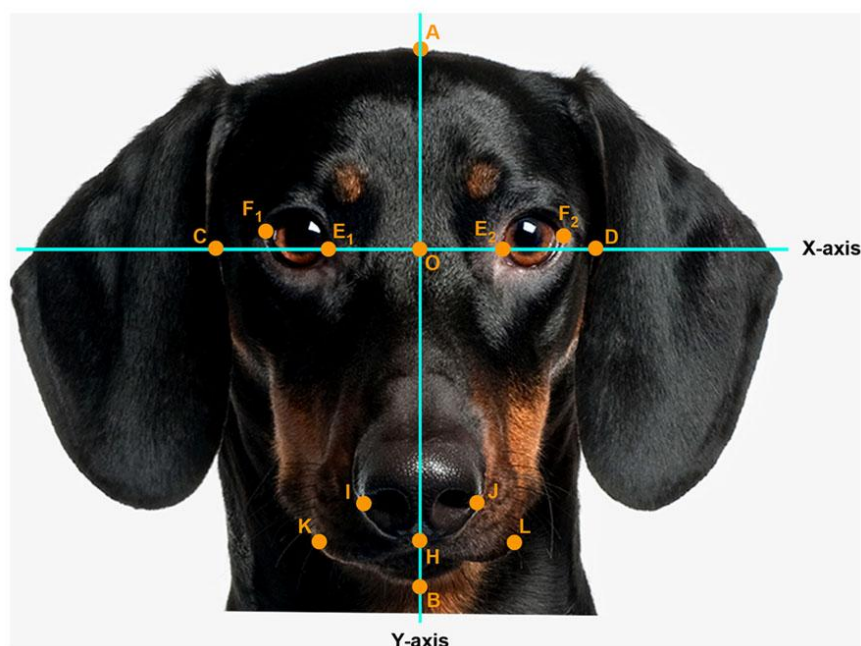
na zdjęciach różnili się od siebie tylko pod względem parametrów występowania cech *baby face* (jak na zdjęciu 1 i 2). Wyniki potwierdzają wcześniejsze doniesienia, że reakcja na schemat dziecka pojawia się we wczesnym etapie rozwoju człowieka, a słodka reakcja towarzyszy preferencji skupiania uwagi na specyficznych cechach *baby face*.

Zdjęcie 1. Przykładowe zdjęcia prezentowane uczestnikom badania Borgi i in. (2014)



[Źródło (Borgi, Cogliati-Dezza, Brelsford, Meints, Cirulli, 2014)]

Zdjęcie 2. Przykładowe zdjęcie prezentowane uczestnikom badania Borgi i in. (2014)



Źródło (Borgi, Cogliati-Dezza, Brelsford, Meints, Cirulli, 2014)]

Słodka reakcja jest przykładem dymorfizmu emocjonalnego, czyli bardzo silnych emocji, w wyniku których produkowane są reakcje zarówno pozytywne takie, jak uśmiech, jak i te, które konwencjonalnie uznawane są za negatywne np. łzy, czy chęć ściskania lub gryzienia źródła słodkości. Niektórzy ludzie płaczą w reakcji na pozytywne bodźce np. przy narodzinach dziecka lub gdy osiągną swoje cele, natomiast inni warczą i wyrażają chęć uszczypnięcia dziecka w policzki. Czyli we wszystkich opisanych wyżej sytuacjach, pozytywne emocje wywołały dymorficzne wrażenia, w wyniku czego jednostka odczuła nie tylko afekty pozytywne, ale też te, które zwykle kojarzone są z emocjami negatywnymi (złość, smutek, strach/lęk); (Ekman, Friesen, 1971; 1996). Obydwa rodzaje ekspresji pojawiają się jednocześnie, występują w sposób zdezorganizowany, co sprawia, że postronni uczestnicy sytuacji komunikacyjnej opierają się na kontekście zachowania (Carroll, Russell, 1996; Zaki, Hennigan, Weber, Ochsner, 2010)

W ramach konsensusu naukowego do opisu przetwarzania dymorficznych emocji przyjmuje się procesowy model emocji (*Proces Model of Emotions*); (Johnson, Richards, 2000; Aragón, Clark, Dyer, Bargh, 2015). Należy jednak pamiętać, że omawiany rodzaj afektu charakteryzuje się wyraźnym wzorcem zdarzenia bodźcowego, jedną oceną sytuacji, jednym określonym doświadczeniem oraz dwoma zachowaniami ekspresyjnymi, należącymi do przeciwstawnych kategorii emocji (pozytywne/

negatywne); (Aragón, Clark, Dyer, Bargh, 2015). Dla przykładu osoba, która wygrała na loterii milion złotych i oceniła to wydarzenie jako coś niesamowitego, może być przepełniona szczęściem, i wyrażać to śmiechem i płaczem jednocześnie. Dymorficzna reakcja pojawia się, gdy jednostka przytłoczona jest np. szczęściem i osiąga punkt, w którym emocje są nie do opanowania. Tak silne afekty przekraczają możliwości przetwarzania na poziomie neuronalnym (Folkman, Moskowitz, 2000). Zachowania powszechnie uznane za negatywne pojawiają się w celu rozładowania napięcia. Na przykład ściskanie źródła słodkości np. uroczego kotka/pieska oraz chęć włożenia sobie ich części ciała do ust. Jest to konieczne, gdyż tak silne emocje mogą być szkodliwe dla organizmu (Colom, Vieta, Martínez-Arán, Reinares, Benabarre, Gastó, 2000). Dymorficzne reakcje pojawiają się także podczas doświadczania negatywnych emocji. Fredrickson i Levenson (1998) zaobserwowali dymorficzną reakcję podczas oglądania filmów: ponad połowa uczestników uśmiechała się podczas najsmutniejszych momentów. Układy krążenia osób wykazujących tę reakcję szybciej wracały do normy, po silnie emocjonalnym smutnym wydarzeniu. To potwierdza już dobrze ugruntowaną w nauce tezę, że reakcja dymorficzna pełni funkcję regulacyjną, polegającą na równoważeniu silnie negatywnych emocji pozytywnymi i odwrotnie, silnie pozytywnych negatywnymi. Istnieją zadania testujące hipotezę, że funkcją reakcji dymorficznej jest regulowanie intensywnych emocji — np. słodka reakcja może być neutralizowana zachowaniami agresywnymi ³⁹ (Fredrickson, Mancuso, Branigan, Tugade, 2000; Schimmack, 2001; Kappas, 2011; Samson, Gross, 2012; Gross, 2013; Aragón, Clark, Dyer, Bargh, 2015). Wyniki potwierdziły, że uczestnicy eksperymentu przejawiali dymorfie w reakcji na silne emocjonalne bodźce.

Socjolog Edward O. Wilson (1984) rozwinął myśl filozofa i psychologa Ericha Fromma — Hipotezę Biofilii. Forman użył terminu greckiego pochodzenia, oznaczającego „miłość do życia”. W jej ramach uważa się, że ludzie posiadają naturalną skłonność do szukania połączeń z naturą (fauną i florą). Za słusznością wspomnianej hipotezy przemawia wiele badań wykazujących tendencję do inicjowania oraz akceptowania kontaktu ze zwierzętami pozaludzkimi, już na wczesnym etapie rozwoju

³⁹ Ciekawym przykładem jest japońska kultura masowa *Kawaii*, która uwidoczniła się w Japonii w latach '80 XX w. Uznaje się, ją za kulminację powojennego kryzysu kultury tego państwa. Słowo *kawaii* tłumaczyć można jako *Kawaii* znaczy: „dziecinny”, „słodki”, „niewinny”, „uczciwy”, „słaby”, „niedoświadczony” (Kinsella, 1995; Bator, 2007). Jednak jak wskazuje Burdzik (2015) złożoność omawianego zjawiska sprawia, że trudno wskazać na jedną definicję oddającą jego znaczenie.

(Deloache Pickard, Lobue, 2011; Lobue, Bloom-Pickard, Sherman, Axford, Deloache, 2012; Borgi, Cirulii, 2013). Istnieją przesłanki, przemawiające za tezą, że osoby z deficytami w sferze społecznej (jak np. osoby ze spektrum autyzmu) preferują bodźce pochodzące od ZINL (Celani, 2002; Prothman, Ettrich, Protthmann, 2009; Grandgeorge Bourreau, Alavi, Lemonnier, Tordjman, Deleau, 2014). Okazuje się, że ludzie zwykle unikający kontaktów z innymi jednostkami własnego gatunku, przejawiają więcej zachowań społecznych w obecności przedstawicieli innych gatunków. Badania potwierdzające ten wniosek opierają się na udostępnieniu uczestnikowi miejsca do zabawy zabawkami lub kontaktu z przedstawicielem innego gatunku. Okazuje się, że u uczestników obserwuje się wzrost ilości zachowań społecznych w kontakcie ze zwierzęciem pozaludzkim w porównaniu z kontaktem z człowiekiem.

Co ciekawe analizy relacji opiekun – zwierzę domowe, pokazują, że jest ona bardzo podobna do relacji rodzic – dziecko. Dla przykładu w komunikacji ze zwierzętami pozaludzkimi używana jest „czuła mowa”, różniąca się od mowy matczynej (*baby talk*)⁴⁰ tylko pod względem hiperartykulacji, która w kontakcie z ludzkimi dziećmi nastawiona jest na nabywanie kompetencji artykulacyjnych, natomiast w kontakcie z ZINL przybiera inną formę (Burnham, Kitamura, Vollmer-Conna, 2002). Dodatkowo psy wykazują efekt bezpiecznej bazy, czyli wykorzystują opiekuna do interakcji z otoczeniem, podobnie do ludzkich dzieci (Horn, Huber, Range, 2013). To sprawia, że możemy wnioskować o możliwości zastosowania teorii przywiązania⁴¹ do opisu interakcji ze zwierzętami pozaludzkimi. Słodka reakcja leży u podstaw tendencji do przypisywania zwierzętom innym niż ludzie zdolności poznawczych specyficznych dla człowieka. Możemy zatem powiedzieć, że jest to mechanizm charakterystyczny dla antropomorfizmu. Jak wcześniej zaznaczałam postawa ta, ma głębokie korzenie kulturowe i religijne. W artykule *Anthropomorphism*

⁴⁰ Jest to rodzaj mowy używany najczęściej przez osoby starsze do komunikacji z osobami młodszymi (dzieckiem czy niemowlęciem). Charakteryzuje się śpiewną intonacją, wolniejszym tempem mowy i krótszymi wypowiedziami (słowa zazwyczaj są skracane i upraszczane); (Lam, Kitamura, 2012).

⁴¹ Psycholożka Mary Dinsmore Ainsworth pionierka Teorii Przywiązania, przeprowadziła badania z udziałem ludzkich dzieci. Eksperyment polegał na postawieniu małych (kilku i kilkunastomiesięcznych) dzieci w „dziwnej sytuacji” (*strange situation*), czyli na zostawieniu ich w obecności nieznanego opiekuna, gdy matka wychodziła z pokoju. Uczestnicy eksperymentu wykazywali zachowania stresowe (*anxiety*) odpowiadające 3 różnym wzorcom — bezpiecznemu, ambiwalentno-lękowemu i unikającemu (Ainsworth, Blehar, Waters, Wall, 2015).

and Anthropomorphic Selection—Beyond the "Cute Response" Jamesa Serpella (2002) znajdziemy ciekawe rozważania dotyczące ewolucyjnych implikacji antropomorfizmu w kontekście relacji człowiek – inne zwierzęta. Gdzie autor argumentuje za słusnością hipotezy, że korzystanie z myślenia antropomorficznego umożliwia człowiekowi interpretację zwierzęcych (np. tych towarzyszących) zachowań społecznych. Jest to ciekawe, biorąc pod uwagę, że większość literatury dotyczącej antropomorfizmu dotyczy jego wagi w formułowaniu interpretacji zachowania ZINL w kontekście eksperymentalnym (McFarland, 1981; Lockwood, 1985; Kennedy, 1992; Mitchell-Tompson, Milesa, 1997).

Korzyść wynikająca z myślenia antropomorficznego człowieka zdaje się oczywista — pozwala ono na skuteczniejszą komunikację ze zwierzętami współzyskującymi z ludzkimi społeczeństwami oraz na zrozumienie i przewidywanie interakcji z osobnikami żyjącymi poza naszymi społecznościami. Uważa się, że na wczesnych etapach rozwoju człowieka omawiany mechanizm umożliwiał np. sprawniejsze polowanie, a później udomowienie wielu gatunków zwierząt. Natomiast po stronie ZINL, antropomorfizm stanowi ogromną presję selekcyjną, która ukształtowała anatomię, morfologię i zachowanie różnych gatunków np. zwierząt towarzyszących. Ludzie karmią ZINL ludzkim jedzeniem, nadają im ludzkie imiona, świętują urodziny, zabierają je do lekarzy i opłakują ich śmierć, a także grzebią je na cmentarzach, przy czym, pochówek zawiera wszystkie rytuały charakterystyczne dla ludzkich pogrzebów (Serpell, 1996a). Ubierają swoje ZINL w ubrania drogich marek, zapisują do przedszkoli oraz jeśli trzeba, opłacają kosztowne zabiegi, jak przeszczepy narządów (Serpell, 2002). Faktem jest, że jeżeli biorąc pod uwagę jedynie liczebność, antropomorfizm zapewnił wielu gatunkom sukces ewolucyjny, jednak patrząc z perspektywy dobrostanu ich przedstawicieli, jego skutki bywają znacznie mniej pozytywne. Selekcja na korzyść cech fizycznych i behawioralnych, tak aby przedstawiciele gatunku (czy konkretnie rasy) posiadały wiele cech atrakcyjnych dla człowieka nazywa się doborem antropomorficznym. Nadmierna selekcja pod kątem cech ułatwiających przypisywanie ludzkich stanów mentalnych innym zwierzętom spowodowała, że niektóre rasy np. psów czy kotów (ale też wielu innych zwierząt towarzyszących czy hodowlanych) krzyżowane były w sposób upośledzający zdrowe funkcjonowanie (mopsy, persy). Dlatego też, rozmnażanie niektórych ras, jest aktualnie nielegalne w wielu miejscach na świecie.

Niektórzy badacze uważają, że antropomorfizm jest jedną z cech charakterystycznych dla współczesnego człowieka (*Homo sapiens sapiens*); (Mithen, 1998). Mithen opiera swoje twierdzenia na przesłankach archeologicznych, świadczących o nagłej zmianie w relacjach człowiek — inne zwierzęta oraz człowiek — świat przyrody. Zmiana ta przypada na środkowo – górny paleolit, czyli na czas, gdy (Mithen, 1998):

- a. Pojawiają się pierwsze oznaki sztuki dekoracyjnej oraz reprezentacyjnej,
- b. Ludzie wynaleźli m.in. strzały i łodzie, oraz wiele innych narzędzi,
- c. Zaczęto stosować pochówki rytualny i in. praktyki religijne (pierwsze świadectwo ich występowania pochodzi z wyznaczonego okresu).

Mithen twierdzi, że zmiany te związane są ze zmianą architektury ludzkiego umysłu a ściślej stosunkowo szybkiej i radykalnej zmianie funkcjonalnej. W ramach teorii wspomnianego autora, umysły gatunków kopalnych ludzi (przed ok. 40 000), tak jak umysły *Homo sapiens sapiens*, wyposażone były w moduły wyspecjalizowane w określonych dziedzinach inteligencji:

- a. Moduł inteligencji społecznej: jego funkcją jest radzenie sobie ze złożonością interakcji społecznych. Zdolny jest do wykorzystania samowiedzy w celu przewidywania zachowań innych.
- b. Moduł historii naturalnej: jego funkcją jest przetwarzanie danych o dostępności i rozmieszczeniu zasobów biologicznych, w tym o zachowaniu innych gatunków zwierząt np. drapieżników czy celu własnego polowania (ofiary).
- c. Moduł inteligencji technicznej: skupiający się na fizycznych aspektach świata materialnego, obejmuje techniki manipulowania i konstruowania obiektów, takich jak, broń czy inne narzędzia.
- d. Moduł wywiadu ogólnego: zajmuje się rozwiązywaniem problemów i podejmowaniem decyzji.

Każdy z modułów stanowi obszar wyspecjalizowany w przetwarzaniu określonego rodzaju informacji. 40 000 lat temu miała miejsce rewolucja poznawcza nazywana płynnością poznawczą (*cognitive fluidity*), w ramach której pojawiła się zdolność różnych modułów do wzajemnej komunikacji. Jak twierdzi autor *The prehistory of the mind: A search for the origins of art, religion and science*, rozwój d. ograniczył tempo ewolucji kulturowej wczesnych ludzi, umożliwiając interakcję między modułami.

U gatunków kopalnych człowieka moduły te działały w dużej mierze niezależnie, a dopiero wykształcenie przepływu informacji między nimi wywołało niespotykaną wcześniej eksplozję kulturową. Aktualnie uznaje się także, że to właśnie wraz z rewolucją neolityczną wykształconych zostało wiele mitów, np. mit króla kapłana ⁴², co świadczy o dalszym rozwoju złożonych form kultury i narracji. (Flannery, Marcus, 2012). Według badań archeologicznych i antropologicznych, mit króla kapłana jest przykładem powstania narracji łączącej władzę świecką i religijną, co świadczy o rozwoju abstrakcyjnego myślenia, symboliki i systemów wierzeń (Flannery, Marcus, 2012; Wills-Moore, 2019). Takie mity pełnią funkcję integrującą społeczność, porządkując świat społeczny i naturalny oraz uzasadniając hierarchie i rytuały. W tym sensie powstanie mitu króla kapłana świadczy o dalszym rozwoju kultury i narracji, ponieważ oznacza ukształtowanie się złożonych struktur społecznych, religijnych i symbolicznych, które wymagają już nie tylko prostych opowieści, ale złożonych mitów wyjaśniających pochodzenie władzy, porządek społeczny i relacje między ludźmi a siłami nadprzyrodzonymi.

2.7. Efekt halo

Efekt halo to tendencyjność widoczna w atrybucji cech, przejawiająca się preferencją punktu odniesienia o charakterze pozytywnym albo negatywnym. Spostrzeżenie przynajmniej jednej cechy pozytywnej, nie tylko powoduje przypisanie kolejnych pozytywnych własności, ale też podnosi atrakcyjność pozostałych cech i na odwrót, spostrzeżenie jednej cechy negatywnej, obniża atrakcyjność kolejnych przypisywanych atrybutów (Marczak, 2016). Zatem zjawisko to pojawia się na etapie tworzenia wrażenia o odbieranym obiekcie. Dla przykładu, jeżeli jednostka posiada wysokie parametry w zakresie *baby schema*, osoba postrzegająca będzie bardziej skłonna przypisać jej inteligencję emocjonalną, wrażliwość czy bycie godnym zaufania (Zebrowitz, Franklin, 2014). Zatem możemy powiedzieć, że efekt halo polega na wytworzeniu preferencji opartej na emocjonalnym odbiorze informacji (Mękarska, 2024). Omawiane zjawisko zidentyfikowane zostało w 1907 roku przez amerykańskiego psychologa Fredrike Wellsa, jednak to Edward Thorndike (1920)

⁴² Mit Króla kapłana (występujący także jako mit wodza czy czarownika) jest przykładem mitu bohatera (Zowczak, M. 1984). Legitymizuje on władzę absolutną, gdyż w jego ramach to właśnie bohater podtrzymuje pomyślność oraz zapewnia stałość swojemu królestwu.

dostarczył dowodu empirycznego na jego istnienie oraz nadał mu nazwę pochodzącą od złudzenia optycznego Halo⁴³. W wyniku tego to właśnie ten badacz oficjalnie uznawany jest za odkrywcę omawianego efektu. Eksperyment Thorndike i współpracowników przeprowadzony został w 1915 roku, a jego uczestnikami byli pracownicy dwóch korporacji przemysłowych. Walter Dilla Scott poprosił uczestników o ocenę osób na zdjęciach, pod względem 4 cech:

- a. Atrakcyjność fizyczna,
- b. Inteligencja,
- c. Przywództwo,
- d. Cechy charakteru.

W instrukcji wyraźnie zaznaczył, że cechy powinny być oceniane/ przydzielane (w przypadku d.) niezależnie od siebie. Okazało się, że ludzie nie potrafili wykonać tego zadania, ich odpowiedzi wyraźnie wykazywały korelacje między ocenami kryteriów: im wyższa ocena w punkcie a., tym wyższe parametry w b. i c. oraz cechy wymieniane w kryterium d. miały charakter pozytywny. Na tej podstawie Thorndike (1920) sformułował wniosek, że ludzie nie postrzegają cech innych jako odrębnych zbiorów stereotypów, nabytych na drodze różnych doświadczeń, za to ich atrybucja zachodzi zgodnie z kategoriami poznawczymi, do których przydzielony zostanie postrzegany obiekt. Autor *A constant error in psychological ratings* zaznacza, że tak wyraźny wpływ efektu halo na ocenę innego, powinien być brany pod uwagę podczas raportowania doniesień naukowych. Podkreśla, że eksperymentatorzy powinni unikać oceniania otrzymanych wyników oraz zamiast tego powinni skupić się na raportowaniu dowodów naukowych. Niestety w artykule nie znajdziemy odpowiedzi na pytanie, *co jest opinią o wynikach a co poprawną interpretacją wyników?*

Z omawianego zjawiska wynikają dwa inne efekty – aureoli/anielski i diabelski/rogów (*horn effect*). Pierwszy z wymienionych pojawia się, gdy cecha determinująca dalsze postrzeganie ma charakter pozytywny, drugi natomiast gdy jej charakter jest negatywny. Efekt halo jest konsekwencją płynącą z opisanego wcześniej mechanizmu kategoryzacji, a jego wpływ widoczny jest w wielu sferach życia społecznego np. w kulturze. Najwyraźniejszym i często podnoszonym wyrazem efektu halo są postaci w bajkach dla dzieci, zarówno tych starych baśniach oralnych, jak

⁴³ Halo złudzenie optyczne obserwowane wokół tarczy słońca lub księżyca.

i nowszych przekazach medialnych np. bajki wytwórni Disney czy Pixar. Postaci pozytywne są przedstawiane nie tylko jako piękne fizycznie, ale także zachwycające dobrocią i bogactwem wewnętrznym. Natomiast postaci negatywne przedstawiane są jako szkaradne, głupie i nierzadko brutalne. Dla przykładu w Kopciuszku, czyli w baśni, o której pierwsze wzmianki pochodzą już z XVII wieku, Anastazja i Gryzelda, czyli przybrane siostry Kopciuszka opisywane były jako brzydkie i grube (ATU 510A; Pullman, 2014; Bednarek, 2021). Główna bohaterka natomiast była piękna, nawet gdy była brudna i ubrana w stare łachmany. W niektórych dziełach kulturowych dostrzegamy wykorzystanie efektu halo do przełamania pewnych społecznych konwenansów. Dla przykładu w słynnej powieści Frankenstein autorstwa Mary Shelley (2013) potwór powołany do życia przez tytułowego doktora odrzuca swoją fizjonomię, jednak jego bogactwo wewnętrzne zachwyca. Jest to wyraźne wykorzystanie efektu rogów w celu podkreślenia, że „syn” Frankensteina nie powstał jako potwór, a stał się nim na drodze odrzucenia i nienawiści wywołanej jego aparycją. Natomiast wykorzystanie efektu aureoli dostrzec możemy w bajce Kraina Lodu (2013), gdzie piękny książę, w którym zakochuje się księżniczka Anna, okazuje się podstępny podłym człowiekiem, który dla władzy zrobi wszystko.

2.8. Podsumowanie

Rozdział II poświęcony został analizie wybranych mechanizmów poznawczych, które leżą u podstaw tendencyjności związanej z uprzedzeniami. Szczegółowo omówiłam w nim, jak ograniczenia poznawcze – takie jak pojemność pamięci czy zasoby uwagi – wymuszają na umyśle selekcjonowanie i grupowanie informacji. Kategoryzacja poznawcza pozwala na redukcję złożoności świata poprzez łączenie obiektów w klasy na podstawie wspólnych cech. Szczególne znaczenie w kontekście tej dysertacji ma kategoryzacja społeczna, prowadząca do automatycznego przypisywania obiektów – zidentyfikowanych jako agenci – do określonych grup. Proces ten nie tylko upraszcza przetwarzanie informacji, lecz także wzmacnia tożsamość i samoocenę poprzez mechanizm autowaloryzacji. Przynależność do cenionych grup staje się narzędziem obrony *ego*, zaś etnocentryzm i faworyzacja grupy własnej znajdują uzasadnienie zarówno w adaptacyjnych mechanizmach ewolucyjnych, jak i społecznych motywatorach.

Omówiłam również mechanizm infrahumanizacji – tendencję do odmawiania przedstawicielom grupy obcej cech uznawanych za wyjątkowo ludzkie, takich jak teoria umysłu czy zdolność do odczuwania emocji wtórnych. Takie postawy wzmacniają mity o ludzkiej wyjątkowości i wpływają na sposób, w jaki interpretujemy zachowania innych. Badania (Leyens i in. 2003) pokazują, że ludzie chętniej przypisują emocje wtórne członkom grupy własnej, a ich wyrażanie przez członków grupy obcej budzi nieufność i negatywny odbiór. W efekcie prowadzi do preferencyjnego traktowania członków grupy własnej kosztem grupy obcej. Mechanizm ten związany jest także z samo-licencjonowaniem, opartym na przekonaniu o *własnej moralnej wyższości*. Może on służyć jako usprawiedliwienie uprzedzeń i sprzyjać ich podtrzymywaniu.

Przedstawione procesy mają charakter automatyczny i uniwersalny – pojawiają się zarówno w odniesieniu do klasycznych kategorii społecznych, takich jak płeć czy narodowość, jak i tych tworzonych sytuacyjnie. Mechanizm akcentuacji sprzyja wyolbrzymianiu różnic międzygrupowych i bagatelizowaniu różnorodności wewnątrzgrupowej, co skutkuje efektem jednorodności grupy obcej. Z kolei członkowie grupy własnej są kodowani jako indywidualni, unikatowi. Tendencje te nasilają się w sytuacjach zagrożenia samooceny i zostały potwierdzone licznymi badaniami przytoczonymi w rozdziale. Choć omawiane mechanizmy mają ogólny charakter społeczno-poznawczy, ich znaczenie okazuje się kluczowe w kontekście interpretacji zdolności umysłowych zwierząt innych niż ludzie. Kategoryzacja społeczna, efekt jednorodności grupy obcej czy infrahumanizacja mogą prowadzić do systematycznego pomijania lub niedostrzegania oznak złożonego przetwarzania poznawczego u zwierząt. Jeżeli przedstawiciele grupy obcej (np. zwierzęta pozaludzkie) są postrzegani jako niezdolni do doświadczenia emocji wtórnych lub formowania intencji, ich zachowania będą interpretowane jako np. sprytne czytanie zachowania innego (*smart behavior reading*). W ten sposób opisywane mechanizmy nie tylko zniekształcają społeczne postrzeganie zwierząt, ale także wpływają na to, czy i jak przypisujemy im zdolności poznawcze. Ma to bezpośrednie konsekwencje zarówno dla metodologii badań nad teorią umysłu, jak i dla etyki relacji międzygatunkowych.

Celem rozdziału było ukazanie poznawczych podstaw kształtowania i utrwalania postaw, podkreślając, jak głęboko zakorzenione i automatyczne mogą być procesy odpowiedzialne za powstawanie i utrzymywanie uprzedzeń. Z uwagi na ograniczenia objętościowe pracy, zdecydowałam się na wybór siedmiu kluczowych mechanizmów, które zostały szeroko opisane w literaturze i mają największe znaczenie dla dalszych

analiz. Tym samym rozdział II stanowi teoretyczne zaplecze dla kolejnych części dysertacji.

Rozdział III: Narracje naukowe konstytuujące porządek społeczny w kontekście relacji człowiek—inne zwierzęta

3.1. Historia badania umysłów zwierząt pozaludzkich

Od początku istnienia gatunku ludzkiego zwierzęta pozaludzkie fascynują człowieka, a fascynacja ta doprowadziła do powstania wierzeń nazywanych animalizmem lub zoolatrią. W niektórych kulturach status ZINL zależał od wierzeń w nich zakorzenionych. I chociaż oddawana im cześć prowadziła niekiedy do objęcia niektórych gatunków zakazem polowania, znacznie częściej nie chroniła ich od śmierci z rąk człowieka (Łukasiewicz, 2010). Należy także pamiętać, że zakaz zabijania to nie wszystko; dla przykładu krowy uznawane za święte w Indiach wiodą raczej smutne życie, blakając się po miastach głodne i spragnione (Margul, 1996; Bator, 2013). Dodatkowo cześć oddawana niektórym gatunkom, nie gwarantowała szacunku w stosunku do pozostałych. Istnieją przekazy mówiące o bardzo brutalnym obchodzeniu się ze zwierzętami wchodzącymi do panteonów wierzeń innych niż własne, które traktowane były jak szkodniki (Łukasiewicz, 2010). Status zwierząt innych niż ludzie zmieniał się z czasem, a wraz z powstaniem religii chrześcijańskiej rozpowszechniło się przekonanie o *wyjątkowości człowieka* (Ryczek, 2013). Ważne jednak, aby pamiętać, że chociaż religia chrześcijańska miała ogromny wpływ na kształtowanie postaw w stosunku ZINL, to zgodnie ze współczesną wiedzą trudno podtrzymać twierdzenie o pierwotności przekazu biblijnego jako opiniotwórczego w tej kwestii (Bator, 2013). Religie świata miały istotne znaczenie dla kształtowania się relacji człowiek – inne zwierzęta, jednak w niniejszym podrozdziale skupiam się na postrzeganiu ZINL w środowisku badawczym. Chociaż nie poświęcam w niniejszej dysertacji szczególnej uwagi roli religii w kształtowaniu postaw społecznych, zdaję sobie sprawę z tego, że niektórzy naukowcy podejmujący temat umysłów ZINL, są osobami wierzącymi, w związku z czym mity religijne mogą wpływać na konstruowane czy wspierane przez nich interpretacje. Uważam jednak, że w interesującym mnie obszarze, temat roli mitologii religijnej zasługuje na osobne opracowanie.

Literatura, wskazuje, że za kształtowanie relacji człowiek – inne zwierzęta w kulturze europejskiej, odpowiada m.in. myśl starożytnej greckiej (Zieliński 1927; Bator, 2013, Ryczek, 2013). Chociaż Tadeusz Zieliński (1927) w swoich pracach podkreśla harmonię między światem ludzi i zwierząt pozaludzkich, która miała panować

w cywilizacji helleńskiej, aktualnie dostrzega się, że świat starożytnej grecji, a nawet kultury przedgreckiej, naznaczony był silną ideą antropocentryczną. Wierzano, że człowiek powstał na podobieństwo boga, którego wyobrażano sobie jako nadczłowieka, przyjmującego czasem formę innego zwierzęcia lub hybrydę zwierząt (Bator, 2013). Obecne było także przekonanie, że naturę należy ujarzmić i podporządkować sobie, a elementy przyrody niedające się okiełznać, należy po prostu zniszczyć (Łyczkowska, Szarzyńska, 1981). Podejście to kształtuje dwa rodzaje postrzegania ZINL: jako dobytek oraz jako dziką bestię, którą należy zgładzić. Jak zaznacza Bator (2013), prawdą jest, że w porównaniu do postaw azjanickich czy semickich, Grecy wydawali się bardziej powściągliwi, jednakże w świetle obserwacji Bator (2013), przekonująca wydaje się konstatacja, że wszystkie te wymienione przeze mnie kultury podkreślały ludzką wyjątkowość. Przykładami wspomnianego „wyważenia” kultury helleńskiej są poglądy filozofów takich jak Pitagoras (ok. 570 – 495 p.n.e.), czy Empedokles z Agrigentu (ok. 494 – 434 p.n.e.), którzy głosili pogląd określany jako metempsychoza, czyli wierzyli w wędrówkę dusz. W konsekwencji takiego poglądu stanowczo stawali w obronie ZINL, propagując wegetarianizm i zaprzestanie składania ich w ofierze. Należy pamiętać, że wspomniana łagodność i tolerancja dla ZINL związane były z przekonaniem, że w ich ciele może znajdować się dusza zmarłego człowieka. Była zatem motywowana dobrem człowieka, a nie samego zwierzęcia. Innym przedstawicielem kultury helleńskiej był Arystoteles. Jego poglądy kształtowały relacje człowiek – inne zwierzęta do początku czasów nowożytnych (Grabowska, 2014). Ten starożytny filozof uznawał, że ZINL są istotami czującymi, przypisując niektórym z nich m.in. pamięć, czy nawet wyobraźnię. Podkreślał jednak *wyjątkowość człowieka*, ponieważ jak twierdził, tylko ten gatunek obdarzony jest rozumem. Rozum wynosił człowieka na szczyt hierarchii bytów, co z punktu widzenia koncepcji Arystotelesa, czyniło inne byty jego własnością oraz upoważniało ludzi do ich wykorzystywania, nawet wówczas, gdy wiązało się to z zadawaniem im cierpienia. Filozof ten uważał jednak, że zdawanie cierpienia ZINL zawsze powinno być uzasadnione, przestrzegając przed nadmiernym stosowaniem wobec nich okrucieństwa. Wiazało się to z przekonaniem, że nadmierne zadawanie cierpienia ZINL prowadzić będzie do okrutnego zachowania człowieka wobec innych ludzi oraz że wyrozumiałość wobec nich będzie skutkować wyższą wrażliwością na potrzeby i krzywdę drugiego człowieka. Co ważne, pogląd Arystotelesa podzielany był przez Tomasza z Akwinu, a jego obecność, zaobserwować można aż do czasów nowożytnych (Grabowska, 2014).

XVII-wieczny *Zeitgeist* naznaczony był nadchodzącą rewolucją przemysłową, która ostatecznie wybuchła w XVIII wieku. Coraz częściej pojawiające się w codziennym życiu urządzenia, wywarły ogromny wpływ na światopogląd ówczesnych ludzi. Jedną z największych sensacji w XVII-wiecznym świecie, był zegar mechaniczny (Schultz, 2008, str. 28). Jego wynalezienie umożliwiło dokonywanie dokładnego pomiaru czasu, co prowadziło do wielu konsekwencji naukowych i praktycznych, dlatego Daniel Boorstin nazywa zegar mechaniczny „matką maszyn” (Boorstin, 1983, str. 71; Schultz, 2008, str.29). Ogólna fascynacja zegarami zaczęła skłaniać filozofów do traktowania zegarów jako fizycznych modeli świata. (Schultz, 2008, str. 28-39). Innym mechanizmem, który wywarł duży wpływ na poglądy XVII-wiecznych ludzi, był automat mnicha, który rozbudził marzenie o stworzeniu maszyny podobnej do człowieka. Zgodnie z duchem epoki, panująca koncepcja człowieka jako maszyny, proponowała metodę naukową, pozwalającą na badanie ciała jako maszyny.

Jednym z myślicieli wykorzystujących ideę mechanicyzmu był Rene Descartés. Jego fascynacja automatami zamontowanymi w Królewskich Ogrodach, obserwowanie ich i wprawianie w ruch, prowadziła do przekonania, że materialne ciało nie musi być kierowane przez niematerialny umysł, a przez czynniki zewnętrzne. Podkreślał przy tym, że jedyną funkcją umysłu jest myślenie (Schultz, 2008, str. 40). Zgodnie zresztą, z ogólną koncepcją miejsca człowieka w świecie, którą niesie ze sobą wiara katolicka – zwierzęta inne niż ludzie nie posiadają duszy, zatem uważano je za automaty. Pogląd ten stanowił uzasadnienie dla bardzo brutalnych zachowań wobec ZINL. Istnieją opisy rozbawienia eksperymentatorów, którzy prowadzili badania na żywych zwierzętach, bez dostępnych metod znieczulenia. Rozbawienie naukowców wywołane było zawrotem głowy torturowanego zwierzęcia. Jak twierdzili dźwięki te, to tylko „hydrauliczne gwizdanie i wibracje maszyn” (Jaynes, 1970). Jest to skrajny przejaw myśli mechanicystycznej, zachowanie ZINL można wyjaśnić w kategoriach maszyny, gdyż jako automaty nie są zdolne do odczuwania. Zgodnie z oczekiwaniami, badania ujawniły uderzające podobieństwo w fizjologii ludzi i innych zwierząt. Paradoksalnie to przyczyniło się do spadku wiarygodności skrajnych twierdzeń takich jak te Kartezjusza (Ryczek, 2013). W XVIII wieku zaczęło pojawiać się coraz więcej głosów za ograniczeniem okrucieństwa wobec ZINL. Szkocki filozof David Hume powiedział nawet „zasady człowieczeństwa kazałyby nam (...) dobrze się z istotami tymi obchodzić” (Hume, 2005). W 1780 Jeremy Bentham pierwszy raz w historii nie nazywa panowania człowieka legalnym rządem, a tyranią (Bentham, 1958). Jest to niewątpliwy

postęp, który w XIX wieku zaowocował realnymi zmianami sytuacji ZINL, gdyż w 1824 roku powstała pierwsza na świecie organizacja broniąca praw zwierząt – *Society for the Prevention of Cruelty of Animal* (ASPCA). Choć XVIII wiek przyniósł pozytywną zmianę w postawach w stosunku do ZINL, nie wszyscy od razu przyjęli nową narrację. Immanuel Kant podczas swoich wykładów z etyki mówił: „Wobec zwierząt nie mamy żadnych bezpośrednich obowiązków. Nie są one istotami samoświadomymi, są więc jedynie środkami do celu. Celem jest człowiek” (Ryczek, 2013, str. 121). Co ciekawe we *Wprowadzeniu do zasad moralności i prawodawstwa* Bentham (1958) odpowiada Kantowi słowami: „należy pytać nie o to, czy zwierzęta mogą rozumować i czy mogą mówić, lecz czy mogą cierpieć” (Ryczek, 2013).

Ludzkie przekonanie o swojej wyjątkowości nie znajdowało istotnego zaprzeczenia, aż do przełomu, jaki dokonał się w nauce za sprawą darwinowskiej tezy o ciągłości ewolucyjnej, która mówi, że różnica pomiędzy zdolnościami umysłowymi ludzi a zwierząt pozaludzkich ma charakter stopnia, a nie rodzaju (Sikorska, Trojan, 2018). Uściślając, zmiana społeczna zaczęła jeszcze przed wydaniem *O powstaniu gatunków drogą doboru naturalnego*. Wraz z odkryciami geograficznymi napotkano gatunki zwierząt, których istnienia ludność europejska wcześniej nie zdawała sobie sprawy. W 1501 roku Amerigo Vespucci powiedzieć miał:

Co mam powiedzieć o mnóstwie dzikich zwierząt, ogromnej ilości pum, lampartów, dzikich kotów, nie takich jak te w Hiszpanii a takie jak te w Antypodach; o tak wielu wilkach, łosiach, kotowatych, licznych rodzajach marmozet i wielu wielkich węzach. [...] Tak wiele zwierząt nie weszłoby do arki Noego (Schultz, 2008, str. 135).

Oznacza to, że ludzie zaczęli poddawać w wątpliwość prawdziwość mitu o arce Noego i o jego uratowaniu wszystkich gatunków zwierząt (Schultz, 2008, str. 135). Teoria ewolucji dostarczyła alternatywnego wyjaśnienia obserwowanej różnorodności. Najnowsze odkrycia archeologiczne gatunków wymarłych, wzbogacały wiedzę o żyjących. Możemy zatem mówić, że na skutek rewolucji przemysłowej, zaszła gruntowna zmiana w ludzkim postrzeganiu swojego miejsca w świecie. Zmiana ta umocniła pozycję nauki, gdyż od tej pory wyjaśnienia świata, te proponowane przez biblię przestały być wystarczające. Dzieło Darwina dostarczyło wielu danych oraz usystematyzowało je, tworząc pełną teorię wypierającą dotychczasowe wyjaśnienia świata odnoszące się do religii (Schultz, 2008).

Obecnie teoria ewolucji jest raczej dobrze znana, dlatego tutaj niezbędne jest tylko powołanie się na wspomnianą wyżej tezę o ciągłości ewolucyjnej, gdyż to w niej dostrzegam główny element kształtujący ówczesznie panujące relacje człowiek – inne zwierzęta. Do *O powstaniu gatunków drogą doboru naturalnego* nie znajdowano powodów, dla których nauka miałaby się zajmować zwierzęcym umysłem, przecież uznawano je za bezduszne automaty. Wraz z teorią darwinowską przyjęto, że nie tylko ludzie zdolni są m.in. do odczuwania bólu, przynajmniej podstawowych emocji, a także, że mają marzenia senne, a zatem, że przynajmniej w jakimś stopniu posiadają wyobraźnię. Co więcej, teoria ewolucji była impulsem do powstania zoopsychologii, której głównym celem było szukanie dowodów na to, że człowiek nie różni się znacząco od innych zwierząt (Schultz, 2008). Większość dorobku tej nauki skupiało się wtedy na inteligencji zwierząt. Jak wskazują autorzy *Współczesnej historii psychologii* w latach '60 i '70 XIX wieku pojawiały się doniesienia o niezwykłych przejawach inteligencji u różnych gatunków zwierząt (Schultz, 2008).

Następcą Dawina był George Ramones (fizjolog), który otrzymał od mentora notatki o zachowaniu ZINL (Schultz, 2008). Ramones stworzył drabinę umysłową zwierząt (tabela 4), która pokazywała stopień funkcjonowania psychicznego osobników różnych gatunków. Warto odnotować, że według autora tej hierarchizującej klasyfikacji, meduzy prezentują zaawansowany stopień rozwoju poznawczego. Jak na tamte czasy, pogląd ten był mocno kontrowersyjny (Schultz, 2008). Prace Ramonesa oparte były na metodach anegdotycznej i introspekcji, co polegało na wykorzystywaniu doniesień dotyczących zachowania zwierząt oraz konstruowaniu na ich podstawie twierdzeń dotyczących ich życia psychicznego. Niestety wiele obserwacji, na podstawie których wyciągał swoje wnioski, pochodziło od nieposiadających przeszkolenia psychologicznego ludzi. Ten istotny błąd metodologiczny prowadził do nadinterpretacji obserwowanych zachowań. Innym podejściem metodologicznym, stosowanym przez omawianego fizjologa, było użycie myślenia analogicznego w połączeniu z introspekcją. Przy takich założeniach osoba analizująca zachowanie zakłada, że obserwowanym zwierzęciem kierują analogiczne procesy psychiczne do tych, które zachodzą u obserwatora. Interpretacje danych uzyskanych metodą introspekcji i analogii lub anegdoty, narażone były na antropomorfizację. W rezultacie Ramones uznał, że ZINL zdolne są m.in. do złożonego rozumowania np. racjonalizacji. Co więcej, jego zwolennicy przypisywali niektórym zwierzętom zdolności i umiejętności umysłowe, które przewyższały zdolności przeciętnego człowieka (Schultz, 2008).

Tabela. 4 . Drabina umysłowa opracowana przez Georga Ramonesa

Gatunki	Poziom rozwoju intelektualnego
Małpy człekokształtne, psy	Nieokreślona moralność
Małpy, słonie	Wykorzystanie narzędzi
Ptaki	Rozpoznawanie obiektów, rozumienie słów
Pszczoły, osy	Komunikowanie pojęć
Gady	Rozpoznawanie osób
Homary, kraby	Zdolność rozumowania
Ryby	Kojarzenie przez podobieństwo
Ślimaki, kałamarnice	Kojarzenie przez styczność
Rozgwiazdy, jeżowce	Pamięć
Meduzy, ukwiały	Świadomość, przyjemność, ból

[Opracowanie własne (Schultz, Schultz, Andruszko, 2008, str. 161)]

Następcą Romanesa był C. Lloyd Morgan, który formalnie rozpoznał słabości stosowania zestawu wyżej opisanych metod oraz ustanowił prawo oszczędności nazywane także zasadą lub kanonem Morgana. Zgadzał się z Ramonesem, że nie należy

ignorować subiektywnych opisów, ale sprzeciwiał się poleganiu na niesprawdzonych anegdotach. Napisał nawet:

Jeśli chodzi o zbiór anegdot Ramonesa [...] uważałem, niewątpliwie tak jak i on, że psychologii porównawczej jako nauki nie można zbudować na anegdotycznych podstawach. Większość historyjek to były po prostu przypadkowe świadectwa, uzupełnione amatorskimi opiniami przygodnych obserwatorów, których wyszkolenie psychologiczne było niemal bez znaczenia. Żywiłem wówczas wątpliwości, czy z umysłów zwierząt można wydobyć [...] dane stanowiące konieczny warunek nauki (Schultz, Schultz, 2008, str. 163).

Zasada Morgana głosi, że „nie wolno tłumaczyć zachowania zwierząt za pomocą wyższych procesów psychicznych, jeśli można je przypisać procesom niższym” (Morgan, 1894, Schultz, Schultz, 2008, str. 162; Bermudez, 2003). Należy też pamiętać, że jej autor był zwolennikiem darwinowskiej tezy o ciągłości ewolucyjnej. Prawo oszczędności miało na celu zredukowanie poziomu ufności społecznej i naukowej do interpretacji zachowań ZINL w perspektywie wyznaczanej przez Romanesa. Obecnie kanon Morgana krytykowany jest za brak odpowiedniego zdefiniowania zakresów pojęć *wyższych* i *niższych* procesów, co prowadzi do zamieszania terminologicznego, gdyż ostatecznie badacze operują na różnych aparatach pojęciowych. Zbieżne stanowisko prezentuje Jose Bermdez (2003), który wskazuje, że choć Morgan nie sformułował spójnej teorii „skali psychologicznej”, jego podejście zostało później zinterpretowane w kategoriach ewolucyjnych i rozwojowych, sugerując, że wyższe procesy psychiczne można zakładać tylko wtedy, gdy nie da się wyjaśnić zachowania przez procesy niższego rzędu. Bermúdez dostrzega jednak trudności i ograniczenia wynikające z literalnego stosowania omawianej reguły. Zwraca uwagę, że współcześnie kanon Morgana bywa używany jako argument przeciwko wyjaśnieniom mentalistycznym – nie tylko w etologii, ale także w psychologii rozwojowej, neurokognitywistyce i archeologii poznawczej. W praktyce prowadzi to do sytuacji, w której psychologiczne wyjaśnienia są traktowane jako wyjaśnienia ostatniej instancji, możliwe do przyjęcia tylko wtedy, gdy zawiodą wszystkie inne (Bermúdez, 2003, str. 7). Bermúdez argumentuje, że taka interpretacja kanonu jest zbyt rygorystyczna – sprzyja uproszczonym wyjaśnieniom, nawet tam, gdzie dostępne dane sugerują większą złożoność. Co ważne, autor zauważa, że Morgan nie uważał, że wyjaśnienia prostsze są zawsze bardziej adekwatne, lecz nawoływał do ostrożności w przyjmowaniu

wyjaśnień mentalistycznych. Zdaniem Bermúdeza zasada ta może pełnić użyteczną funkcję metodologiczną, jeśli nie blokuje z góry możliwości interpretowania zachowań w kategoriach psychologicznych – szczególnie tam, gdzie zachowanie nie daje się sprowadzić do prostych reguł typu bodziec – reakcja. Wskazuje również na konieczność lepszego określenia, czym dokładnie są psychologiczne wyjaśnienia i jak różnią się one od niepsychologicznych. W tym kontekście zwraca uwagę na rolę tzw. *input–output links* – psychologiczne wyjaśnienia stają się niezbędne dopiero wtedy, gdy nie da się już przewidzieć zachowania tylko na podstawie stałych mechanizmów reakcji na bodźce. Mimo wszystko należy pamiętać, że omawiany paradygmat pozwolił wypracować stosunkowo wysokie standardy metodologiczne w dziedzinie badań nad umysłem i zachowaniem, jaką jest psychologia porównawcza.

Pod wpływem koncepcji ewolucji konwergentnej w XIX wieku powszechne było stanowisko, że ZINL posiadają takie same zdolności psychiczne jak ludzie (Trojan, 2013, min. 2.50 - 9.08). Za powszechnością takiego stanowiska przemawia przypadek inteligencji Mądrego Hansa, czyli berlińskiego konia, który znany był na całym świecie np. ze swoich zdolności matematycznych⁴⁴. W roku 1900 berliński emerytowany nauczyciel matematyki, wiedziony przekonaniem ugruntowanym przez teorię Darwina, kupił ogiera, aby wyuczyć go kompetencji matematycznych. Istotny w kontekście niniejszej części jest wpływ, jaki miała owa anegdota na kształtowanie relacji człowiek – ZINL. Literatura wskazuje, że odkrycia Oskara Pfungsta stanowiły punkt zwrotny, w którym stało się jasne, że metoda introspekcji nie jest skuteczna w opisywaniu wielu złożonych zachowań zwierząt (Trojan, 2013, str. 11). Stało się tak, ponieważ wskazano wyjaśnienie, w ramach którego zachowanie konia było tylko reakcją na obserwowane przez niego sygnały pozawerbalne, także te niezależne od obserwatora (np. ruchy żrenic). Warto przypomnieć, że od tej właśnie historii pochodzi nazwa opisanego wyżej efektu mądrego Hansa (w literaturze możemy znaleźć go także pod nazwą efekt Roshentala). Jest to określenie na sytuację, w której obserwator nieświadomie i poza werbalnie wpływa na zachowanie obserwowanego zwierzęcia. Zjawisko to zaburza wyniki obserwacji, dlatego od wskazanej historii, badacze wielu dziedzin korzystają z wiedzy o tym zjawisku (wykorzystując je np. w kryminalistyce (Grzelak i in., 2023)).

Na początku XX wieku wyłonił się nowy nurt naukowy – nazywany

⁴⁴ Jego przypadek wspomniałam wcześniej, w części 2.4. *Efekty oczekiwania*.

behawioryzmem klasycznym, który w założeniu był kontynuacją zoopsychologii (Schultz, 2008, str. 255). Jego powstanie związane było ze sprzeciwem wobec stosowania ówczśnie przyjmowanych metod introspekcyjnych. Uważa się, że nurt ten jest odpowiedzią na nienaukowe podejście proponowane przez psychoanalizę (Kokocińska, Kaleta, 2015). W literaturze znajdziemy rozróżnienie na behawioryzm np. klasyczny, metodologiczny, radykalny oraz filozoficzny. Ze względu na cel niniejszej dysertacji, nie uważam za konieczne umieszczanie w niej szczegółowych opisów najważniejszych badań związanych ze wszystkimi wyżej wymienionymi stanowiskami. Pominę zagadnienie behawioryzmu filozoficznego (logicznego), czyli nurtu zajmującego się naturą umysłu i jego relacją do ciała (Szubaka, 2004). Przedmiotem zainteresowania wspomnianego nurtu, są zagadnienia dotyczące językowego opisu stanów mentalnych. W jego ramach uznaje się, że terminy dotyczące umysłu, to jedynie językowa forma opisu zachowania, a jego głównym postulatem jest empiryczna weryfikacja twierdzeń psychologicznych. W takiej perspektywie nie obserwuje się rozróżnienia na zachowania i stany mentalne, gdyż o tym drugim mówi się tylko w kategoriach pierwszego (Kokocińska, Kaleta, 2015, str. 4). Ważne, aby zauważyć, że stanowisko to nie spełnia wymogów empiryczności stawianych przez behawiorystów innych nurtów, a w konsekwencji, nie znajduje obecnie wielu zwolenników.

Za oficjalną datę powstania nurtu behawiorystycznego, uznaje się rok 1913, a dokładnie – publikację manifestu behawiorystycznego Johna Watsona. W związku z faktem, że był to pierwszy nurt przyjmujący wskazaną nazwę, ówczśnie nazywany jest klasycznym. Historia Hansa miała także silne znaczenie w kształtowaniu myśli watsonowskiej, prace wspomnianego już Oskara Pfungsta (Trojan, 2013 str.11). W pracach Watsona obserwujemy wyraźny wpływ teorii ewolucji (nacisk na celowość) oraz mechanicyzmu (całkowite odrzucenie terminów związanych z umysłem). Tutaj należy zaznaczyć, że sama idea behawioryzmu powstała wcześniej. Za czołowych przedstawicieli tego nurtu uznaje się np. Edwarda Thorndike (1874-1949), który sformułował prawo efektu⁴⁵, czy Iwana Pawłowa (1849-1936) – odkrywce warunkowania klasycznego. Nowy nurt psychologiczny stanowczo odrzucał zarówno prace syna Charlesa Darwina – Francisa, który obszernie opisywał kwestię świadomości u roślin, czy Alfreda Bineta, który twierdził, że organizmy jednokomórkowe mają

⁴⁵ Prawo efektu głosi, że zachowanie zwierząt jest efektem wzmocnienia.

zdolność do spostrzegania, jak i różnicowania obiektów⁴⁶. Watson podkreślał, że umysł ma naturę ciągłą i dynamiczną oraz że każda próba rozbicia go na elementy pierwsze, jest zniekształcaniem. Celem nowo powstałego nurtu było podniesienie poziomu psychologii jako nauki, poprzez skupienie uwagi na dających się zmierzyć obserwowalnych zmiennych. Początkowo behawioryzm nie negował istnienia umysłu, a jedynie podkreślał niemożliwość naukowego zbadania subiektywnych stanów mentalnych, ze względu na brak dostępnych narzędzi. Korzystając z metafory czarnej skrzynki, Watson całkowicie przekierował rozważania psychologiczne na zachowanie.

Rozwój psychologii ewolucyjnej udowodnił, że genetyczna transmisja pokoleniowa jest odpowiedzialna nie tylko za kształtowanie struktur biologicznych (takich jak narządy), ale także za przynajmniej niektóre zachowania. Tutaj należy zaznaczyć, że twierdzenia behawioryzmu klasycznego, np. dotyczące zasady łuku odruchowego, czy warunkowania klasycznego i instrumentalnego, świetnie opisują zachowania prostszych organizmów i podsystemów poznawczych. Z czasem jednak stało się jasne, że zwierzęta przetwarzają informacje napływające ze środowiska – pisali o tym badacze tacy jak np. Edward Tolman (1886-1959), czy Clark Hull (1884-1952) doprowadzając, do wzrostu zainteresowania procesami zachodzącymi w czarnej skrzynce. Pierwszy z wymienionych badaczy, uznawany jest za twórcę tzw. behawioryzmu celowościowego, w ramach którego uważa się, że zachowanie zwierząt zorientowane jest na cel. Tolman uzupełnił schemat bodziec – reakcja, poprzez wprowadzenie zmiennych pośredniczących (S-O-R), czyli niedających się obserwować czynników wewnętrznych, które są rzeczywistymi determinantami zachowania (Schultz, Schultz, 2008, str. 319). Klasycznym już przykładem zmiennej pośredniczącej jest głód. Chociaż nie można go zaobserwować, daje się on opisać za pomocą innych zmiennych empirycznych takich jak np. ilość spożytego pokarmu, szybkość spożycia pokarmu czy długość czasu, jaki upłynął od ostatniego posiłku. Twórca omawianego nurtu odrzucił sformułowane przez Thorndike prawo efektu oraz stwierdził, że wykonywanie określonych zachowań, utrwała relację między oczekiwaniami organizmu i sygnałami środowiskowymi. Perspektywa Tolmana przeciwstawiała się założeniom behawioryzmu klasycznego, jednak zachowała metodologię opartą na behawiorze,

⁴⁶ Warto zaznaczyć, że przytoczone stanowisko Bineta, znajduje potwierdzenie we współczesnych eksperymentach. Dla przykładu komórki układu odpornościowego: leukocyty i limfocyty, bardzo dobrze rozpoznają obiekty chorobotwórcze.

dlatego badacz ten wymieniany jest jako jeden z pierwszych przedstawicieli behawioryzmu metodologicznego.

W tym miejscu chcę zaznaczyć, że tradycje badania dotyczące zachowania zwierząt obecne na różnych kontynentach różniły się od siebie. W niniejszym opracowaniu skupię się na szkołach amerykańskiej i europejskiej, które rozwijały się równolegle, a ich przedstawiciele wyznaczali ramy badania zachowania zwierząt. Behawioryzm metodologiczny to szeroko stosowana w psychologii metoda badawcza, rozwijana na gruncie obydwu wybranych szkół. Może być realizowana za pośrednictwem eksperymentów i obserwacji, a jej funkcją jest formułowanie i weryfikowanie hipotez dotyczących praw oraz zasad rządzących zachowaniem (Kokocińska, Kaleta 2015). W tym samym czasie, na gruncie amerykańskim, rozwijał się nurt nazywany behawioryzmem radykalnym. Za twórcę niniejszej koncepcji uznaje się Burrhusa Skinera, który zorientowany był na opisywanie, a nie wyjaśnianie zachowania. Chociaż jego podejście pod wieloma względami stanowiło powrót do teorii Watsona – historyk MacLeod napisał nawet: „duch Watsona jest niezniszczalny. Odnowiony i oczyszczony, manifestuje się w pismach B.F. Skinera” (1959, str. 34). Jednak behawioryzmu radykalnego nie należy traktować jako kontynuacji nurtu klasycznego. Skinner posuwał się w swoich twierdzeniach znacznie dalej od Watsona, prezentując stanowisko, które z czasem zyskało miano koncepcji „pustego organizmu” (Schultz, Schultz, 2008, str. 329). Należy tutaj zaznaczyć, że biografowie omawianego badacza podkreślają, że nie negował on istnienia stanów mentalnych, a jedynie zaprzeczał ich wartości eksplanacyjnej (Richelle, 1993, str. 10). Odkrywca warunkowania instrumentalnego odrzucił klasyczny schemat S-R i wzorując się na koncepcji behawioryzmu metodologicznego, zaproponował ulepszoną koncepcję: bodziec → reakcja (zachowanie) → konsekwencje (wzmocnienie) → reakcje w przyszłości. Celem przyświecającym omawianemu nurtowi było wyeliminowanie wpływu badacza na wyniki eksperymentu. Choć podejścia neobehawiorystyczne doceniały już rolę zaskoczenia u uczących się zwierząt, czy chociażby wagę konsekwencji własnego zachowania, to mimo wszystko tradycyjny, oparty na idei „czarnej skrzynki” behawioryzm wciąż silnie wpływa na współczesne rozumienie psychiki zwierząt (Peirce, Hall 1980; 1982; 2010; Roesh i inni, 2012).

Metoda, jaką postulował behawioryzm metodologiczny stanowi kluczowe narzędzie dla powstałej w pierwszych dziesięcioleciach XX wieku etologii, czyli nauki o zachowaniu. Za jej początek uznaje się klasyczne już prace Jakoba Johann von

Uexküll (1864 - 1944), Konrada Lorenza (1903 - 1989) oraz Nikolaasa Tinbergena (1907- 1988). Polega na systematycznym obserwowaniu i wprowadzaniu uogólnień dotyczących zachowania zwierząt. W ramach metody behawioralnej zachowanie rozumiane jest jako ruchowa odpowiedź zwierzęcia na sytuację (Kokocińska, Kaleta 2015). Podczas gdy behawioryści klasyczni koncentrowali się na przeprowadzaniu ściśle kontrolowanych eksperymentów laboratoryjnych, etolodzy przyjęli metody obserwacji zwierząt w naturalnym środowisku. Etologia kładła nacisk na opisywanie spontanicznych zachowań typowych dla gatunku oraz nie stroniła od korzystania z pojęć mentalistycznych (Wróblewski, 2016). Mimo że nowy nurt wywodzący się z biologii zaczął kształtować się jeszcze przed II Wojną Światową, jego rozkwit nastąpił po opublikowaniu książki *Badania nad instynktem* (Tinbergen, 1951). Zasadą etologii było niejako uchylenie wieka watsonowskiej czarnej skrzynki. W tym, miejscu warto skupić się na kilku głównych założeniach behawioryzmu metodologicznego, ponieważ wiążą się one z rozważaniami zawartymi w rozdziale *V: Tendencyjność w badaniu teorii umysłu*.

Uexküll znany z wprowadzenia pojęcia *Umweltu*, które miało istotne znaczenie w zakresie biologii, psychologii oraz filozofii⁴⁷. Przeciwstawiał się dominującemu wówczas w psychologii behawioryzmowi klasycznemu, a jego koncepcja stanowiła alternatywę do badań z udziałem ZINL (Uexküll, 2013). Podkreślał, że nie istnieje jedna obiektywna rzeczywistość, lecz wiele różnych rzeczywistości, zależnych od perspektywy i zdolności danego organizmu. Twierdzenia Uexküll pomogły lepiej zrozumieć, zagadnienia percepcji i zachowania zwierząt. Głosiły, że każdy organizm odbiera i rozumie świat w sposób, który jest dostosowany do jego struktury anatomicznej, zmysłów i zachowań (Uexküll, 2013). Uexküll skupiał się na tym, że każde zwierzę posiada unikalne pojęcie świata, wynikające z jego zdolności percepcyjnych i interakcji ze środowiskiem. Przyjmując, że każdy organizm ma swoje własne subiektywne spojrzenie na świat, które wynika z jego biologicznych właściwości, należy przyjąć, że posiada także własny, indywidualny sposób interpretowania świata, który jest oparty na jego zdolnościach percepcyjnych, zmysłach i specyficznych zachowaniach. To podejście zainspirowało naukowców XX w. do analizowania reakcji zwierząt z ich własnej perspektywy, zamiast próby projektowania ludzkich oczekiwań na ich zachowanie. Koncepcja *Umweltu* miała istotny wpływ na

⁴⁷ Więcej tym autorze pojawiło się w części 2.3. *Myślenie analogiczne*.

poglądy i postawy związane z relacją człowiek – inne zwierzęta w nauce, poprzez zwrócenie uwagi na subiektywne spojrzenie zwierząt na świat oraz ich unikalne sposoby postrzegania i interakcji ze środowiskiem.

Innym kluczowym w kontekście niniejszego opracowania badaczem, jest opisywany wcześniej holenderski ornitolog, uznawany za jednego z ojców etologii — Tinbergen. Holenderski zoolog, w 1949 roku rozpoczął pracę w Oxfordzie, to właśnie za jego pośrednictwem etologia jako nauka, przedostała się do krajów anglojęzycznych. Najważniejsze elementy jego myśli przedstawione zostały w części 1.2. *Podstawowe kwestie terminologiczne*, dlatego tutaj pragnę jedynie przypomnieć, że wyznaczył on ramy wyjaśniania zachowań zwierząt, formułując cztery pytania, na które musi odpowiedzieć badacz konstruujący interpretację zachowania zwierzęcia. Pytania, pogrupowane są w perspektywie rodzaju wyjaśnienia (*ultimate, proximate*).

Z kolei Konrad Lorenz był austriackim zoologiem, etologiem i ornitologiem, uważanym za jednego z ojców nowoczesnej etologii. Początkowo studiował medycynę na Uniwersytecie Wiedeńskim, jednak później zainteresował się zoologią i otrzymał tytuł doktora w tej dziedzinie. Jego prace kładły nacisk na badanie zachowania zwierząt w naturalnym środowisku, co odróżniło go na tle behawiorystów klasycznych, a tym bardziej radykalnych. Uważał, że aby zrozumieć mechanizmy zachowania zwierząt, należy obserwować ich pełen zakres w naturalnych warunkach. Często używał metody antropomorfizacji, aby wyobrazić sobie stany psychiczne zwierząt, uważając, że są one zdolne do przeżywania emocji podobnych do ludzkich. Zachowywał przy tym ostrożność, która chroniła go przed nadmierną antropomorfizacją. Lorenz jest najbardziej znany ze swoich badań nad wdrukowaniem (*imprinting*), procesem, w którym młode zwierzęta, takie jak gęsi lub kaczki, tworzą silne więzi z pierwszym obiektem, który widzą po wykluciu się (Lorenz, 1935). Jest on kluczowy dla rozwoju zachowań społecznych i rozpoznawania gatunku⁴⁸. Udowodnił, że gęsi mogą związać

⁴⁸ Mechanizm wdrukowania zaproponowany przez Konrada Lorenza obecnie poddawany jest krytyce i reinterpretacji (Robledo, Cross, Boada-Bayona, Demogeot, 2022). Jego klasyczny eksperyment z młodymi gęsiami wykazał, istnienie okresu krytycznego, w którym młode osobniki przywiązują się do pierwszego dużego ruchomego obiektu, który zobaczą. Współczesne badania wskazują, że granice tego okresu nie są tak ostre i jednoznaczne, jak pierwotnie zakładano. Dodatkowo, koncepcja wdrukowania jako procesu całkowicie nieodwracalnego i niezależnego od dalszych doświadczeń jest obecnie kwestionowana. Badania wykazują, że *imprinting* może być bardziej plastyczny i zależny od kontekstu środowiskowego niż pierwotnie sądzono. Po dalsze informacje dotyczące krytyki mechanizmu

się nawet z człowiekiem, jeśli ten jest pierwszym poruszającym się obiektem, jaki widzą po wykluciu się. Co istotne w kontekście niniejszej pracy, wprowadził pojęcie *Fixed Action Patterns* (FAPs), które odnosi się do genetycznie uwarunkowanych zachowań, wyzwalanych przez określone bodźce. Te zachowania są niezbędne dla przetrwania i rozrodu zwierząt. Ich przykładem są złożone zachowania godowe (np. taniec ptaków), obrona terytorium i opieka rodzicielska.

Podsumowując, koncepcje dotyczące umysłów innych zwierząt obecne w nauce XX wieku dostarczyły cennych narzędzi metodologicznych oraz wywarły znaczący wpływ na współczesne perspektywy badawcze. Chociaż to behawioryzm metodologiczny przetrwał do dzisiejszych czasów, badania dotyczące wpływu eksperymentatora prowadzone na gruncie behawioryzmu radykalnego, dostarczają istotnych wskazówek dotyczących metod konstruowania eksperymentów. To w ramach behawioryzmu zaproponowano najważniejsze metody badawcze etologii, co pozwoliło na wyjaśnienie wielu aspektów zwierzęcego poznania.

Przełom nastąpił pod koniec XX wieku, kiedy ponownie wrócono do korzystania z terminów mentalistycznych (w tym kontekście odnoszących się szeroko do poznania⁴⁹ u zwierząt pozaludzkich), aby opisać ich zachowania (Wróblewski, 2016). Etolodzy zarzucali behawiorystom, że Ci nie patrzą na zwierzę jako całość, a na poszczególne części, które podlegają testowaniu (Lorenz, 1981, str. 38.). Przykładem może być stanowisko Bruce Moore i Susan Stuttard, którzy w 1979, roku skrytykowali odkrycia Edwina Guthrie i George’a Hortona (1946) zainspirowane pracami Edwarda Thorndike (1898) nad prawem efektu. Klasyczny eksperyment dotyczący wspomnianego prawa, polegał na umieszczeniu kota wewnątrz klatki. Jak opisywali behawiorysty – zwierzę nerwowo poruszało się we wnętrzu oraz ocierało się o pręty. Zachowanie to powodowało przesunięcie zasuwki otwierającej klatkę. Koty mogły wyjść i zjeść nieopodal leżący pokarm. Wykazano, że im częściej powtarzano badanie, tym szybciej koty wydostawały się z zamknięcia. Uznano, że jest to dowód, że nawet z pozoru złożone zachowanie związane ze zdolnością do rozwiązywania problemów, jak otwieranie klatki, może zostać wyjaśnione przez proste mechanizmy uczenia się

proponowanego przez Lorenza zachęcam do lektury artykułu *Back to basics: A re-evaluation of the relevance of imprinting in the genesis of Bowlby's attachment theory* (Robledo, Cross, Boada-Bayona, Demogeot, 2022).

⁴⁹ Przyjmując, że poznanie to „mentalne przetwarzanie sygnału zmysłowego w wiedzę na temat środowiska oraz stosowanie tej wiedzy” (de Waal, 2018, str. 21).

poprzez warunkowanie instrumentalne. Moore i Stuttard stwierdzili, że wiele badań dotyczących warunkowania instrumentalnego pozbawiona jest kontekstu zawierającego istotne informacje o naturalnym zachowaniu określonych gatunków. Uważają, że brak kontekstu utrudnia rozpoznanie typowych dla gatunku reakcji na sytuację. To duże niedociągnięcie, gdyż jest to kluczowe, aby zrozumieć procesy uczenia się. Prezentują replikację klasycznego eksperymentu oraz argumentują, że badacze nie uwzględnili typowego dla kotowatych zachowania. Behawior polegający na wykonywaniu charakterystycznego ruchu głową związanego z ocieraniem się jest zachowaniem charakterystycznym dla powitania i zalotów. Zatem, zwierzęta badane przez Guthrie i Hortona (1946) witały się z osobami je obserwującymi, co oznacza, że motywacją dla tego zachowania była obecność człowieka, a nie chęć uzyskania nagrody w postaci pożywienia. Gdy zostawiono koty w pomieszczeniu same – bez obserwatora, nie były w stanie wydostać się z klatki (Moore, Stuttard, 1979). Analizując zachowanie kota Guthrie i Horton (1946) nie uwzględnili zachowań typowych dla gatunku, co doprowadziło do stworzenia nietrafnej interpretacji badania (Moore, Stuttard, 1979).

Innym ciekawym i bardziej współczesnym przykładem wpływu eksperymentatora na wyniki badania, jest odkrycie reakcji stresowej u myszy na badaczy płci męskiej (Sorge, Martin, Isbester, 2014). W przytoczonym badaniu odkryto, że bodźce zapachowe, których źródłem jest męski eksperymentator, mogą wywoływać analgezję stresową. Analgezja stresowa to zjawisko, w którym organizm doświadcza zmniejszenia odczuwania bólu w odpowiedzi na stresujące sytuacje. Co ciekawe reakcja ta występuje, nawet gdy mysz ma kontakt tylko z koszulką, która wcześniej noszona była przez mężczyznę. Biorąc pod uwagę, że stres jest czynnikiem wpływającym na kondycję poznawczą zwierząt badania we wskazanym zakresie, stawiają pod znakiem zapytania wiele wyników badań dotyczących zdolności poznawczych myszy, które prowadzone były z udziałem męskich eksperymentatorów (Sorge, Martin, Isbester, 2014).

Zatem możemy powiedzieć, że w historii nauki wyróżniamy dwie grupy postaw związane z relacją człowiek – ZINL. Pierwszą nazwać można ogólnie antropomorfizmem, drugą antroponegacjonizmem (Wróblewski, 2016). Pierwszy z wymienionych może odnosić się do preferencji przypisywania zwierzętom innym niż ludzi zdolności poznawczych uznawanych za dystynktywne dla człowieka. Natomiast antroponegacjonizm, związany jest z tendencją do negowania *a priori* człowiekopodobnych cech u ZINL lub zwierzęcopodobnych cech u ludzi (de Waal,

2018, str. 39). Pierwszy z wymienionych, ma głębokie korzenie kulturowe, przejawiające się np. przez postawy powodowane przekonaniami religijnymi (takimi jak np. chrześcijaństwo). Dodatkowo wraz z odrzuceniem wyjaśnień zachowania zwierząt, odnoszących się do zdolności poznawczych i stanów mentalnych, oparta była na myśleniu analogicznym, opisanym w poprzednim rozdziale. Aktualnie w ramach nurtów kognitywnych korzysta się z analogii zachowań ZINL do człowieka, zawsze jednak należy brać pod uwagę kontekst w postaci *Umweltu* uczestników badania.

3.2. Dwie perspektywy interpretowania eksperymentów z udziałem ZINL

W psychologii porównawczej z zakresu poznania społecznego wyróżnić można dwie kluczowe perspektywy, które stanowią przeciwległe bieguny spektrum — mentalistyczną i behawioralną (Povinelli, 1996; Whiten, 1996; Shettleworth, 2010a, 2010b). Pierwsza z wymienionych zakłada, że interpretacje eksperymentów z udziałem zwierząt pozaludzkich, w uzasadnionych przypadkach mogą odwoływać się do stanów mentalnych uczestników badania. Druga zakłada, że wyniki eksperymentów z udziałem ZINL powinny być wyjaśniane w kategoriach interpretowania zachowania innych (*smart behaviour-reading*), czyli bez odniesienia do stanów mentalnych (Whiten, 1996). Jej zwolennicy uważają, że teoria umysłu jest konsekwencją pewnych unikalnych dla człowieka własności, jak kultura czy język⁵⁰. Jej rozwój możliwy jest zatem dzięki wyjątkowym ludzkim praktykom socjalizacyjnym (Heyes Frith, 2014; Krupenye, Call, 2019). Innym stanowiskiem przyjmowanym w nurcie behawioralnym, jest to opierające się na twierdzeniach głoszących, że eksperymenty konstruowane w paradygmatach odwołujących się do reakcji na obiekty charakteryzujące się samowzbudzającym ruchem oraz śledzeniem kierunku patrzenia, można wyjaśnić prostszymi procesami, niewymagającymi zaangażowania przypisywania agencyjności. Tym samym, podważają dobrze ugruntowane historycznie i merytorycznie interpretacje eksperymentów z zakresu psychologii rozwojowej (Wimmer, 1983; Baron-Cohena, 2009). Koncepcje powszechnie uznawane za najlepsze dostępne wyjaśnienia określonych zjawisk są wynikiem długotrwałych badań i konsensusu wśród ekspertów, a ich podważanie bez solidnych dowodów może prowadzić do chaosu w dyskursie naukowym. W psychologii rozwojowej eksperymenty są projektowane tak, aby

⁵⁰ Szczegółowy opis argumentacji opartej o kompetencje językowe, znajduje się w części 4.3.1. *Argument dotyczący braku reprezentacji językowych u ZINL*.

kontrolować zmienne i ustalać związki przyczynowo-skutkowe, co pozwala na sformułowanie precyzyjnych twierdzeń. Zatem każde podważanie tych ustaleń powinno być poparte równie solidnymi dowodami, aby nie naruszać zasad rzetelności i dokładności naukowej. Niezależnie od argumentacji, w ramach perspektywy behawioralnej interpretacji o charakterze mentalistycznym w zakresie teorii umysłu, uznaje się, za przykłady nadmiernego antropocentryzmu, łamiące zasadę Morgana.

Podstawowym argumentem wysuwany w ramach perspektywy mentalistycznej, jest twierdzenie głoszące, że niektórych procesów nie da się zredukować do prostszych mechanizmów, takich jak warunkowanie czy proste asocjacje. Zwolennicy tego stanowiska wskazują, że pewne reakcje, takie jak rozwiązywanie problemów w nowym kontekście czy elastyczne dostosowywanie do zmieniających się warunków, sugerują istnienie reprezentacji mentalnych. Podkreślają również, że zakładanie istnienia stanów mentalnych u zwierząt pozaludzkich może być zgodne z zasadą parsymonii, o ile prowadzi do spójniejszych modeli wyjaśniających zachowanie, niż wyłącznie odwołanie się do warunkowania czy prostych asocjacji. Tym samym, badacze przyjmujący perspektywę mentalistyczną zwracają uwagę na analogie między eksperymentami przeprowadzanymi z udziałem dzieci ludzkich, a badaniami dotyczącymi zwierząt pozaludzkich. Dla przykładu śledzenie kierunku patrzenia oraz inne mechanizmy przypisywania agencyjności innym, stanowią kluczowy punkt odniesienia dla argumentacji mentalistycznej (Baron-Cohen, 2009). W eksperymentach z udziałem niemowląt wykazano, że już na bardzo wczesnym etapie rozwoju dzieci, zwracają uwagę na kierunek patrzenia innych osób oraz używają go jako podstawowej wskazówki dotyczącej intencji obserwującego. Podobne badania przeprowadzono w odniesieniu do naczelnych i innych gatunków zwierząt. Eksperymenty z zakresu śledzenia kierunku patrzenia pokazały, że niektóre naczelne, w tym szympansy, potrafią podążać za wzrokiem innych osobników, co może sugerować, że posiadają zdolność do reprezentowania cudzych stanów. W eksperymentach tego typu badacze manipulują bodźcami wzrokowymi, np. poprzez zmianę punktu, na który patrzy inny osobnik, a następnie sprawdzają, czy uczestnicy badania również patrzą na ten sam punkt. Jeśli zwierzęta reagują w prawidłowy sposób, sugeruje to, że rozumieją relację między kierunkiem patrzenia, a potencjalnym obiektem zainteresowania. Mentalistyczna interpretacja tych wyników zakłada, że uczestnicy nie tylko reagują na widoczne wskazówki behawioralne (*smart behaviour-reading*), ale także posiadają pewien poziom reprezentacji stanów mentalnych innych (*mindreading*).

Jak postaram się wykazać nieco później⁵¹ przypisywanie prostych stanów wolicjonalnych innym oraz śledzenie kierunku patrzenia u ludzkich dzieci interpretowane są jako wczesne przejawy zdolności czytania umysłów innych. Stanowią one naturalne wyposażenie ludzkiego systemu poznawczego, co sugeruje, że są filogenetycznie młodsze niż człowiek. Dlatego nie sprowadza się ich do jeszcze prostszych mechanizmów asocjacyjnych. Teorie behawioralne sugerują, że zwierzęta uczą się tych reakcji poprzez powtarzalne skojarzenia, jednak dane empiryczne podważają to stanowisko (Johnson Dziurawiec, Ellis, Morton, 1991). Na przykład, niemowlęta wykazują zdolność do śledzenia wzroku dorosłych już w pierwszych miesiącach życia, czyli zanim zdążą nabyć wystarczające doświadczenie asocjacyjne. Podobnie, w badaniach na zwierzętach innych niż ludzie zaobserwowano przypadki podążania za spojrzeniem w sposób kontekstowo elastyczny, co sugeruje, że nie jest to jedynie efekt prostych reguł skojarzeniowych, lecz skutek bardziej złożonych procesów poznawczych (Tomasello, Call, Hare, 1999). Co więcej, gdyby opisywane mechanizmy opierały się wyłącznie na schematach warunkowania behawioralnego, to w sytuacjach nowych i niestandardowych ich skuteczność powinna drastycznie spadać. Eksperymenty pokazują, że zarówno dzieci, jak i zwierzęta potrafią dokonywać złożonych generalizacji⁵² znanych schematów w nowych kontekstach. Ciekawym przykładem są badania nad oszustwem (*deceptive behavior*), które wykazały, że osobniki różnych gatunków potrafią ukrywać informacje przed rywalami, co wymaga rozumienia, że inna istota posiada odrębną perspektywę poznawczą (Hare, Call, Tomasello 2001). Wykazano, że zwierzęta potrafią unikać pokazywania innym osobnikom lokalizacji pożywienia lub celowo wprowadzać ich w błąd udając fałszywe intencje, lub zmieniając lokalizację ukrywanego jedzenia, co wymaga dostosowania zachowania do tego, czy inny nie tylko widział jednostkę, ale też w zależności od jego statusu społecznego (Abbott, Abbott, Grant, 1975; Lima, 1985; Emery, Clayton, 2001; Bugnyar, Kotrschal, 2002). Takie działania wskazują na zdolności z zakresu *mindreadingu* na poziomie wyższym niż proste przypisywanie stanów wolicjonalnych czy percepcji. Wymagają manipulowania swoim zachowaniem, tak aby zmylić obserwatora, co trudno byłoby wytłumaczyć jedynie za pomocą prostych mechanizmów asocjacyjnych. Gdyby szympansy jedynie reagowały na powtarzające się schematy, ich

⁵¹ Wspomniane opracowanie znajduje się w rozdziale IV: *Kluczowe koncepcje teorii umysłu*.

⁵² Wspomniane opracowanie znajduje się w rozdziale V: *Tendencjonalność w badaniu teorii umysłu*.

zdolność do elastycznego dostosowywania strategii do zmieniających się warunków byłaby znacznie ograniczona. Co więcej, biorąc pod uwagę złożoność teorii umysłu, w narracji behawioralnej widoczny jest wyraźny przeskok jakościowy między człowiekiem a innymi gatunkami (Dunbar, 2013). Luka ta jest na tyle duża, że generuje zarzut o braku ciągłości ewolucyjnej, na który narażone jest omawiane stanowisko. W takiej perspektywie interpretacje behawioralne w zakresie teorii umysłu postrzega się jako przykłady nadmiernego antroponegacjonizmu, które, poprzez konsekwentne odniesienie do zasady Morgana łamią kluczową zasadę teorii ewolucji⁵³.

W tym miejscu warto także przypomnieć, o nieścisłościach terminologicznych w rozważaniach dotyczących teorii umysłu. Wstępne wprowadzenie do tej problematyki umieściłam w części 1.2. *Podstawowe rozróżnienie terminologiczne*, jednak rozważania dotyczące wskazanego zagadnienia wymagają dłuższego wyjaśnienia. W literaturze tematu pojawia się pewnego rodzaju utożsamienie ToMM z inną obszerniejszą kategorią, określaną *mindreadingiem* (Baron-Cohen, 2009). Chociaż Baron-Cohen (2009) dopatruje się źródła wskazanego problemu w artykule Premacka i Woodrooffa (1978), uważam, że praca ta bardziej wyprzedziła własne czasy, niż przyczyniła się do omawianego błędu. Podstawowym źródłem wskazanej nieścisłości jest złożoność *mindreadingu* oraz komplementarność mechanizmów leżących u podstaw poznania społecznego. Należy pamiętać, że w czasach, gdy autorzy *Does the chimpanzee have theory of mind* przeprowadzali swoje eksperymenty, uznawano, że tylko człowiek zdolny jest do wyjaśniania zachowania innych osobników w oparciu o nieobserwowalne byty (stany mentalne). Dlatego twierdzenia Premacka i Woodruffa (1978) odnoszą się do przypisywania prostych stanów wolicjonalnych, takich, jak intencje czy cele. Podejmują także nieudaną próbę uzyskania wiarygodnych wyników w zakresie wiedzy, jednak, jak wyjaśnię niżej⁵⁴, ostatecznie nie uda im się wyeliminować kluczowych problemów proceduralnych. W późniejszych latach Call i Tomasello (2011), w swoim przełomowym artykule *Does the chimpanzee have theory of mind. 30 years later* wprowadzą przytaczane już rozróżnienie na teorię umysłu w sensie szerokim i wąskim⁵⁵. W tym momencie chcę podkreślić, że przeddefiniowali oni omawianą zdolność tak, aby kluczowe w niej było wnioskowanie na podstawie

⁵³ Szczegółowa polemika między stanowiskami antropomorficznymi — antropoegecjonistycznymi, znajduje się w rozdziale V: *Tendencjonalność w badaniu teorii umysłu*.

⁵⁴ Wspomniane opracowanie znajduje się w części 5.1.1. *Przełomowe badania Premacka i Woodruffa*.

⁵⁵ Szczegółowy opis tego artykułu znajdzie się w części 5.1.3. *Badania Calla i Tomasello*.

falszywych przekonań. Tym samym, w swoich rozważaniach przekierowali uwagę na to co Baron-Cohen określa mianem ToMM. Autor artykułu *Rozwój zdolności czytania innych umysłów: cztery etapy* zapożycza pojęcie teorii umysłu od Alana Leslie'go (1993), który w swojej książce *A theory of agency* zaproponował model zdolności do tworzenia metareprezentacji (Baron-Cohen, 2009). Ściślej mówiąc, Leslie (1993), korzystając z terminu *teoria umysłu* ma na myśli, to samo co Baron-Cohen opisuje, jako komponent zdolności *mindreadingu*: ToMM. Wydaje się zatem, że zabieg terminologiczny Calla i Tomasello, zgodny był z ogólnym trendem w psychologii, zapoczątkowanym m.in. przez przełomowe badania Baron-Cohana (2009). Jego celem było lepsze zrozumienie tego, na czym dokładnie polega czytanie umysłów sprowadzone do zdolności do posiadania teorii umysłu.

Chociaż problemy terminologiczne dotyczące zdolności z zakresu poznania społecznego są szeroko omawiane na gruncie badań dotyczących empatii, przyjmowania perspektywy innego czy teorii umysłu, to ich konsekwencji nieustannie można upatrywać w szeroko komentowanych pracach z zakresu psychologii. Dlatego już teraz chcę czytelników uczulić na wieloznaczność nomenklatury wykorzystywanej w niniejszej dysertacji. Warto zachować szczególną ostrożność do języka, w którego kategoriach przywoływani przeze mnie autorzy formułują swoje tezy i teorie. W następnych częściach niniejszego opracowania, korzystając z terminu *teoria umysłu*, będę się odnosić do określonego aspektu zdolności do *mindreadingu*.

3.3. Komponent poznawczy

Przejdźmy teraz do prezentacji i analizy stereotypów i mitów, które mogą przyczyniać się do powstawania i utrzymywania uprzedzeń wpływających na interpretację wyników badań dotyczących procesów poznawczych ZINL. W szczególności chodzi o badania z zakresu zdolności do tworzenia teorii umysłu, czyli procesu przez dekady przypisywanego wyłącznie człowiekowi. Tendencyjność interpretacji badań naukowych jest dobrze udokumentowanym zjawiskiem, w kontekście takich mechanizmów jak efekty oczekiwań czy dysonans poznawczy, szczegółowo opisanych wcześniej.

Dla przypomnienia postawa rozumiana jest jako „osąd lub uczucie względem osoby, idei lub obiektu: pozytywne, lub negatywne” (Openstax, 2022). Główną osią moich zainteresowań jest spektrum, którego bieguny można określić jako:

antropomorfizm —antroponegacjonizm, a każda ze wskazanych postaw może przyjmować formę uprzedzenia o charakterze kulturowym i ukrytym. Uważam, że konstruowanie lub komentowanie interpretacji wyników badań z udziałem ZINL, jest ściśle związane z przyjmowaniem postaw wobec badanych zwierząt, gdyż narracje, w których to się odbywa, mają charakter opowieści, konstytuujący porządek społeczny badacza. Przyjmowanie perspektywy mentalistycznej dystynktywne jest dla postaw antropomorficznych, u podstaw których leży pierwotny mit o *jedności człowieka z naturą*, natomiast perspektywa behawioralna, ściśle związana jest z narracją o *wyjatkowości człowieka*. Określone mity mogą funkcjonować w umyśle jednostki jednocześnie, niezależnie od siebie, nie wywołując przy tym sprzeczności w systemie poznawczym.

Przekonania zgodne z różnymi stereotypami biorącymi udział w procesach kategoryzacji społecznej, kształtują postawy jednostki wobec obiektów obecnych w codziennym życiu, takich jak produkty w sklepach czy oferty gastronomiczne. Ich formowanie jest procesem złożonym, u podstaw którego leży wiele mechanizmów poznawczych, w tym tych omówionych w rozdziale *II: Wybrane mechanizmy poznawcze leżące u podstaw tendencyjności związanej z uprzedzeniami*. Jak wspomniałam wcześniej, przynależność do określonej kategorii, w tym przypadku *człowiek*, przeciwstawiana jest kategorii *zwierzęta*. Grupa *my* charakteryzuje się pozytywnymi atrybutami, takimi jak: najinteligentniejszy gatunek, jedyny posługujący się językiem⁵⁶, czy zdolny do tworzenia teorii umysłu. W następnych częściach poddam analizie mity i stereotypy, leżące u podstaw oceniania wiarygodności interpretacji wyników eksperymentów z udziałem ZINL. Co będzie stanowić punkt wyjścia do późniejszych rozważań, zawartych w rozdziale *V: Tendencjność w badaniu teorii umysłu*.

3.3.1. Mity i inne narracje

We współczesnej kulturze popularnej mity zajmują wyjątkowo trwałe miejsce. Narracje o bogach (np. Kratos w serii gier God of War), herosach (jak Herkules czy

⁵⁶ Aspekt roli języka w zdolności do tworzenia teorii umysłu zostanie rozwinięty w rozdziale *V: Tendencjność w badaniu teorii umysłu*, gdzie pokrótce omówię argument dotyczący braku reprezentacji językowych u zwierząt innych niż ludzie oraz w *VI: niesprawiedliwość wynikająca z uprzedzeń wobec zwierząt innych niż ludzie*, gdzie szerzej omówię kwestię traktowania zwierząt innych niż ludzie, jako podmioty wiedzy.

Kirke), bohaterach baśni (np. Kopciuszek, Śpiąca Królewna, Alicja z Krainy Czarów) czy duchach (Dziady, Wiedźmin) stanowią nie tylko rozrywkę, ale także swoisty sposób wyjaśniania świata. Jak wskazują badacze mitów, myślenie mitologiczne nie jest wyłącznie domeną społeczeństw archaicznych. Przeciwnie, we współczesnych, świeckich i rozwiniętych społeczeństwach odgrywają nadal istotną rolę, choć ich funkcjonowanie przybiera bardziej złożone i często trudniejsze do uchwycenia formy (Czeremski, 2021). Wyjaśnienie omawianych struktur pozostaje niepełne bez odwołania do ustaleń kognitywistyki. Kluczowe jest tu rozumienie „naturalności” jako poznawczej łatwości i automatyzmu (Barrett, 2011; Czeremski, 2021, s. 16). Tę automatyzację należy z kolei rozumieć w kontekście szeroko opisywanego mechanizmu skąpcia poznawczego — dążenia do minimalizacji wysiłku, gdyż mit oferuje gotowe schematy pozwalające zredukować złożoność świata otaczającego jednostkę. Jego charakter może być wrodzony albo nabyty. W pierwszym przypadku chodzi o pojawienie się określonych form aktywności (np. mitotwórstwa) w toku rozwoju jednostki (ontogenezy), jako wyrazu wrodzonych predyspozycji. Dzieci, które konstruują osobiste opowieści wyjaśniające otaczającą je rzeczywistość, stanowią przykład tego zjawiska. W sensie nabytym mit rozumiany jest jako forma narracyjna, przyswajana kulturowo, jednak jego poznawcza naturalność przejawia się w łatwości, z jaką zostaje zinternalizowany. Czyli przetwarzanie mitów nie angażuje świadomego, refleksyjnego systemu drugiego, lecz odbywa się automatycznie, zgodnie z tendencjami poznawczymi charakterystycznymi dla systemu pierwszego (Kahneman, 2012).

Religie — jako kulturowo usankcjonowane systemy przekonań — wciąż pełnią ważną funkcję w kształtowaniu wyobrażeń o świecie, a tym samym, wciąż stanowią jądro wielu systemów kulturowych. Jak wskazywałam już w pierwszej części niniejszego rozdziału, religia wpływa na kształtowanie zarówno indywidualnych, jak i kulturowych postaw — w tym także uprzedzeń wobec ZINL. Oznacza to, że nawet osoby, które nie identyfikują się jako wierzące, mogą funkcjonować w ramach poznawczych struktur wzmacnianych przez mitologiczne narracje religijne obecne w ich kulturze. Interesujące w kontekście niniejszego opracowania, wydaje się zagadnienie świeckich form mitu — choć pozbawione sakralnego statusu, nadal skutecznie organizują przekonania jednostki. Obecne są w teoriach spiskowych,

legendach miejskich czy internetowych *copypastach*⁵⁷. Funkcjonują w literaturze, kinie i innych produktach kultury masowej, ale również wykorzystywane są w socjotechnikach — np. narracja oblężonej twierdzy⁵⁸ — kształtując postawy polityczne czy konsumenckie. Istnieją także mity indywidualne, szczególnie widoczne u dzieci, które spontanicznie tworzą wyjaśnienia porządkujące ich rzeczywistość.

Wielu badaczy uważa, że myślenie mityczne jest naturalnym dla człowieka sposobem interpretowania rzeczywistości. Dlatego możliwe jest określenie jego ewolucyjnych korzyści i wad. W tym celu chcę odwołać się do zasady głoszącej, że nie trafność przekonań, ale skuteczność zachowań decyduje o ich adaptacyjnej wartości (Tinbergen, 1963). Nawet jeśli przekonania oparte są na fałszywych założeniach, mogą prowadzić do skutecznych działań (Czeremski, 2021). David Sloan Wilson wskazuje kilka powodów, dla których mity mogą być skuteczniejsze niż racjonalne wyjaśnienia, a najistotniejsze z nich to (za: Czeremski, 2021, s. 14):

- a. Opóźnione konsekwencje – skutki wielu działań są odroczone w czasie i trudne do zaobserwowania, co sprawia, że jednostka potrzebuje fikcyjnych instrukcji, które zastępują brak wiedzy.
- b. Korzyści zbiorowe – pozytywne konsekwencje niektórych działań (np. altruistycznych) obejmują społeczność, a nie jednostkę, która może nawet ponieść koszty. Aby skłonić ją do działania, potrzebne są narracje, które nadają sens jej poświęceniu.
- c. Optymalizacja poznawcza – mit jako krótka i emocjonalnie angażująca narracja lepiej odpowiada ograniczeniom poznawczym człowieka niż złożone, ale trudne do przyswojenia wyjaśnienia.

Co ciekawe, w swojej książce Czeremski (2021) łączy modularne podejście z założeniami teorii ucieleśnienia — tradycyjnie przeciwstawianej modelowi komputacyjnemu (Hohol, 2017; Czeremski, 2021, s. 21). W jego ujęciu podejścia te nie

⁵⁷ Jest to rodzaj memu w formie pisemnej, przybierają formę wymyślonych historii, kopiowanych i rozpowszechnianych w internecie. Ma humorystyczny charakter, a jej częstymi elementami są m.in. wulgaryzmy czy błędy ortograficzne oraz gramatyczne.

⁵⁸ Konsekwencją socjotechniki oblężonej twierdzy jest stan nazywany syndromem oblężonej twierdzy. Polega na wywołaniu poczucia zagrożenia płynącego z grupy obcej (Ziółkowski, 2016). W takim rozumieniu socjotechnika rozumiana jest jako zbiór dyrektyw określających zasady skutecznego zarządzania społeczeństwem.

są wcale sprzeczne, lecz wzajemnie uzupełniające, co znajduje swój wyraz m.in. w koncepcjach metafory pojęciowej i teorii amalgamatu, omówionych w książce *Mit w umyśle: Ewolucyjno-kognitywne podstawy form mitycznych*.

Ramy narracyjne funkcjonują jako schematy, które nadają znaczenie zdarzeniom i pomagają w ich interpretacji. Są nie tylko narzędziami poznawczymi, ale także elementami tożsamości – pozwalają jednostkom i grupom rozumieć siebie, swoją przeszłość i swoje miejsce w świecie. Ich zaletą jest fakt, że nie podlegają one warunkom prawdziwości, a zatem nie wymagają weryfikacji – są przyjmowane jako oczywiste, a tym samym odporne na zmianę. Jak zauważa Marcin Napiórkowski (2019; 2022), narracje odgrywają kluczową rolę w kształtowaniu społecznego obrazu świata. To właśnie poprzez opowieści – zarówno te codzienne, jak i zakorzenione w tradycji kulturowej – porządkujemy rzeczywistość, określamy, co jest ważne, oraz konstruujemy tożsamość indywidualną i zbiorową. Co ciekawe w literaturze przedmiotu znaleźć można koncepcję tożsamości narracyjnej (Ricoeur, 2005), która wnosi istotny wkład w rozumienie *ja* jako dynamicznego i osadzonego w czasie konstruktu. W przeciwieństwie do ujęć esencjalistycznych, traktujących tożsamość jako niezmienny rdzeń podmiotu, teoria tożsamości narracyjnej zakłada, że jednostka konstruuje siebie poprzez opowiadanie historii o sobie samej. Tożsamość nie jest tu dana raz na zawsze, lecz tworzona i przekształcana w procesie opowiadania i reinterpretowania własnych doświadczeń. Zatem opowiadanie pełni funkcję pośrednika między tym, co konceptualne a tym, co przeżywane. Narracja pozwala jednostce zintegrować rozproszone, często sprzeczne doświadczenia w względnie spójną strukturę znaczeniową. Ricoeurowskie *ja* nie istnieje poza narracją – kształtuje się w niej i poprzez nią. Tym samym tożsamość – mówiąc nieco skrótowo i upraszczając tę koncepcję – staje się procesem interpretacyjnym, w którym doświadczenie nabiera sensu dzięki wpisaniu go w opowieść o własnym życiu. W omawianej koncepcji to właśnie narracyjność umożliwia jednostce zachowanie ciągłości tożsamości mimo zmienności doświadczeń, a jednocześnie otwiera przestrzeń dla autorefleksji, reinterpretacji i przemiany.

Przykładem badań nad mitami kulturowymi są interpretacje baśni Braci Grimm autorstwa Pullmana (2012), który pokazuje, że nawet opowieści uznawane za neutralne czy uniwersalne są silnie osadzone w światopoglądach i przekazują określone wartości. W takim ujęciu mity określają ramy interpretacyjne, wpływające na postawy, przekonania i decyzje społeczne. W swoich pracach Napiórkowski zaprezentował

analizę mitów, związanych z postawami politycznymi np. (softpatriotyzm — turbopatryjotyzm) oraz ciekawy przegląd współczesnych narracji obecnych w myśleniu o przyszłości (technooptymizm — technopesymizm); (Napiórkowski, 2019). W książce *Naprawić przyszłość: Dlaczego potrzebujemy lepszych opowieści, żeby uratować świat* (2022) analizuje dwa dominujące sposoby myślenia o technologii. Obie te postawy są w istocie dwoma biegunami jednego spektrum – opowieściami, które próbują nadać sens zmianom technologicznym i ich wpływowi na rzeczywistość jednostki. Technooptymizm to postawa, cechująca się zbiorami przekonań głoszącymi, że technologia uratuje świat. Zatem osoby prezentujące wskazane stanowisko wierzą, że dzięki postępowi naukowemu i technologicznemu ludzkość przezwycięży najpoważniejsze problemy – od kryzysu klimatycznego po choroby i śmierć. Dlatego nazywana jest narracją postępu. W skrajnej wersji przyjmuje formę technologicznej utopii: świat, w którym sztuczna inteligencja, nanotechnologia, genetyka i robotyka rozwiązują wszystkie trudności. Typowymi przykładami tego rodzaju narracji są te obecne w transhumanizmie czy dziełach kultury masowej takich jak książki, czy filmy *science-fiction*. Afektywny komponent omawianego uprzedzenia manifestuje się m.in. w redukowaniu dysonansu poznawczego, może prowadzić do powierzchownych lub zniekształconych interpretacji doniesień naukowych – zwłaszcza w przekazach medialnych tworzonych przez osoby bez odpowiedniego przygotowania merytorycznego. W skrajnej formie postawa ta przejawia się w narracjach głoszących, że nie należy martwić się o przyszłość, ponieważ technologia rozwiąże wszystkie problemy człowieka, a jej rozwój zagwarantuje bezpieczną przyszłość. Z kolei technopesymyści głoszą, że technologia zniszczy świat, tym samym postrzegają ją jako siłę prowadzącą do alienacji, degradacji środowiska, pogłębiania nierówności i utraty człowieczeństwa. W tym nurcie pojawiają się narracje apokaliptyczne – o sztucznej inteligencji, która przejmie kontrolę nad światem i w najlepszym wypadku, zniewoli człowieka, czy o zautomatyzowanym społeczeństwie pozbawionym sensu i pracy oraz zdegradowanym moralnie. Omawiana postawa wiąże się często z nostalgią za przeszłością i pragnieniem powrotu do „prostszej życia”, co sprawia, że współwystępuje z inną postawą określaną w literaturze jako *juvenoia* (Vsauce, 2016; Finkelhor, 2022). Cechuje się ona strachem lub wrogością żywioną przez przedstawicieli starszych pokoleń względem młodszych, lub w stosunku do kultury młodzieżowej w ogólności. Napiórkowski podkreśla, że zarówno technooptymizm, jak i technopesymizm są narracjami upraszczającymi złożoną rzeczywistość. Zamiast

namysłu nad technologią, sprowadzają ją do mitu: albo jako zbawienia, albo jako katastrofy. To – zdaniem autora – osłabia sprawczość społeczną. W wyniku braku dyskusji o konkretnych rozwiązaniach, np. etycznych, prawnych czy politycznych, związanych z technologią, społeczeństwo przyjmuje tendencje, które można sprowadzić do biernego oczekiwania na wybawienie lub zagładę.

W badaniach nad poznaniem ZINL wyróżnia się opisane we wcześniejszej części moich rozważań spektrum antropomorfizm — antroponegacjonizm (de Wal, 2018). Obejmuje ono zakres postaw, które ludzie przyjmują w odniesieniu do nadawania zwierzętom innym niż ludzie cech, uznanych za typowe dla człowieka (*human-like*). Osoby prezentujące pierwszą z wymienionych postaw, wykazują tendencje do dokonywania wskazanej atrybucji. Natomiast antroponegacjonizm to tendencja do odrzucania lub niedowierzania w możliwość istnienia złożonych procesów poznawczych poza gatunkiem ludzkim, nawet mimo mocnych przesłanek. Obie postawy pełnią funkcję strukturalnie podobną do technooptymizmu i technopesymizmu – są to narracyjne skrajności, które służą uporządkowaniu złożonej rzeczywistości i ograniczeniu niepewności poznawczej. Tak, jak technooptymizm zakłada, że technologia rozwiąże problemy ludzkości, tak antropomorfizm zakłada, że wszystkie zwierzęta są do nas bardzo podobne, więc możemy (i powinniśmy) interpretować ich zachowania przez analogię do ludzkich. W takich opowieściach człowiek nie jest samotnym wyjątkiem, lecz częścią większej całości – wspólnoty istot obdarzonych umysłem i emocjami. Zgodnie z tą narracją, badania dotyczące poznania zwierząt innych gatunków stanowią lustrzane odbicie, w którym odnajdujemy człowieka w innej formie (Safin, 2018). Pasuje to do mitycznej opowieści *o jedności człowieka z naturą* — narracji o wspólnym rodowodzie, empatii i pokrewieństwie. Co więcej, tak jak technopesymizm postrzega technologię jako zagrożenie, tak antroponegacjonizm postrzega przypisywanie zwierzętom ludzkich cech jako błąd poznawczy, który przyczynia się do łamania reguły parsymonii, jaką jest zasada Morgana. Narracja ta opiera się na przekonaniu o radykalnej *odrębności człowieka od innych gatunków zwierząt* – tylko ludzie posiadają moralność, język czy potrafią czytać umysły innych. W tym ujęciu zwierzęta stają się biologicznymi automatami, zdeterminowanymi przez instynkt i reakcje warunkowe. Jest on ściśle związany z mitem *o wyjątkowości człowieka*, który konstytuuje porządek społeczny i legitymizuje dominującą pozycję naszego gatunku.

W kontekście poznania zwierząt i postaw prezentowanych wobec nich, warto zauważyć, że mity i narracje mają istotny wpływ na sposób formułowania wniosków w badaniach naukowych. Uprzedzenia opierają się na zakorzenionych kulturowo strukturach narracyjnych, które funkcjonują jako poznawcze ramy interpretacyjne (Pullman, 2012). Mity te mogą funkcjonować zarówno na poziomie kulturowym, jak i indywidualnym i właśnie one organizują nasze rozumienie relacji międzygatunkowych. Dlatego w nauce również można odnaleźć mity, które zyskały status „prawdy”. Przykładem może być idea „walki o byt” wywodząca się z darwinizmu społecznego, spopularyzowana przez Herberta Spencera, która do dziś wpływa na postrzeganie relacji międzygatunkowych jako konfliktu i rywalizacji (Sikorska, Trojan, 2022). Ten sposób myślenia obecny jest w popularnych i akademickich interpretacjach zachowań zwierząt, mimo że alternatywne modele – np. oparte na współpracy czy empatii – również znajdują poparcie empiryczne. Jak zauważył Kanouse (1971), ludzie nie zawsze dążą do najbardziej trafnych i racjonalnych wyjaśnień; często zadowolają się interpretacjami wystarczającymi. Ma to związek z ograniczeniami poznawczymi i potrzebą uproszczeń – czyli właśnie z ramami narracyjnymi, które pozwalają organizować doświadczenie.

Dobrym przykładem takiej narracji jest twierdzenie, że *koty są naturalnymi samotnikami*, co wpływa na sposób interpretowania ich zdolności społecznych i poznawczych. Takie przekonanie podzielała także Simona Kossak (2019, str. 119-133), która w swoich gawędach *O zwierzętach i przyrodzie* pisze:

Trudniejsza sprawa jest ze zwierzętami żyjącymi samotnie. A jednak kot, to typowy samotnik, zawarł z człowiekiem pakt o nieagresji, który czasem my wypowiadamy (Kossak, 2019, str. 125).

Tego rodzaju potoczne przekonania mają swoje odpowiedniki także w dyskursie naukowym. W artykule *Zarys procesów kognitywnych kota domowego* Marcin Urbaniak (2024) przyjmuje m.in., że koty są gorzej uspołecznione niż psy oraz wykazują słabą zdolność rozpoznawania ludzkich intencji, a ich zachowania społeczne ograniczają się do „niższych” komponentów *mindreadingu*. Faktycznie nie ma dowodów na zdolności z zakresu czytania umysłów innych u kotów, jednak współczesne badania i obserwacje etologiczne podważają takie ujęcie. Po pierwsze, koty domowe nie są typowymi samotnikami w sensie biologicznym. Choć polują

samotnie, to posiadają złożoną mimikę⁵⁹ czy wokalizacje oraz reagują na gesty wskazywania. Potrafią rozpoznawać swoje imię, odpowiadać na emocje opiekunów (Saito, Shinozuka, 2013; Saito Takagi, Arahori, Fujita, 2019; Vitale Shreve, Udell, 2015). Szczególnie w środowisku miejskim i domowym, gdzie żyją w grupach, nawiązują relacje społeczne i angażują się w zachowania afiliacyjne zarówno z innymi kotami, jak i z ludźmi (Crowell-Davis, Curtis, Knowles, 2004; Bradshaw, 2016). Co więcej koty wykształciły elastyczność społeczną, co oznacza, że ich zachowania są znacznie bardziej złożone niż sugeruje stereotyp samotnika. Badania wskazują również, że koty potrafią rozpoznawać ludzkie słowa, kojarzyć je z obrazami, co świadczy o ich zaawansowanych zdolnościach poznawczych i społecznym rozumieniu otoczenia (Mäses, Wascher, 2022). Co więcej, reagują na niezgodności między słowami a obrazami, co sugeruje, że przetwarzają informacje na bardziej zaawansowanym poziomie niż dotąd sądzono (Takagi, Saito, Arahori, Chijiiwa, Koyasu i in. 2023). W tym kontekście zasadne staje się pytanie, czy negatywne wyniki testów *mindreadingu* u kotów wynikają raczej z niedostosowania paradygmatów badawczych do specyfiki *Umweltu* tych zwierząt — ich sposobu przetwarzania informacji i form komunikacji — niż z rzeczywistego braku zdolności we wskazanym zakresie. Koty, w przeciwieństwie do psów i naczelnych, reprezentują odmienny typ przystosowania ewolucyjnego – są gatunkiem mniej uległym, bardziej niezależnym i nastawionym na samotne polowanie. Ich zachowania społeczne i poznawcze rozwijały się w innym kontekście ekologicznym, co sprawia, że nie odpowiadają one bezpośrednio na typowe paradygmaty eksperymentalne. Może to prowadzić do błędnych wniosków na temat ich rzekomych deficytów poznawczych.

Ostatnia kwestia, którą pragnę poruszyć, dotyczy rozróżnienia terminologicznego, które warto doprecyzować na tym etapie pracy, aby uniknąć potencjalnych nieporozumień. Klasyczny i metodologiczny behawioryzm to tradycyjne nurty, których głównym celem było ugruntowanie psychologii jako nauki empirycznej, opartej na obserwowalnych zachowaniach oraz danych możliwych do obiektywnej weryfikacji. Tymczasem w części 3.2. *Dwie perspektywy interpretowania eksperymentów z udziałem zwierząt pozaludzkich* omawiam perspektywę behawioralną obecną we współczesnej psychologii poznawczej, porównawczej oraz etologii. Nie stanowi ona kontynuacji behawioryzmu metodologicznego, lecz funkcjonuje

⁵⁹ Zwierzęta samotnicze nie wykształciły mimiki twarzy, co dobrze widać na przykładzie niedźwiedzi.

równolegle jako odrębne podejście badawcze, skoncentrowane na analizie zachowania. Dlatego narzędzia wypracowane w obrębie behawioryzmu metodologicznego uznawane są przez badaczy przyjmujących perspektywę behawioralną, lecz również wśród przedstawicieli podejść mentalistycznych. Innymi słowy, zarówno badacze sceptyczni wobec przypisywania zwierzętom stanów mentalnych, jak i ci, którzy takie przypisywanie uznają za uprawnione, często posługują się zbliżonym zestawem technik badawczych, opartych na obserwacji i analizie zachowania, wypracowanych na gruncie behawioryzmu metodologicznego. Zatem nurt ten stanowi wspólną podstawę wielu współczesnych podejść badawczych w stosowanych w różnych dziedzinach psychologii — niezależnie od przyjmowanych założeń ontologicznych dotyczących istnienia stanów mentalnych. Taka klarowność terminologiczna sprzyja bardziej obiektywnej i rzetelnej interpretacji wyników badań z udziałem zwierząt innych niż ludzie.

Podsumowując tę część rozważań, pragnę zauważyć, że w interesującym mnie obszarze, funkcjonują przynajmniej dwa mity: *o wyjątkowości człowieka* oraz *o jedności człowieka z naturą*. Przekonania zgodne z pierwszym z wymienionych służą legitymizacji uprzedmiotawiania ZINL i ich instrumentalnego traktowania. Widoczne są w nurcie behawioralnym, opisanym w części 3.2. *Dwie perspektywy interpretowania eksperymentów z udziałem zwierząt pozaludzkich*. Jak wykażę w rozdziale V: *Tendencyjność w badaniu teorii umysłu*, autorzy przyjmujący postawę antroponegacjonistyczną — mimo deklaratywnej obiektywności — niekiedy powielają antropocentryczne założenia. Z kolei narracje antropomorficzne, które przypisują zwierzętom cechy ludzkie, wynikają z mitu *jedności człowieka z naturą* i bywają obecne w perspektywie mentalistycznej oraz np. w nurtach proekologicznych czy weganizmie etycznym.

3.3.2. Stereotypy

Walter Lippmann (1922) jako pierwszy zastosował termin „stereotyp” w kontekście społecznym, opisując tendencję do postrzegania osób lub przedmiotów jako podobnych, to znaczy przypisywania im tych samych cech (atrybutów), (Nelson, 2003, s. 25). Choć pierwotnie pojęcie to odnosiło się do techniki kopiowania matryc w druku wypukłym, to z czasem zostało zaadaptowane przez nauki społeczne jako narzędzie analizy procesów poznawczych i społecznych. Ujęcie Lippmanna było przełomowe, ponieważ określiło funkcję stereotypów jako mechanizmów poznawczych — stanowią one swojego rodzaju filtr, który ukierunkowuje uwagę, selekcionującą

informacje z otoczenia, decydując, co zostanie uznane za istotne, a co pominięte. Omawiane struktury można więc rozumieć jako kryteria przyporządkowywania obiektów do kategorii poznawczych – określają sposób ich postrzegania poprzez przypisywanie im cech zgodnych z założeniem, że członków danej grupy (zwłaszcza obcej) łączą określone właściwości (Nelson, 2003, s. 22). Proces ten określa się mianem stereotypizacji, a jego podstawowym mechanizmem jest generalizacja⁶⁰.

Podkreśla się, że podstawową funkcją adaptacyjną omawianych kryteriów jest ograniczenie zużycia zasobów poznawczych (Macrae, Milne, Bodenhausen, 1994). Przeprowadzono serię eksperymentów, których wyniki wspierają stanowisko głoszące, że stereotypy upraszczają przetwarzanie informacji w złożonym środowisku społecznym, zatem stanowią narzędzie do oszczędzania zasobów poznawczych. Autorzy *Stereotypes as energy-saving devices: A peek inside the cognitive toolbox* (1994), przeprowadzili serię trzech eksperymentów z wykorzystaniem paradygmatu podwójnego zadania⁶¹, aby sprawdzić, czy stosowanie stereotypów rzeczywiście prowadzi do oszczędności zasobów umysłowych. W pierwszym badaniu uczestnicy mieli jednocześnie formułować wrażenia o osobach na podstawie cech oraz monitorować tekst prozatorski. Wyniki wykazały, że gdy w zadaniu pojawiały się etykiety stereotypowe, uczestnicy osiągalni lepsze wyniki w monitorowaniu tekstu, co sugeruje, że stereotypy upraszczają i przyspieszają proces przetwarzania informacji o innych, pozostawiając więcej zasobów na zadania równoczesne (Macrae, Milne, Bodenhausen, 1994). W drugim eksperymencie autorzy zastosowali subliminalną

⁶⁰ Więcej o stereotypizacji i generalizacji znaleźć można w częściach 2.1. *Mechanizm kategoryzacji społecznej* 2.3 *Myślenie analogiczne*, 2.6. *Słodka reakcja* czy 3.4.2. *Heurystyki i wynikające z nich tendencje w zachowaniu i podejmowaniu decyzji istotne w kontekście uprzedzeń międzygatunkowych*.

⁶¹ Paradygmat podwójnego zadania (*dual-task paradigm*) to metoda badawcza stosowana w psychologii poznawczej, która polega na jednoczesnym wykonywaniu przez uczestnika dwóch zadań. Celem tego podejścia jest zbadanie, jak obciążenie poznawcze wpływa na wydajność w każdym z tych zadań oraz jak zasoby uwagi i przetwarzania informacji są dzielone między różne procesy umysłowe. Uczestnicy wykonywali jednocześnie zadanie polegające na formułowaniu wrażeń o ludziach oraz monitorowaniu tekstu. Lepsza wydajność w monitorowaniu tekstu w sytuacji aktywacji stereotypów sugerowała, że stereotypy działają jak „narzędzia” upraszczające przetwarzanie informacji i zmniejszające obciążenie poznawcze. Zatem paradygmat podwójnego zadania to skuteczna metoda pozwalająca badać, jak różne procesy poznawcze współdzielą ograniczone zasoby umysłowe oraz czy pewne mechanizmy (np. stereotypy) działają automatycznie i oszczędzają energię poznawczą.

aktywację stereotypów ⁶² , uzyskując podobny efekt oszczędności zasobów poznawczych, co wskazuje, że korzystanie ze stereotypów może zachodzić automatycznie, bez świadomej intencji. Trzecie badanie, wykonano w paradygmacie reakcji na bodziec ⁶³, co potwierdziło wcześniejsze doniesienia w innym kontekście pomiarowym. Autorzy podkreślają, że stereotypy pełnią ważną funkcję adaptacyjną, pozwalając radzić sobie z nadmiarem informacji społecznych poprzez uproszczenie i automatyzację procesów poznawczych. Jednak, jak zauważają, choć stereotypy są efektywnym narzędziem oszczędzania energii poznawczej, ich automatyczne stosowanie może prowadzić do uproszczeń, uprzedzeń i błędów w ocenie innych.

Jednym z kluczowych badaczy zajmujących się omawianym zagadnieniem jest Richard Ashmore, który w swoich pracach z lat 70. XX wieku podkreślał, że stereotypy należy rozumieć jako struktury poznawcze, które porządkują wiedzę o członkach określonych grup społecznych i umożliwiają szybkie, choć uproszczone, wnioskowanie na temat ich cech oraz zachowań. W artykule z 1979 roku napisanym wspólnie z Frances'em Del Bocą, Ashmore zwracał uwagę, że postęp w badaniach nad stereotypami płci był utrudniony przez brak jasnych definicji tego, czym stereotyp jest oraz jaką pełni rolę w procesach poznawczych i społecznych. Autorzy postulowali, aby stereotypy traktować jako zbiory powiązanych ze sobą cech, które są przypisywane określonym kategoriom społecznym. Oznacza to, że ludzie mają tendencję do automatycznego łączenia pewnych cech i atrybutów z przynależnością do danej grupy, co wpływa na sposób postrzegania i oceniania jej członków. Ashmore i Del Boca (1979) podkreślali, że stereotypy nie są jedynie prostymi etykietami, lecz stanowią złożone schematy poznawcze, które mają wpływ na procesy percepcji, interpretacji i zapamiętywania informacji o innych ludziach. Wskazywali również, że stereotypy

⁶² Subliminalna aktywacja stereotypów odnosi się do procesu, w którym stereotypowe treści są wywoływane automatycznie i nieświadomie, zwykle poprzez prezentację bodźców podprogowych. W praktyce oznacza to, że bodźce, takie jak słowa, obrazy czy inne przekazy, są prezentowane w taki sposób, iż osoba nie zdaje sobie z nich sprawy, ale mimo to mogą one wpływać na jej myśli, emocje i zachowania. Na przykład, bodźce podprogowe mogą aktywować stereotypy płciowe lub rasowe, powodując automatyczne skojarzenia i uprzedzenia, które mogą się ujawniać w późniejszych reakcjach czy ocenach, mimo braku świadomości ich ekspozycji.

⁶³ Zadanie reakcji na bodziec to metoda badawcza, w której uczestnik musi jak najszybciej zareagować na pojawiający się na ekranie lub w otoczeniu bodziec, np. kliknąć przycisk, nacisnąć klawisz, podnosząc rękę lub wykonując inną reakcję motoryczną. Celem takich badań jest zmierzenie czasu reakcji – czyli jak szybko uczestnik zareaguje na dany bodziec.

mogą być zarówno pozytywne, jak i negatywne, a ich treść i funkcje są silnie związane z organizacją grup społecznych, a konkretnie z relacjami władzy oraz strukturą społeczną. W odniesieniu do stereotypów płciowych autorzy proponowali analizowanie ich w kategoriach szerszych teorii dotyczących relacji międzygrupowych i procesów poznawczych, co pozwala lepiej zrozumieć mechanizmy ich powstawania i utrwalania. Wcześniejsze prace Ashmore'a (1970) dotyczące uprzedzeń międzygrupowych również podkreślały rolę stereotypów jako kluczowych narzędzi organizujących wiedzę o innych oraz wpływających na postawy i zachowania wobec członków grup obcych. Stereotypy, jako poznawcze uproszczenia, mogą prowadzić do dyskryminacji, zwłaszcza w sytuacjach rywalizacji lub konfliktu interesów między grupami.

Zatem ich podstawową funkcją jest gromadzenie i przekazywanie zbiorowej wiedzy społeczności na temat obcych grup, co pozwala na uproszczenie przetwarzania informacji o nowo napotkanych jednostkach (Mirośławska, Kofta, 2007). Ich zaletą jest oszczędność energii oraz umożliwienie uniknięcia ryzyka związanego z potencjalnymi zagrożeniami ze strony grupy obcej. Stereotypizacja ma jednak istotne konsekwencje negatywne: prowadzi do zniekształcania reprezentacji poszczególnych grup, m.in. poprzez homogenizację grupy obcej (redukcję różnic wewnątrz kategorii) oraz akcentuację różnic międzygrupowych. Omawiane kryteria pełnią także funkcje afektywne i motywacyjne – służą obronie samooceny i samousprawiedliwieniu. Umożliwiają podtrzymywanie pozytywnego obrazu grupy własnej, a także, redukują dysonans poznawczy związany z nieetycznym postępowaniem wobec członków innych grup. Okazuje się, że ludzie wkładają wiele wysiłku w zachowanie spójności między swoimi przekonaniami a informacjami z nimi sprzecznymi (Sherman i współpracownicy, 2005). Poświęcają w tym celu znacznie więcej uwagi informacjom sprzecznym z przekonaniami składającymi się na istniejące struktury poznawcze niż informacją z nimi zgodnym. W ramach eksperymentu poproszono heteroseksualnych studentów o ocenę homoseksualnego mężczyzny Roberta, któremu przypisano 8 aktywności zgodnych ze stereotypem typowego geja (np. uczy się tańca interpretacyjnego) oraz 8 aktywności niezgodnych z tym stereotypem (np. w niedzielę obejrzał mecz footballu). Studenci uprzedzeni do homoseksualistów zniekształcali informacje niezgodne ze stereotypem, dla przykładu twierdzili, że Robert obejrzał mecz tylko dlatego, że przyszedł do niego kuzyn Frank. Tavis i Aronson (2007) podkreślają, że stereotypy pozwalają utrzymać kluczowe elementy tożsamości społecznej, a zarazem usprawiedliwiają działania, które mogłyby być uznane za moralnie problematyczne. Co

ciekawe, omawiane kryteria same nierzadko wymagają usprawiedliwienia – co przez lata sprawiało, że były one oceniane jednoznacznie negatywnie (Duckitt, 1992, s. 19). Jak pokazują badania, proces stereotypizacji pojawia się już we wczesnym dzieciństwie i nie ogranicza się wyłącznie do ludzi – występuje również u zwierząt społecznych, u których kryteria kategoryzacji stanowią podstawę funkcjonowania poznania społecznego (Duckitt, 1992).

W kontekście tematu niniejszej dysertacji nie może zabraknąć odniesienia do stereotypów odnoszących się do procesów kategoryzowania zwierząt innych niż ludzie oraz ich potencjalnego wpływu na tworzenie i wspieranie interpretacji eksperymentów dotyczących teorii umysłu u ZINL. Opisana wyżej teoria stereotypów Ashmore’a i Del Boca, która traktuje stereotypy jako złożone schematy poznawcze organizujące wiedzę o grupach społecznych, może być bezpośrednio odniesiona do uprzedzeń pojawiających się w badaniach z zakresu *mindreadingu*. Badacze interpretujący zachowania zwierząt często nieświadomie posługują się stereotypami – przypisując określonym gatunkom zbiór cech (np. inteligentny, agresywny, przyjazny), które wpływają na sposób postrzegania i oceniania tych zwierząt oraz na decyzje badawcze i interpretacje wyników eksperymentów (Sevillano, Fiske, 2019). Zgodnie z ujęciem Ashmore’a i Del Boca, stereotypy są strukturami poznawczymi, które porządkują i upraszczają wiedzę o innych, prowadząc do automatycznego przypisywania cech na podstawie przynależności do danej kategorii. W kontekście badań nad zwierzętami takie schematy mogą prowadzić do uprzedzeń – na przykład do przeceniania zdolności kognitywnych niektórych gatunków lub pomijania złożoności zachowań innych, co z kolei może wpływać na projektowanie eksperymentów, interpretację wyników czy nawet decyzje o publikacji badań. Przykładem konsekwencji opisywanej stereotypizacji może być zawyżenie ocen zdolności kognitywnych gatunków uznawanych kulturowo za „wyjątkowe” (np. małe człokształtne, delfinów czy psów) lub pomijaniem złożoności zachowań zwierząt postrzeganych jako „prymitywne” czy „dzikie”. Co istotne, to, z jakich stereotypów jednostka (w tym badacz) korzysta w danym momencie, zależy od szerszych ram interpretacyjnych, jakimi się posługuje – a więc od mitu, który w danym kontekście organizuje poznanie. Mity – rozumiane tu jako kulturowe opowieści zawierające podstawowe znaczenia i wartości społeczeństwa – pełnią funkcję nadrzędnych struktur poznawczych, wpływających na to, jakie cechy i kategorie zostaną uznane za relewantne, a jakie pominięte. Przykładowo, mit o *człowieku jako władcy natury*, wpisujący się w opisaną wyżej narrację o *wyjątkowości*

człowieka może prowadzić do stereotypizacji zwierząt jako istot niższych, przewidywalnych i podporządkowanych człowiekowi, podczas gdy mit o *jedności człowieka z naturą* może sprzyjać nadawaniu niektórym gatunkom, cech romantyzujących je, co prowadzi — jak pisałam w części 2.7. *Efekt halo* — do infantylnizacji natury. W ten sposób stereotypy, choć zakotwiczone w strukturach poznawczych, są jednocześnie podatne na wpływ treści kulturowych, które determinują ich aktywizację i funkcje w danym kontekście poznawczym.

3.4. Behawioralny komponent uprzedzeń

Komponent behawioralny postawy jest rozumiany jako gotowość do określonego działania wobec obiektu oraz jego powiązań ze stereotypami i mitami wpływającymi na relacje człowieka z ZINL. Szczególne znaczenie mają tu praktyki badawcze — w tym projektowanie eksperymentów, wybór gatunków do badań, interpretacja wyników, a nawet decyzje dotyczące publikacji — które, choć uznawane za neutralne, w rzeczywistości są osadzone w określonym porządku społecznym, normach kulturowych i indywidualnych postawach. Te zaś, jak pokazałam we wcześniejszych częściach rozprawy, uwarunkowane są działaniem mechanizmów poznawczych, takich jak stereotypizacja, torowanie czy dysonans poznawczy. Postawy wobec ZINL, choć często nieuświadomiane, przekładają się na konkretne działania: badacze nie tylko przypisują określone cechy przedstawicielom różnych gatunków (komponent poznawczy), ale też w sposób mniej lub bardziej konsekwentny podejmują wobec nich działania — takie jak wykluczanie z badań, czy uprzednie zakładanie ich „niższości poznawczej” lub odwrotnie: przecenianie ich zdolności na podstawie antropomorfizujących narracji. W tym sensie komponent behawioralny postawy przejawia się w czynnościach, które kształtują całościowy obraz zwierząt w nauce i kulturze.

Zachowania determinowane są przez stereotypy oraz głębiej zakorzenione mity kulturowe, które dostarczają ram interpretacyjnych dla relacji człowiek — ZINL. Mity te mogą funkcjonować równolegle, często nie powodując sprzeczności poznawczej, ponieważ są aktywowane sytuacyjnie. Badacz może zatem w jednej publikacji opowiadać się za mentalistycznym ujęciem głoszącym, że zwierzęta inne niż ludzie zdolne są do teorii umysłu, w jakiejś formie (oczywiście mniej złożonej, niż występujące u człowieka), a w innej wyraźnie dystansować się od przypisywania zwierzętom

zdolności do atrybucji innym intencji czy stanów percepcyjnych – nie dostrzegając w tym braku konsekwencji. Komponent behawioralny jest więc kluczowy, ponieważ ujawnia, jak przekonania, determinują działania podejmowane wobec ZINL – zarówno w przestrzeni badawczej, jak i w szerszym kontekście społecznym. Działania te konstytuują określony porządek – zarówno epistemologiczny, jak i normatywny – w którym określone gatunki uznaje się za godne poznania, ochrony lub współodczuwania, a inne za nieistotne, dzikie lub nieracjonalne. W kolejnych podrozdziałach przeanalizuję, jak mityczne narracje wpływają na konkretne zachowania wobec zwierząt, a także jak komponent behawioralny ujawnia się w relacjach między badaczem a obiektem jego badań. Takie ujęcie stanowi istotny krok w kierunku zrozumienia, w jaki sposób ukryte uprzedzenia międzygatunkowe przekładają się na wybory praktyczne i epistemiczne w badaniach nad teorią umysłu.

3.4.1. Tendencyjność

Przyjmując perspektywę wywodzącą się z biologii ewolucyjnej, zakładam, że każda komórka i narząd tworzące organizm dostosowane są do własnego lokalnego środowiska. Ludzie, jak i inne zwierzęta wyewoluowali tak, aby funkcjonować w świecie najskuteczniej, jak to tylko możliwe, a celem nadrzędnym całego organizmu–systemu jest zmaksymalizowanie szans na własne przetrwanie. W literaturze znaleźć możemy metaforę głoszącą, że lata ewolucji wykształciły maszyny do wnioskowania, którymi to maszynami są umysły i mózgi (OpenStax, 2022). Często występujące w niniejszej pracy pojęcie przetwarzania informacji⁶⁴ rozumiane jest jako wszystkie operacje, dzięki którym odbierane przez jednostkę bodźce (zewnętrzne i wewnętrzne) są: transportowane, modyfikowane, redukowane, wzmacniane, zachowywane, lub przywoływane. Procesy te przebiegają zgodnie z wielokrotnie już omawianym w niniejszej dysertacji paradygmatem skąpca poznawczego. Konsekwencją założenia istnienia tego mechanizmu jest twierdzenie, że ludzkie myślenie, tu rozumowanie, opiera się na heurystykach, schematach, stereotypach i mitach, a nie regułach logiki formalnej (Nęcka, 2013, str. 25-29). W literaturze psychologicznej i filozoficznej pojęcia rozumowania (*reasoning*) i wnioskowania (*inference*) często używane są

⁶⁴ Termin ten wywodzi się z cybernetyki, czyli nauki o sterowaniu z wykorzystaniem sprzężenia zwrotnego oraz matematycznej teorii informacji. W takiej terminologii informacja to redukcja informacji – gdy rzucamy monetą czujemy niepewność, natomiast wynik rzutu zredukuje tę niepewność o połowę (Nęcka, 2013, s. 23).

zamiennie, choć ich znaczenia nie zawsze są tożsame. W zależności od przyjętej perspektywy — filozoficznej, psychologicznej czy kognitywnej — akcenty mogą być rozłożone inaczej. W psychologii poznawczej rozumowanie traktowane jest jako ogólny proces umysłowy, w którym jednostka operuje na informacjach — przesłankach i przekonaniach, aby dojść do wniosków, decyzji lub rozwiązania problemu. W tym ujęciu wnioskowanie stanowi jeden z elementów rozumowania, tj. etap, w którym na podstawie przesłanek dochodzi się do określonej konkluzji, co czyni je bardziej precyzyjnym, operacyjnym aspektem szerszego procesu. Z filozoficznego punktu widzenia różnica między rozumowaniem a wnioskowaniem bywa ujmowana jeszcze wyraźniej. Wnioskowanie jest wówczas traktowane jako szczególny typ rozumowania — ścisły, formalny, podlegający regułom logiki, często utożsamiany z dedukcją lub indukcją w sensie ścisłym. Rozumowanie natomiast to szerszy, nie zawsze formalny proces intelektualny, który może zawierać zarówno poprawne, jak i błędne operacje poznawcze. W tym sensie rozumowanie może przebiegać intuicyjnie (np. wiedza przez wgląd), podczas gdy wnioskowanie wymaga większej dyscypliny logicznej. Niektórzy autorzy (np. w analizach kognitywnych) utożsamiają te pojęcia lub rozpatrują je na kontinuum, przy czym wskazują na możliwość ujęcia rozumowania/wnioskowania w sensie szerokim (jako każde przejście od przesłanek do wniosków), węższym (jako wnioskowanie dedukcyjne i indukcyjne), oraz najwęższym (jako czysto formalną dedukcję), (Floridi, 2011).

Umysł wyciąga wnioski dotyczące świata na podstawie spostrzeżeń, emocji, a także istniejących już w systemie struktur. W kontekście niniejszego opracowania interesujące jest, jak system–organizm przetwarza dane w obrębie komponentu poznawczego, czyli w sferze przekonań i mitów, a także to jak na ich podstawie generuje zachowanie (komponent behawioralny), w postaci faworyzowania określonych narracji naukowych w kontekście badań z zakresu *mindreadingu*. Jak wskazałam w części 1.2.3. *Klasyfikacja uprzedzeń*, omawiane postawy mogą przyjmować formę jawną albo ukrytą, natomiast tendencyjność rozumiana jest jako preferencja systemu poznawczego wynikająca z użycia określonej heurystyki, czyli uproszczonej reguły wnioskowania. Uprzedzenia jawne, wiążą się z otwartym wyrażaniem negatywnych opinii wobec członków określonych grup, często z towarzyszącym poczuciem zagrożenia (Zakrzewska, 2015, s. 1). Postawy ukryte manifestują się w sposób bardziej złożony – poprzez zaprzeczanie przeżywanym emocjom i silne przywiązanie do norm oraz wartości grupy własnej, a także podkreślanie różnic międzygrupowych

(Zakrzewska, 2015, s. 79). Uprzedzenia ukryte charakteryzują się zaprzeczaniem przeżywanym emocjom, a ich konsekwencją jest silna tendencja do obrony tradycyjnych wartości grupy własnej (realnej lub wyobrażonej) oraz podkreślaniem różnic międzygrupowych (Zakrzewska, 2015, str. 79).

W odniesieniu do badań dotyczących zdolności poznawczych ZINL, w szczególności z zakresu *mindreadingu*, uprzedzenia (zarówno jawne, jak i ukryte) mogą wpływać na sposób interpretowania zachowania podczas eksperymentu. Komponent poznawczy – rozumiany jako zestaw przekonań o cechach i możliwościach różnych gatunków – może determinować interpretację danych w postaci zachowań zwierząt podczas eksperymentu. Na przykład, jeśli badacz nieświadomie zakłada, że określony gatunek nie posiada zdolności do przypisywania stanów percepcyjnych innym, może interpretować zachowania tego zwierzęcia wyłącznie w kategoriach behawioralnych, pomijając możliwość alternatywnego, bardziej złożonego wyjaśnienia. Tym samym, poznawcze uprzedzenie skutkuje konkretnym działaniem – wyborem jednej interpretacji spośród możliwych – a zatem bezpośrednio przekłada się na komponent behawioralny. Zjawisko to może być szczególnie silne w przypadku uprzedzeń ukrytych, które pozostają poza bezpośrednią kontrolą jednostki, a ich działanie objawia się w preferencyjnym stosowaniu określonych typów wyjaśnień czy projektowaniu eksperymentów w sposób potwierdzający wcześniejsze założenia. W konsekwencji, interpretacje wyników badań nad teorią umysłu u ZINL mogą być nie tylko odbiciem danych empirycznych, ale również ekspresją postaw poznawczo-afektywnych, które — choć nie zawsze rozpoznawane — kształtują granice poznania naukowego.

3.4.1.1. Tendencyjność komponentu poznawczego

Za przetwarzanie w obrębie komponentu poznawczego, odpowiada złożony zbiór procesów obejmujących dostęp do istniejących już struktur, generowanie i testowanie hipotez, interpretację otrzymanych dowodów oraz ich weryfikację. Wszystkie te elementy mają charakter silnie zintegrowanego systemu, w którym przekonania początkowe (np. ogólne informacje o tym, jak działa świat) wpływają na interpretację bieżących danych (Klayman, 1995). Natomiast testowanie możliwych hipotez dostarcza informacji o efektywności generowania alternatywnych wyjaśnień, a także wpływa na procesy ich wyszukiwania. Rozważania z zakresu rozwoju czynności poznawczych są w dużej mierze domeną psychologii rozwojowej, w ramach której

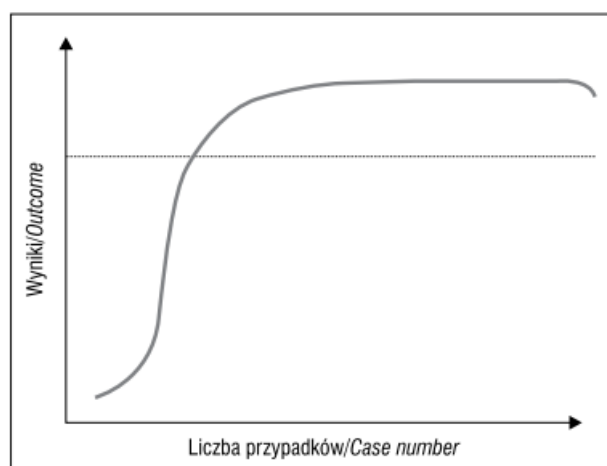
uznaje się, że pewne czynności ujawniają się w zachowaniu, a z czasem (wraz z doświadczeniem) ulegają mechanizmowi interioryzacji (inaczej nazywanemu internalizacją), którego funkcją jest *uwewnętrznienie* czynności poznawczych (Nęcka 2013, str. 21-29). W ten sposób nowo nabywane wzorce zachowań zostają utrwalone w postaci np. ram. Mechanizm interioryzacji pozwala na automatyzację czynności poznawczych lub motorycznych, co wiąże się z uwolnieniem ich spod kontroli poznawczej⁶⁵, dzięki temu wykonywane są szybko, bez wysiłku, jednak schematycznie (Nęcka, 2013, str. 229-277). Podstawową wartością adaptacyjną wymienionych mechanizmów, jest zmniejszenie kosztów związanych z kontrolą czynności jednoczesnych (podzielność)⁶⁶ i przerzutności uwagi⁶⁷. Wykres przyrostu kompetencji (krzywa uczenie się) przedstawiony został na zdjęciu 4. Interesującym zjawiskiem z tego zakresu jest *flow*, czyli stan możliwy do osiągnięcia tylko dla kogoś, kto posiadał bardzo wysoki poziom biegłości w wykonywaniu zadania (Nakamura, Csikszentmihalyi, 2009). Charakteryzuje się on intensywną koncentracją i poczuciem kontroli czy sprawczości osobistej. Zatem towarzyszące mu odczucia określane są jako pozytywne czy przyjemne. Zjawisko to może zostać zilustrowane przez krzywą uczenia się, opisującą jak nowicjusz doświadcza przyrostu kompetencji wraz z doświadczeniem, stając się ekspertem.

⁶⁵ Kontrola poznawcza to zdolność systemu poznawczego do nadzorowania i regulowania własnych procesów a także do planowanego sterowania ich przebiegiem (Nęcka, 2013, str. 230).

⁶⁶ Podzielność uwagi to zdolność pozwalająca na równoczesne wykonywanie więcej niż jednego procesu. Jeżeli te są proste lub zautomatyzowane, jednoczesne wykonywanie różnych czynności poznawczych, nie niesie negatywnych konsekwencji dla żadnego z nich (np. nie spowalnia przetwarzania).

⁶⁷ Przerzutność uwagi to zdolność uwagi do przełączania się między zadaniami, wymaga zahamowania jednego procesu poznawczego i uruchomienia alternatywnego (Nęcka, 2013, str. 184).

Zdjęcie 4. Krzywa przyrostu kompetencji



[Źródło (Kurek, Król, 2016, str. 16-23)]

Rozważania dotyczące roli reguł logiki formalnej i heurystyk w procesach przetwarzania informacji mają długą tradycję, w ramach której filozofowie starają się odpowiedzieć na pytanie, *czy człowiek jest istotą racjonalną?* Teoria racjonalności ekologicznej, głoszonej przez Leda Cosmidesa i Johna Tooby'ego (2013), zakłada, że umysł ludzki nie jest w pełni dostosowany do myślenia abstrakcyjnego, zamiast tego, jest bardziej praktyczny (Nęcka, 2013, str. 432). Uproszczenia, a nawet błędy mogą mieć charakter racjonalny, chociaż niezgodny z zasadami logiki formalnej (Urbański, 2009). Przez lata badania z obszaru zagadnień dotyczących natury myślenia i rozumowania skupione były na formalnych właściwościach myślenia, mierzonych przez obserwację błędów w rozumowaniu uczestników eksperymentów rozwiązujących sylogizmy (Nęcka, 2013, str. 422). Abstrakcyjna treść zadania takiego rodzaju pozwala na wgląd w procesy oderwane od doświadczenia jednostki. Co ciekawe, treść zadania ma istotne znaczenie w omawianych procesach: może je facylitować lub utrudniać, co widać najlepiej w paradygmacie zadań izomorficznych⁶⁸ (Carrahera, Schliemanna i in. 2006). W badaniu przeprowadzonym w 1985 roku, przez Davida Williama Carrahera, Analucia Schliemanna i współpracowników (2006), podzielono uczestników na trzy grupy: pierwsza rozwiązywała równania matematyczne, druga zadania z treścią, natomiast trzecia wykonywała transakcję kupna i sprzedaży. Wyniki pokazały, że najniższą rozwiązywalność mają zadania, które otrzymała gr. 1 – wykonywane były poprawnie tylko w 37%, przy czym, transakcje kupna i sprzedaży (gr. 3) wykonywane

⁶⁸ Problemy izomorficzne mają tę samą strukturę, ale inną treść (Nęcka, 2013, str. 423).

były prawidłowo przez aż 96%. Innym ciekawym przykładem wspierającym teorię racjonalności ekologicznej jest porównanie wyników eksperymentu stosującego dwa warianty (zadania izomorficzne) testu selekcji Wasona (Griggs, Cox, 1982). Pierwszy stanowił klasyczną wersję, w której, uczestnikom prezentowano 4 karty: „A”, „G”, „6”, „9” oraz proszono ich o odwrócenie 2 kart tak, aby przetestować hipotezę, głoszącą, że gdy po jednej stronie karty znajduje się samogłoska, to po drugiej nadrukowana jest liczba parzysta. W znacznej mierze respondenci poprawnie odwracali kartę „A”, jednak mylili się przy drugim ruchu, wskazując kartę „6”. Wariant drugi zmieniony jest pod względem treści implikacji — uczestnikom prezentowano następujące zadanie (Griggs, Cox, 1982):

Wyobraź sobie, że jesteś policjantem na służbie. Twoja praca polega na kontroli respektowania obowiązujących norm. Karty leżące przed Tobą zawierają informację o czterech osobach siedzących przy stole. Na jednej stronie kart podany jest wiek osoby, a na drugiej stronie — to co dana osoba pije. Obowiązująca reguła jest następująca: jeżeli osoba pije piwo, to musi mieć ukończone 18 lat. Wskaż kartę albo karty, które koniecznie musisz odwrócić, aby ustalić, czy reguła został naruszona (Nęcka, 2013, str. 459).

Na kartach znajdowały się kolejno napisy o treści: „pije piwo”, „pije colę”, „22”, „17”. W tym wariantcie aż 73% osób odpowiedziało poprawnie (wskazując pierwszą i ostatnią kartę), to znacznie więcej, niż w klasycznej wersji, gdzie kartę z nr „9” wskazywało tylko 16% respondentów. Innymi słowy, uczestnicy radzą sobie lepiej z logicznym rozumowaniem, gdy ma ono dla nich praktyczny wymiar. Steven Pinker (2002) wyjaśnia takie wyniki, przyjmując stanowisko, że początkowo mózg i umysł wyewoluowały na afrykańskiej sawannie do rozwiązywania problemów charakterystycznych dla społeczności zbieracko-łowieckich, jednak z czasem na skutek egzaptacji ⁶⁹ przystosowały się do myślenia abstrakcyjnego. Za słuszością przytoczonego twierdzenia przemawiają badania prowadzone na społecznościach żyjących w trybie pierwotnym. Michael Cole i John Gay, prowadzili badania z udziałem przedstawicieli plemienia Kpelle z Liberii (Cole, Gay, i.in., 1971). W ramach zadania uczestnikom zadawano pytania, na które mogli odpowiedzieć na podstawie uprzednio

⁶⁹ Egzaptacja to proces dostosowania struktur mózgowych do nowych wymagań (Nęcka, 2013, str.422).

podanych informacji. Innymi słowy, problem przedstawiany był w formie prostych sylogizmów, takich jak:

Flumo i Yakapalo zawsze piją razem sok trzciny.
Flumo pije sok trzciny

Pytanie: czy Yakapalo pije sok trzciny?

(Nęcka, 2013, str. 422)

Odpowiedzi uczestników były zaskakujące, gdyż te miały charakter literalny np. jeden z przedstawicieli Kpelle odparł, że gdy ostatnio widział jak Flumo pił sok trzciny, to Yakapalo nie było w wiosce, inny zaś oznajmił, że Yakapalo nie ma aktualnie obok oraz zapytał dlaczego interlokutor sam do niego nie podejdzie i go nie zapyta. Jak twierdzi Neisser (1976), odpowiedzi te nie świadczą o niskim poziomie inteligencji uczestników badania, a raczej pokazują wpływ warunków środowiskowych na rozwiązywanie problemów. To jakie zachowania są cenione i wzmacniane (np. społecznie) jest ściśle związane z potrzebami i celami kultury, w której żyje jednostka. Dlatego zaskakujące odpowiedzi przedstawicieli Kpelle, są w rzeczywistości bardzo inteligentne, gdyż wyzwania życia codziennego plemienia, z którego pochodzą, nie wymagają rozwinięcia wysokich umiejętności myślenia abstrakcyjnego.

Uważam jednak, że wniosek wyciągnięty z przytoczonego badania jest zbyt stanowczy, aby uzasadnić swoje stanowisko odwołam się do krótkiej dygresji dotyczącej wpływu kultury na kształtowanie procesów poznawczych, w tym tych związanych z myśleniem i wnioskowaniem. Wyniki uzyskane w przytoczonym badaniu mogą być po prostu rezultatem błędu w jego konstrukcji. Eksperymentatorzy wykorzystali bowiem w sylogizmach imiona przedstawicieli plemienia, co mogło sugerować uczestnikom, że pytanie dotyczy rzeczywistych Flumo i Yakapalo. Przedstawiciele plemienia Kpelle nie posiadają w swoich strukturach poznawczych pojęcia eksperymentu naukowego, dlatego mogli nie rozumieć pytania. Każda osoba, żyjąca w cywilizacji technologicznej, wie, że treść sylogizmu nie dotyczy rzeczywistych ludzi żyjących w świecie, a raczej Flumo' i Yakapalo', czyli fikcyjnych ludzi żyjących w wyobrażonym plemienu Kpelle'. Jedną z zasad przedstawionego w sylogizmie świata jest to, że Flumo' i Yakapalo' zawsze piją razem sok trzciny.

Taki problem z interpretacją zadania przez jego uczestników tłumaczyłby literalne odpowiedzi przedstawicieli plemienia Kpelle. W celu jego redukcji należałoby poprzedzić sylogizmy treścią, w której uczestnik proszony jest o wyobrażenie sobie, że istnieje plemię X⁷⁰, w którym żyją ludzie o imionach⁷¹ Z i Y. Taka konstrukcja eksperymentu rodzi natomiast inne znaczące problemy. Przyjmując perspektywę racjonalności ekologicznej, należy uznać założenie o różnicach w dostosowaniu systemów poznawczych kultur żyjących w trybie pierwotnym oraz tym technologicznym, dlatego prośba o *wyobrażenie sobie czegoś*, może stanowić dużą przeszkodę. Należy podejrzewać, że zadanie to wymagałoby szczegółowego wyjaśnienia koncepcji *wyobrażania sobie*, gdzie pierwszą trudnością byłoby pokonanie bariery językowej i przetłumaczenie słowa „wyobrażenie” na podobne określenie w języku Kpelle. Podejrzenia te oparte są na badaniach antropologicznych Daniela Everetta, który był misjonarzem pełniącym swoje obowiązki w plemieniu Pirahã zamieszkującym tereny Amazonii. Jego celem było wykonywanie misji chrystianizacyjnej, a w swojej książce *Nie śpij, tu są węże! Życie i język w amazońskiej dżungli* przytacza liczne rozmowy pokazujące, że faktycznie przedstawiciele ludów pierwotnych mają problemy ze zrozumieniem pewnych abstrakcyjnych koncepcji. Ze względu na swoją misję skupia się przede wszystkim na koncepcji Boga. Wspomina, że gdy próbował ją wyjaśnić, często spotykał się z pytaniem, *czy ten Bóg jest teraz z „nami”*⁷²? W odpowiedzi misjonarz próbował tłumaczyć, że Bóg *tak jakby jest teraz z nami*, bo jest on wszędzie, jednak nie można go zobaczyć, ani poznać za pomocą zmysłów. Jednak tłumaczenia te przerywane były przez inne pytania, dotyczące tego, *czy Bóg jest członkiem rodziny Everetta, może jego bratem, albo ojcem?* Odpowiedź negatywna spotykała się z brakiem zrozumienia i z pytaniami o to, *po co w ogóle w takim razie o nim rozmawiać?* Badacz zauważa, że jest to bardzo trafne pytanie w kontekście życia codziennego plemienia. Faktycznie jego członkowie nie tworzą kategorii na takim poziomie abstrakcji, jednak nie znaczy to, że nie są oni zdolni do myślenia abstrakcyjnego. Jak wiadomo wierzenia kultur funkcjonujących w trybie pierwotnym, są bardzo bogate, a ich wytworzenie wymaga wysokiego poziomu

⁷⁰ Zakładając, że X jest inne niż Kpelle.

⁷¹ Zakładając, że imiona przedstawicieli fikcyjnego plemienia są inne niż jakiegokolwiek imię określające przedstawicieli plemienia Kpelle.

⁷² Pytanie to dotyczy lokalizacji obiektu w świecie. Można je sparafrazować pytaniem o to, czy Bóg jest teraz na terenach należących do Pirahã.

abstrakcji. Abstrakcja ta osadzona jest jednak w ich własnym kontekście kulturowym, a bóstwa jawią im się np. jako zjawiska przyrodnicze czy w postaci innych zwierząt. Ich bogowie są często namacalni, gdyż, wierzący spotykają ich⁷³, podczas wizji spowodowanych różnymi substancjami oddłużającymi (takimi jak *ayahuasca*, typowymi dla ich środowiska naturalnego (Olszewski, 2015).

3.4.1.2. Tendencyjność komponentu afektywnego

Procesy będące domeną afektywnego komponentu uprzedzeń związane są z przetwarzaniem emocji, uczuć i innych stanów afektywnych. Jak wskazałam w rozdziale II: *Wybrane mechanizmy poznawcze leżące u podstaw tendencyjności związanej z uprzedzeniami*, a szczególnie w podrozdziałach 2.4. *Efekty oczekiwań*, 2.5. *Teoria dysonansu poznawczego*, 2.6. *Słodka reakcja* oraz 2.7. *Efekt halo* afekty pełnią istotną rolę w ferowaniu interpretacji sytuacji oraz atrybucji stanów mentalnych innym.

W swoim przełomowym artykule *The Rationality of Emotions* (1979), Ronald de Sousa podejmuje się obrony tezy, że emocje nie są irracjonalnymi przeszkodami w działaniu poznawczym, lecz mają własną, wewnętrzną racjonalność. Wbrew dominującej tradycji filozoficznej – od stoików, przez Kartezjusza, aż po klasyczne ujęcia w psychologii – która postrzega emocje jako coś, co należy podporządkować rozumowi lub wręcz jako jego zaprzeczenie, de Sousa argumentuje, że emocje stanowią integralny element ludzkiego poznania i działania, a ich funkcjonowanie można – i należy – oceniać w kategoriach racjonalności. Punktem wyjścia dla wskazanego autora jest rozróżnienie trzech warunków, które musi spełniać stan psychiczny, aby mógł być oceniany jako racjonalny:

- a. Po pierwsze, racjonalność jest zawsze teleologiczna – oznacza skuteczność w realizacji jakiegoś celu.
- b. Po drugie, racjonalność odnosi się tylko do stanów intencjonalnych – czyli takich, które są „o czymś”, mają swój przedmiot.
- c. Po trzecie, każdy stan oceniany jako racjonalny musi spełniać warunek minimalnej racjonalności: musi istnieć przynajmniej jeden kontekst, w którym jego istnienie może być sensownie uzasadnione.

⁷³ Przykładem mogą być złożone wierzenia indian Matsigenka żyjących w południowej Ameryce (Świerk, 2017).

Autor podkreśla, że emocje pełnią rolę struktur organizujących uwagę, kierujących rozumowaniem i wpływających na to, co jednostka uzna za istotne. Właśnie one określają elementy, na które kierujemy uwagę, jakie pytania zadajemy i jakie strategie inferencyjne wybieramy, czyli tam, gdzie logika nie daje jednoznacznej odpowiedzi. W tym sensie emocje wypełniają luki pozostawione przez „czysty rozum”, pozwalając jednostce funkcjonować w świecie pełnym nieokreśloności, wieloznaczności i konfliktów wartości. De Sousa zaznacza, że już od wczesnego dzieciństwa poprzez doświadczenia, opowieści czy literaturę ludzie uczą się rozpoznawać typowe sytuacje (np. zagrożenie, zdrada, strata) i wiązać je z określonymi reakcjami emocjonalnymi. To właśnie te scenariusze stanowią ramy, w których emocje nabierają sensu i mogą być oceniane jako adekwatne lub nieadekwatne. W tym kontekście racjonalność emocji oznacza ich dopasowanie do sytuacji, która odpowiada znanemu scenariuszowi.

Zagadnieniem wartym bliższego przyjrzenia się w koncepcji de Sousa, jest kwestia zmienności emocji oraz ich roli w rozwoju moralnym. Skoro emocje są oparte na społecznie nabytych scenariuszach, to możliwa jest ich racjonalna transformacja – w procesie dojrzewania moralnego czy zmian światopoglądowych. Autor podkreśla, że niektóre scenariusze mogą stać się przestarzałe lub zostać zastąpione nowymi, bardziej adekwatnymi – co jest podstawą zjawisk takich jak np. zmiana przekonań pod wpływem feminizmu czy krytyki rasizmu. Tym samym emocje nie tylko są racjonalne same w sobie, ale również mogą się rozwijać i podlegać ewaluacji z perspektywy bardziej złożonej etycznie czy społecznie. Podkreśla, że emocje są silnie zindywidualizowane i osadzone w doświadczeniu jednostki, jednocześnie argumentuje na rzecz twierdzenia, że emocje mogą mieć obiektywny wymiar. Wyjaśnienie obiektywności emocji odnosi się do pojęcia „trybu obiektywności” (*mode of objectivity*), w jakim mogą znajdować się omawiane stany mentalne. Określa on raczej charakter relacji jednostki do ludzkiej natury, jej doświadczeń i społecznych scenariuszy. W tym rozumieniu obiektywność emocji nie ma charakteru absolutnego czy matematycznego. Poprzez odniesienie do formalnych obiektów emocji (*formal objects of emotions*); takich jak zagrożenie dla strachu czy utrata dla smutku wskazuje, że emocje można oceniać pod kątem ich trafności w odniesieniu do sytuacji. De Sousa (1979) uważa, że analogicznie do percepcji wzrokowej rozróżniamy zniekształcenie od trafnego spostrzeżenia i, tak samo można mówić o emocjach adekwatnych i nieadekwatnych – czyli o emocjonalnych złudzeniach i ich korekcie.

Inną teorią dobrze opisującą wpływ emocji na rozumowanie, w tym wnioskowanie, stworzona została przez Antonio Damasio (1994), który postulował, że nie można oddzielić emocji od rozumu (rozumowania/myślenia). W jego koncepcji emocje rozumiane są jako zorientowana na cel reakcja organizmu, przygotowująca go do wygenerowania określonych reakcji. W *Błędzie Kartezjusza* (1994) autor argumentuje za twierdzeniem, że emocje pomagają w racjonalnym myśleniu, gdyż jednostka najlepiej rozwiązuje problemy, gdy jej na tym zależy (ma motywację). Jako osoba pracująca w dziedzinie neurobiologii behawioralnej dopatruje się głównych funkcji emocji, w procesach prowadzących do utrzymania równowagi w organizmie-systemie. Przetwarzanie tego rodzaju ma charakter allostatyczny i homeostatyczny. Damasio pokazuje, że procesy emocjonalne również ulegają sprzężeniu zwrotnemu, w wyniku czego np. obraz (zdjęcie złamania ręki) lub np. dźwięk (krzaków w ciemnym lesie), może generować reakcję fizjologiczną (podniesione tętno) oraz emocjonalną (np. lęk). Utrzymanie homeostazy odbywa się za pośrednictwem neuroprzekazników. W takiej perspektywie proces ten jest biologiczną tendencją do utrzymywania stanu równowagi w organizmie. Odbywa się w oparciu o pętlę sprzężenia zwrotnego, w której bieżący stan organizmu jest monitorowany i modyfikowany w celu przywrócenia ciała do stanu równowagi. Większość procesów homeostazy jest nieświadoma, jednak niektóre emocje dostają się do świadomości. Dla przykładu, gdy poziom cukru w organizmie spada, zachodzi szereg procesów biologiczno-chemicznych, wywołujących uczucie głodu. Co więcej, jak opisałam w części 2.5. *Teoria dysonansu poznawczego* nieprzyjemny afekt może być redukowany świadomie lub nie. Stan afektywny stanowi motywację do działania, w celu zaspokojenia podstawowej, fizycznej lub emocjonalnej potrzeby. Zatem główną rolą emocji jest motywacja, która pozwala na promowane zachowania zwiększającego prawdopodobieństwo przetrwania. Dodatkowo mózg przewiduje przyszłe zdarzenia, projektując prawdopodobne scenariusze przebiegu sytuacji na podstawie zgromadzonych zestawów doświadczeń i ram wyznaczanych przez normy i interakcje społeczne. Proces przygotowywania organizmu do realizacji przewidywanych potrzeb nazywa się allostazą (Damasio, 1994). Lisa Feldman Barrett (2017) wyjaśnia, że w mózgu przechowywane są ścieżki neuronowe, zawierające schematy zachowań. Ich celem jest zwiększenie szans przeżycia organizmu-systemu. Przykładem zachowania allostatycznego jest m.in. zwiększenie ilości uderzeń serca na minutę. Wspomniane ścieżki aktywowane są przez

zewewnętrzne lub wewnętrzne bodźce, a ich funkcją jest adaptacja organizmu do bieżących warunków.

Innymi słowy, afektywne reakcje na bodźce (zewewnętrzne i wewnętrzne) mają charakter homeostazy, co oznacza, że organizm ma tendencję do przywracania optymalnego dla niego stanu równowagi oraz allostatyczny, ponieważ przygotowują organizm na przyszłe zdarzenia, tworząc prawdopodobne modele mentalne sytuacji oraz optymalizując rytm funkcjonowania organizmu-systemu.

3.4.2. Heurystyki i związane z nimi tendencje w zachowaniu i podejmowaniu decyzji istotne w kontekście uprzedzeń międzygatunkowych

W rozważaniach nad rozumowaniem zakłada się istnienie dwóch form przetwarzania informacji, za które odpowiadają dwa systemy: pierwszy (szybki) i drugi (wolny)⁷⁴. System szybki jest automatyczny, czyli nie podlega mechanizmom kontroli poznawczej oraz działa bez ingerencji uwagi, co za tym idzie, wymaga niewielkiego nakładu energii. Jego rolą jest przetwarzanie np. sygnałów mimicznych i posturalnych innych agensów. Czyli zadania cechujące się łatwością poznawczą. Natomiast system wolny angażuje uwagę i organizuje ją między szczegółowymi zadaniami. Wiąże się z subiektywnym poczuciem skupienia, będącym skutkiem koncentracji. Do jego zadań należy przetwarzanie trudnych problemów jak np. wyliczenia matematyczne, czyli podczas rozwiązywania zadań wymagających nakładu energii, nazywanego wysileniem poznawczym (*cognitive ease*). Na wskaźnik wysilenia składa się zarówno aktualny poziom wysiłku umysłowego, jak i obecność nieukończonych zadań. Kahneman w swoich licznych badaniach (2012) argumentuje, że domeną systemu 1 jest generowanie emocji i innych stanów mentalnych, które stanowią podstawowe źródło świadomych przekonań i intencjonalnych wyborów. Natomiast system 2 angażuje racjonalne rozumowanie, a co za tym idzie, wymaga uwagi. Przy czym należy pamiętać, że procesy uwagowe wymagają wysokiego nakładu energii, dlatego ich udział jest ograniczony do absolutnego minimum. Co więcej, myślenie racjonalne potrzebne do rozwiązywania niektórych zadań może być niemożliwe do wykonania dla poszczególniej jednostki, ponieważ może ono wymagać wykorzystania np. reguł logiki formalnej. Ta forma przetwarzania wykorzystywana jest w sytuacjach nowych, a co za tym idzie, potencjalnie niebezpiecznych lub zwyczajnie wykraczających poza

⁷⁴ Terminologię systemu pierwszego i drugiego wprowadzili Keith Stanovich i Richard West (2000).

codzienną rutynę. Klasycznym eksperymentem obrazującym konflikt między dwoma systemami przetwarzania jest test Stroopa (Hoshikawa, Yamamoto, 1997). Co więcej, domeną systemu pierwszego jest także tworzenie spostrzeżeń, w konsekwencji niektóre reprezentacje percepcyjne generowane przez umysł są nieadekwatne do rzeczywistości. W psychologii poznawczej błędy tego rodzaju noszą nazwę zniekształceń np. iluzji lub złudzeń. Różnica między wymienionymi polega na tym, że pierwsze doświadczane są w taki sam sposób przez ludzi żyjących w obrębie określonych kultur, drugie natomiast mają charakter indywidualny. Przykładem popularnych iluzji są linie Müllera-Lyera, czy iluzja Ebbinghausa. Drugie natomiast mają charakter indywidualny i są wynikiem chwilowego zniekształcenia lub tendencji do postrzegania schematów/wzorów, nawet tam, gdzie ich nie ma (myślenie autystyczne (Nęcka, 2013, str. 297-301)). Takim zniekształceniem są pareidolie lub sytuacje, w których człowiek wykonuje zachowanie lękowe na skutek np. bodźca, jakim jest wąż ogrodowy, ponieważ agens myli go z prawdziwym wężem (Wardle, 2023). System 2 wymaga dużego nakładu energii, dlatego jego rozwój jest długi i kształtuje się przez całe życie. Istnieją różnice indywidualne w prędkości kształtowania się opisywanego systemu, które widać już we wczesnym dzieciństwie.

Dla lepszego zrozumienia powyższego stwierdzenia powołałam się na najbardziej popularny paradygmat we wskazanym obszarze – test pianek (*marshmallow test*). Nagrania z jego przebiegu cieszą się dużą popularnością na platformach społecznościowych, takich jak np. *youtube* (Igniter media, 2010). Eksperyment polega na umieszczeniu dzieci do 4 roku życia w pomieszczeniu z lustrem weneckim. Wewnątrz stoi tylko stół, a na nim talerzyk z jedną pianką oraz krzesło. Na ścianach nie ma żadnych obrazów czy innych rozpraszaczy uwagi, nie ma także żadnych zabawek. Dziecko proszone jest o zajęcie miejsca na krześle oraz informuje się je, że eksperymentator musi wyjść na 15 min, jednak prosi, aby dziecko nie zjadło pianki. Jeżeli uda mu/jej się wytrzymać, po powrocie dorosłego dostanie kolejną piankę. Kahneman (2012) twierdzi, że powtórne badania (w formie wywiadów) ujawniły, że dzieci, którym udało się wytrzymać 15 min w wieku 4 lat podczas eksperymentu, przejawiały m.in. mniejszą skłonność do uzależnień w przyszłości, co więcej dzieci, które nie poddają się pokusie, uzyskują w przyszłości wyższe wyniki w testach na inteligencję (Kahneman, 2012, str. 66). Takie wyniki można wytłumaczyć faktem, że *marshmallow test* jest narzędziem do pomiaru procesów kontroli poznawczej, opisanych w poprzedniej części. Dzieci, które w wieku 4 lat mają wykształcone procesy

hamowania na tyle, aby zatrzymać pokusę otrzymania szybkiej nagrody, w przyszłości są bardziej cierpliwe w realizacji zadań z odroczoną nagrodą, czyli są skuteczniejsze w długoterminowych działaniach.

Metafora, z której korzysta się, do opisu umysłu w omawianym zakresie głosi, że ten jest maszyną skojarzeniową. Mechanizm jej działania został opisany wcześniej jako rozprzestrzeniająca się aktywacja. Polega on na tym, że aktywowanie w umyśle pojęcia, konceptu czy innej struktury poznawczej, wywołuje aktywację innych elementów sieci połączonych ze sobą relacją skojarzeniową. Proces ten odbywa się szybko i równolegle, tworząc samowzmacniający się schemat zintegrowanych i różnorodnych reakcji poznawczych, który jest skojarzeniowo spójny (Nęcka, 2013; Kahneman, 2012). Podejście to ma początek w latach '80 XX wieku, gdy doszło do rewolucji w psychologii. Jej przyczyną było odkrycie, że bodziec w postaci słowa (etykiety werbalnej) natychmiastowo skutkuje łatwością przypominania sobie wyrazów pokrewnych⁷⁵. Zjawisko to nazywane zostało torowaniem, a badania go dotyczące prowadzone są do dzisiaj w paradygmacie prymowania. Prymowanie jest techniką eksperymentalną polegającą na badaniu wpływu bodźca prezentowanego wcześniej (pryma), na przetwarzanie bodźca docelowego. Pryma i bodziec docelowy połączone są różnymi relacjami np. o charakterze semantycznym lub sensorycznym. Wyróżnia się sześć rodzajów omawianej techniki (Nęcka, 2013):

- a. Poprzedzanie podprogowe/ subliminalne: gdy pryma jest bodźcem podprogowym,
- b. Poprzedzanie repetytywne: polega na krótkiej ekspozycji na bodziec, który po chwili pojawia się ponownie i ma być przetworzony przez uczestnika badania,
- c. Poprzedzanie sensoryczne: gdy pryma jest np. synonimem bodźca docelowego. W takiej sytuacji ułatwia przetwarzanie bodźca docelowego,
- d. Poprzedzanie pozytywne/torowanie: paradygmat, w którym pryma ułatwia przetworzenie bodźca docelowego,
- e. Poprzedzanie peryferyczne: gdy bodziec znajduje się na peryferiach pola percepcyjnego,
- f. Poprzedzanie afektywne: rolę prymu pełnią emocje np. zdjęcia uśmiechniętych twarzy.

⁷⁵ Pokrewnych, czyli np. należących do tej samej kategorii poznawczej.

Mechanizm torowania uczestniczy w ferowaniu interpretacji i narracji przeszłej, aktualnej i potencjalnej sytuacji agensa. Czyli dotyczy on nie tylko etykiet werbalnych i pojęć, ale też emocji, czy złożonych struktur poznawczych takich jak stereotypy, a co za tym idzie, towarzyszy atrybucji cech obiektom percepcyjnie dostępnym czy wyobrażeniowym. Czyli zjawisko torowania jest ściśle związane z rozważaniami dotyczącymi mechanizmów kategoryzacji poznawczej czy procesami egotyzmu atrybucyjnego opisanymi wcześniej.

Inny klasyczny eksperyment dotyczący torowania został przeprowadzony przez Ericka Tompsona i jego współpracowników (Thompson i.in., 1994). W jego ramach zadaniem uczestników było ułożenie zdania składającego się z czterech słów, mając do dyspozycji pięć wyrazów. Grupa badawcza otrzymała słowa, które w założeniu kojarzyły się ze starością, natomiast kontrolna — słowa związane z innymi kategoriami poznawczymi. Następnie uczestnicy proszeni byli o przejście korytarzem z punktu A do punktu B. Wykonywano pomiar prędkości pokonania dystansu, a jego wyniki ujawniły, że przedstawiciele grupy badawczej pokonywali określony dystans znacznie wolniej niż Ci z grupy kontrolnej. Innym ważnym wnioskiem wyciągniętym z badań przeprowadzonych we wskazanym zakresie, jest fakt, że pojęcie śmierci pojawiające się w środowisku (np. w przekazie medialnym) człowieka, zwiększa dla niego atrakcyjność narracji autorytarnych (Kahneman, 2012, str.78).

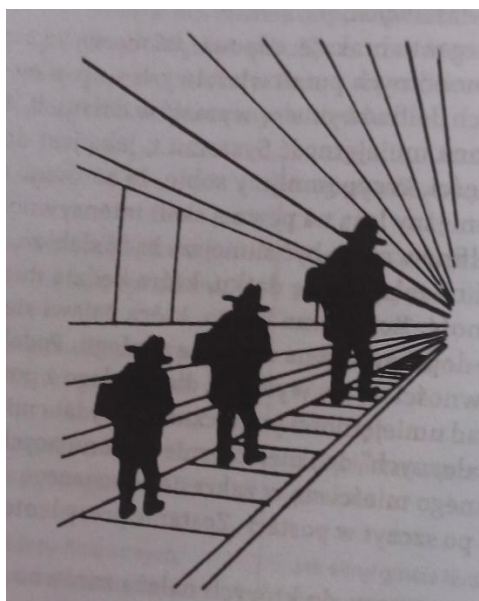
System pierwszy automatycznie generuje opinie dotyczące złożonych aspektów świata, proces ten odbywa się z udziałem mechanizmu polegającego na zastąpieniu (*substitution*) głównego problemu (pytania leżącego u jego podstaw) innym, łatwiejszym do przetworzenia (Kahneman, 2012, str. 132-41). Dlatego autor *Pułapek myślenia. O myśleniu szybkim i wolnym* wyróżnia dwa rodzaje pytań: właściwe i heurystyczne. Pierwsze z wymienionych jest pytaniem leżącym u podstaw problemu, natomiast drugie stanowi uproszczenie pierwszego, na które system odpowiada, zastępując nim pytanie właściwe, znacznie trudniejsze do przetworzenia, gdyż wymaga uwzględnienia znacznie większej bazy danych, przy czym niektórych informacji koniecznych do przeprowadzenia takiego procesu może brakować. Pionierami badań nad zastępowaniem pytania jest wielokrotnie wspomniany już Kahneman oraz jego wieloletni współpracownik, mentor — Amos Tversky. Początkowo definiowali heurystykę jako „prostą procedurę, która pozwala udzielić adekwatnych, niedoskonałych odpowiedzi na trudne pytania” (Kahneman, 2012, str. 133). Ich rozważania rozpoczęły się od próby opisanie potocznego rozumienia

prawdopodobieństwa. Autorzy *Judgment under uncertainty: Heuristics and biases* zauważyli, że ludzie szacują prawdopodobieństwo, pomimo że nie wiedzą, czym ono dokładnie jest (Kahneman, Tversky, 1974). Na podstawie tej obserwacji założyli, że pojęcie to musi być w pewien sposób upraszczane, dlatego postanowili zbadać, na czym owo uproszenie polega. Badania raportowane przez autorów oraz wielokrotnie powielane pozwoliły udzielić odpowiedzi na przyjęty problem badawczy, pokazując, że w rzeczywistości, gdy ludzie myślą, że szacują prawdopodobieństwo, przeprowadzają inne procesy (Kahneman, Tversky, 1974, Kahneman, 2012; Aronson, Travis, 2007; Thaler, Sunstein, 2023 i wiele innych).

W celu zobrazowania omawianych procesów, pokrótce omówię dwie heurystyki. Pierwsza — heurystyka trójwymiarowości — pokazuje, że przetwarzanie spostrzeżeń i innych reprezentacji percepcyjnych jest domeną systemu pierwszego, a co za tym idzie, są one podatne na błędy. Druga — heurystyka afektu — sprawia, że większość poglądów dotyczących rzeczywistości zależna jest od chęci i niechęci agensa. Zatem jest ona silnie związana z opisanym we wcześniejszym rozdziale efektem potwierdzenia. Należy pamiętać, że wymienione heurystyki nie działają pojedynczo czy sekwencyjnie, a funkcjonują równolegle, wyniki rozumowania w obrębie jednej z nich mogą wpływać na obliczenia innej. Dla przykładu łatwo dostępne dane, doświadczenia, wspomnienia, przekazy medialne i kulturowe mogą stanowić kotwicę oraz określać cechy reprezentatywne dla określonej grupy.

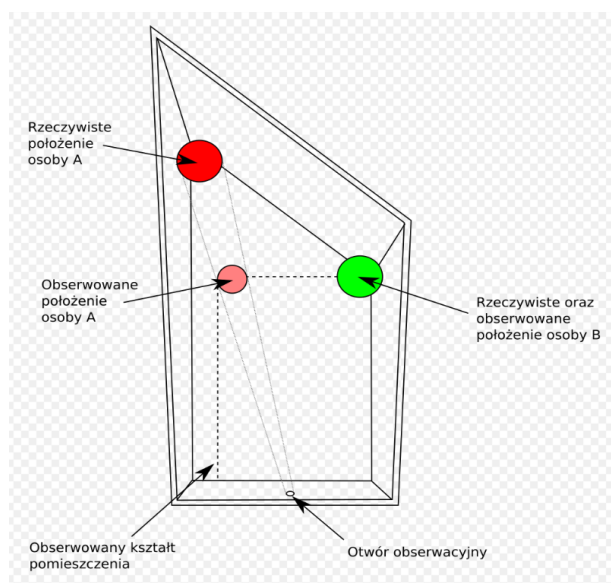
Heurystyka trójwymiarowości (zdjęcie 5.) wiąże się z iluzjami takimi jak pokój Amesa (zdjęcie 6.). W swojej książce Kahneman (2012) przedstawia ciekawą iluzję, sprawiającą, że widoczne na Zdjęciu 7. postaci wydają się różnego wzrostu, chociaż w rzeczywistości ich rozmiar jest ten sam.

Zdjęcie 5. Iluzja Trójwymiarowości



[Źródło (Kahneman, 2012, str. 136)].

Zdjęcie 6. Iluzja Pokoju Amesa



[Źródło: (Wikipedia, pokój Amesa)]

Wskazany błąd wynika z tego, że system poznawczy interpretuje dwuwymiarowy obrazek na papierze jako trójwymiarową scenę. W ramach tej interpretacji postać znajdująca się najwyżej jest zarazem największa oraz widoczna jest odległość między nią a pozostałymi postaciami, znajdujących się niżej. Zastąpienie w tym wypadku polega na dominacji trójwymiarowego spostrzeżenia, względem dwuwymiarowego.

Heurystyka ta dobrze pokazuje, że zastąpienie nie jest związane ze złym zrozumieniem pytania (Kahneman, 2012, str. 137). Warto pamiętać o tym, że mechanizm zastąpienia, nie jest bezpodstawny, pewne elementy kontekstowe sugerują opisaną interpretację. W rzeczywistości wskazówki te są nieistotne do udzielenia odpowiedzi na pytanie właściwe, jednak zastąpienie jest mechanizmem automatycznym, angażowanym w sytuacjach, które „wydają się” go wymagać, w celu uproszczenia problemu.

Termin „heurystyka afektu” zaproponowany został przez Paula Slovic, Howarda Kunreuthera i Gilberta White’a (2016), którzy definiowali ją, jako „pozwolenie sobie” na to, aby własne poglądy i opinie kształtowane były przez osobiste upodobania i niechęci (Kahneman, 2012, str. 139). Czyli to m.in. to ta heurystyka leży u podstaw efektu halo oraz bierze udział w redukcji dysonansu poznawczego. Rolą systemu 2 jest monitorowanie i korygowanie procesów zachodzących w ramach przetwarzania szybkiego. Dlatego to właśnie system wolny odpowiada za samokrytykę i autorefleksję. Jednakże jak wielokrotnie już wspomniałam, ze względu na wysilenie poznawcze konieczne do wykonania wskazanych procesów, to właśnie omawiany schemat rozumowania angażowany jest do usprawiedliwiania własnych działań, nawet w sposób niezgodny z podzielanymi poglądami.

3.4.2.1. Heurystyki oraz tendencyjność związana z uprzedzeniami wobec ZINL

Przyjmując założenie, że system pierwszy przetwarza informacje korzystając z heurystyk zamiast z reguł logiki formalnej, warto przytoczyć koncepcję Toma Parkera (1983) opisaną w książce *Zasady praktyczne [Rules of Thumb]*, w której autor opisuje tytułowe reguły rządzące wnioskowaniem. Wskazuje on, że ze względu na ograniczenia systemu–organizmu schematy pełną funkcję adaptacyjną, redukując ilość zużywanych podczas przetwarzania zasobów. Zatem heurystyki rozumiane są jako praktyczne zasady rządzące rozumowaniem, będące wynikiem racjonalności ekologicznej, wykształconej na przestrzeni lat w procesie egzaptacji. Najważniejsze z nich leżące u podstaw uprzedzeń prezentowanych w stosunku do ZINL opisane zostaną szczegółowo w kolejnych podrozdziałach niniejszej części. W tym miejscu chciałabym jednak przytoczyć kilka heurystyk i tendencji, które jak uważam, również pełnią istotną rolę w tworzeniu interpretacji wyników badań z udziałem ZINL, także tych z zakresu teorii umysłu.

Jak wielokrotnie wskazałam, system pierwszy jest bardzo skutecznym mechanizmem adaptacyjnym, który bezwysiłkowo i automatycznie wykrywa związki

przyczynowo-skutkowe, nawet gdy w rzeczywistości taka relacja nie zachodzi. Należy także pamiętać, że sama wiedza o *czymś* nie jest wystarczająca do umiejętności korzystania z niej. Jak wielokrotnie wskazywałam wyżej, ludzie nauki również ulegają tendencyjności podczas interpretowania i konstruowania eksperymentów. Jedną z heurystyk leżącą u podstaw wskazanych błędów jest heurystyka małych liczb, która odnosi się do wnioskowania na podstawie założeń opartych o małą liczbę danych, w wyniku czego generalizacja⁷⁶ uzyskanych wyników jest niemożliwa. Kahneman (2012, str. 151) argumentuje, że psychologowie często wybierają przeprowadzanie badań na małych grupach badawczych, gdyż wtedy istnieje mniejsza szansa na uzyskanie wyników negatywnych (niepotwierdzenie hipotez założonych w badaniu).

Wspomniana już w niniejszej dysertacji książka *7 Deadly Sins of Psychology. A Manifesto for Reforming the Culture of Scientific Practice*, autorstwa Chrisa Chambersa stanowi krytyczne spojrzenie na współczesną psychologię i jej praktyki badawcze. Autor identyfikuje siedem głównych „grzechów”, które osłabiają naukową wiarygodność i rzetelność badań we wskazanej dyscyplinie. Zwraca uwagę na przynajmniej nieostrożne metody projektowania badań naukowych, które prowadzą do błędnych wniosków. Podkreśla także problem z replikacją eksperymentów, w którego konsekwencji niemożliwa jest weryfikacja wyników oraz, co za tym idzie, postęp naukowy. Chambers krytykuje również tendencje w systemie wydawniczym, polegające na preferencji wyników pozytywnych, względem tych prezentujących wyniki negatywnych. Te ostatnie są ignorowane i praktycznie nie mają szans na publikację. Ciekawą praktyką opisaną przez autora jest HARKing (*Hypothesizing After the Results are Known*), polega on na formułowaniu hipotez po uzyskaniu wyników eksperymentu, co zniekształca prawdziwy proces badawczy. Taki zabieg sprawia, że wyniki stanowią kotwicę determinującą hipotezę. Kolejną powszechnie stosowaną praktyką jest p-hacking, czyli manipulowanie danymi eksperymentalnymi w celu uzyskania istotnych statystycznie wyników. Zabieg ten skutkuje brakiem przejrzystości w udostępnianiu danych i metod, co utrudnia weryfikację wyników przez innych badaczy. Chambers krytykuje również system recenzowania w ramach czasopism naukowych, który prowadzi do publikacji prac niskiej jakości. W praktyce oznacza to, że artykuły naukowe są czasem recenzowane niedbale lub przez osoby, które nie są

⁷⁶ W tym kontekście generalizacja rozumiana jest jako rozszerzenie wniosków płynących z badania, na szerszą populację.

wystarczająco kompetentne w określonym zakresie. Taki system może prowadzić do akceptacji lub publikacji prac, które zawierają błędy metodologiczne, są nieodpowiednio zaprojektowane, a ich założenia teoretyczne są nieprzejrzyste. Wymienione błędy publikowanych prac nie są korygowane na wczesnym etapie, w wyniku czego w literaturze naukowej pojawiają się prace, które nie spełniają wysokich standardów naukowych (Chambers, 2019).

Jak wskazałam w części 2.5. *Teoria dysonansu poznawczego* ludzie przejawiają tendencje do przesadnej wiary w wyniki własnych rozumowań. Jest to związane z niechęcią do odrzucania istniejących już struktur poznawczych, pozwalających utrzymać pozytywny wizerunek własnego *ja*. W konsekwencji systemy pamięciowe cechują się egotyzmem atrybucyjnym, a także efektem powyżej przeciętnej (Thaler, Sunstein, 2023, str. 51-54). W celu wyjaśnienia drugiego z wymienionych odwołam się teraz do obserwacji popartych doświadczeniami empirycznymi, wykazującymi, że większość (90%) kierowców pojazdów (samochodów, ciężarówek, motorów) uważa siebie za ponadprzeciętnego, czyli uważa, że jest lepszy w prowadzeniu pojazdów niż większość ludzi posiadających prawo jazdy. Co ciekawe, badanie to przeprowadzone zostało także na grupie uczestników noszących tytuł naukowy profesora, wyniki ujawniły, że ok. 94% przedstawicieli wskazanej grupy, jednej z uczelni wyższych uważa się za ponad przeciętnych w obszarze wykonywanej pracy (Thaler, Sunstein, 2023, str. 52). Oczywiście wyniki uzyskane z badania prowadzonego w obrębie jednego ośrodka naukowego, nie są wystarczające do tego, aby uznać je za adekwatne dla populacji uwzględniającej większą liczbę instytucji, jednak mając na uwadze fakt, że podobne rezultaty uzyskano w wielu innych eksperymentach, wydaje się, że wyniki można poddać generalizacji dla większej populacji (Thaler, Sunstein, 2023; Kahneman, 2012; Nęcka, 2013; Aronson, Travis, 2007). Heurystyka ta obserwowana jest u uczestników licznych badań, bez względu na ich przynależność do kategorii społecznej. Zatem omówiona tendencja polegająca na przecenianiu szans na zajście negatywnych wydarzeń, których *ja* jestem podmiotem. Leży także u źródła nieostrożnych zachowań, których przykładem jest decyzja podjęta przez wielu ludzi podczas pandemii Covid-19, w latach 2020 i 2021 — o niekorzystaniu z maseczek ochronnych w przestrzeni publicznej (Thaler, Sunstein, 2023).

Poczucie pewności dotyczące własnych wniosków jest konsekwencją postrzegania rzeczywistości jako bardziej spójnej, niż jest faktycznie, dlatego odkrywca

omawianej heurystyki podkreśla, że prawo małych liczb jest częścią szerszej heurystyki polegającej na preferowaniu przetwarzania cechującego się pewnością, względem tego budzącego dalsze wątpliwości (Kahneman, 2012, str. 154). Preferencja ta skutkuje faworyzowaniem prostych rozwiązań, które często nie uwzględniają złożoności problemu. Tendencja do dostrzegania związków przyczynowo-skutkowych nawet tam, gdzie rzeczywiście nie zachodzą, wynikają z architektury podsystemów poznawczych, takich jak percepcja oraz tych odpowiadających za przetwarzanie odgórne. Przyjmując metaforę głoszącą, że umysł jest maszyną do wnioskowania, która zawiera w swoim repertuarze reguł heurystykę, w wyniku której pytanie, *co mogło się wydarzyć?* zostaje zastąpione pytaniem o przyczyny zjawiska, pomimo że w większości przypadków brakuje konkretnego rzeczywistego powodu, przez który doszło do określonych wydarzeń. Omawiany schemat rozumowania wiąże się z poczuciem domknięcia poznawczego, czyli naturalną potrzebą życia w świecie uporządkowanym, pełnym regularności, a co za tym idzie, relatywnie bezpiecznym, gdyż istnieje w nim możliwość przewidzenia przyszłych zdarzeń. Natomiast, sytuacje nieznane i niezrozumiałe generują niepewność i niepokój, będące źródłem stresu. To właśnie domknięcie poznawcze pozwala nam na redukcję wymienionych afektów, a także w ramach domknięcia poznawczego generowane są narracje, w tym opisane niżej mity.

Przyjrzyjmy się teraz heurystyce, nazywanej mianem *farming*, określanej również efektem kontekstu sytuacyjnego (Thaler, Sunstein, 2023). Jest ona wynikiem architektury czynności pamięciowych, a także organizacji materiału w systemie poznawczym, zatem jego źródłem są przekonania o świecie, które podziela agens. Rozpatrywana jest w obszarze badań nad czynnościami pamięciowymi, np. w kontekście elaboracji, czyli opracowaniu materiału, polegającym na dodawaniu do zestawu informacji, które zawarte są w podsystemie pamięci nowych, czyli np. danych, których nie było we wcześniejszym materiale eksperymentalnym (Stein, Bransford, 1979). Takie opracowanie może przyjmować formę powiązania kodowanego materiału z jednym lub wieloma podzielanymi przekonaniami. Celem wskazanego mechanizmu jest doprecyzowywanie i uszczegółowienie posiadanych informacji, można więc powiedzieć, że elaboracja służy nadaniu głębszego sensu kodowanemu materiałowi. Szczególnie interesującym w kontekście niniejszej pracy, rodzajem omawianego mechanizmu jest poszukiwanie wyjaśnień, które polega na przeszukiwaniu zestawu dostępnych narracji lub stereotypów, mogących dostarczyć informacji o przyczynach i potencjalnych skutkach spostrzeganych wydarzeń.

Kolejnym kluczowym czynnikiem determinującym kodowanie i odpamiętywanie informacji jest kontekst (Tulving Thomson, 1973; Godden, Baddeley, 1975). Konsekwencje jego wpływu stanowiły przedmiot licznych badań naukowych z zakresu psychologii poznawczej. Eksperymenty te ujawniły szereg efektów jak np. specyficzności kodowania czy grupowania kategorialnego, które są elementem architektury podsystemu pamięci. Drugi z wymienionych opisany został szerzej w części 2.1. *Mechanizm kategoryzacji społecznej*. Pierwszy natomiast odnosi się do procesów odpamiętywania informacji, Endel Tulving i Donald Thomson (1973) pokazali, że tego rodzaju wyszukiwanie jest łatwiejsze, jeżeli sygnały (bodźce) pojawiające się podczas odtwarzania informacji są zgodne z sygnałami obecnymi podczas jej kodowania. W innym eksperymencie potwierdzono wskazany wniosek. Jego uczestnikami byli płetwonurkowie, którzy podzieleni zostali na dwie grupy (Godden, Baddeley, 1975). Część respondentów zapamiętywała (kodowanie) materiał, znajdując się pod wodą, natomiast innych proszono o to samo na lądzie. Następnie część z nich zmieniała kontekst t.j. ci, którzy kodowali informacje pod wodą, wychodzili na ląd, natomiast ci, którzy pierwszy etap eksperymentu przechodzili na lądzie, schodzili pod wodę, gdzie odtwarzali zapamiętany wcześniej materiał (czyli odpamiętywanie zachodziło w kontekście niezgodnym z warunkami kodowania). Wyniki eksperymentu pokazały, że uczestnicy odtwarzający informacje, w kontekście zgodnym z warunkami kodowania, uzyskiwali znacznie lepsze efekty.

3.4.2.2. Heurystyka zakotwiczenia i dopasowania

Heurystyka zakotwiczenia i dopasowania związana jest z procesami szacowania opartymi na strategii dostosowywania wartości wyjściowej aż do uzyskania wartości ostatecznej (Kahneman, 2012, str. 571). Czyli jest to zjawisko, w którym początkowa informacja (kotwica) wpływa na późniejsze decyzje i osądy. Określona wartość wyjściowa może zostać wybrana podczas procesu formułowania problemu lub przyjęta jako przesłanka we wstępnej fazie rozwiązywania problemu. Zakotwiczenie jest ściśle związane z opisanym wcześniej zjawiskiem torowania, gdyż można powiedzieć, że wartość, do której odnosi się agens, podczas szacowania, stanowi prymę dla dalszych procesów rozumowania/myślenia. Co ciekawe jednym ze skutków zakotwiczenia jest zjawisko fiksacji opisane szerzej w częściach 2.4. *Efekty oczekiwań* oraz 2.6. *Słodka reakcja*. Uważam, że omawiana heurystyka związana jest z prezentowaniem uprzedzeń i podzielaniem sądów, także tych wobec zwierząt pozaludzkich. W przypadku postaw

kotwicą mogą być doświadczenia, stereotypy, a nawet pojedyncze przekonania, zatem kultura, media, pierwsze wrażenie i tendencja do potwierdzania początkowych przekonań, kształtują postawy, jakie przyjmują jednostki funkcjonujące w ich zasięgu.

Agens korzysta z kotwic, gdy musi odpowiedzieć na trudne pytanie, nie znając na nie odpowiedzi. W takiej sytuacji następuje poszukiwanie informacji, która może stanowić sugestię (prymę) wspomagającą znalezienie rozwiązania. Wiele badań w omawianym obszarze wskazuje, że przypadkowe kotwice, torują rozumowanie/myślenie tak samo mocno, co te potencjalnie dostarczające informacji. Jednym z klasycznych eksperymentów uczestnicy, proszeni byli o odpowiedź na jedno z dwóch pytań kotwiczących, które dotyczyły temperatury: *czy średnia temperatura roczna w Niemczech jest wyższa, czy niższa niż 20 stopni?* oraz *czy średnia temperatura w Niemczech jest wyższa, czy niższa niż 5 stopni?* (Strack i współpracownicy, 1988). Wyniki ujawniły, że ci, którzy odpowiadali na pierwsze z wymienionych pytań, łatwiej przypominali sobie (w paradygmacie swobodnego odtwarzania) pojęcia związane z latem, natomiast uczestnicy odpowiadający na drugie z wymienionych, łatwiej przywoływali pojęcia związane z zimą. Oznacza to, że informacja o temperaturze stanowiła prymę, która na drodze selektywnej aktywacji utorowała odtwarzanie pojęć i wspomnień z odpowiedniej kategorii [lato]; [zima]. Inny klasyczny eksperyment przeprowadzony został na niemieckich studentach, którzy podzieleni zostali na dwie grupy (Strack i współpracownicy, 1988). Przedstawiciele jednej z nich udzielali odpowiedzi na dwa pytania: *jak szczęśliwy szczęśliwa jesteś w ostatnim czasie?* oraz *ile razy w ubiegłym miesiącu byłeś/byłaś na randce?*, w drugiej grupie odwrócono kolejność pytań. Odpowiedzi na pytania we wskazanej kolejności nie wykazywały silnej zależności statystycznej, co sugeruje, że osoby kształcące się w tej placówce, nie traktowały randkowania jako najważniejszego czynnika determinującego szczęście w życiu. Jednak w grupie ze zmienioną kolejnością pytań zależność ta była znacznie wyższa. Zjawisko to można wyjaśnić mechanizmem torowania, gdyż pytanie o randki prymowało odpowiedź na pytanie dotyczące poziomu szczęścia w życiu. Czyli punkt odniesienia, jakim jest życie uczuciowe (mierzone przez ilość randek), rzutował na ocenę poziomu zadowolenia z życia (Kahneman, 2012).

Znaleźć możemy jeszcze jedno badanie, szczególnie istotne w kontekście prowadzonych przeze mnie analiz (Haggaga, Pacci, 2014). W jego ramach porównywano napiwki kierowców taksówek pracujących w dwóch różnych korporacjach. Jedna z nich oferowała klientom wpłatę wysokości 15-20-25% od kwoty

zapłaconej za przejazd, druga natomiast umożliwiała wpłatę napiwku w wysokości 20-25-30%. Wydaje się, że taksometr z wyższymi kwotami powinien znacznie zwiększyć zarobki pojedynczych taksówkarzy, tak się stało. Jednak wzrost zarobków poszczególnych kierowców, zwiększył także liczbę pracowników, a co za tym idzie, pociągną za sobą wzrost liczby kierowców nieotrzymujących napiwków wcale. Było to odpowiedzią na zwiększające się niezadowolenie niektórych klientów, wynikające z mechanizmu oporu psychicznego⁷⁷. W jego konsekwencji pracownicy korporacji oferującej wyższe kwoty napiwków w skali rocznej zarabiali znacznie mniej. Udział heurystyki zakotwiczenia i dopasowania w procesach wnioskowania jest szeroko omawiany w literaturze (Arronson, Travis, 2007; Kahneman, 2012; Nęcka, 2013; Stroh, 2022; OpenStax, 2015; Thaler, Sunstein, 2023). Ciekawych danych w omawianym obszarze dostarcza artykuł *How the Implicit Bias, Representative Heuristic, and Anchoring Effect Embed Institutionalized Racism into the American Judicial System* (Stroh, 2022), w ramach którego analizowano podejmowanie decyzji przez sędziów i ławników w sytuacji procesu sądowego. Badania w nim przedstawione wskazują, że wyroki obarczone są licznymi zniekształceniami poznawczymi determinującymi ostateczny werdykt. Eksperyment miał na celu zbadanie wpływu uprzedzeń na wnioskowanie wykonywane przez sędziów. Aby zrealizować wskazany cel, sędziowie otrzymywali rozprawy, które dotyczyły bardzo podobnych spraw, jednak pierwsza połowa podejrzanych była osobą czarnoskórą, druga — białą. Eksperyment ujawnił, że wyroki w sprawach osób czarnoskórych były bardziej zbliżone do żądań prokuratora, niż w przypadku białych sprawców. To sugeruje, że kotwica w postaci wyroków proponowanych przez prawników zakłóca proces ewaluacji ich żądań. Autorzy wskazują, że zależność ta widoczna jest także podczas rozpraw prowadzonych przez sędziów deklarujących, że obywateli należy traktować równo, a co za tym idzie, w sądzie nie ma miejsca na wyroki motywowane uprzedzeniami rasowymi. Oznacza to, że pomimo świadomych deklaracji dyskryminują czarnoskórych obywateli poprzez wydawanie tendencyjnych wyroków, czyli sami ulegają zniekształceniom wynikającym ze stereotypów, mitów, a także afektów związanych z określonymi grupami społecznymi. Na podstawie tych obserwacji autorzy wysuwają twierdzenie, że u podstaw systemowych uprzedzeń w sądownictwie leżą stereotypy i mity o charakterze

⁷⁷ Pojawiającego się, gdy podmiot czuje, że jest do czegoś zmuszany, może pojawić się w sytuacji, gdy wydawany jest fortunny rozkaz, jak i wtedy, gdy dochodzi do sugestii (Thaler, Sunstein, 2023, str. 48).

ukrytym. Czyli na skutek kategoryzacji społecznej, która sprzyja atrybucji cech zgodnych ze stereotypami, sędziowie wydają wyroki będące wynikiem wnioskowania motywowanego, nie zaś racjonalnego.

Istnieje wiele badań z zakresu psychologii społecznej i rozwojowej, wskazujących, że powszechnych kotwic dostarczają: wczesne doświadczenia, wrażenia, stereotypy, edukacja, tradycje, kultura, literatura, media oraz mechanizm potwierdzenia. Jedną z teorii wspierających wskazane stanowisko jest teoria uczenia się społecznego Alberta Bandury i Richard Waltersa (1977), która zakłada, że ludzie uczą się zachowań społecznych (tych akceptowalnych i nieakceptowalnych) oraz przyjmują postawy na podstawie obserwacji innych: rodziców, nauczycieli czy rówieśników. Wczesne interakcje z innymi agensami mogą wpływać na ferowanie przekonań i postaw w dorosłym życiu. Dla przykładu, jeżeli ludzie w otoczeniu dziecka dzielą przekonania, głoszące, że przedstawiciele pewnych gatunków zwierząt są niebezpieczni (np. wilki) lub brudni (np. świnie), jednostka będzie miała tendencję do postrzegania ich w ten sposób, także w dorosłym życiu. Omawiane stanowisko wspierane jest także przez autorów takich, jak Frances Aboud (1988) czy Patricia'ę Devine (1989) wykazujących w swoich badaniach, że stereotypy i uprzedzenia często przekazywane są przez rodziców i otoczenie społeczne, czyli ich nabywanie odbywa się we wczesnym dzieciństwie. Fenomen przejmowania przekonań i postaw od innych agensów wyjaśniany jest na gruncie wspomnianej wcześniej teorii przywiązania, w ramach której argumentuje się, że na proces kształtowania schematu przywiązania jednostki, wpływają m.in. relacje i postawy interpersonalne, a co za tym idzie, dzieci, które rozwijają bezpieczny styl przywiązania, są statystycznie bardziej pewne siebie oraz otwarte na eksplorowanie świata wokół nich (Bowlby, 1969). Taka otwartość może prowadzić do większej akceptacji i pozytywnych postaw wobec różnorodnych form życia, w tym zwierząt pozaludzkich. Badania pokazują, że dzieci z bezpiecznym stylem przywiązania są bardziej empatyczne i lepiej rozumieją emocje innych, co może przekładać się na empatię wobec ZINL (Cassidy, 1994).

Inną teorią, w ramach której podkreśla się rolę zakotwiczenia w heurystycznym przetwarzaniu informacji, jest teoria rozwoju moralnego Lawrence Kohlberga (1981). Jej autor wskazuje, że proces rozwoju zdolności podejmowania decyzji etycznych zachodzi w kilku etapach. Początkowo, dzieci kierują się prostymi informacjami opartymi na karze i nagrodzie, na tym etapie postawy wobec zwierząt formowane są głównie na podstawie obserwacji zachowań dorosłych. Na wyższym poziomie rozwoju

moralnego, jednostki zaczynają dostrzegać wartość życia zwierząt oraz ich zdolność do odczuwania cierpienia. W tym etapie dzieci mogą rozwijać empatię wobec zwierząt i zrozumienie ich potrzeb. Badania sugerują, że osoby na wyższych poziomach rozwoju moralnego są bardziej skłonne do sprzeciwu wobec okrucieństwa wobec zwierząt i wspierania ochrony ich praw (Kohlberg, 1981). Co więcej, Kohlberg podkreślał znaczenie edukacji moralnej w kształtowaniu postaw, promując powstawanie programów edukacyjnych, które propagują empatię i zrozumienie dla zwierząt, co według niego, wspiera rozwój moralny dzieci i przyczynia się do redukcji uprzedzeń. Autor *Essays on Moral Development* podkreślał rolę edukacji w uwrażliwianiu na potrzeby i prawa zwierząt pozaludzkich.

Jak podkreśliłam wcześniej, uprzedzenia wobec ZINL nabywane są już w dzieciństwie oraz mogą wpływać na to, jakie informacje staną się kotwicą dla przetwarzania informacji w procesach rozumowania wykonywanych w dorosłym życiu. Badania wskazują, że otoczenie społeczne, w którym dziecko dorasta, ma kluczowy wpływ na formowanie postaw, które mogą utrzymywać się w dorosłym życiu (Harris, 1995). Rodzice i nauczyciele, jako główne wzorce, przekazują dzieciom normy i wartości, które mogą stać się punktami odniesienia dla przyszłych osądów dotyczących zwierząt pozaludzkich (Harris, 1995). W szczególności, ich negatywne przedstawienia powtarzające się w komunikatach do dzieci (np. w formie przysłów), wpływają na przyszłe osądy i postawy wobec zwierząt, a także prowadzą do ich dalszego utrwalania w dorosłym życiu (Tizard, Hughes, 1984). Dzieci, które są eksponowane na negatywne stereotypy, często używają tych informacji jako kryteriów do późniejszych ocen, co prowadzi do ich utrwalenia.

Innym źródłem kotwic poznawczych mogą być media, kultura, tradycje i literatura, które nierzadko umacniają negatywny obraz różnych gatunków już u najmłodszych odbiorców. Negatywne przedstawienia zwierząt pozaludzkich w bajkach, filmach, opowieściach takich, jak przekazy historyczne, a także stereotypy kulturowe, mogą prowadzić do utrwalenia negatywnych postaw i uprzedzeń wobec tych zwierząt. Kultura i tradycje, w jakiej żyje jednostka, określają, które zwierzęta są „wartościowe”, a które nie. Dobrym przykładem jest opisana wyżej różnica kulturowa w postrzeganiu psów i gęsi. Co więcej, jak wskazałam w części 3.1. *Historia badania umysłów zwierząt pozaludzkich*, w wielu kulturach pewne zwierzęta są historycznie postrzegane jako nieczyste lub przynoszące pecha, co wpływa na to, jak dzieci uczą się reagować na ich przedstawicieli. Przykładowo w kulturze zachodniej czarne koty często

są kojarzone z pechem, co determinuje negatywne postawy wobec nich (Puckett, 1991). Procesy przekazywania z pokolenia na pokolenie wzorców zachowań, wartości kształtujących postawy i osądy opisywany jest w ramach teorii kulturowych skryptów rozwiniętej przez Rogera Schanka i Roberta Abelsona (2013). Skrypty rozumiane jako specyficzne rodzaje struktur poznawczych, reprezentujących typowe sekwencje zdarzeń w określonych kontekstach. Na przykład, skrypt dla wizyty w restauracji obejmowałby etapy takie jak wchodzenie do restauracji, siadanie, zamawianie jedzenia, jedzenie, płacenie rachunku, wyjście. Co więcej, są to struktury o charakterze dynamicznym, które mają wyraźny porządek czasowy, czyli zawierają informacje o tym, jakie zdarzenia powinny mieć miejsce w określonych sytuacjach oraz w jakiej kolejności zdarzenia te następują. Skrypty opisują szczegółowe działania i interakcje, które są typowe dla danej sytuacji i kontekstu. Nabywane są w trakcie socjalizacji i mają silny wpływ na sposób postrzegania świata, w tym także funkcjonujących w nim zwierząt pozaludzkich. W kontekście uprzedzeń wobec ZINL kluczową rolę odgrywają skrypty kulturowe, które nabywane są poprzez interakcje z rodzicami, nauczycielami oraz mediami, przedstawiającymi określone zwierzęta w pozytywnym lub negatywnym świetle. Scenariusze te przekazywane są za pomocą języka i tradycji ustnych, czyli przez powtarzane opowieści czy przysłowia, których treść kształtuje reprezentacje pojedynczych gatunków zwierząt, kreując stereotypy i mity podzielane przez jednostkę. Na przykład, opowieści o lisie jako sprytnym oszuście mogą prowadzić do generalizacji i negatywnych postaw wobec przedstawicieli tego gatunku. Inny przykład może stanowić wierszyk Jana Brzechwy o treści:

*Dzik jest dziki, dzik jest zły,
Dzik ma bardzo ostre kły,
Kto zobaczy w lesie dzika,
Niech na drzewo szybko zmyka.*

(Wiersze dla Juniora: https://wiersze.juniora.pl/brzechwa/brzechwa_z07.html)

Teoria kulturowych skryptów pokazuje zatem, że przekazywane z pokolenia na pokolenie wzorce zachowań i wartości mają silny wpływ na nasze postawy i osądy, także te odnośnie do zwierząt innych niż ludzie. Kulturowe skrypty kształtują sposób,

w jaki postrzegamy i reagujemy na zwierzęta, co może prowadzić do trwałych uprzedzeń, które są trudne do zmienienia.

Ważną rolę w kształtowaniu skryptów pełnią również media, takie jak filmy, programy telewizyjne, literatura i inne formy przekazu przedstawiające ZINL w określony sposób, co wpływa na przyjmowane postawy społeczne. Źródła te dostarczają licznych kotwic w formie stereotypów czy narracji (często mitów), legitymizujących dyskryminację różnych gatunków zwierząt. George Gerbner i jego współpracownicy przedstawili teorię kulturowej akumulacji, która opisuje jak długotrwała ekspozycja na określone treści medialne, prowadzi do konwergencji postaw odbiorców, w tym często negatywnych przedstawień zwierząt jako niebezpiecznych lub złośliwych (Gerbner, 1998). Badania te pokazują, że negatywne przedstawienia zwierząt w bajkach i opowieściach mogą wpływać na sposób, w jaki dzieci postrzegają te zwierzęta w rzeczywistości. Na przykład, w klasycznych bajkach (np. *Trzy świnki*, *Czerwony Kapturek*) wilki często są przedstawiane jako groźne i złośliwe, co może prowadzić do utrwalenia negatywnych stereotypów już w najmłodszych latach jednostki (Bettelheim, 1976). Te stereotypy mogą utrzymywać się w dorosłości i wpływać na postrzeganie poszczególnych zwierząt w różnych kontekstach. Także wspomniany przy okazji opisu tego, jak wczesne kontakty interpersonalne dostarczają kotwic, do przetwarzania informacji w późniejszych etapach życia Albert Bandura (2001) twierdzi, że ludzie uczą się postaw przez naśladowanie modeli, które obserwują w mediach i literaturze.

Wymienione teorie nie są jednak jedynymi wspierającymi omawianą tezę o wpływie mediów i literatury na kształtowanie postaw. Teoria perswazji kulturowej głosi, że te mogą kształtować ludzkie postawy poprzez powtarzające się komunikaty i narracje, pełniące funkcję perswazyjną (Gerbner, Gross Morgan, Singnorielli, 1986). Stanowisko to wspierane jest przez model propagacji innowacyjnej (Rogersa, 2003). Pozwala on na analizę tego, jak nowe idee i zachowania promowane przez media przyswajane są przez jednostki żyjące w społeczeństwie. Badania pokazują, że różne formy przekazu mogą utrzymywać stereotypy składające się na postawy ich odbiorców (Entman, Rojecki, 2001). Co więcej, literatura wskazuje także, że przekazy medialne mogą pełnić funkcję prymy, a co za tym idzie, konstytuują one akceptowalne społecznie reakcje (Bargh, Pietromonaco, 1982). Dodatkowo jak wskazałam w kontekście wpływu kultury i tradycji, narracje i opowieści, także te kreowane przez liczne środki przekazu,

od najmłodszych lat kreują przekonania i wartości odbiorców (Green, Brock, 2000). Dla przykładu reprezentacja różnych grup społecznych w mediach i literaturze wpływa na postawy przyjmowane wobec przedstawicieli tych grup (Mastro, Tropp, 2004).

Innym źródłem kotwic są pierwsze wrażenia, które znacząco wpływają na późniejsze osądy względem np. innych agensów. Przykładem mechanizmu, który wykorzystuje kotwice w postaci pierwszego wrażenia, jest opisany w rozdziale II efekt halo. Badania Thorndike (1920) oraz Tversky'ego i Kahnemana (1974), pokazują, że ogólne wrażenie o jednym aspekcie zwierzęcia może wpływać na ocenę jego innych cech. Na przykład, jeśli ktoś uważa psy za przyjazne, może automatycznie zakładać, że wszystkie psy są inteligentne i lojalne, nawet jeśli konkretne dowody tego nie potwierdzają. Co więcej, konfiguracyjny model percepcji osób (*Configural Model of Person Perception*) stworzony przez Solomona Ascha (1946) wyjaśnia siłę pierwszego wrażenia poprzez odwołanie do efektu pierwszeństwa⁷⁸. W swoim artykule *Forming impressions of personality* wykazuje, że pierwsze informacje, jakie jednostka zdobędzie o określonym agensie, mogą wpływać na jej ogólne wrażenie dotyczące tego agensa.

Ostatnim mechanizmem poznawczym, jaki trzeba przytoczyć w kontekście omawianej heurystyki, jest efekt potwierdzenia (*confirmation bias*). Został on szeroko opisany w części 2.5 *Teoria dysonansu poznawczego*, gdzie omówiona została tendencja do wyszukiwania i interpretowania informacji w sposób, który potwierdza już istniejące w systemie poznawczym struktury, także te oparte na pierwszych wrażeniach i stereotypach (Nickerson, 1998). W kontekście zwierząt innych niż ludzie, jeżeli jednostki żyjące w społeczeństwie dzielają określone stereotypy lub mity dotyczące poszczególnych gatunków zwierząt, będą one zapamiętywać, interpretować lub odpamiętywać zdarzenia, potwierdzające te struktury poznawcze, nawet kosztem ignorowania albo odrzucania kontrargumentów. Co więcej, jak podkreśliłam wcześniej, uprzedzenia mogą być silniejsze, gdy wzmacnianie są przez selektywne przetwarzanie informacji i potwierdzanie istniejących przekonań (Nickersonem, 1998).

Podsumowując, badania dotyczące roli heurystyki zakotwiczenia w przetwarzaniu stereotypów, wskazują, że dzieci, które internalizują negatywne przekonania o zwierzętach pozaludzkich, mające swoje źródło w mediach, literaturze,

⁷⁸ Efekt pierwszeństwa (*primacy effect*) polega na łatwiejszym przywoływaniu informacji, które jednostka zapamiętała w pierwszej kolejności oraz ich wpływie na dalsze procesy przetwarzania i tworzenia ogólnych wrażeń i osądów (Nęcka, 2013; Anderson, 1971).

tradycjach, czy pierwszym wrażeniu, selektywnie zapamiętują te informacje, które potwierdzają przyjmowane przez nich postawy, co prowadzi do ich wtórnego wzmacniania.

3.4.2.3. Heurystyka dostępności

Heurystyka dostępności po raz pierwszy została opisana przez, wielokrotnie już wspomnianych w niniejszej dysertacji, Khanemana i Tversky'ego. Jak twierdzi pierwszy z wymienionych, rozpoczęli badania nad omawianym zagadnieniem od próby odpowiedzi na pytania takie jak: *w jaki sposób ludzie szacują wielkość określonych kategorii poznawczych (np. ilość ludzi, którzy rozwodzą się po 60, lub niebezpieczne rośliny?)*. Wyniki doprowadziły do zdefiniowania heurystyki dostępności jako poleganie na przesłankach, łatwo dostępnych oraz ignorowanie tych, których wydobywanie wymaga wysiłku poznawczego (Necka, 2013, str. 555). Czyli mechanizmem leżącym u podstaw wskazanego procesu jest przywoływanie z pamięci egzemplarzy należących do określonej kategorii. Jak zaznaczyłam wcześniej, heurystyki są regułami, którymi posługuje się system pierwszy, natomiast drugi na bieżąco koryguje jego przewidywania. Jeżeli proces przywoływania jest łatwy i szybki, system 1 uznaje, że kategoria, w której znajdują się egzemplarze, jest obszerna, jeżeli jednak przypominanie wymaga wysiłku, a co za tym idzie, przebiega wolno, system 1 wnioskuję, że określona kategoria poznawcza nie jest liczna. Co ciekawe okazuje się, że samo oszacowanie wielkości nie wymaga przywołania nawet jednego egzemplarza, wystarczą szacunki systemu szybkiego dotyczące łatwości poznawczej zadania (Kahneman, 2012, str. 176).

Klasyczne już badanie z zakresu dostępności, przeprowadzone przez Tversky'ego i Kahnemana (1983), polegało na zadaniu respondentom pytania dotyczące częstotliwości występowania liter w słowach. Chociaż oryginalne badanie jest bardzo ciekawe, paradygmat użyty w eksperymencie, chciałabym wyjaśnić na podstawie jego polskojęzycznej wersji (Nęcka, 2013, str. 555). W tej wersji uczestnicy proszeni byli o oszacowanie częstotliwości występowania litery „K” jako przedostatniej litery oraz końcówki „SKI” w polskich nazwiskach. Pomimo faktu, że nazwiska z końcówką „SKI” stanowią podzbiór nazwisk, których przedostatnią literą jest „K”, to respondenci szacowali pierwszą kategorię jako znacznie liczniejszą od drugiego wymienionego zbioru. Można więc powiedzieć, że heurystyka dostępności jest

mechanizmem do wyszukiwania przykładów określonych kategorii, czyli gdy jednostka dysponuje ogólną kategorią (np. końcówka „SKI”) oraz wyszukuje w niej reprezentatywne dla kategorii egzemplarze. W tym sensie omawiana heurystyka polega na traktowaniu przykładów łatwych do zapamiętania jako bardziej reprezentatywnych niż są w rzeczywistości. Jeżeli przykłady tej kategorii przywoływane są z łatwością, pojawia się tendencja do przeszacowywania liczebności tej kategorii, natomiast gdy elementy przypominane są z trudnością, tendencja jest odwrotna, czyli liczebność kategorii jest niedoszacowana. Oznacza to, że tak jak w przypadku innych heurystyk podczas przetwarzania na drodze dostępności dochodzi do zastąpienia pytania — liczebność zmieniana jest na łatwość przywoływania określonych elementów.

Innym klasycznym już eksperymentem z zakresu badań nad heurystyką dostępności jest, przeprowadzone na początku lat 90, badanie zespołu pod kierownictwem Schwarza (1991). Polegało ono na poproszeniu uczestników o wymienienie przykładów określonych kategorii (zachowania asertywne), a następnie sprawdzano jak szybkie lub wolne przywoływanie, wpłynęło na szacunki dotyczące liczebności kategorii. Uczestnicy podzieleni zostali na dwie grupy. Osoby z gr. 1 proszono o przypomnienie sobie i wymienienie 6 sytuacji, w których zachowali się asertywnie, natomiast przedstawicieli gr. 2 o wymienienie aż 12 takich sytuacji. Po tym zadaniu wszyscy uczestnicy mieli ocenić własną asertywność. Ze względu na fakt, że większość ludzi ma problem z wymienieniem aż 12 elementów określonej kategorii, zakładano, że naturalne trudności w przywołaniu tak licznego zbioru, będą obniżać szacowania dotyczące własnej asertywności. Badanie potwierdziło założenie eksperymentatorów, gdyż okazało się, że liczba wymienionych przykładów wpływa na późniejszą ocenę własnej asertywności. Osoby, które przywoływały 12 wspomnień, oceniały siebie jako mniej asertywne niż ludzie z grupy 1. Co ciekawe, okazało się także, że najistotniejszym czynnikiem determinującym takie szacunki jest nie tyle liczba przywołanych egzemplarzy, ile łatwość tego procesu. Nawet jeżeli agens bardzo szybko przypomni sobie 4 lub 7 przykładów asertywnego zachowania, pozostałe 5 będzie wymagało więcej wysiłku, a co za tym idzie, płynność wykonywania zadania spadnie, a to spowoduje obniżenie szacunków. Co więcej, gdy następnie poproszono uczestników z gr. 2 o przywołanie 12 przykładów swojego braku asertywności, pod wpływem trudności przypominania, zmieniali zdanie, twierdząc, że nie są jednak tak nieasertywni, jak wydawało im się na początku. Zespół Schwarza zauważył jeszcze jedną ciekawą zależność. Okazało się, że heurystyka dostępności może zostać

zablokowana, poprzez wstępne nastawienie. Tym razem uczestnicy wykonywali zadanie przy akompaniamencie muzycznym. Przed jego wykonaniem uprzedzono respondentów o wskazanym elemencie, jednak osoby w jednej grupie poinformowano, że muzyka powinna ułatwiać im wykonywanie zadania, natomiast ludziom z grupy drugiej, że jej funkcją jest utrudnienie wykonywania polecenia eksperymentalnego. Wykazano, że uczestnikom, którym wyjaśniono brak płynności w przywoływaniu przykładów przez obecność muzyki, oceniali swoją asertywność niezależnie od szybkości przywoływanych elementów. Oznacza to, że gdy istnieje inne prawdopodobne wyjaśnienie braku płynności w wykonywaniu zadania, łatwość przywoływania przestaje wpływać na szacunki jednostki.

Zespół Schwartz (1998) odkrył również, że jeżeli osądy dotyczą tematów, które traktowane są przez jednostkę osobiście, ludzie są bardziej skłonni do kierowania się ilością przywoływanych przykładów niż płynnością przypominania. Co oznacza, że wpływ heurystyki dostępności na procesy pamięciowe i dokonywanie osądów jest znacznie mniejszy. W pierwszej części eksperymentu informowano uczestników, że badanie dotyczy chorób układu krążenia. Tutaj należy nadmienić, że respondenci byli dobrani w taki sposób, że tylko połowa z nich miała historię chorób serca w rodzinie — zakładano, że Ci uczestnicy potraktują zadanie bardziej osobiście. Następnie wszystkie osoby biorące udział w eksperymencie poproszono o przypomnienie sobie własnych rutynowych zachowań (w liczbie 3 albo 8), które mogą wpływać na układ krążenia. Jedna grupa osób odpowiadała na pytanie o zachowania podnoszące ryzyko, druga o te zmniejszające ryzyko wystąpienia chorób serca. Studenci z rodzin nieposiadających historii schorzeń układu krążenia w rodzinie podeszli do zadania swobodnie, kierując się heurystyką dostępności. Autor *Accessible Content and Accessibility Experiences* interpretuje w ten sposób wyniki, które pokazywały, że osoby, którym trudno było przypomnieć sobie 8 przykładów zachowań ryzykownych dla układu krążenia, uznawały się za stosunkowo bezpieczne. Natomiast osoby, które nie mogły sobie przypomnieć 8 zachowań bezpiecznych, uznawały, że znajdują się w grupie ryzyka. Jednak w grupie studentów posiadających w rodzinie historię chorób serca prawidłowość była odwrotna. Po przypomnieniu sobie 8 zachowań ryzykownych czuli się bardziej bezpiecznie oraz częściej deklarowali, że ocena ryzyka, której dokonali podczas badania, wpłynie na ich zachowanie po jego zakończeniu.

Dla lepszego zrozumienia omawianej tu uproszczonej reguły wnioskowania powołałam się także na przykład zaczerpnięty z książki Thaler i Sunsteina (2023, str.

48). Autorzy wskazują, że obywatele Stanów Zjednoczonych zapytani: *jakie przestępstwo jest bardziej popularne: zabójstwo z użyciem broni, czy samobójstwo z użyciem broni?*, wskazują pierwsze z wymienionych. Fakt, że ludzie częściej wskazują zabójstwo, jest związany z „medialnością” określonej grupy czynów — w mediach częściej mówi się o odebraniu komuś życia siłą, niż o zakończeniu własnego istnienia, co sprawia, że ludzie mają wrażenie, że w rzeczywistości zdarza się więcej zabójstw. Okazuje się, że w USA, gdy człowiek kupuje broń, w celu samoobrony lub obrony członków swojej rodziny przed potencjalnym napadem, istnieje większa szansa, że któryś członek rodziny użyje jej do odebrania sobie życia, niż że zginie w wyniku napadu z bronią palną.

Ciekawej perspektywy na heurystykę dostępności dostarczają badania z zakresu ekonomii behawioralnej. W swoje pracy Howard Kunreuther i Geoffrey Heal (2012) wskazują na rosnącą ilość strat w wyniku katastrof (np. naturalnych), autorzy wyjaśniają, że źródłem tego problemu jest niepodejmowanie odpowiednich środków ochronnych przez osoby objęte ryzykiem takich wydarzeń. Przedstawiają analizę danych wskazującą, że ludzie częściej ubezpieczają się od klęsk żywiołowych po przeżyciu takiej klęski, zarówno wtedy, gdy są jej bezpośrednimi ofiarami, a także, gdy ta nie dotknęła ich osobiście, jednak w przestrzeni publicznej jest o niej głośno. Autorzy artykułu z 2012 roku wskazują, że u źródła takiego schematu zachowania leży heurystyka dostępności, ponieważ ryzyko zaczyna im się wydawać realne, dopiero gdy są w stanie łatwo i szybko wydobyć jego przykład z pamięci trwałej o takim wydarzeniu (ponieważ jest to niedawne wydarzenie). Oznacza to, że omawiana uproszczona reguła wnioskowania ściśle związana jest ze zjawiskiem nazywanym efektem miejsca w szeregu (Glanzer, Cunitz, 1966). Występuje on podczas oceny prawdopodobieństwa zdarzeń oraz charakteryzuje się schematem, w którym kolejność ekspozycji na informacje, wpływa na stopień zapamiętania informacji, np. podczas czytania tekstu najlepiej zapamiętujemy początek i koniec, najgorzej środek. W przypadku gdy mowa jest o przywoływaniu informacji z pamięci trwałej, widoczne są trzy efekty: świeżości, pierwszeństwa oraz pewności wstecznej. Pierwszy z wymienionych określa tendencję do sprawniejszego (łatwiejszego, szybszego) przywoływania z pamięci elementów, które jednostka kodowała niedawno. Natomiast drugi, polega na sprawniejszym przywoływaniu informacji, które jednostka zapamiętywała w pierwszej kolejności. Ostatnim wymienionym zjawiskiem poznamy jest efekt pewności wstecznej, czyli tendencja do przecenienia własnych

zdolności przewidywania przyszłych zdarzeń. Dochodzi do niego, gdy jednostka zna już przebieg wydarzeń. Przykłady tego zjawiska możemy zaobserwować m.in. w medialnych wywiadach z ludźmi, np. znającymi osoby, które popełniły jakieś przestępstwo. Często usłyszeć można w nich zdania, takie jak np. „zawsze wiedziałem, że coś jest z nim nie tak”.

Efekt miejsca w szeregu nie jest jedynym zjawiskiem poznawczym uczestniczącym w procesie generowania osądów opartych na dostępności. Dlatego istotne jest, aby pamiętać, że żadna heurystyka nie działa autonomicznie, a wszystkie stanowią zintegrowany system reguł, w którym każdy mechanizm ściśle współpracuje z innymi. Dla przykładu Margaret Matlin (1994), w swoim podręczniku *Cognition* przywołuje przypadek lekarza, który nie zastosował terapii koniecznej do wyleczenia pacjenta (P1), ponieważ w jego pamięci wciąż świeży był przypadek innego pacjenta (P2), u którego stosowana była ta konkretna terapia oraz P2 nie przeżył leczenia. Autorka zaznacza, że śmierć została poniesiona pomimo zastosowania terapii, a nie w jej wyniku, jednak użyta terapia i śmierć, postrzeganie były w perspektywie relacji przyczynowo-skutkowej. Oznacza to, że w wyniku przetwarzania opartego na dostępności ekspert podjął decyzję o niewdrażeniu leczenia, ryzykując życie P1. Wydaje się, że w przytoczonym przykładzie przetwarzania lekarza wystąpiły także złudzenia nazywane w literaturze iluzoryczną korelacją i złudzeniem doświadczenia osobistego. Pierwsze z wymienionych polega na dostrzeganiu relacji skojarzeniowej łączącej dwa elementy sieci, tam, gdzie w rzeczywistości go nie ma. Jeżeli relacja ta jest silna, to jednostka będzie postrzegać te elementy jako często współwystępujące, co spowoduje dostrzeganie np. związków przyczynowo-skutkowych tam, gdzie nie zachodzą. Natomiast złudzenie osobistego doświadczenia polega na przedkładaniu własnych doświadczeń, nad bardziej wiarygodnymi statystykami. Jak wskazują dane medyczne, heurystyka dostępności powoduje, że dane dotyczące relacji między paleniem tytoniu a zachorowaniami na nowotwory przekonują ludzi, którzy w jakiś sposób doświadczyli negatywnych skutków zażywania tytoniu (Marchewka, 2016). Ludzie, którzy w swoim otoczeniu posiadają osoby wypalające po 2 paczki papierosów dziennie oraz nie mają problemów ze zdrowiem, przedłożą doświadczenie osobiste nad statystykami. Zdanie, które jest klasycznie przytaczane we wskazanym kontekście jako przykład językowego przejawu omawianego złudzenia to „mój dziadek palił 2 paczki dziennie przez 50 lat i dożył 90 lat” (Nęcka, 2013, str. 556). Tak silny wpływ doświadczenia osobistego na szacowanie dotyczące liczebności kategorii wynika

z faktu, że informacje zapisane w pamięci autobiograficznej wzmacnianie są wyraźnymi szczegółami (nie zawsze prawdziwymi (Nęcka, 2013, str. 372-418)) i emocjami, co sprawia, że są łatwiej dostępne poznawczo, niż informacje płynące ze statystyki. Jak zaznacza autor jednego z najpopularniejszych podręczników z zakresu psychologii poznawczej, to właśnie złudzenie doświadczenia osobistego ułatwia powstawanie uprzedzeń (Nęcka, 2013, str. 556). Powołuje się on na przykład, w którym londyńczyk wychodzący rano do pracy, zobaczy kilka razy Polaków spożywających alkohol, jego doświadczenia będą wzmacniać przekonania, o tym, że „wszyscy Polacy piją alkohol od rana”. Co ciekawe, tak jak w przypadku heurystyki zakotwiczenia i dopasowania, w przypadku przetwarzania opartego na dostępności, istotną rolę odgrywają media. Informacja przekazywana odbiorcom, kreowana jest tak, aby Ci traktowali je bardzo osobiście, co sprawia, że przekaz ten jest istotnym czynnikiem determinującym decyzje np. polityczne czy konsumenckie. Aronson i Aronson w *Człowiek istota społeczna* (2007) zaznaczają, że u podstaw pojawiania się wszystkich opisanych w tej części błędów poznawczych, jest fakt, że na dostępność składa się wiele czynników np. znajomość przywoływanych przykładów, ich obrazowość (wyrazistość/szczegółowość), odległość czasową przypominanych zdarzeń (szybkość wydobywania) oraz łatwość w wyobrażeniu przykładów, oraz ich możliwości kojarzeniowe, co sprawia, że najłatwiej dostępne egzemplarze często nie są typowe dla określonej kategorii.

Innym mechanizmem poznawczym związanym z przetwarzaniem opartym na dostępności są efekty oczekiwań (opisane w części 2.4. *Efekty oczekiwań*, (Kahneman, 2012, str. 182)). W odniesieniu do opisywanego wcześniej eksperymentu zespołu Schwartza (1991), w którym uczestnicy wymieniali przykłady własnych zachowań asertywnych i nieasertywnych. Ci przywołujący aż 12 egzemplarzy doświadczali naturalnego spadku płynności wykonywania zadania oraz na jego podstawie wyciągali wnioski dotyczące poziomu swojej asertywności. Autor *Pulapek myślenia. O myśleniu szybkim i wolnym* (Kahneman, 2012) zauważa, że respondenci wykazywali oczekiwania dotyczące płynności wykonywania zadania, w wyniku czego spadek sprawności wymieniania kolejnych przykładów zachowań, obserwowali u siebie szybciej, niż się spodziewali. To właśnie element zaskoczenia sprawia, że osoby wyliczające 12 elementów uznają się za mało asertywne, jednak, gdy szybko znajdzie się wyjaśnienie takiej sytuacji (obecność muzyki), niska płynność przestaje wpływać na szacunki. Za słuszością takich twierdzeń przemawiają badania donoszące o zdolności systemu 1 do generowania oczekiwań i wywoływania afektu zdziwienia (i nie tylko, innym

przykładem jest słodka reakcja opisana szerzej w części 2.6. *Słodka reakcja*), gdy te oczekiwania nie zostają spełnione. Należy także pamiętać, że niezgodność stanu rzeczywistego z oczekiwaniami jednostki generuje dyskomfort, który jest dodatkowym czynnikiem motywującym, do przywrócenia spójności w systemie-organizmie (czyli homeostazy w rozumieniu Damasio (1994)). Dlatego, system pierwszy stara się wyjaśnić zaskoczenie przez szybkie wygenerowanie jego prawdopodobnych przyczyn, w wyniku czego, na drodze heurystyki dostępności przywołuje informację podaną przed eksperymentem lub wyciąga wniosek, o poziomie asertywności jednostki.

Heurystyka dostępności ściśle związana jest także z opisaną w części 3.4.2 *Heurystyki i wynikające z nich tendencje w zachowaniu i podejmowaniu decyzji istotne w kontekście uprzedzeń międzygatunkowych*, tendencją do wnioskowania w oparciu o afekt, polegającej, na kierowaniu się preferencjami m.in. takimi jak: lubię — nie lubię, podoba mi się — nie podoba mi się. Czyli jak wskazałam wcześniej, heurystyka ta sprawia, że przekonania/opinie/postawy zależne są od chęci czy niechęci agensa. W wyniku współpracy wskazanej heurystyki z generowaniem osądów na podstawie dostępności, jeżeli jakiś obiekt jest lubiany przez jednostkę (wywołuje w niej pozytywne afekty), to łatwo dostępne będą tylko pozytywne atrybuty (zalety) tego obiektu, w wyniku czego te postrzegane będą jako nieobarczone wadami. Jest także odwrotnie, obiekty, które wywołują w jednostce negatywne afekty, będą postrzegane jako pozbawione zalet, natomiast argumenty dotyczące związanego z nim ryzyka, będą łatwo dostępne.

Zależność ta wykazana została przez zespół Slovic (2007), który przeprowadził badanie opinii respondentów dotyczącej różnych technologii (np. fluoryzowanie wody pitnej, produkcja chemikaliów, konserwanty w żywności). Następnie uczestnicy proszeni byli o wymienienie przykładów potencjalnego ryzyka i korzyści płynących z poszczególnych technologii. Okazało się, że osoby pozytywnie oceniające określone technologie przywoływały wiele przykładów korzyści, ale niewiele skutków związanych z potencjalnym ryzykiem. Natomiast gdy jednostka oceniała określoną technologię jako niebezpieczną, sytuacja wyglądała odwrotnie — przywoływała liczne przykłady ryzyka oraz niewiele korzyści płynących z określonej technologii. W następnym etapie wszyscy uczestnicy otrzymywali teksty, niektóre z nich przedstawiały zalety inne wady określonych technologii. Osoby, które wcześniej krytycznie podchodziły do określonych rozwiązań, zapoznawały się z tekstami przedstawiającymi ich zalety. Po tym, gdy respondenci zapoznali się

z kontrargumentami, ponownie sprawdzono opinię wszystkich uczestników. Wyniki wykazały, że osoby, które w pierwszej części oceniały określone technologie jako niebezpieczne, zmieniały swoje stanowisko odnośnie do ryzyka z nim związanego, pomimo że na ten temat nie otrzymały żadnych dodatkowych informacji. Zespół Slovic (2007) zauważył także, że istnieje silny związek między łatwością, z jaką przypominane są poszczególne czynniki ryzyka a emocjonalnymi reakcjami na nie, przy czym szczególnie łatwo przypomnieć sobie wspomnienia czy informacje budzące lęk. Badanie przeprowadzone zostało przy użyciu kwestionariusza, w którym wypisano pary przyczyn zgonów np. udar mózgu – wypadek samochodowy, zadaniem respondentów była ocena, który element każdej pary zdarza się częściej oraz jak często dochodzi do każdego z nich. Następnie eksperymetatorzy odnieśli szacowane przez uczestników liczby do aktualnych statystyk. Pomimo że udary mózgu statystycznie wywołują dwa razy więcej zgonów, niż wypadki komunikacyjne razem wzięte, to 80% respondentów uznała za bardziej prawdopodobny ten rodzaj śmierci. Ogólny konsensus dotyczący takich wyników wskazuje, że postawy przyjmowane przez ludzi są kreowane przez przekazy medialne. Jak wykazałam wcześniej, informacje te są zniekształcone w taki sposób, aby zwiększyć oglądalność, dlatego np. przekazywane są głównie treści o silnym zabarwieniu emocjonalnym.

Wszystkie doniesienia badawcze zespołu pod przewodnictwem Slovic, znajdują odniesienie w opisaną wcześniej teorii Antonio Damasio (1994) oraz Ronalda Sousa (1979), tym samym uznaje się, że dostarczają one empirycznego zaplecza wskazanej koncepcji. A także, stanowią potwierdzenie tezy głoszącej, że afektywna ocena potencjalnych konsekwencji decyzji oraz związane z tym bodźce wewnętrzne (tzn. stany cielesne (Kahneman, 2012, str. 189)) i tendencje odgrywają główną rolę w procesie generowania osądów oraz podejmowania decyzji. Tak jak wszystkie inne heurystyki, rozumowanie oparte na dostępności jest skutecznym schematem rozumowania w rozwiązywaniu niektórych prostych/codziennych problemów. Przykładem są szacunki prawdopodobieństwa wystąpienia opadów w określonym terminie, w których informacja o tym, kiedy ostatnio padało, jest ważną przesłanką, ponieważ warunki pogodowe mają tendencję do utrzymywania się (OpenStax, 2022). Zatem omawiana heurystyka jest efektywna podczas rozwiązywania problemów związanych z powtarzającymi się zjawiskami, co wynika z faktu, że wykorzystuje zapisane w pamięci trwałe schematy (związane z behawiorem) oraz narracje (komponent poznawczy), próbując je zaimplementować do rozwiązywanego problemu. Łatwość

przywoływania powtarzających się wzorców sytuacji, przydatna jest nie tylko w szacunkach dotyczących nadchodzących warunków pogodowych. Jak wskazuje autor książki *Mit w umyśle. Ewolucyjno-kognitywne podstawy form mitycznych* (Czeremski, 2021, str. 14), często zdarza się, że zachowania wynikające z nietrafnych, a czasami wręcz bezsensownych przekonań, są bardziej adaptacyjne, niż te oparte na racjonalnych regułach wnioskowania. Dzieje się tak, ponieważ z perspektywy ewolucyjnej ocenie użyteczności przystosowawczej podlegają zachowania, a nie koncepcja czy wyjaśnienia sytuacji. To sprawia, że wyjaśnienia (narracje) zbudowane na uproszczonych regułach wnioskowania, generują trafniejsze modele mentalne — przedstawiane przez nie instrukcje postępowania, są prostsze i bardziej przejrzyste niż wyjaśnienia oparte na bardziej wiarygodnych przesłankach. Jak już wielokrotnie wspominałam w niniejszej dysertacji, wnioskowania oparte na regułach logicznego rozumowania w wysokich warunkach niepewności mogą być niemożliwe do przetworzenia. Wyjaśnia się to w odniesieniu do paradygmatu skąpca poznawczego, w ramach którego system poznawczy nie ma możliwości obliczeniowych do ujęcia w swoich szacunkach wielu możliwych modeli rzeczywistości — ze względu na brak danych, nie jest w stanie ich trafnie skonstruować.

Czyli innymi słowy, szacunki oparte na łatwo dostępnych (poznawczo) danych są skuteczne, gdy informacje takie jak wcześniejsze doświadczenia, ładunek emocjonalny lub kolejność odbioru informacji są adekwatne do problemu. Co więcej, należy wskazać na szczególną efektywność heurystyki dostępności w sytuacjach, gdy konieczne jest szybkie podjęcie decyzji, czyli m.in. w sytuacjach kryzysowych, w których czas na analizę możliwych do zastosowania modeli mentalnych jest znacznie ograniczony. W takich przypadkach np. łatwość przywołania wcześniejszych doświadczeń może prowadzić do wykonania odpowiedniej reakcji, zwłaszcza jeśli podobne okoliczności miały miejsce wcześniej i zakończyły się pozytywnym/neutralnym skutkiem. Omawiana heurystyka sprawdza się również w zadaniach rutynowych – jeżeli osoba wielokrotnie korzystała z określonego rozwiązania z korzystnym efektem, łatwość przypomnienia sobie tej strategii może pomóc w szybkim rozstrzygnięciu problemu. Ponadto, omawiana heurystyka pełni kluczową funkcję w bardziej codziennych decyzjach, np. tych konsumenckich. Konsument może łatwo przywołać pozytywne doświadczenia związane z daną marką i na tej podstawie dokonać trafnego wyboru. Skuteczność tej heurystyki zwiększa się również, gdy decyzje podejmowane są w dobrze znanych kontekstach. W takich

przypadkach osoby z dużym doświadczeniem mogą szybko przypomnieć sobie odpowiednie informacje i z sukcesem rozwiązać problem – jak np. lekarz, który szybko diagnozuje chorobę na podstawie podobnych przypadków z przeszłości.

Powyżej zaznaczyłam już, że heurystyka dostępności pełni kluczową funkcję w kształtowaniu prezentowanych przez jednostkę postaw oraz wykazałam, w jaki sposób przekaz medialny wpływa na generowanie osądów opartych na dostępności. Wspomniany już Napiórkowski wskazuje, że jednym z mechanizmów kluczowych dla generowania wyjaśnienia sytuacji jest myślenie analogiczne, opisane szerzej w części 2.3. *Myślenie analogiczne*. W swoim wykładzie *Mity polityczne wczoraj i dziś* wygłoszonym w 2018 roku koncentruje się przede wszystkim na analogii historycznej, co związane jest z podjęciem przez niego tematu perswazji w celach politycznych. Abstrahując od najważniejszych tematów podjętych przez tego autora (choć bardzo ciekawe, nie mieszczą się w obszarze zainteresowań niniejszej dysertacji). W tym miejscu skupię się jednak na tym, jak heurystyka dostępności wykorzystywana jest do kreowania postaw społecznych obywateli. W literaturze znaleźć można wiele opracowań socjotechnik⁷⁹, które wykorzystują generowanie osądów w oparciu o dostępność, aby skłonić przedstawicieli społeczeństwa np. do przeszacowywania ryzyka związanego z potencjalnymi działaniami rządu. Jedną z kluczowych strategii uzyskania wskazanego celu, jest nagłośnienie podobnych przypadków działań, które zakończyły się niepowodzeniem, a ich konsekwencje były silnie negatywne. A także przeciwnie, dobrym sposobem zwiększenia poczucia bezpieczeństwa wobec określonego rozwiązania (politycznego/ gospodarczego/ ekonomicznego /społecznego), jest rozpowszechnienie informacji dotyczących innego podobnego działania, którego zastosowanie przyniosło pozytywne dla obywateli konsekwencje. Jak wskazuje Mateusz Bartosiewicz (2017), istnieje wiele czynników facylitujących większe prawdopodobieństwo zastosowania wniosku opartego na dostępności, są to:

- a. Posiadanie władzy,
- b. Wysoki wskaźnik zaufania intuicji,

⁷⁹ Socjotechnika może być definiowana na dwa sposoby: wąski i szeroki. W pierwszym przypadku rozumiana jest jako mechanizm kształtowania postaw i wpływania na opinie społeczne, jednak jak wskazuje Ziółkowski (2016), definicja ta nie pozwala na odróżnienie socjotechniki od propagandy czy marketingu politycznego. W sensie szerokim rozumiemy socjotechnikę jako zbiór dyrektyw skutecznego rządzenia, czyli sterowania systemem społecznym.

- c. Wykonywanie, w tym samym czasie innego zadania, wymagającego wysilenia poznawczego.

Rola wszystkich wskazanych czynników w przetwarzaniu informacji przez system pierwszy została opisana we wcześniejszych częściach dysertacji. Co ciekawe, okazuje się, że opieranie się na intuicji nie musi być cechą osobowości, wystarczy przypomnieć jednostce okres, w którym posiadała władzę, aby jej zaufanie do własnej intuicji się zwiększyło (Ana Guinote, 2003). Fakt ten pokazuje, że między wyżej wymienionymi czynnikami zwiększającymi prawdopodobieństwo użycia heurystyki dostępności zachodzi korelacja, co oznacza, że nie należy analizować ich niezależnie od siebie.

Innym wartym odnotowania przykładem wpływu wnioskowania w oparciu o dostępność wynikającą z afektu, na przyjmowanie postaw, jest stanowisko obywatelskie przejawiające się sprzeciwem (komponent behawioralny) wobec budowy elektrowni jądrowych (Eko-Unia.pl, 2023). Jej komponent poznawczy budują narracje sugerujące, że istnieje duże ryzyko awarii takiej elektrowni. Natomiast jej komponent afektywny stanowi trwały lęk wywołany przez potencjał skojarzeniowy kategorii poznawczej „elektrownie jądrowe”, związany z łatwo dostępnymi egzemplarzami wspomnianej kategorii, takimi jak Czarnobyl czy Fukushima. Skala katastrof wymienionych elektrowni nadal budzi silne afekty w obywatelach państw wschodu (Polska, Białoruś, Ukraina, Rosja), co związane jest z łatwą dostępnością konsekwencji wydarzeń z Prypeci i Fukushima.

Zatem heurystyka dostępności ściśle związana jest z prezentowaniem uprzedzeń, także tych wobec ZINL. Jak już wspominałam wcześniej, kategoria „zwierzęta pozaludzkie” jest ogólnym zbiorem złożonym z wielu podkategorii (np. gatunki). Rozpatrywanie roli wnioskowania opartego na dostępności w kreowaniu tak złożonych postaw, jak uprzedzenia, wymaga m.in. określenia językowego obrazu wybranych kategorii zwierząt (gatunki/ rola (domowe/hodowlane)) oraz analizy przekazów społecznych (medialnych, literackich, oralnych) dotyczących tych gatunków, a następnie dokładnego opisu wszystkich komponentów postaw, których przejawy widoczne są w tak powstałej analizie. Chociaż rozważania te wydają się bardzo ciekawe, wykraczają poza ramy tematycznie niniejszej pracy, Literatura pełna jest jednak podobnych opracowań, co więcej, pośrednio temat ten poruszałam kilkakrotnie w niniejszej dysertacji, przede wszystkim, gdy wspominałam o postawach, jakie ludzie żywią wobec wilków (Woźniak, 2016; Grabowska, 2011). Dla przypomnienia Henryk Okrama zabiera głos w kwestii polityki leśnej, a konkretnie wypowiada się na temat

problemu społecznego, jakim są coraz częstsze doniesienia nie tylko o obecności wilków na terenach wiejskich i miejskich, ale też o atakach tych zwierząt na ludzi. Zwolennicy Okramy, uważają, że zachowanie w postaci podchodzenia coraz bliżej ludzkich zabudowań, jest spowodowane zbyt dużą liczebnością populacji tego gatunku. Proponowanym przez nich rozwiązaniem wskazanego problemu jest redukcja populacji wilków poprzez kontrolowany odstrzał⁸⁰. Czyli można powiedzieć, że komponentem behawioralnym opisywanej postawy jest tendencja do podejmowanie zachowań społecznych (np. głosowanie w wyborach na odpowiednich kandydatów), mających na celu redukcję populacji wilków. Jej wyraźnym komponentem afektywnym jest lęk spowodowany m.in. przekazem medialnym informującym o obecności wilków w przestrzeni miejskiej i o ich agresywności wobec ludzi. Należy także pamiętać, że na skutek efektu potwierdzenia, nawet gdy jednostka prezentująca opisywaną postawę, znajdzie informację (np. w zaproponowanym przez media społecznościowe poście), sprostowującą przekaz o ostatnim ataku wilka na człowieka, z którego wynika, że agresywne zwierzę, nie było wilkiem, a psem (często mylonej z wilkiem rasy wilczak). Co więcej, wskazany efekt, będzie wywoływał tendencję do niechęci w przyjmowaniu alternatywnych wyjaśnień zachowań wilków np. tych odwołujących się do sytuacji ekologicznej przedstawicieli tego gatunku: rosnąca populacja, wylesianie terenów, trudności w zdobyciu pożywienia. Ignorowanie wymienionych wyżej czynników leżących u podstaw zachowań grupy obcej, wynika z akcentuacji, homogenizacji i innych mechanizmów charakterystycznych dla uprzedzeń mechanizmów. Skutkują one także bagatelizowaniem merytorycznych uwag dotyczących tego, jak człowiek powinien się zachowywać wobec wilków, aby zminimalizować ryzyko ataku. Co więcej, odwołując się do przykładu z wcześniejszej części, uprzedzenia rodziców mogą być przejmowane przez dzieci. Jeżeli rodzice i osoby z bliskiego otoczenia dziecka (dziadkowie/rówieśnicy/przyjaciele) prezentują postawę, której komponentem

⁸⁰ Badania wykazują, że wilki potrafią dostosowywać rozmnażanie do warunków środowiskowych, szczególnie do dostępności pożywienia. Klasyczne już analizy Messiera i Crête'a (1985) potwierdziły, że przy niskiej gęstości ofiar, takich jak łosie, tempo rozrodu wilków ulega ograniczeniu, co prowadzi do naturalnego spadku liczebności populacji. Nowsze badania Wikenrosa i współpracowników (2021) wykazały natomiast, że gęstość populacji oraz warunki bytowe wpływają na wiek, w którym osobniki osiągają dojrzałość rozrodczą – szczególnie u samic – co również stanowi mechanizm regulujący dynamikę liczebności gatunku. Oznacza to, że populacje wilków nie wymagają sztucznej kontroli poprzez odstrzał, ponieważ ich liczebność jest ściśle uzależniona od warunków ekosystemowych.

poznawczym są mity o wilku jako *dzikiej bestii*, głoszące, że wilki są niebezpieczne, bo przestały bać się człowieka. Należy się przed nimi bronić, ponieważ wchodzi na „nasze” terytoria i atakują przedstawicieli naszego gatunku (mieszkańców wsi, obrzeży miast lub osoby spacerujące w lasach). Bez względu na to, czy narracje te są prawdziwe, ich silne nacechowanie emocjonalne (heurystyka afektu) będzie powodowało, że przypadki ataków wilków na ludzi będą łatwiej dostępne i chętniej nagłaśniane przez media. Jak pokazuje wiele analiz dotyczących uprzedzeń, u ich podstaw leży niedostateczna wiedza o dzikich gatunkach (Orosz, Krekó Paskuj, Tóth-Király, Bóthe, Roland-Lévy, 2016). Takie deficyty wynikają nie tylko z braku bezpośrednich doświadczeń z przedstawicielami tych gatunków, ale też narracji przyjmowanych w określonej społeczności (zawartej np. w tradycyjnych przekazach, dziełach kulturowych, czy przekazach medialnych). Ponadto, efekt potwierdzenia, sprawiający, że informacje zgodne ze wskazanymi strukturami poznawczymi są lepiej zapamiętywane, niż te z nimi niezgodne, czyli jest on mechanizmem, na którego drodze przekonania, stereotypy i afekty składające się na tendencję do podejmowania zachowań mających na celu redukcję populacji wilka, są stale wzmacniane. Bodźcami wzmacniającymi tę postawę mogą być doświadczenia osobiste nabyte podczas odbioru przekazu medialnego (Bartosiewicz, 2017). Media społecznościowe/telewizja/internet, kształtują przekaz tak, aby odbierany był osobiście, czyli aby odbiorca odczuwał, że komunikowane mu treści dotyczą go w jakiś sposób osobiście. Oczywiście taki przekaz nie zawsze przystaje do faktycznych warunków i zmian ekologicznych wilków. Innym mechanizmem wpływającym na powstawanie nieadekwatnych do rzeczywistości reprezentacji, jest myślenie antropomorficzne. Jak argumentowałam w pierwszym rozdziale niniejszej dysertacji, wspominając o wskazanej postawie, ludzie z większą trafnością przypisują przekonania oraz doświadczenia emocjonalne innym ludziom, niż zwierzętom pozaludzkim. Atrybucja stanów mentalnych przedstawicielom innych gatunków oparta jest często na myśleniu analogicznym, co skutkuje rozumowaniem opartym o projekcje dotyczące potrzeb i innych stanów mentalnych, nawet najbliższych jednostce ZINL: zwierząt domowych. Empatyzowanie i przyjmowanie perspektywy innego jest prostsze w relacji człowiek – człowiek, jednak nie jest tak powszechne w relacji człowiek – ZINL. Nieznajomość sygnałów komunikacyjnych zwierząt pozaludzkich sprawia, że ludzie mogą mieć problemy z trafnym rozpoznaniem stanu mentalnego, co zakłóci mechanizm empatii oraz uniemożliwi trafne przyjęcie perspektywy innego. Wracając do przykładu wilka ludzie, którzy odczuwają silny lęk

wobec przedstawicieli wskazanego gatunku, będą wykazywać tendencję do ignorowania informacji dotyczących dobrostanu tych zwierząt, gdyż postrzegają ich w kategoriach *dzikiej bestii*. Przekaz medialny, nastawiony na przedstawienie informacji w formie takiej, aby jednostka odebrała je personalnie, czyli facylitowanie przetwarzania opartego na afekcie, sprawia, że ataki wilków na ludzi są łatwiej dostępne poznawczo niż pozytywne, codzienne interakcje⁸¹, co sprawia, że jednostki żyjące w społeczeństwie będą cechować się uprzedzeniami negatywnymi takimi jak szowinizm gatunkowy czy karnizm.

3.4.2.4. Heurystyka reprezentatywności

Heurystyka reprezentatywności jest uproszczoną regułą rozumowania, w ramach której odbywa się szybka kwalifikacja obiektów dostępnych poznawczo, do przynajmniej jednej kategorii (K). Wskazana klasyfikacja odbywa się na podstawie częściowego podobieństwa (np. morfologiczne) poznawanego obiektu (O) do przypadków typowych dla (K); (Baker, 2010; Polański, 2013). Omawiana reguła angażowana jest podczas wykonywania wszystkich procesów identyfikacji obiektów oraz szacunków prawdopodobieństwa pozwalających odpowiedzieć na pytania dotyczące tego: *w jakim stopniu A jest reprezentatywne dla B?* (Kahneman, 2012; Mękarska, 2024, str. 115). Przykładami takich problemów jest szacowanie prawdopodobieństwa, że:

- a. Obiekt A należy do klasy B?;
- b. Zdarzenie A bierze początek z procesu B?;
- c. Proces B wygeneruje zdarzenie A?

Udzielanie odpowiedzi na powyższe pytania odbywa się na drodze procesów bardziej szczegółowo opisanych w częściach 2.1. *Mechanizm kategoryzacji społecznej*

⁸¹ Piszac o „codziennych interakcjach”, mam na myśli interakcje, z których ludzie często nie zdają sobie sprawy. Należy pamiętać, że zmysłem dominującym wilka jest węch, a co za tym idzie, przedstawiciele tego gatunku czują człowieka znacznie szybciej, niż ten może go zobaczyć. Wilki nie są skore do atakowania człowieka — takie stanowisko wynika z obserwowanych zmian tras migracji, których źródłem jest ślad obecności zapachu człowieka na szlaku watahy. Co ciekawe, zwierzęta żyjące na określonym terenie faktycznie przyzwyczajają się do zapachu niektórych ludzi (tak, że nie wywołuje w nich lęku). Są to przykłady leśniczych, którzy na co dzień rejestrują życie leśnych zwierząt (przykładem może być dziennikarz (Leśny Kawaler, link do kanału na platformie youtube: <https://www.youtube.com/c/Le%C5%BnyKawaler>), który na co dzień rejestruje życie leśnych zwierząt za pomocą wielu kamer zamontowanych w lasach).

i 2.3 *Myślenie analogiczne*. Jednym z najważniejszych elementów procesu kategoryzacji w kontekście heurystyki reprezentatywności jest generalizacja, która pozwala na atrybucję potencjalnych cech (O), wynikających z przynależności do (K). Myślenie analogiczne polegające na uproszczeniu relacji analogii do relacji podobieństwa pozwala na porównanie (O) z innymi egzemplarzami poszczególnych kategorii. Jak wspomniałam we wskazanej części, pojęcie prawdopodobieństwa jest powszechnie nieostre, szczególnie dla ludzi, którzy nie mają w swoim życiu doświadczeń z kursami ze statystyki, co sprawia, że pytanie o prawdopodobieństwo zastąpione zostaje pytaniem o reprezentatywność, czyli podobieństwo do egzemplarza danej (K); (Tversky, Kahneman, 1974; Bilek, 2018; Polański, 2013; Grether, 1992). Przyporządkowanie obiektu np. jednostki społecznej do określonej kategorii poznawczej, odbywa się w ramach kryteriów nazywanych stereotypami (Garcia-Marques, Mackie, 1999). Oznacza to, że w przypadku heurystyki reprezentatywności mamy do czynienia z zastąpieniem pytania właściwego: *jakie jest prawdopodobieństwo, że (O) należy do (K)?*, pytaniem heurystycznym o treści: *w jakim stopniu (O) jest reprezentatywne dla (K)?*. Dlatego w literaturze przetwarzanie na drodze heurystyki reprezentatywności nazywane jest często opartym na stereotypie lub reprezentatywności (Grether, 1992; Polański, 2013; Kahneman, 2012).

W ramach omawianego rozumowania ludzie szybciej i skuteczniej przypominają sobie informacje zgodne z kryteriami przyporządkowywania obiektów do kategorii. Jak w przypadku innych opisywanych heurystyk, badania pokazują, że zmęczenie (deficyt zasobów), a co za tym idzie, obniżenie kondycji poznawczej jest główną przyczyną korzystania z myślenia heurystycznego opartego o stereotyp (Wojciszke, 2011, str. 89-92; Polański, 2013, str. 3). Innym ważnym czynnikiem sprzyjającym zastosowaniu omawianej uproszczonej reguły rozumowania, jest opisany wyżej paradygmat skąpca poznawczego. Doprecyzowując, gdy przetwarzana sytuacja, wymaga przeanalizowania zbyt dużej ilości danych, aby możliwe było zastosowanie reguł logicznych, jednostka przetworzy informacje na drodze heurystycznej. Kolejnym czynnikiem wymuszającym rozumowanie oparte na reprezentatywności, jest deficyt wiedzy, czyli sytuacja, w której jednostka nie posiada istotnych przesłanek potrzebnych do przeprowadzenia rozumowania logicznego. Co więcej, przypominając wielokrotnie już podkreślaną w niniejszej dysertacji tezę, głoszącą, że heurystyki funkcjonują raczej, jako zintegrowany zestaw reguł rozumowania, dlatego zastosowanie np. zakotwiczenia i dostosowania sprawia, że np. procesy podejmowania decyzji oparte na określonych

kotwicach, będą prymować korzystanie z konkretnych stereotypów. Fakt ten pokazuje jak trudno miarodajnie określić stopień, w jakim poszczególna heurystyka wpływa na cały proces.

Literatura z zakresu reprezentatywności jest bardzo bogata, a temat ten podejmowany jest przez specjalistów różnych dziedzin, jednak istnieją dwa eksperymenty, które zasługują na szczegółowy opis w niniejszej części, gdyż odniesienie do nich znajduje się w przynajmniej większości prac omawianego obszaru. Chodzi o tzw. eksperymenty z Tomaszem W. i Lindą przeprowadzone przez Kahnemana i Tversky'ego. Pierwszy z wymienionych składał się z 3 zadań. Na początku uczestnicy otrzymywali informację, że Tomasz W.⁸² realizuje studia II stopnia na dużej uczelni, a także listę kierunków studiów przez nią oferowanych. Zadaniem respondentów było uszeregowanie pozycji na liście według prawdopodobieństwa (1-najbardziej prawdopodobne, 9- najmniej prawdopodobne), że fikcyjny Tomasz W. realizuje studia określonego kierunku. Wyniki eksperymentu wykazały, że we wskazanym zadaniu uczestnicy szacowali popularność poszczególnych studiów (określali wartość bazową), czyli ustalali jaki procent wszystkich magistrów stanowią studenci poszczególnych kierunków oraz na podstawie tych założeń formułowali odpowiedź. W ramach zadania drugiego respondenci otrzymywali szczegółowy opis z informacją, że został on przygotowany przez profesjonalnego psychologa, kiedy Tomasz W. chodził do ostatniej klasy liceum:

Tomasz W. jest osobą o wysokiej inteligencji, choć brakuje mu autentycznej kreatywności. Odczuwa potrzebę porządku i przejrzystości oraz funkcjonowania w uporządkowanych systemach, w których każdy szczegół jest na swoim miejscu. Píše dość nudnym i mechanicznym stylem, który miejscami ożywiają wysilone żarty językowe i przebłyśki wyobraźni w stylu *science fiction*. Cechuje się silnym dążeniem do kompetencji. Nie wydaje się odczuwać większego zainteresowania innymi ani współczucia dla nich, nie lubi interakcji międzyludzkich. Jest skupiony na sobie, ale ma głębokie poczucie moralności (Kahneman, 2012 str. 199).

Jak w przypadku wcześniejszego eksperymentu, uczestnicy otrzymywali listę kierunków studiów, a ich zadaniem było wskazanie najbardziej i najmniej

⁸² W rzeczywistości Tomasz W. jest postacią fikcyjną stworzoną przez Kahnemana z myślą o reprezentacji typowego studenta informatyki.

prawdopodobnych pozycji na liście (1-najbardziej prawdopodobne, 9- najmniej prawdopodobne). Jednak w tym przypadku zadanie wymagało ze skorzystania z istniejących kryteriów kategoryzacji lub wytworzenia stereotypów magistrantów poszczególnych kierunków znajdujących się na liście. Co pozwalało na porównanie opisywanej postaci do innych egzemplarzy kategorii oraz na tej podstawie określenie jej reprezentatywności względem niej. Opis Tomasza W. celowo był reprezentatywny dla kategorii studentów kierunków ścisłych, nie zaś humanistycznych (czyli najbardziej licznych). Wyniki pokazały, że w tym zadaniu respondenci dokonywali szacunków opartych o stereotyp, w wyniku czego nie uwzględnili istotnych danych – wspomnianych wyżej wartości bazowych.

W ramach ostatniego, trzeciego, zadania uczestnicy ponownie otrzymali wcześniej wskazany opis Tomasza W., stworzony przez Kahnemana oraz listę kierunków studiów z informacją o wartości bazowej, czyli respondenci wiedzieli jaki procent wszystkich studentów, stanowią Ci poszczególnych kierunków. Co więcej, eksperymentatorzy poinformowali ich, że prezentowany opis nie jest w pełni miarodajny. Jak w przypadku wcześniejszych zadań uczestnicy dostali polecenie uszeregowania określonych kierunków pod względem prawdopodobieństwa, że Tomasz W. realizuje na nich studia. Wyniki pokazały, że pomimo dostępności istotnych przesłanek, uczestnicy ignorowali wartości bazowe, na rzecz rozumowania opartego na stereotypie.

Drugim klasycznym dla heurystyki reprezentatywności jest wielokrotnie powielony eksperyment, w ramach którego uczestnicy pytani byli o fikcyjną kobietę imieniem Linda. Badanie było bardzo podobne do opisanego wyżej zadania z Tomaszem W. Pierwszy wariant wskazanego eksperymentu przeprowadzony został w paradygmacie grup równoległych, gdzie każdemu uczestnikowi zaprezentowano krótki opis kobiety:

Linda ma trzydzieści jeden lat, jest niezamężna, wygadana i bardzo inteligentna. Skończyła filozofię. Na studiach bardzo się angażowała w zwalczanie dyskryminacji i promowanie sprawiedliwości społecznej, a także brała udział w demonstracjach przeciw energii jądrowej (Kahneman, 2012, str. 211).

Eksperymentatorzy przygotowali także dwie listy składające się z 7 pozycji w postaci twierdzeń dotyczących prawdopodobnych zawodów, których Linda może być

przedstawicielką. Na każdej z nich znajdował się jeden kluczowy punkt — „kasjerka bankowa” albo „kasjerka bankowa, działająca w ruchu feministycznym”. Oznacza to, że zadaniem osób znajdujących się w jednej grupie, tej która otrzymywała listę z pozycją „kasjerka bankowa”, było szacowanie prawdopodobieństwa w oparciu o założone przez uczestników wartości bazowe. Natomiast w grupie, gdzie umieszczona była pozycja „kasjerka bankowa, działająca w ruchu feministycznym” miała wykonać rozumowanie oparte na reprezentatywności. Zgodnie z oczekiwaniami, porównanie wyników w grupach równoległych pokazało, że uśredniona ocena zgodności ze stereotypem, była równa uśrednionej ocenie szacunków prawdopodobieństwa.

Kolejnym etapem było przeprowadzenie eksperymentu w paradygmacie grup naprzemiennych. W tym wariancie uczestnicy otrzymywali listy, na których znajdowało się 8 pozycji:

Linda jest nauczycielką w szkole podstawowej.

Linda pracuje w księgarni i ćwiczy jogę.

Linda działa w ruchu feministycznym.

Linda jest pracownicą opieki społecznej i pomaga osobom z zaburzeniami psychicznymi.

Linda jest członkinią Wyborczej Ligi Kobiet.

Linda jest kasjerką bankową.

Linda jest agentką ubezpieczeniową.

Linda jest kasjerką bankową i działa w ruchu feministycznym. (Kahneman, 2012, str. 211).

Respondenci proszeni byli o oszacowanie, która z podanych możliwości jest najbardziej prawdopodobna. Wyniki ujawniły, że aż 89% wszystkich biorących udział w badaniu uznało za najbardziej prawdopodobną pozycję koniunkcyjną (*Linda jest kasjerką bankową i działa w ruchu feministycznym*). Oznacza to, że uczestnicy w znacznej większości dokonywali szacowania opartego o stereotyp, zgodnie, z którym opis kobiety pasował do osoby działającej w ruchu feministycznym. W rzeczywistości nie jest możliwe, aby zbiór kasjerek bankowych feministek, był liczniejszy od zbioru kasjerek bankowych, jednak opis Lindy stworzony był z myślą o stereotypie działaczki w ruchu feministycznym, dlatego uzupełnienie pozycji „kasjerka bankowa” o „działanie w ruchu feministycznym” sprawiło, że opcja ta wydała się uczestnikom bardziej prawdopodobna. Ustalono, że u podstaw wyboru respondentów leży tendencja nazywana błędem koniunkcji, polegająca na tym, że koniunkcja dwóch zdarzeń

oceniania jest jako bardziej prawdopodobna niż jakikolwiek pojedynczy element tej koniunkcji, oznacza to, że w bezpośrednim porównaniu koniunkcja dwóch cech, uznana zostaje za bardziej prawdopodobną niż każdy z jej elementów (Kahneman, 2012; Nęcka, 2013). Co ciekawe badanie to powtórzone zostało w grupie doktorantów realizujących studia w jednostce Stanford Graduate School specjalizującej się w obszarze statystyki i teorii decyzyjności. Wyniki ujawniły, że aż 85% uczestników popełniło błąd koniunkcji, pomimo ukończenia, co najmniej kilku kursów ze statystyki, czy samego prawdopodobieństwa. Studenci popełniali ten sam błąd, co osoby bez wskazanego wykształcenia, tj. znacząca większość uznawała, za najbardziej prawdopodobną koniunkcję pozycji *kasjerka bankowa* i *działaczka w ruchu feministycznym*.

Przeprowadzenie kolejnego wariantu eksperymentu z Lindą wynikało z obserwacji, że kluczowe pozycje, oddzielone zostały stwierdzeniem, *Linda jest agentką ubezpieczeniową*. Uznano, że taki zabieg spowodował, że uczestnicy oceniali prawdopodobieństwo każdego scenariusza niezależnie od siebie. Zakładano, że umieszczenie kluczowych scenariuszy jeden po drugim sprawi, że system 2 dokona korekty błędu wynikającego z myślenia heurystycznego. W związku, z czym uproszczono zadanie, prezentując uczestnikom pytanie w następującej formie:

Która możliwość jest bardziej prawdopodobna:

Linda jest kasjerką bankową.

Linda jest kasjerką bankową i działa w ruchu feministycznym? (Kahneman, 2012, str. 214).

Przeprowadzono kilka eksperymentów z krótkim wariantem Lindy, grupę badawczą stanowili studenci aż czterech renomowanych uczelni, którzy w trakcie studiów realizowali kursy dające podstawę teoretyczną pozwalającą uniknąć błędu koniunkcji. Wyniki ujawniły, że tylko doktoranci kierunku socjologia z uniwersytetów Stanforda i Berkeley odpowiedzieli w większości (64%) poprawnie. W pozostałych grupach procent ten wahał się między 10-20. Uzyskane wyniki skłoniły badaczy do powtórzenia eksperymentu w celu sprawdzenia jak grupa studentów z zakresu socjologii, wypadnie w dłuższym wariantcie Lindy z zastosowaniem paradygmatu grup naprzemiennych. Zakładano, że oddzielenie kluczowych pozycji innym stwierdzeniem, sprawi, że uczestnicy ocenią wszystkie scenariusze niezależnie od siebie. Wskazana hipoteza

została potwierdzona, gdyż procent poprawnych odpowiedzi spadł do 15. Oznacza to, że w związku z tym, że nie dochodziło do porównania poszczególnych elementów na liście, system 2 nie dokonywał korekty. To doprowadziło autorów do wniosku, że błąd koniunkcji nie przestaje się narzucać, nawet gdy jednostka zna prawidłową odpowiedź (Gould, 2010; Kahneman, 2012). Interpretację tą potwierdza także Stephen Jay Gould⁸³, autor książki *The panda's thumb: More reflections in natural history* (2010), który zauważa, że nawet osoby dobrze zaznajomione z zasadami logiki i rachunkiem prawdopodobieństwa, mają trudności z zahamowaniem intuicyjnego wniosku, że *Linda jest feministką i kasjerką w banku* stanowi bardziej prawdopodobną opcję niż sama *kasjerka w banku* (Gould, 2010). Gould uważa, że ta silna intuicja wynika z ewolucyjnie wykształconej tendencji do tworzenia spójnych narracji, nawet jeśli wymaga to zignorowania zasad statystycznych takich, jak reguła koniunkcji⁸⁴. Stanowisko to podkreśla związek heurystyki reprezentatywności z efektem potwierdzenia opisanym szczegółowo w części 2.5. *Teoria dysonansu poznawczego*. To wyraźnie prezentuje, że heurystyka reprezentatywności pełni kluczową rolę w tworzeniu najbardziej wiarygodnych (nie zaś, prawdopodobnych) scenariuszy. Opisane stanowisko wspierane jest przez wyniki eksperymentów izomorficznych, gdzie uczestnikom prezentowano zadania o następującej treści (Gould, 2010):

Która z poniższych opcji jest bardziej prawdopodobna

- a. Marek ma włosy
- b. Marek ma włosy koloru blond

albo:

⁸³ Amerykański paleontolog, biolog ewolucyjny i historyk nauki. Znany był jako jeden z najbardziej wpływowych popularyzatorów nauki XX wieku. Jego badania koncentrowały się na ewolucji i paleontologii, szczególnie na teorii równowagi przerywanej (*punctuated equilibrium*), którą opracował wspólnie z Nilesem Eldredgem. Teoria ta sugeruje, że ewolucja zachodzi w okresach szybkich zmian oddzielonych długimi okresami względnej stabilności. Gould napisał wiele książek popularnonaukowych, w tym *Bully for Brontosaurus* i *The Panda's Thumb* (2010). Był też krytykiem socjobiologii i deterministycznych interpretacji genetyki.

⁸⁴ Reguła koniunkcji głosi, że jednoczesność występowania dwóch elementów zawsze jest mniej prawdopodobna, niż zajście pojedynczych zdarzeń (Gould, 2010).

Która z poniższych opcji jest bardziej prawdopodobna

- a. Jane jest nauczycielką
- b. Jane jest nauczycielką i chodzi do pracy piechotą

Wyżej przedstawione przykłady zadań izomorficznych dla problemu Lindy zawierają koniunkcję, w których kolejny element stanowi jedynie uszczegółowienie, niepodnoszące stopnia wiarygodności czy spójności narracji, co sprawia, że nie wywołują błędu koniunkcji.

Istnieje specyficzny związek między heurystyką reprezentatywności, a dostępności, któremu warto się przyjrzeć. Pomimo odmienności tych mechanizmów w niektórych przypadkach prowadzą do tych samych błędów – przykładem jest opisywany w poprzednim akapicie błąd koniunkcji. W przypadku heurystyki dostępności, koniunkcja wydaje się bardziej prawdopodobna niż jej poszczególne elementy, ponieważ jednostce łatwiej przypomnieć sobie obydwa zdarzenia, które zaszły jednocześnie, niż którekolwiek z nich osobno. Jest to związane z mechanizmem rozprzestrzeniającej się aktywacji. W przypadku koniunkcyjnym uaktywniane są dwa węzły, co sprawia, że aktywacja jest silniejsza. Reguły te uzupełniają się szczególnie w przypadku przyjmowania postaw takich jak uprzedzenia (Sternberg, 2013). Dla przykładu, jeżeli jednostka posiada przekonanie, że osoby należące do określonej grupy, np. partii politycznej cechują się określonymi atrybutami, to przykłady reprezentatywne, czyli te, które ujawniają te cechy, będą bardziej dostępne, a co za tym idzie, łatwiejsze do przywołania, niż przypadki niezgodne z typowymi egzemplarzami tej kategorii.

Ważną cechą rozumowania heurystycznego jest poszukiwanie wystarczającego wyjaśnienia, nie zaś optymalnego (Gigerenzer, Brighton, 2009). W związku tym, podczas rozumowania na rzecz utrzymania wiarygodności i spójności narracji ignorowane są istotne przesłanki, takie jak wartości bazowe. Jedną z tendencji wynikających z rozumowania opartego na reprezentatywności jest m.in. wspomniana już preferencja do wymiennego postrzegania prawdopodobieństwa i podobieństwa. Dla przypomnienia, pierwsza z wymienionych relacji, została szczegółowo opisana w części 2.3. *Myślenie analogiczne*. Wskazałam wtedy, że w wyniku wspomnianej tendencji dochodzi do błędów w postaci np. stosowania reguł rozumowania analogicznego do relacji o charakterze homologicznym. Inną preferencją wynikającą z heurystyki reprezentatywności jest tendencja do przeceniania prawdopodobieństwa wystąpienia

wydarzenia, którego wartość bazowa jest niska, jednak jest ono reprezentatywne dla np. obecnego kontekstu (Kahneman, 2012). W takiej sytuacji jednostka podejmuje decyzję w oparciu o mało istotne przesłanki, ignorując jednocześnie te istotne. Mechanizm ten działa także w przypadkach przeciwnych — gdy jednostki nie doszacowują prawdopodobieństwa wydarzeń o wysokiej wartości bazowej, gdyż samo wydarzenie nie jest zgodne (podobne) z typowymi egzemplarzami (K). Co więcej, literatura wskazuje, że rozumowanie oparte na stereotypie jest ściśle związane z opisanym wyżej, mechanizmem polegającym na dostrzeganiu schematów tam, gdzie nie występują, czyli np. w procesach losowych (Bílek, Nedoma, Jirásek, 2018). Wskazuje na to fakt, że niektóre sekwencje zdarzeń o charakterze losowym nie są reprezentatywne dla swojej kategorii, w wyniku czego ludzie odrzucają wyjaśnienie o jej losowości (Stroh, 2022). Innym mechanizmem zaangażowanym w przetwarzanie w oparciu o stereotyp jest opisany w części 2.5. *Teoria dysonansu poznawczego* efekt potwierdzenia. Współpraca tendencji wynikających ze wskazanych mechanizmów leży u podstaw ignorowania jakości danych, które są spójne z ogólną narracją czy innymi strukturami poznawczymi (Gilovich, 1991; Nickerson, 1998; Kahneman, 2012). Oznacza to, że ludzie często akceptują informacje zgodne z ich wcześniejszymi przekonaniem bez krytycznej analizy ich wartości merytorycznej. Skutkuje to uznaniem za prawdziwe danych potwierdzających popularny stereotyp lub znany schemat, jeśli te dane są nieprecyzyjne lub wątpliwe. Zjawisko to wynika z potrzeby utrzymania spójności poznawczej (teoria dysonansu poznawczego) i ograniczenia zużycia zasobów (skąpiec poznawczy). W wyniku opisywanej tendencji nie tylko podejmowane są błędne decyzje, ale także formułowane opinie odbywa się na podstawie niekompletnych lub mylnych informacji. Dla przykładu, jeśli jednostka wierzy, że grupa społeczna (G) jest bardziej skłonna do popełniania przestępstw, może ignorować statystyki temu przeczące i zapamiętywać jedynie przypadki potwierdzające to przekonanie. Oznacza to, że zarówno heurystyka reprezentatywności, jak i efekt potwierdzenia współpracują ze sobą, wzmacniając poznawczą stronniczość, co prowadzi do utrwalania błędnych przekonań i odporności na zmianę opinii, nawet w obliczu nowych, bardziej wiarygodnych przesłanek. Co więcej, jak wskazuje Hong Liu oraz Sariana Du (2016) zastosowanie heurystyki reprezentatywności prowadzi do błędu wartości bazowej (*base-rate fallacy*). Błąd ten polega na przecenieniu znaczenia bieżących informacji, które wydają się bardziej reprezentatywne lub typowe dla określonego schematu poznawczego, przy jednoczesnym ignorowaniu ogólnej wiedzy i danych statystycznych. Dla przykładu, gdy

jednostka szacuje prawdopodobieństwo, że dany człowiek jest inżynierem, robi to na podstawie stereotypowych cech takich, jak np. introwertyzm czy niskie kompetencje społeczne, jednak obiektywnie rzecz ujmując, kryteria te nie uwzględniają rzeczywistego odsetka inżynierów w populacji.

Heurystyka reprezentatywności leży u podstaw generowania wielu fałszywych wspomnień. Fakt, że ludzie często opierają się na podobieństwie sytuacji lub informacji do znanych schematów poznawczych, zamiast na dokładnej analizie faktów, prowadzi do tego, że informacje, które wydają się zgodne ze schematami lub stereotypami, są łatwiej zapamiętywane, nawet jeśli nigdy się nie wydarzyły. Wydobywanie informacji z pamięci długotrwałej w oparciu o określone przez kategorię kryteria skutkuje zwiększoną ilością intruzji (Nęcka, 2013). Czyli gdy pewne elementy wspomnienia (reprezentacji zapisanej w pamięci autobiograficznej) są niejasne lub niepełne, umysł automatycznie wypełnia te luki na podstawie najbardziej prawdopodobnego scenariusza, generując w ten sposób fałszywe wspomnienia. Podanie sugestywnych informacji zgodnych z reprezentatywnymi schematami może prowadzić do modyfikacji istniejących wspomnień, zwłaszcza jeśli są one świeże lub słabo utrwalone. Ponadto, jeśli nowe doświadczenie jest podobne do wcześniejszego, może dojść do pomieszania wspomnień z przeszłości z nowymi, co prowadzi do powstania zniekształconych wspomnień. Najbardziej znanym eksperymentem prezentującym omawiane zjawisko jest przeprowadzony przez Loftusa i Palmera (1974), gdzie uczestnikom przedstawiano nagranie, na którym dwa auta ulegają kolizji. Następnie podzielono ich na dwie grupy, respondenci jednej z nich odpowiadali na pytanie (a) „z jaką prędkością jechały samochody, gdy zderzyły się ze sobą?”, natomiast pozostałych pytano (b) „z jaką prędkością jechały samochody, gdy roztrzaskały się o siebie?”. Wyniki ujawniły, że osoby odpowiadające na pytanie (b) szacowały prędkość wyżej niż uczestnicy odpowiadający na pytanie (a). Oznacza to, że heurystyka reprezentatywności pełni rolę filtra poznawczego, którego funkcją jest dopasowywanie nowych informacji do istniejących wzorców myślenia, co zwiększa liczbę intruzji np. w postaci nieadekwatnych do rzeczywistości reprezentacji.

Istnieje wiele rodzajów fałszywych wspomnień, które mogą powstać na skutek rozumowania opartego o stereotypy. Jednym ze szerzej opisanych zniekształceń wspomnień nazywany błędem atrybucji źródła, polega na błędnym przypisaniu źródła wspomnienia, a co za tym idzie osadzeniu przywoływanych treści w złym kontekście, zjawisko to może skutkować całkowitą zmianą znaczenia określonych treści. Często

przytaczanym przykładem występowania wskazanego błędu są zeznania naocznych świadków, którzy nieświadomie zniekształcają swoje wspomnienia (Stroh, 2022; Aronson, Travis, 2007). Szczegółowy opis błędu atrybucji źródła znaleźć można w pracy Daniela Schactera *Siedem grzechów pamięci* (2001). Autor wskazanej pozycji powołuje się na drastyczny przypadek błędu atrybucji źródła, w którym reprezentacja wydarzenia zawierała elementy, które zostały dodane z innego źródła, co doprowadziło do znacznego zniekształcenia wspomnienia. Chodzi o przykład kobiety, która oskarżyła znanego prezentera telewizyjnego o dokonanie gwałtu, jednak ten miał mocne alibi — gdy dokonano przestępstwa, mężczyzna był w studio, ponieważ właśnie brał udział w programie na żywo. Schacter wyjaśnia, że ofiara widziała prezentera w telewizji podczas, przed lub po przeżyciu traumy. Choć temat błędów pamięci jest bardzo ciekawy, wykracza poza ramy tematycznie niniejszej dysertacji, jednak po więcej informacji na ten temat odsyłam do wielokrotnie już przywoływanej książki popularnonaukowej autorstwa Aronsona i Trávisa (2007).

Jak wielokrotnie starałam się wykazać, heurystyka reprezentatywności jest ściśle związana z uprzedzeniami, ponieważ prowadzi do szacowania prawdopodobieństwa zdarzeń na podstawie ich zgodności ze stereotypami, a co za tym idzie, podobieństwa do najbardziej typowych egzemplarzy (K). Zastosowanie omawianej heurystyki związane jest z atrybucją potencjalnych cech czy prawdopodobnych zachowań określonemu (O) dostępnemu poznawczo. Takie rozumowanie może prowadzić do błędnych założeń o charakterze (O), a także do ignorowania cech, czy zachowań niezgodnych ze stereotypem. Oznacza to, że w kontekście społecznym, heurystyka reprezentatywności może wspierać powstawanie i podtrzymywanie negatywnych stereotypów na temat mniejszości, ras, czy różnych grup społecznych, wzmacniając istniejące nierówności (Stroh, 2022). Mechanizm ten dotyczy także uprzedzeń wobec zwierząt innych niż ludzie. Przedstawiciele różnych gatunków oceniani są przez pryzmat stereotypów zakorzenionych w kulturze jednostki przejawiającej omawiane postawy. Dobrym przykładem jest koncepcja karnizmu Melanie Joy (2010), wyjaśnia ona, że przedstawicielom różnych gatunków przypisuje się kulturowe role społeczne mieszczące się w kategoriach np. „do jedzenia” lub „do kochania”. Osobnikom należącym do drugiej z wymienionych kategorii często przypisuje się więcej możliwości poznawczych (np. uznaje się, że zdolne są one do odczuwania emocji). Gatunki stereotypowo postrzegane jako „do jedzenia” postrzegane są jako „maszyny” do pozyskiwania mięsa/jajek/mleka (Joy, 2010; King, 2017; Singer, 2018). W jednej ze

swoich książek Barbara King (2017, str. 10-11) antropolożka, która przez lata pracowała na Wydziale Antropologii *College of William & Mary*, wspomina rodzinną wizytę w restauracji The Carnivore Restaurant zlokalizowanej w Nairobi. W tym czasie jako absolwentka antropologii brała udział w badaniach terenowych dotyczących tego, jak młode pawiany uczą się, które rośliny lub elementy roślin są jadalne. Podkreśla, że interesowały ją także inne gatunki zwierząt, które obserwowała w czasie wolnym, dla własnej satysfakcji. Jej fascynacja i poczucie jedności ze zwierzętami pozaludzkimi nie stanowiły przeszkody do odwiedzenia restauracji serwującej różne nietypowe rodzaje mięs zwierząt, które z zachwytem obserwowała w Amboseli. W swojej książce *Osobowość na talerzu* (King, 2017, str. 11) autorka przyjmuje metaforę „osobliwego dualizmu”, który sprawia, że z osoby podziwiającej i szanującej inne gatunki zwierząt zamieniała się w osobę ignorującą fakt, że mięso, które spożywa, jest częścią ciała jakiegoś zwierzęcia. Na talerzu miała mięso, nie zaś czującego, posiadającego bogate życie wewnętrzne osobnika. King podkreśla powszechność wskazanego dualizmu, przytaczając treść reklamy zachęcającej turystów do podziwiania wolnożyjących krokodyli i aligatorów, a także jednocześnie do spróbowania mięsa tych zwierząt⁸⁵.

Zagadnienie dotyczące karnizmu i gatunkowizmu poruszone zostały także w klasycznej już książce psychologa Hala Herzoga (2010). Co prawda we wskazanej pozycji nie znajduje się bezpośrednie odniesienie do heurystyki reprezentatywności, jednak wiele z omawianych przez autora zagadnień można interpretować w kontekście jej działania. Jedno z najistotniejszych twierdzeń zawartych we wskazanej książce głosi, że relacje człowiek–inne zwierzęta nie są oparte na spójnych zasadach moralnych, lecz na subiektywnych kryteriach (stereotypach). Herzog argumentuje, że różnice w kulturalnym postrzeganiu określonych gatunków ZINL są arbitralne i nie wynikają z obiektywnych cech tych zwierząt, lecz z lokalnych tradycji i wierzeń. Powołuje się na popularne przykłady, jak różnica w kulturowym postrzeganiu krów — w Indiach krowy są czczone, co wyklucza ich spożywanie, podczas gdy w krajach zachodnich są one postrzegane głównie jako źródło mięsa i mleka. Podejmuje także temat lęku, który ludzie czują przed ZINL postrzeganymi kulturowo jako groźne takimi, jak wilki, mimo że statystycznie są one mniej niebezpieczne niż wiele innych zwierząt (np. psy). Autor zaznacza, że stereotypy kulturowe i indywidualne skojarzenia wpływają nie tylko na

⁸⁵ Treść reklamy przytoczonej przez King: „Zobacz piękne aligatory rano! Rozkoszuj się „krokodylonuggetsami w południe!” (King, 2017, str.12).

ludzkie decyzje w kwestiach takich jak ochrona środowiska, dieta czy badania naukowe, ale także determinują poziom empatii odczuwanej do przedstawicieli różnych gatunków zwierząt, dla przykładu ludzie są bardziej skłonni do odczuwania współczucia wobec delfinów (nie tylko kojarzą się z inteligencją, ale także są ssakami, jak ludzie) niż ryb (które stereotypowo są postrzegane jako „zimne” i „prymitywne”).

Wyżej przytoczone dane, ale i wiele innych wskazują, że heurystyka reprezentatywności pełni kluczową rolę w powstawaniu i utrzymywaniu uprzedzeń wobec ZINL (Joy 2010; (Plamer 2010; Herzog, 2010; Marino, 2011; King, 2017; Barash, 2018). Postawy te wpływają na prawo, ochronę środowiska i decyzje konsumenckie, co czyni je istotnym problemem społecznym.

3.5. Podsumowanie

Rozdział trzeci stanowi pogłębioną analizę epistemicznych oraz historycznych czynników, które determinują interpretacje zdolności poznawczych zwierząt innych niż ludzie, ze szczególnym uwzględnieniem przypisywania im teorii umysłu. Omówione zostały zasadnicze przemiany w myśleniu o umyśle zwierząt – od koncepcji starożytnych, przez mechanicyzm kartezjański i rewolucję darwinowską, aż po rozwój współczesnej etologii oraz psychologii porównawczej. Wskazano, że oceny dotyczące poznawczych kompetencji zwierząt są nierozzerwalnie związane z dominującymi w danej epoce przekonaniami kulturowymi, religijnymi i filozoficznymi, jak również z obowiązującymi paradygmatami metodologicznymi. W tym kontekście szczególnego znaczenia nabiera kanon Morgana, który – choć sformułowany jako zasada ostrożności interpretacyjnej – bywa nadinterpretowany, na jego podstawie formułowane są argumenty o strukturze *ad ignorantiam*, skutkiem czego złożone zdolności poznawcze zwierząt pozaludzkich, zostają odrzucone.

W dalszej części rozdziału przeanalizowałam współczesne ramy interpretacyjne, ujawniające napięcia pomiędzy podejściami mentalistycznymi, zakładającymi istnienie stanów umysłowych u zwierząt, a stanowiskami behawioralnymi, które takich stanów im odmawiają. Zasygnalizowałam wtedy istnienie asymetrii w interpretowaniu wyników uzyskiwanych w eksperymentach z udziałem dzieci ludzkich oraz zwierząt – identyczne dane empiryczne oceniane są odmiennie w zależności od badanego podmiotu, co wskazuje na obecność ukrytych uprzedzeń interpretacyjnych, osadzonych w kulturowych narracjach i oczekiwaniach badaczy. Więcej uwagi temu zagadnieniu poświęcę w następnych rozdziałach.

Istotnym elementem analizy są również mity oraz stereotypy funkcjonujące jako struktury poznawczo-kulturowe, które kształtują zarówno antropomorficzne, jak i antroponegacjonistyczne postawy wobec ZINL. Wskazałam dominujące mity w interesującym mnie obszarze oraz określiłam ich rolę, jako epistemicznych schematów wpływających na odbiór oraz interpretację danych empirycznych. Wskazano ponadto, że uprzedzenia – zarówno poznawcze, jak i afektywne – przekładają się na konkretne praktyki społeczne, w tym na racjonalizację przemocy wobec przedstawicieli grup obcych, co dodatkowo wzmacnia kulturowe oraz językowe mechanizmy usprawiedliwiające ich uprzedmiotowienie.

Wreszcie, omówiłam wybrane heurystyki i błędy poznawcze, które przyczyniają się do przyjmowania określonych postaw wobec zwierząt pozaludzkich. Ich

przykładami są automatyczna strategia redukcji napięcia pomiędzy deklarowanymi wartościami a rzeczywistymi praktykami wobec zwierząt. Całość prowadzi do konkluzji, iż interpretacje zdolności poznawczych ZINL nie stanowią neutralnego odzwierciedlenia danych empirycznych, lecz są uwarunkowane czynnikami kulturowymi, epistemicznymi oraz afektywnymi, które determinują zarówno naukowe, jak i społeczne postawy wobec zwierząt. Tym samym, rozdział podkreśla konieczność krytycznej refleksji nad tym, kto i w jaki sposób interpretuje dane, oraz jak postawy wobec zwierząt, kształtują zarówno społeczne oraz naukowe narracje.

Rozdział IV: Kluczowe koncepcje teorii umysłu

4.1. Współczesne filozoficzne koncepcje mindreadingu

W literaturze filozoficznej, psychologicznej i kognitywistycznej z zakresu poznania społecznego toczy się szeroka dyskusja między zwolennikami dwóch dominujących nurtów wyjaśniających, jak zwierzęta interpretują stany mentalne innych. Podejście symulacyjne zakłada, że rozumienie stanów mentalnych innych osób polega na odtwarzaniu lub naśladowaniu stanów innych w swoim własnym umyśle. Badania wykazują, że właśnie takie mechanizmy leżą u podstaw empatii i zjawisk pokrewnych, jak zarażenie emocjonalne⁸⁶ czy ziewanie przez naśladownictwo. Przykładowo, badania nad ewolucją empatii wskazują, że szeroki zakres reakcji empatycznych, w tym zarażenie emocjonalne, można wyjaśnić poprzez symulację poznawczą (Mafessoni, Lachmann, 2019). Podobnie, prace z zakresu neuronauki i psychologii pokazują, że symulacja jest podstawowym mechanizmem konstruowania emocji i empatycznych reakcji na poziomie mózgu, angażując rozległe sieci sensoryczno-motoryczne (Wilson-Mendenhall, 2017). Zwolennicy teorii symulacyjnych, tacy jak Alvin Goldman czy Karsten Stueber, utożsamiają symulację z empatią, traktując ją jako „wczuwanie się” w sytuację drugiej osoby (Burns, 2017; Stjernberg, n.d.). Zaletą wskazanych wyjaśnień jest ich zgodność z intuicyjnymi doświadczeniami empatii, a konkretnie z jej automatycznym charakterem. Wydaje się jednak, że wskazane teorie mają trudności w wyjaśnieniu bardziej złożonych stanów mentalnych, takich jak rozumienie fałszywych przekonań innego, rozpoznanie, że przekonania własne od tych innego różnią się od siebie oraz rozumienie ukrytych intencji, które często nie są wyrażane wprost (nie ujawniają się w zachowaniu). Dlatego można przypuszczać, że symulacja jest jednym z wielu mechanizmów biorących udział w przetwarzaniu informacji z zakresu poznania społecznego.

Drugim nurtem wyjaśniającym, jak zwierzęta interpretują stany mentalne innych, są teorie metareprezentacyjne. Proponują one znacznie bardziej złożony mechanizm tworzenia metareprezentacji i wnioskowania na ich podstawie. To podejście bardziej przypomina użycie modelu teoretycznego, który pozwala przewidywać zachowania innych na podstawie ich przekonań, pragnień i celów. Jego zastosowanie

⁸⁶ Opisanego szerzej w części 1.2.1 *Zdolności z zakresu poznania społecznego*.

pozwała wyjaśnić wnioskowania zachodzące w sytuacjach wymagających złożonej analizy intencji, ukrytych motywów oraz intencjonalnego wprowadzania w błąd (np. kłamstwa), oraz działaniach strategicznych. Co więcej, to właśnie z modeli metareprezentacyjnych korzysta się do konstruowania wyjaśnienia zdolności językowych umożliwiających generowanie i odczytywanie ironicznych lub sarkastycznych aktów komunikacji intencjonalnej. W tym ujęciu kluczowe jest posiadanie „teorii”⁸⁷ na temat tego, jak działają umysły innych, co pozwala przewidywać ich zachowania na podstawie dzielanych przez nich przekonań i intencji, nawet jeśli różnią się one od naszych własnych. Zaletą wskazanego nurtu jest precyzyjność i możliwość wyjaśnienia, jak ludzie potrafią przewidywać zachowania innych na podstawie ich przekonań i intencji. Dyskusja ta stanowi istotny punkt wyjścia do rozważań na gruncie ewolucyjnym, próbujących odpowiedzieć na pytanie, czy teoria umysłu – rozumiana jako zdolność do przypisywania innym stanów mentalnych, takich jak przekonania (także te fałszywe), intencje czy emocje – występuje także poza gatunkiem ludzkim? Natomiast jej wadą jest fakt, że zakłada ona udział złożonych procesów⁸⁸ poznawczych, których istnienie przypisuje się głównie człowiekowi.

Zarówno modele symulacyjne, jak i metareprezentacyjne znajdują zastosowanie w badaniach nad poznawczymi zdolnościami zwierząt innych niż ludzie, oferując różne

⁸⁷ Termin „teoria umysłu” został ujęty w cudzysłów, ponieważ nie odnosi się on do teorii w ścisłym, naukowym sensie – jako systemu spójnych i empirycznie weryfikowalnych twierdzeń – lecz raczej do zbioru racjonalnych przekonań formułowanych na podstawie obserwacji, analizy i wnioskowania. W dalszej części pracy, dla uproszczenia zapisu, cudzysłów będzie pomijany.

⁸⁸ Warto odnotować również alternatywne ujęcie rozpoznawania stanów mentalnych, zaproponowane przez Ludwiga Wittgensteina, który odrzucał zarówno model metareprezentacyjny, jak i symulacyjny. W jego ujęciu poznanie stanów wewnętrznych innych osób nie wymaga ani dedukcji, ani wnioskowania, lecz opiera się na bezpośrednim postrzeganiu – dostrzeganiu oznak, które są nierozdzielnie związane z danym stanem. Analogicznie do obserwacji dymu, który wskazuje na obecność ognia, zachowanie drugiej osoby (np. śmiech, płacz) postrzegane jest jako widoczna oznaka stanu mentalnego, a nie jako zewnętrzny przejaw czegoś ukrytego. Choć ujęcie to dobrze wyjaśnia rozpoznawanie emocji czy doznań cielesnych (jak ból czy głód), jego zastosowanie w przypadku bardziej złożonych form poznania społecznego – takich jak identyfikacja fałszywego przekonania czy przypisywanie wiedzy w kontekście pozorów i udawania – wydaje się ograniczone. Przypadki tego rodzaju, jak np. rozpoznawanie intencjonalnego wprowadzania w błąd podczas zabawy lub testy fałszywego przekonania, wydają się wymagać mechanizmów bardziej złożonych niż to, co implikuje wittgensteinowskie „widzenie” stanu umysłu. Ciekawe ujęcie podobnego pomysłu znaleźć możemy w koncepcji *showing* i *signaling* stworzonej przez Mitchella Greena (2019).

narzędzia interpretacyjne do analizy ich zachowań. Modele symulacyjne opierające się na założeniu, że jednostka odtwarza „w sobie” stany emocjonalne lub poznawcze innych poprzez mechanizmy zarażenia emocjonalnego, rezonansu afektywnego lub empatycznej symulacji, znajdują szczególne zastosowanie w interpretacji zachowań społecznych u ssaków naczelnych, psowatych czy niektórych gatunków ptaków (Preston, de Waal, 2002; Decety, Jackson, 2004). W tym ujęciu empatia nie wymaga dostępu do złożonych reprezentacji mentalnych, lecz opiera się przede wszystkim na zdolności do automatycznego współodczuwania z innym osobnikiem. Przykładowo, u szczurów wykazano tendencję do uwalniania współwięzionego towarzysza z pojemnika, co bywa interpretowane jako przejaw elementarnej formy empatii motywowanej dyskomfortem wywołanym przez obserwację cierpienia innego (Bartal, Decety, Mason, 2011). Podobnie u makaków i szympanсів obserwuje się przejawy zachowań wskazujących na rozpoznawanie stanów emocjonalnych innych osobników — takich jak agresja, frustracja czy strach — które mogą być natychmiastowo odwzorowywane we własnym układzie afektywnym, co skutkuje odpowiednimi reakcjami behawioralnymi (de Waal, Preston, 2017). Podejście symulacyjne dobrze wpisuje się zatem w wyjaśnienie mechanizmów automatycznych, szybkich i afektywnie naładowanych reakcji społecznych, jakie obserwuje się u wielu gatunków. Jego ograniczeniem bywa trudność w wyjaśnianiu bardziej złożonych form rozumienia społecznego, wymagających dystansu wobec przeżywanego afektu oraz refleksyjnego podejścia do cudzych intencji (Goldman, 2006). W prostszych wersjach teorii symulacyjnych zakładano, że jednostka nie rozróżnia siebie od innego jako odrębnego źródła stanów mentalnych, a sama symulacja zachodzi nieświadomie i bezpośrednio. Jednak nowsze ujęcia, takie jak koncepcja symulacji kontrolowanej, podkreślają istnienie mechanizmów metapoznawczych, umożliwiających odróżnienie własnych od symulowanych stanów (Goldman, 2006; Decety, Jackson, 2004). To rozróżnienie pozwala uniknąć nadmiernego utożsamiania się z innymi oraz umożliwia bardziej elastyczne reagowanie na cudze emocje i intencje. Niemniej jednak złożone zachowania społeczne, takie jak intencjonalne wprowadzanie w błąd, manipulacja informacją czy odczytywanie aluzji — choć trudno uchwytnie wśród ZINL — wymagają wyjścia poza afektywne współodczuwanie.

Teoria symulacji oferuje również wyjaśnienie zdolności do symulowania cudzych przekonań, intencji i myśli. Argumentując, że jednostka tymczasowo przyjmuje punkt widzenia innego, symulując, jak on mógłby rozumować w danej sytuacji,

wiedząc to co wie i biorąc pod uwagę własne cele. Symulacja poznawcza zachodzi więc „z pierwszej osoby” oraz „z perspektywy innego”, ale jest wewnętrznie rekonstruowana przez symulującego. Takie wyjaśnienie zgodne jest z innym mechanizmem z zakresu poznania społecznego, jakim jest przyjmowanie perspektywy innego, które dla przypomnienia, polega na procesie aktywnego wyobrażania sobie, jak odczuwany i postrzegany jest świat przez inną jednostkę. Jak wykazałam, w części 1.2.1. *Zdolności z zakresu poznania społecznego* mechanizm ten nie jest tożsamy, z interesującą mnie teorią umysłu. Oczywiście przyjmowanie perspektywy innego, korzysta z ustaleń teorii umysłu, jednak procesy teorii umysłu nie wymagają tak złożonej reprezentacji, a także nie wykorzystują empatii. Zdolność ta umożliwia jednostce rozumienie, że inni mają własne myśli, przekonania, intencje, pragnienia i emocje, które mogą różnić się od jej własnych. Zamiast reagowania emocją, bardziej stosowną do sytuacji cudzej niż własnej, teoria umysłu pozwala na przetwarzanie informacji dotyczących stanów poznawczych często nieznajdujących bezpośredniego odzwierciedlenia w zachowaniu. Dlatego w tym obszarze istotne staje się zastosowanie modeli metareprezentacyjnych. Chociaż przez wiele lat przypisywano zdolność do tworzenia metareprezentacji wyłącznie człowiekowi, nowsze doniesienia dotyczące badań z udziałem zwierząt innych niż ludzie, kwestionują to rozróżnienie (Perner, 1991). Kluczowym testem dla takich zdolności pozostają zadania oparte na rozumieniu fałszywych przekonań (*false belief tasks*), które przez dekady uznawano za niedostępne dla zwierząt. Tymczasem eksperymenty, takie jak te przeprowadzone przez Krupenye i współpracowników (2016), wykorzystujące okulografię (*eye-tracking*) w badaniach z udziałem szympanów, bonobo i orangutanów, sugerują, że niektóre naczelne są zdolne do przewidywania działania innych na podstawie fałszywych przekonań. Zwierzęta te nie tylko potrafią rozpoznać cel działania innego osobnika, ale również przewidzieć, jakie będą jego zachowania w sytuacji, gdy jego wiedza o świecie jest niezgodna z rzeczywistością. Zastosowanie modeli metareprezentacyjnych w badaniach nad ZINL otwiera więc nową przestrzeń do analizowania ich zdolności inferencyjnych, elastyczności poznawczej i możliwości posługiwania się złożonymi strategiami społecznymi (Call, Tomasello, 2008). Dotyczy to nie tylko naczelnych, ale także niektórych ptaków — np. srok i papug — które wykazują zdolność do ukrywania pożywienia w sposób sugerujący, że potrafią przewidzieć, co „wiedzą” ich potencjalni obserwatorzy i modyfikować własne zachowanie w zależności od tych założeń. Przykładowo, badania polegające na obserwacji sójek wskazują, że ptaki te są w stanie

zmienić miejsce ukrycia pożywienia, jeśli uprzednio same działały jako obserwatorzy i tym samym mogą „wnioskować”, że inny ptak może je teraz obserwować (Clayton, Dally, Emery, 2007). Chociaż wyniki te nie dają jednoznacznej odpowiedzi na pytanie, czy inne zwierzęta faktycznie tworzą złożone, sformalizowane metareprezentacje cudzych stanów mentalnych to jednak wskazują na istnienie bardziej elastycznego niż dotąd sądzono mechanizmu poznania społecznego. Wydaje się zatem, że bardziej adekwatne byłoby mówienie o kontinuum zdolności metapoznawczych i symulacyjnych niż o sztywnym rozróżnieniu między obecnością a brakiem teorii umysłu (Andrews, 2012).

Podsumowując, modele symulacyjne i metareprezentacyjne oferują różne, lecz komplementarne perspektywy w badaniach nad poznaniem zwierząt. Pierwsze pozwalają uchwycić afektywne podstawy zachowań społecznych i ich automatyzm, drugie umożliwiają analizę strategii poznawczych i potencjalnych zdolności do rozróżniania cudzych przekonań, także tych fałszywych. Ich zestawienie pozwala nie tylko lepiej zrozumieć różnorodność funkcji poznawczych u ZINL, ale również dostarcza argumentów dla hybrydowych stanowisk w sporze o naturę teorii umysłu, zarówno w odniesieniu do ludzi, jak i zwierząt (Lurz, 2011; Buckner, 2013). Chociaż udział mechanizmów symulacyjnych, stanowi bardzo ciekawe zagadnienie, w niniejszej dysertacji chciałabym poświęcić uwagę teorii metareprezentacyjnym, przyjmowanym przez badaczy takich jak Simon Baron-Cohen, David Premack czy Guy Woodruff. Decyzja o skupieniu się na modelach metareprezentacyjnych wynika z ich szczególnej przydatności w analizie złożonych form poznania społecznego, które wykraczają poza bezpośrednie i automatyczne reakcje afektywne. Ujęcia reprezentowane przez takich badaczy jak Simon Baron-Cohen, David Premack czy Guy Woodruff umożliwiają precyzyjne modelowanie zjawisk wymagających refleksji, dystansu wobec przeżywanych emocji oraz zdolności do operowania złożonymi reprezentacjami stanów mentalnych innych. Kluczowym przykładem pozostaje tutaj rozumienie fałszywych przekonań — traktowane jako „test kamienia milowego” w badaniach nad teorią umysłu — które trudno wyjaśnić wyłącznie na gruncie modeli symulacyjnych. Co więcej, podejścia metareprezentacyjne pozwalają lepiej uchwycić funkcje języka, humoru czy kłamstwa, które zakładają zdolność do świadomego i intencjonalnego odróżniania cudzych stanów wiedzy od własnych. W kontekście badań nad zwierzętami innymi niż ludzie, zastosowanie modeli metareprezentacyjnych pozwala również wyznaczyć granice poznawcze tych zdolności oraz sformułować bardziej precyzyjne

kryteria ich obecności lub braku. Dzięki temu możliwe jest uniknięcie nadmiernie antropomorfizujących interpretacji i oparcie się na kryteriach empirycznych, które lepiej odpowiadają wymaganiom współczesnych nauk kognitywnych. Z tego powodu, chociaż nie neguję znaczenia modeli symulacyjnych w opisie afektywnego komponentu zachowań społecznych, to właśnie modele metareprezentacyjne uznaję za bardziej adekwatne w kontekście interesującej mnie analizy poznawczo-refleksyjnych aspektów teorii umysłu.

4.2. Model zdolności czytania umysłów innych Simona Baron-Cohena

Chociaż teoria empatii Simona Baron-Cohena (2009) nie znajduje zastosowania w niniejszej dysertacji, jego model teorii umysłu stanowi istotny punkt odniesienia dla zawartych w niej rozważań. W rozdziale *Rozwój zdolności czytania innych umysłów: cztery etapy* (2009) autor przedstawia model zdolności *mindreadingu*. Natomiast interesującą mnie zdolność definiuje jako rozumienie, że inny może mieć własne przekonania (odmienne od moich) i przynajmniej podejrzewam, że niektóre z tych przekonań są fałszywe. Co oznacza, że przyjęta przez niego definicja, zgodna jest z wąskim, rozumieniem teorii umysłu (Call, Tomasello, 2008). Baron-Cohen wymienia cztery mechanizmy zaangażowane we wskazane procesy i podkreśla, że każdy z nich stanowi odrębny komponent opisywanej zdolności (Baron-Cohen, 2009). Pierwszy w rozwoju ontologicznym pojawia się detektor intencjonalności (ID). To narzędzie do kategoryzowania obiektów na ożywione —nieożywione. Jego funkcją jest spontaniczne przypisywanie prostych stanów wolicjonalnych (np. cele czy pragnienia) spostrzeganym obiektom na podstawie ich zachowania. To pozwala na nadanie sensu prostym czynnościom ruchowym np. unikaniu i zbliżeniu oraz prowadzi do identyfikacji obiektu jako agenta. Detektor intencjonalności jest automatyczny, oraz reaguje na określone cechy takie, jak samowzbudzający ruch, czyli za każdym razem, gdy wykryty zostanie bodziec agentopodobny. Takie obiekty są bardziej szczegółowo analizowane, co pozwala na potwierdzenie lub odrzucenie wstępnych założeń o ich sprawczości. Należy także pamiętać, że ID może być angażowany nie tylko w reakcji na bodźce wzrokowe, jego aktywacja może być wywołana przez każdą z możliwych modalności. Zatem mechanizm ten pozwala na tworzenie prostych reprezentacji diadycznych takich, jak *agent chce pożywienia*. Paradygmat używany w badaniach dotyczących omawianego mechanizmu polega na zaprezentowaniu uczestnikom filmu, na którym znajdują się

figury geometryczne, poruszające się w różnych kierunkach względem siebie (Heider, Simmel, 1944; Dasser, Ulbaek, Premack, 1989). Po obejrzeniu nagrania, respondenci proszeni są o opisanie tego, co widzieli. Wyniki pokazują, że ludzie mają naturalną skłonność do wykonywania takiego zadania, przypisując obserwowanym obiektom różnorodne stany mentalne, a niektórzy uczestnicy tworzyli złożone narracje wyjaśniające sytuację z filmu (Simon Baron-Cohen, 2003, str. 151). Doniesienia empiryczne sugerują natomiast, że mechanizm ten aktywny jest już w wieku niemowlęcym. Jak już wspomniałam aktywacja detektora intencjonalności jest automatyczna, a co za tym idzie, stany mentalne takie, jak sprawczość, mogą zostać przypisane obiektom, które w rzeczywistości nie są agentami (Dasser, Ulbaek, Premack, 1989; Simon Baron-Cohen, 2003, str. 150)). Zatem odpowiadając na pytanie zawarte w ostatnim akapicie części 2.6 *Słodka reakcja* antropomorfizacja, nie wymaga teorii umysłu (ToMM), do wykonania takiej atrybucji wystarczą procesy ID.

Kolejnym mechanizmem zaangażowanym w tworzenie teorii umysłu jest detektor kierunku patrzenia (EED) aktywowany przez percepcję oczu czyli w przeciwieństwie do ID, funkcjonuje w obrębie modalności wzrokowej. Jego działanie polega na wykryciu bodźca okopodobnego, a następnie monitorowaniu kierunku patrzenia (Blest, 1957). Takie rozumowanie przeprowadzane jest w oparciu o doświadczenie⁸⁹ jednostki – na ich podstawie wyciągany jest wniosek, że element, na który zwrócony jest bodziec okopodobny, widziany jest przez agensa. Czyli za każdym razem, gdy detektor kierunku patrzenia wykryje bodziec, dający się interpretować jako oko, fiksuje się na nim i zaczyna monitorować jego ruchy. Oznacza to, że omawiany mechanizm pozwala na określenie tego, co widzi agens oraz na przypisywanie mu prostych stanów percepcyjnych. Jego podstawowa wartość adaptacyjna to umożliwienie oceny, czy *ja* znajduję się w polu widzenia obserwowanego organizmu. Okazuje się, że już w wieku 9 miesięcy, dzieci wykazują objawy monitorowania patrzenia innych, natomiast od 14 miesiąca życia EED wykrywane jest u wszystkich badanych (zdrowych) niemowlaków. Zatem, działanie detektora kierunku patrzenia prowadzi do powstania prostych reprezentacji diadycznych, przyjmujących formę np.

⁸⁹ Dziecko od urodzenia uczy się, że gdy ma otwarte oczy to widzi świat zewnętrzny, czyli otwarte oczy pozwalają na dostrzeganie obiektów, natomiast zamknięte dają doświadczenie braku widzenia. Co więcej, mają doświadczenia płynące z wykonywania ruchów własnym oczami, co pozwala na odróżnienie stanu *widzenia A*, od stanu *widzenia B*. Projekcja wskazanych doświadczeń na innego, najprawdopodobniej możliwa jest dzięki mechanizmowi generalizacji.

agens patrzy na x. Badania pokazują, że oczy to element, który przyciąga najwięcej uwagi niemowląt – patrzą na niego tak długo, jak na całą twarz, poświęcając znacznie mniej uwagi jej pozostałym elementom⁹⁰ (Haith, Bergman, Moore, 1977; Hainline, 1978; Flavell, Green, Flavell, 1981). Złapanie kontaktu wzrokowego z innym agensem powoduje przyjemne pobudzenie fizjologiczne, które stanowi naturalny mechanizm wzmacniający takie zachowanie (Nichols, Champness, 1971). Co więcej, popularna zabawa w „a kuku”, pojawiająca się w wielu kulturach niezależnie od siebie stymuluje rozwój omawianego mechanizmu na wczesnym etapie życia (Bruner, 1983; Trevarthen, 1979; Super, Harkness, 2010).

Trzecim etapem rozwoju zdolności do *mindreadingu*, jest uwspólnianie uwagi (SAM), który integruje informacje z ID oraz EED. To pozwala na wytworzenie reprezentacji triadycznej, czyli takiej, która uwzględnia *ja*, innego oraz obiekt, na który wspólnie kierujemy uwagę. Uznaje się, że jest to możliwe dzięki zagnieżdżeniu jednej reprezentacji diadycznej, w drugiej. Autor, opisywanego modelu wskazuje, że SAM potrzebuje informacji EED, jednak może budować reprezentacje w oparciu o inne modalności, jest to jednak trudniejsze (Baron-Cohen, 2003). Natomiast oznacza to, że uwspólnienie uwagi może zostać aktywowane także przez dane pochodzące od ID. Należy jednak zaznaczyć, że oczy mają szczególne znaczenie dla czytania umysłów innych, badania pokazują, że są one podstawowym punktem odniesienia w rozumowaniu dotyczącym stanów mentalnych innych (Baron-Cohen, Campbell, Karmiloff-Smith, Grant, Walker, 1995). Już 18-miesięczne dzieci używają obserwowanego kierunku patrzenia jako wskazówki dotyczącej zamierzeń agensa. Okazuje się, że gdy obserwowany obiekt wykonuje niezrozumiałe zachowanie, zwierzęta mają tendencję, do szukania odpowiedzi patrząc w jego oczy. Jeden z klasycznych eksperymentów w tym obszarze polegał na zaprezentowaniu uczestnikom rysunku schematycznej twarzy otoczonej różnymi czekoladami (Baron-Cohen, Campbell, Karmiloff-Smith, Grant, Walker, 1995). Oczy były skierowane na jedną z tabliczek, a inna czekolada wskazywana była przez dużą strzałkę (dystraktor), gdy osoby (zarówno dzieci, jak i dorośli) biorące udział w badaniu pytane były, „którą czekoladę chce Charlie?”, ignorowały nienaturalną wskazówkę, na rzecz kierunku patrzenia.

⁹⁰ Okazuje się, że 6 miesięczne niemowlaki patrzą około 10 razy dłużej niż na usta (Haith, Bergman, Moore, 1977).

Czwartym mechanizmem wyróżnianym przez Baron-Chohena (2003) jest teoria umysłu (ToMM), która pełni dwie funkcje. Po pierwsze odpowiada za tworzenie reprezentacji epistemicznej na podstawie zachowania obserwowanego innego, po drugie integruje wszystkie informacje dotyczące stanów wolicjonalnych, percepcyjnych i epistemicznych przekształcając je w użyteczną teorię umysłu. Za Leslie'm (1984) argumentuje, że postawy propozycyjne zorientowane są na twierdzenia, przyjmując formę: [Królowna Śnieżka-myśli - „kobieta, która sprzedała mi jabłka jest miłą osobą”]. ToMM umożliwia referencyjną nieprzezroczystość, generowanych stanów, co oznacza, że normalne relacje prawdziwości zostają zawieszone (Baron-Cohen, 2003, str. 165). W tym sensie nawet jeśli zdanie [Królowna Śnieżka-myśli-„kobieta, która sprzedała mi jabłka, jest miłą osobą” jest prawdziwe, to twierdzenie [Królowna Śnieżka-myśli-„moja podła macocha jest miłą osobą”], będzie fałszywe⁹¹, gdyż Królowna Śnieżka nie wie, że *kobieta, która sprzedała mi jabłka to moja podła macocha*. Referencyjna nieprzezroczystość jest kluczową cechą epistemicznych stanów mentalnych i chociaż jest to złożony rodzaj rozumowania, badania pokazują, że ludzie nabywają ją już we wczesnym okresie rozwoju (Simon Baron-Cohen, 2003). Leslie (1984) twierdzi, że pierwszym rodzajem stanów epistemicznych, pojawiających się w rozwoju ontologicznym omawianej zdolności, jest udawanie. Badania z zakresu *mindreadingu*, wskazują, że około 3 roku życia w zabawach dzieci pojawia się obserwowalna zmiana jakościowa (Leslie, 1987; Dunn, Dale, 1984). Dzieci same zaczynają przejawiać charakterystyczne dla wskazanego stanu zachowania oraz rozpoznają je u innych (Hadwin, Perner, 1991; Sodian, Frith, 1992). Około czwartego roku życia dzieci zaczynają rozumieć opowieści oparte na oszustwie (*tales of deception*), co jest przejawem wykształcenia u jednostki czwartego komponentu (ToMM). Oznacza to, że mniej więcej w tym czasie dziecko zaczyna pojmować, że inni posiadają inną (niż własna) perspektywę na rzeczywistość, oraz że czasami ich przekonania są fałszywe (nie są wiedzą); (Pratt, Bryant, 1990; Wellmn, 1990; Perner, 1991). Konstruowanie sądów o stanach epistemicznych innych pozwala im zrozumieć narracje takie, jak wyżej wskazana opowieść o Królownie Śnieżce (Pullman, 2014).

⁹¹ Oczywiście „kobieta, która sprzedała mi jabłka” oraz „moja podła macocha” mają wspólny desygnat.

4.3. Test fałszywych przekonań

Test fałszywych przekonań jest jednym z kluczowych narzędzi badawczych wykorzystywanych w psychologii rozwojowej i kognitywistyce. Służy do oceny, czy dana jednostka posiada zdolność do przypisywania innym stanów mentalnych, które mogą być niezgodne z rzeczywistym stanem rzeczy. Jego klasyczna wersja nazywana zadaniem Maxi'ego (*Maxi task*) opracowana została przez Wimmera (1983). Celem jego eksperymentów było zbadanie kompetencji ludzkich dzieci do reprezentowania przekonań innej osoby, które różnią się od tego, co wie uczestnik. Należy zaznaczyć, że chociaż autor *Beliefs about beliefs: Representation and constraining function of wrong beliefs in young children's understanding of deception* korzysta z terminu *teoria umysłu* w sensie szerokim, to jego eksperymenty dotyczyły przypisywania innym fałszywych przekonań, zatem zgodne są z tym, co w niniejszej dysertacji rozumie się za wskazanym pojęciem. Wimmer (1983) zaznacza, że zdolność ta wymaga nie tylko zwykłego wnioskowania o zachowaniu innego, czyli wytworzenia reprezentacji stanu rzeczy (x), ale też uwzględnienia pewne relacje, w których obserwowany pozostaje z tymi obiektami (np. chce x, wierzy w x, itd.). Czyli angażuje zdolność do tworzenia metareprezentacji odgrywającą ważną rolę w dyskusji filozoficznej jako warunek konieczny i wstępny odpowiedzialności moralnej i interakcji społecznych (Dennett, 1978a, 1978b). Zatem, praktyczne znaczenie reprezentowania fałszywych przekonań innej osoby, polega na wykorzystaniu metareprezentacji jako ramy odniesienia do interpretowania lub przewidywania jej działań.

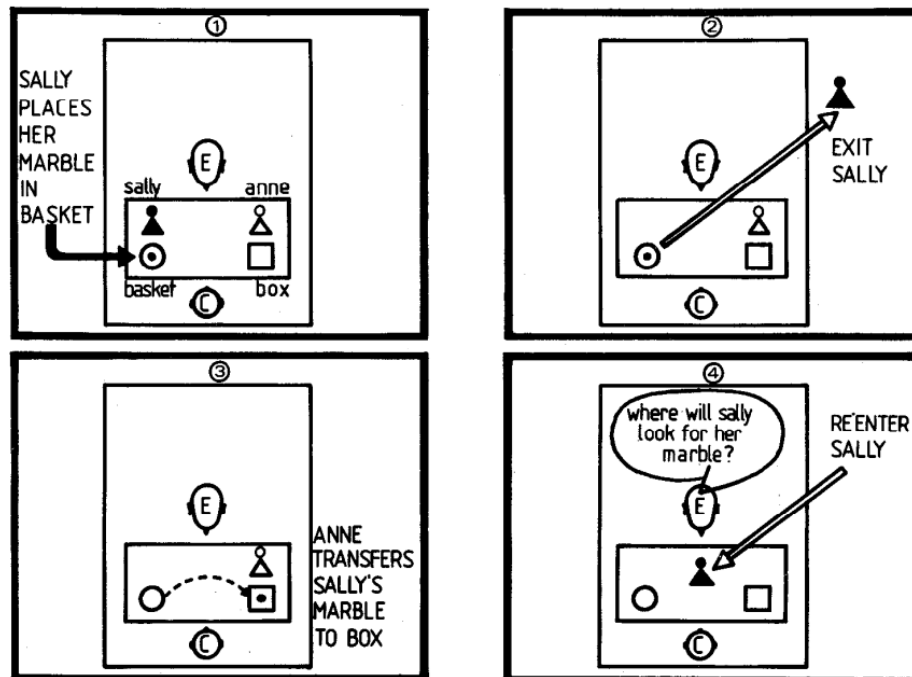
Chociaż można odnaleźć wcześniejsze rozważania teoretyczne dotyczące fałszywych przekonań, to właśnie zadanie Maxi'ego⁹² stworzone przez Wimmera (1983), było szeroko stosowane w badaniach nad teorią umysłu u ludzkich dzieci (Bennett, 1978; Dennett, 1978; Harmon, 1978). W swoim przełomowym artykule odniósł się do tematu zdolności do przypisywania stanów mentalnych innym przez

⁹² W ramach tego testu uczestnicy słuchali historii o chłopcu imieniem Maxi, który wraz z mamą przebywał w kuchni. W pewnym momencie Maxi schował swoją tabliczkę czekolady do zielonej szafki i opuścił pomieszczenie. Pod jego nieobecność matka chłopca przełożyła czekoladę do niebieskiej szafki, po czym eksperymentator pytał dzieci, *gdzie Maxi będzie jej szukał po powrocie?* Jeśli dziecko wskazało zieloną szafkę, uznawano, że uczestnik rozumie, iż Maxi ma fałszywe przekonanie, ponieważ nie widział zmiany i wciąż uważa, że czekolada znajduje się tam, gdzie ją zostawił. Natomiast jeśli dziecko wskazało niebieską szafkę, sugerowało to brak zdolności do przypisywania fałszywych przekonań, ponieważ zakładało, że Maxi posiada tę samą wiedzę co ono samo.

szympansy, jednak prezentuje sceptycznie stanowisko do twierdzeń zawartych w artykule Premacka i Woodruffa (1978). Zgodnie z panującym w jego czasach trendem, podkreśla, że najlepszym sposobem sprawdzenia, czy nieposługujące się językiem szympansy zdolne są do *mindreadingu*, są sytuacje kompetytywne, oparte na oszustwie. Oszustwo posiada wysoką wartość adaptacyjną oraz wymaga konceptualizacji fałszywych przekonań oszukiwanego, a także wykorzystanie ich w strategii planowania zachowania. Sam Wimmer (1983) zauważa, że chociaż oszustwo naturalnie występuje wśród dzikich zwierząt, podkreśla trudności, jakie wiążą się z wyizolowaniem konkretnych aspektów czytania umysłów innych, szczególnie w warunkach laboratoryjnych. Zwraca także uwagę, że długi proces trenowania szympanów do wykonywania zadań w eksperymentalnych w paradygmacie zwodniczego wskazania (*deceptive pointing*), świadczy przeciwko poprzednim doniesieniom Premacka i Woodruffa (1978). Co więcej, autor podnosi argumenty głoszące, że zdolność do teorii umysłu możliwa jest jedynie dzięki językowi, jednak argument ten opiszę szerzej w następnej części.

Artykuł *Beliefs about beliefs: Representation and constraining function of wrong beliefs in young children's understanding of deception* (Wimmer, 1983) miał ogromny wpływ na dalsze badania nad teorią umysłu, stanowiąc podstawę do stworzenia bardziej uproszczonej wersji testu – paradygmatu Sally-Ann, który został dostosowany do badania dzieci znajdujących się na spektrum autyzmu (Baron-Cohen, Leslie, Frith, 1985). Oryginalne obrazkowe przedstawienie tego warunku znajduje się na zdjęciu 7, jednak ze względu na jego złożoność, postanowiłam opisać przebieg eksperymentu w oparciu o inną znaną i bardziej intuicyjną ilustrację, przedstawiającą wskazany warunek (zdjęcie 8). Widać na nim dwie lalki Sally i Ann. Dziewczynki stoją w pomieszczeniu, a między nimi znajduje się koszyk oraz pudełko. Następnie Sally wkłada piłkę do pierwszego z wymienionych miejsc i wychodzi z pomieszczenia, wtedy Ann wstaje i przekłada ją do pudełka. Pytanie zadawane uczestnikom brzmi: *gdzie Sally będzie szukać piłki, gdy wróci do pokoju?* Jeśli dziecko wskaże koszyk – miejsce, w którym, ostatnio Sally widziała przedmiot po raz ostatni – oznacza, że rozumie jej fałszywe przekonanie i potrafi przyjąć jej perspektywę. Jeśli natomiast odpowie, że lalka będzie szukać piłki w pudełku, wskazuje na brak zdolności, do odróżnienia własnej wiedzy o świecie od przekonań innych osób.

Zdjęcie 7. Problem Sally i Ann w teście fałszywych przekonań przedstawiony w artykule *Does the autistic child have a "theory of mind"?*



[Źródło (Baron-Cohen, Leslie, Frith, 1985)]

Zdjęcie 8. Popularne przedstawienie problemu lalek Sally i Ann w warunkach prezentowanym przez Baron-Cohena, Leslie'ego i Fritha w 1985.



Wyniki licznych badań pochodzących z różnych eksperymentów prowadzonych w paradygmacie fałszywych przekonań, wskazują, że ludzkie dzieci odpowiadają poprawnie na postawione im pytania około 3-4 roku życia (Wimmer, 1983; Baron-Cohen, Leslie, Frith, 1985; Wellman, Cross, Watson, 2001; Baron-Cohen, 2009). Istnieją głosy sugerujące, że mechanizm ToMM wykształca się wcześniej, jednak z powodów pragmatycznych w niniejszym opracowaniu, przyjmuję czas jego rozwoju zgodny z bardziej powszechnym stanowiskiem, szczególnie że głosy te oparte są na skrócie myślowym utożsamiającym teorię umysłu ze zdolnością do *mindreadingu* (Onishi, Baillargeon, 2005). Jednak zagadnienie to szerzej omówię w części 5.1.4. *Najnowsze dane empiryczne*.

Zwolennicy nurtu behawioralnego krytykują test fałszywych przekonań, twierdząc, że mechanizmy, które bada, tak jak, inne paradygmaty wykorzystywane do testowania zdolności w zakresie *mindreadingu*⁹³, nie są markerami przypisywania innym stanów mentalnych. Wyniki tych badań wyjaśniane są w odniesieniu do prostszego mechanizmu nazywanego sprytnym interpretowaniem zachowania (*smart behavior-reading*); (Whiten, 1996; Povinelli, Eddy, 1996; Shettleworth, 2010a, 2010b). Z kolei zwolennicy perspektywy mentalistycznej traktują paradygmaty wykorzystywane do badania *mindreadingu*, w tym test fałszywych przekonań jako wiarygodne markery zdolności do konstruowania prostych i bardziej złożonych reprezentacji stanów mentalnych innych. W badaniach z udziałem ludzkich dzieci wykazano, że około czwartego roku życia zaczynają one konsekwentnie udzielać poprawnych odpowiedzi w klasycznych zadaniach tego typu, co w tej perspektywie interpretowane jest jako przejaw jakościowej zmiany w sposobie, w jaki rozumieją one inne umysły (Wellman, Cross, Watson, 2001). Ponadto, jak zauważa Baron-Cohen (2009), rozwój teorii umysłu jest poprzedzony stopniowym przyrostem kompetencji w zakresie prostszych mechanizmów, takich jak system wykrywania oczu (EED) czy system intencjonalności (ID), które stanowią podstawę do późniejszego rozwinięcia zdolności przypisywania innym istotom przekonań. W związku z tym nie ma powodu, by eksperymenty przeprowadzane z udziałem innych zwierząt w tych samych paradygmatach interpretować odmiennie – jeśli w przypadku ludzi wyniki testu fałszywych przekonań uznaje się za dowód na zdolność do reprezentowania stanów mentalnych, analogiczne

⁹³ Takie jak takie jak śledzenie kierunku patrzenia (EED) oraz detektor intencjonalności (ID).

rozumowanie powinno odnosić się także do innych gatunków. Szczegółowa analiza tego problemu znajduje się w *V. Tendencyjność w badaniu teorii umysłu*, a konkretnie w części 5.2.2. *Stanowisko Sary Shettleworth*.

4.4. Stopniowalność teorii umysłu w ujęciu Robina Dunbara i José Luisa Bermúdeza

Teoria umysłu jako zdolność do tworzenia i przetwarzania metareprezentacji (także sprzecznych ze sobą nawzajem), to złożony mechanizm biorący udział w procesach *mindreadingu*. Jego złożoność doprowadziła do wypracowania stanowisk proponujących wprowadzenie stopniowalności omawianej zdolności. Za najciekawsze w kontekście niniejszej dysertacji uznałam teorię Robina Dunbara oraz José Luis Bermúdeza. Pierwszy z wymienionych jest brytyjskim antropologiem biologicznym i psychologiem ewolucyjnym, uzyskał doktorat na Uniwersytecie Oksfordzkim. W czasie, gdy rozwijała się jego kariera naukowa, na tej samej uczelni pracował również opisany wcześniej Niko Tinbergen – laureat Nagrody Nobla i współtwórca etologii. Jego badania koncentrują się na ewolucji poznania społecznego, szczególnie w kontekście teorii umysłu oraz mechanizmów warunkujących powstawanie i utrzymanie więzi społecznych. Jest on autorem koncepcji tzw. „liczby Dunbara”, która odnosi się do maksymalnej liczebności stabilnych relacji społecznych, jakie jednostka jest w stanie efektywnie utrzymać. Choć zagadnienie to jest bardzo ciekawe, znajduje się poza obszarem zainteresowań niniejszej dysertacji.

Bardziej interesująca zdaje się jego koncepcja dotycząca stopni intencjonalności, która w rzeczywistości odnosi się do złożoności metareprezentacji epistemicznej. W tym miejscu chcę jednak podkreślić, że jego użycie słowa *intencjonalność*, różni się od tego, co klasycznie rozumie się pod tym pojęciem filozoficznym (Dennett, 1983). Dlatego w niniejszym opracowaniu będę posługiwać się raczej określeniami *rzędu* czy *poziomu* zdolności do czytania umysłów innych. W swojej szeroko komentowanej książce *Nowa historia ewolucji człowieka* (2014), autor przedstawia swoją propozycję mechanizmu tworzenia metareprezentacji epistemicznej, jako hierarchicznie zorganizowany system, gdzie:

Poziom [1]: organizmy tworzą proste metareprezentacje epistemiczne przyjmujące formę „x chce y” lub „p czuje z”.

Poziom [2]: organizmy posiadają przekonania o intencjach oraz przekonaniach innych, takie jak „x wie, że y widzi”. Ten poziom stanowi kryterium teorii umysłu.

Badania prowadzone na dorosłych ludziach ujawniają, że kompetencje człowieka w tworzeniu metareprezentacji epistemicznej umożliwiają tworzenie wielu rzędów teorii umysłu. Dunbar (2014) zakłada, że istnieje limit ludzkich ograniczeń w tym obszarze, jednak przyznaje, że nie udało się go jednoznacznie określić za pośrednictwem danych eksperymentalnych. Wspomina jednak, że możliwe są metareprezentacje epistemiczne rzędu [9] (Dunbar, 2014, str. 69). Dla porównania przyjrzyjmy się dwóm przykładom, które zobrazują poziom złożoności ludzkiej zdolności do tworzenia teorii umysłu:

Poziom [2]: „Ann jest przekonana [1], że Sally myśli [2], że piłka znajduje się w koszyku”. (Dunbar, 2014)

Rząd [6]: „Piotr wierzy [1], że Ania myśli [2], że Sally chce [3], aby Piotr przypuszczał [4], że Ania ma zamiar [5] przekonać [6] Sally, że piłka jest w koszyku”. (Dunbar, 2014)

Warto zaznaczyć, że Dunbar (2014) przyjmuje szeroką definicję teorii umysłu, zgodnie z którą zdolność ta — podobnie jak pierwotnie twierdzili Premack i Woodruff (1978) — może występować w różnych formach u wielu gatunków. Zróznicowanie to miałyby dotyczyć przede wszystkim poziomu złożoności metareprezentacji epistemicznej, a nie samej obecności lub braku teorii umysłu jako takiej (Mękarska, 2024). Jednakże, w praktyce, przyjęcie tej definicji nie znajduje odzwierciedlenia w kryteriach, według których autorzy oceniają empiryczne przejawy teorii umysłu, co znacząco ogranicza użyteczność takiej deklaracji. Okazuje się, że jego stanowisko w tym kontekście jest wewnętrznie niespójne. Autor ten deklaruje, że teoria umysłu nie jest unikalna dla gatunku ludzkiego, a różnice międzygatunkowe mają charakter ilościowy, nie jakościowy. Jednocześnie jednak zakłada, że tylko ludzie są zdolni do tworzenia metareprezentacji drugiego rzędu, co sam uznaje za minimalne kryterium posiadania teorii umysłu. W efekcie, mimo deklarowanej ciągłości, jego ujęcie faktycznie prowadzi do binarnego podziału, w którym tylko człowiek spełnia warunki właściwej teorii umysłu, a inne gatunki co najwyżej operują uproszczonymi, mechanizmami

mindreadingu, a nawet jeszcze prostszym sprytnym czytaniem zachowań innych (Whiten, 1996).

Przyjęcie takiego stanowiska ilustruje szerszy problem obecny w wielu narracjach behawioralnych. Autorzy tacy jak Shettleworth (2010b) czy sam Dunbar (2014) deklarują uznanie teorii umysłu za zdolność wspólną dla wielu gatunków, po czym stosują skrót myślowy polegający na utożsamieniu tej zdolności z minimalnym poziomem [2], który przypisują wyłącznie ludziom. W ten sposób język deklaracji staje się niespójny z przyjmowanymi kryteriami, prowadząc do pośredniego wykluczenia ZINL z grona istot zdolnych do mentalizacji. Co prawda, niektórzy autorzy funkcjonujący w nurcie behawioralnym dopuszczają, że ZINL mogą być zdolne do czytania umysłów innych na niższych poziomach — detektor intencjonalności czy detekcja kierunku patrzenia (EED) (Dunbar, 2014). W obrębie tego nurtu można jednak znaleźć także bardziej skrajne formy antroponegacjonizmu, całkowicie odmawiające ZINL zdolności do jakichkolwiek form przypisywania agencyjności (Povinelli Eddy, 1996; Shettleworth, 2010a, 2010b). Dla przykładu, poniżej szerzej opisuję koncepcję Sary Shettleworth (2010a, 2010b), która uważa, że zwierzęta inne niż ludzie nie są zdolne do tworzenia reprezentacji fałszywych przekonań, a tym samym nie osiągają nawet poziomu [2]. Oznacza to, że według tego stanowiska ZINL nie posiadają pełnoprawnej teorii umysłu, rozumianej jako zdolność do przypisywania innym stanów mentalnych, nawet gdy odmienne są od własnych. Tego rodzaju przeskok jakościowy między umysłami ludzi i innych zwierząt stanowi podstawę jednego z najczęściej podnoszonych zarzutów wobec teorii nurtu behawioralnego — brak ciągłości ewolucyjnej. Jeśli bowiem zdolność do tworzenia metareprezentacji epistemicznej miała rozwijać się w toku ewolucji stopniowo, należałoby oczekiwać istnienia form pośrednich — zarówno w sensie filogenetycznym, jak i ontogenetycznym. Perspektywa behawioralna często jednak zakłada binarny podział: ludzie posiadają pełnoprawną teorię umysłu, natomiast ZINL postrzegane są jako całkowicie jej pozbawione. Prowadzi to do postulowania istnienia trudnej do uzasadnienia luki pomiędzy prostymi mechanizmami percepcyjno-behawioralnymi, a w pełni rozwiniętymi zdolnościami metareprezentacyjnymi. W najbardziej skrajnych wersjach tej perspektywy mechanizmy takie jak ID (detektor intencjonalności) czy EED (detektor kierunku patrzenia) zostają całkowicie pozbawione funkcji mentalistycznych, co sprawia, że luka ta staje się jeszcze bardziej problematyczna. Zwolennicy tego stanowiska utrzymują, że nawet klasyczne paradygmaty wykorzystywane w badaniach z udziałem ludzkich dzieci nie

dostarczają miarodajnych informacji o rozwoju zdolności mentalistycznych. W związku z tym postulują konieczność reinterpretacji eksperymentów powszechnie uznawanych za markery rozwoju *mindreadingu*. Tego rodzaju podejście uznać można za poważne naruszenie dobrej praktyki epistemicznej, a zwłaszcza normy wrażliwości. Co więcej, podważanie ustabilizowanego dorobku naukowego bez dostatecznych podstaw osłabia dialog teoretyczny. Tymczasem dane empiryczne sugerują, że niektóre gatunki wykazują zachowania, które trudno wyjaśnić wyłącznie przez pryzmat uczenia skojarzeniowego. Badania z udziałem naczelnych (Call, Tomasello, 2008), kruków (Bugnyar, Heinrich, 2005) czy delfinów (Herman, 2006) dostarczają dowodów na występowanie zachowań, które mogą implikować wyższe poziomy przetwarzania społecznego. Z perspektywy mentalistycznej przemawia to za hipotezą, że niektóre gatunki zwierząt pozaludzkich mogą być zdolne do formowania metareprezentacji epistemicznej, być może nawet powyżej rzędu [1].

Jak już wspominałam innym badaczem, postulującym stopniowalność teorii umysłu, jest José Luis Bermúdez. W swoich pracach argumentuje, że zdolność do przypisywania stanów mentalnych nie jest cechą binarną, lecz rozwija się etapami, które można opisać jako kolejne poziomy złożoności metareprezentacji (Bermudez, 2003, 2005, 2009). Uważa zatem, że teoria umysłu rozwija się stopniowo, od prostych reprezentacji własnych stanów mentalnych, przez rozumienie stanów innych, aż po wielopoziomowe metareprezentacje, które umożliwiają, uczestnictwo w złożonych interakcjach społecznych (Dudkowska, 2023). Każdy poziom wymaga coraz bardziej zaawansowanych mechanizmów poznawczych i językowych. Bermúdez (2003) podkreśla, że stopniowalność ta jest kluczowa dla zrozumienia zarówno rozwoju poznawczego, jak i różnic indywidualnych oraz międzygatunkowych.

Zanim przejdę do wyróżnienia poszczególnych poziomów teorii umysłu, chcę uściślić, że w ramach koncepcji wspieranej przez Bermúdeza pojęcia racjonalności i teorii umysłu są ze sobą ściśle związane i często się przenikają. Pierwsze z wymienionych, rozumiane jest jako zdolność do działania na podstawie przekonań i pragnień w sposób spójny, celowy i zgodny z zasadami logicznymi lub praktycznymi. W jego książkach, zwłaszcza *Philosophy of Psychology and Cognitive Science* (2005), dużo miejsca poświęca różnym poziomom racjonalności, od prostych zachowań zwierzęcych po złożone decyzje człowieka. Natomiast teoria umysłu to zdolność do przypisywania innym istotom przekonań, pragnień, intencji i innych stanów mentalnych oraz wyjaśniania i przewidywania ich zachowań na tej podstawie. Oznacza to, że

podobnie jak Dunbar, omawiany autor korzysta z teorii umysłu w sensie szerokim. Bermúdez pokazuje, że stopniowalność racjonalności i teorii umysłu to w istocie dwa aspekty tego samego zjawiska. Analizuje, również to, jak różne zwierzęta różnych gatunków wykazują poziomy racjonalności, zależnie od tego, jak złożone są ich np. przekonania czy pragnienia. Najprostsze formy (np. reakcje zwierząt na bodźce) nie wymagają rozbudowanej teorii umysłu, ale już bardziej złożone (np. planowanie, przewidywanie cudzych działań) – tak. Im wyższy poziom racjonalności, tym bardziej zaawansowana teoria umysłu. Procesy te wymagają nie tylko rozumienia własnych przekonań czy pragnień, ale także przekonania i pragnienia innych (a nawet przekonania o przekonaniach innych). Co więcej, Bermúdez wyróżnia w swojej koncepcji dwa zasadnicze poziomy rozumienia stanów mentalnych: poziom niejawny (*implicit*) oraz poziom jawny (*explicit*). Pierwszy z wymienionych odnosi się do intuicyjnego i automatycznego rozpoznawania stanów mentalnych innych osób, które zachodzi bez świadomej refleksji czy językowego ujęcia. Takie rozumowanie jest szybkie i spontaniczne, co umożliwia natychmiastowe reakcje społeczne — badania wykazują, że jest ono obecne już u niemowląt. Z kolei poziom jawny wymaga rozwiniętych zdolności językowych oraz refleksji, dzięki czemu jednostka potrafi świadomie i w sposób symboliczny wyrażać przekonania o przekonaniach innych. Ten poziom pojawia się później w rozwoju poznawczym i pozwala na bardziej złożone, świadome operacje na stanach mentalnych, co jest niezbędne do pełnego funkcjonowania teorii umysłu.

W literaturze filozoficznej i kognitywistycznej istnieje szeroko rozpowszechnione przekonanie, że zdolność do tworzenia złożonych metareprezentacji – kluczowego elementu teorii umysłu – wymaga kompetencji językowych. Jednym z autorów, którzy podkreślają centralną rolę języka w tym procesie, jest Bermúdez. W swoich pracach argumentuje, że choć niektóre podstawowe formy reprezentacji stanów intencjonalnych mogą zachodzić również u zwierząt nieposiadających języka, to jednak tylko dzięki językowi możliwe staje się tworzenie złożonych, metareprezentacji zagnieżdżonych w innych metareprezentacjach, charakterystycznych dla w pełni rozwiniętej teorii umysłu (Bermudez, 2003; 2005; 2009). Przyjmuje rozróżnienie między tzw. „protonarracjami” (*proto-narratives*), które mogą mieć charakter pozajęzykowy, a złożonymi strukturami narracyjnymi, które umożliwiają intencjonalne rozumienie złożonych zachowań innych podmiotów. W tym kontekście język nie jest jedynie środkiem komunikacji, ale pełni funkcję strukturalną – umożliwia

organizowanie informacji w sposób hierarchiczny, sekwencyjny i referencyjny. Zgodnie z tym stanowiskiem, język wspiera zdolność do teorii umysłu na wyższych rzędach metareprezentacji — takich jak czwarty, piąty czy szósty, opisany przez Dunbara (2014). Bermúdez podkreśla, że taka forma przetwarzania wymaga zdolności do operowania pojęciem przekonania o przekonaniu (*belief about a belief*), co bez języka miałoby być poznawczo niewykonalne. Z tej perspektywy język stanowi warunek *sine qua non* pełnej teorii umysłu. O ile więc podstawowe formy rozpoznawania celowości, takie jak detektor intencjonalności czy detektor kierunku patrzenia, mogą funkcjonować u wielu gatunków zwierząt pozaludzkich, to zdolność do ich wtórnej metareprezentacji wymaga systemu językowego, który umożliwia takie zagnieżdżenia — uznaje się bowiem, że język stanowi swego rodzaju token, pozwalający na zagnieżdżenie jednej reprezentacji w drugiej (Wimmer, 1983; Sperber, 2000).

Warto jednak zauważyć, że powoływanie się na język jako warunek konieczny teorii umysłu bywa wykorzystywane także w bardziej restrykcyjnych, behawioralnych ujęciach. Autorzy tego nurtu, tacy jak Povinelli, Eddy (1996) czy Shettleworth (2010a, 2010b), traktują brak języka jako wystarczający powód do odrzucenia możliwości, że zwierzęta inne niż ludzie posiadają teorię umysłu. W tym sensie argument językowy ulega uproszczeniu i zostaje wprzęgnięty w narrację binarną: posiadanie języka = posiadanie teorii umysłu; brak języka = brak teorii umysłu. Dla przykładu Wimmer (1983) argumentuje, że to język umożliwia tworzenie metareprezentacji oraz powinien być najważniejszym ich wskaźnikiem. U ludzi zdolność do tworzenia metareprezentacji mierzona jest przez obserwację i analizę języka metareprezentacyjnego (*meta-representational language*) w postaci zdań takich jak „x chce y”. Badania pokazują, że dzieci około 2-3 roku życia często używają takiego słownictwa w odniesieniu do siebie i innych, gdy odpowiadają na pytania „dlaczego?” (Hood, Bloom, Brainerd, 1979). Już 3-5 latkowie prawidłowo odróżniają działania zamierzone od tego niezamierzonego (np. odruchy). Natomiast ok. 4 roku życia zaczynają rozumieć niektóre relacje związane z wiedzą, takie jak wiedzieć — zgadywać czy pamiętać — zapominać (Johnson, Maratsos, 1977; Miscione, Marvin, O'Brien, Greenberg, 1978; Johnson, Wellman, 1980, Macnamara, Olson, 1976). Wyraźny problem dzieci wykazują w zadaniach, gdzie trzeba uwzględnić różnicę między relacją własną a kogoś innego, do tej samej treści propozycjonalnej. Czyli np. w sytuacjach, których pojawia się brak wiedzy. Taki stan nazywany jest egocentryzmem epistemicznym (Chandler, Greenspan, 1972; Marvin Greenberg, Mossier, 1976; Mossler, Marvin, Greenberg, 1976). Chociaż zgadzam się

z Wimmerem (1983), że wskazane przez niego narzędzia poznawcze są kluczowe dla rozważań nad teorią umysłu u ludzi, nie uważam za uzasadnione automatyczne odmawianie zwierzętom innym niż ludzie zdolności mentalizacyjnych wyłącznie z powodu braku kompetencji językowych. Obecnie nie dysponujemy bowiem wiarygodnym wyjaśnieniem mechanizmu kompensacyjnego, który umożliwiałby rozwój teorii umysłu u istot pozbawionych języka – co nie oznacza, że taki mechanizm nie istnieje, lecz raczej, że pozostaje on niedostatecznie zbadany. Warto również zauważyć, że choć ogólny konsensus naukowy wskazuje, iż tylko człowiek posługuje się językiem, najnowsze badania sugerują, że systemy komunikacyjne innych zwierząt są znacznie bardziej złożone, niż wcześniej sądzono. Na przykład, ostatnie badania wykazały, że bonobo łączy różne wokalizacje w kombinacje, które przekazują złożone znaczenia, co wcześniej uważano za unikalne dla ludzkiego języka (Berthet, Surbeck, Townsend, 2025). Odkryto 19 różnych kombinacji siedmiu podstawowych typów dźwięków, z których cztery wykazują struktury przypominające składnię ludzką. Przykładowo, połączenie dźwięków *yelp* (oznaczającego „zróbmy to”) i *grunt* („spójrz, co robię”) tworzy kombinację „*yelp-grunt*”, interpretowaną jako „zróbmy to, co ja robię”. Te odkrycia kwestionują przekonanie, że zdolność do tworzenia fraz jest unikalna dla ludzi i sugerują, że poznawcze prekursorzy języka sięgają co najmniej 7 milionów lat wstecz, dzielone przez naszych wspólnych przodków z bonobo i szympanсами (Berthet, Surbeck, Townsend, 2025). Badania te podkreślają, że niektóre nie-ludzkie naczelne wykazują zaawansowaną komunikację wokalną, co dostarcza wglądu w ewolucję ludzkiego języka. W świetle tych danych, argument oparty na kompetencjach językowych jako warunku koniecznym dla posiadania teorii umysłu wydaje się zbyt ograniczający. Przyjęcie takiego stanowiska może prowadzić do błędów teoretycznych oraz do wykluczenia całych klas istot żywych z obszaru moralnej troski.

Ponieważ argument odwołujący się do kompetencji językowych w kontekście teorii umysłu jest często przywoływany przez zwolenników podejścia behawioralnego, a co za tym idzie, został już szczegółowo opisany w literaturze, nie uważam za konieczne jego ponowne, wyczerpujące omówienie w tym miejscu. Zdecydowałam się jednak pokrótce rozwinąć swoje stanowisko w omawianym zakresie — w rozdziale *VI: Niesprawiedliwość wynikająca z uprzedzeń względem zwierząt pozaludzkich*. Chcę także podkreślić, że pogląd tego rodzaju, pozostaje w sprzeczności z danymi empirycznymi. Jak zauważają Call i Tomasello (2008), wiele zachowań naczelnych trudno wyjaśnić bez odwołania się do przynajmniej prostych form reprezentacji stanów

mentalnych innych. Co więcej, wyniki badań nad krukowatymi (Bugnyar i Heinrich, 2005) czy delfinami (Herman, 2006) sugerują, że niektóre gatunki wykazują zdolności, które przekraczają zakres sprytnego czytania zachowań innych i wymagają uznania pewnego stopnia mentalistycznej interpretacji. W tym kontekście argument odwołujący się do języka jako kryterium posiadania teorii umysłu może być krytykowany jako przykład epistemicznej nadmiernej oszczędności, a nawet jako naruszenie zasady epistemicznej wrażliwości. Odrzucanie dobrze ugruntowanych badań empirycznych jedynie dlatego, że ich rezultaty nie mieszczą się w ramach językowo zorientowanej teorii umysłu, oznacza ograniczenie zakresu poznania do jednego, uprzywilejowanego modelu.

Stanowisko Bermúdeza wydaje się jednak bardziej wyważone niż u przedstawicieli skrajnie behawioralnego nurtu. W jego ujęciu język nie jest tożsamy z umysłowością, ale stanowi narzędzie, które umożliwia pewien typ przetwarzania informacji – zwłaszcza w zakresie złożonych zagnieżdżeń metareprezentacji. W tym sensie można uznać, że omawiany autor proponuje gradacyjny model teorii umysłu, w którym językowe metareprezentacje stanowią szczyt możliwości poznawczych, ale nie są jedyną ich postacią. Jednocześnie jednak przyjęcie takiej roli języka może – choć nie musi – prowadzić do odtworzenia klasycznego antropocentryzmu. Jeśli bowiem złożone metareprezentacje zostają utożsamiane z pełną teorią umysłu, a te z kolei z językiem, to nawet najbardziej zaawansowane procesy społeczne zachodzące u innych gatunków zostają zredukowane do „niższego rzędu” zdolności. Takie podejście może prowadzić do implikacji teoretycznych, które są trudne do pogodzenia z danymi pochodzącymi z badań nad komunikacją i interakcją społeczną u zwierząt. Warto zauważyć, że problem relacji między językiem a teorią umysłu jest nie tylko empiryczny, ale również normatywny. To, jakie cechy uznamy za konieczne dla posiadania teorii umysłu, decyduje o tym, jakie istoty moralnie uwzględniamy jako „posiadające umysł”.

Najmniej złożona forma tworzenia metareprezentacji wyróżniana przez Bermúdeza to poziom zerowy, który nie stanowi teorii umysłu *sensu stricto*. Organizmy na tym poziomie, kierują się prostym rozpoznaniem wzorców zachowania i odpowiadają na nie automatycznie. Oznacza to, że jest to proces umożliwiający reagowanie na zachowanie innych bez przypisywania im stanów mentalnych (sprytnie czytanie zachowania innego). Kolejny poziom umożliwia zdawanie sobie sprawy z własnych stanów umysłowych, takich jak przekonania, pragnienia czy emocje.

Podmiot potrafi reprezentować świat w określony sposób, np. „ja chcę wody” czy „wierzę, że pada deszcz”, ale nie ma jeszcze świadomości, że inne istoty mogą mieć własne przekonania, odmienne od jego własnych. Na tym etapie nie występuje jeszcze zdolność do refleksji nad cudzymi stanami umysłowymi. Myślenie pierwszego rzędu ogranicza się wyłącznie do własnych przeżyć i stanów mentalnych. Obserwuje się je u bardzo małych dzieci, a także u wielu zwierząt. Jednostki na poziomie drugim zdolne są do przypisywania stanów mentalnych innym – podmiot rozumie, że ktoś inny może mieć określone przekonania, intencje czy pragnienia, np. „on myśli, że pada deszcz” albo „ona wie, że ja chcę wody”. Jest to zdolność kluczowa dla rozumienia zachowania innych jako uwarunkowanego ich wiedzą i przekonaniami, które mogą różnić się od rzeczywistości oraz od przekonań podmiotu. Na tym poziomie możliwe jest wyjaśnianie i przewidywanie cudzych działań na podstawie ich wewnętrznych stanów mentalnych. Teoria umysłu drugiego rzędu rozwija się u dzieci zazwyczaj około czwartego roku życia i jest niezbędna do poprawnego wykonania testów fałszywych przekonań. Jeszcze bardziej zaawansowane są poziomy trzeciego i wyższego rzędu, które zakładają reprezentację stanów mentalnych dotyczących innych istot. Przykładowo, zdanie „Ona myśli, że on wie, że ja chcę wody” zawiera trzy poziomy zagnieżdżenia metareprezentacji. Każdy kolejny poziom wymaga dodania nowej warstwy refleksji nad przekonaniem lub pragnieniem innej osoby. Takie myślenie leży u podstaw złożonych sytuacji społecznych, takich jak np. intrygi, żarty, kłamstwa czy negocjacje. Wymaga ono znacznych zasobów poznawczych, a badania pokazują, że ludzie rzadko przekraczają cztery do pięciu poziomów intencjonalności w codziennych warunkach (Dunbar, 2010). Bermúdez, podobnie jak Robin Dunbar, podkreśla, że liczba poziomów metareprezentacji, jakie dana jednostka jest w stanie przetwarzać, różni się indywidualnie i może być powiązana z innymi kompetencjami społecznymi.

Warto zauważyć, że przejście między poziomami nie jest jedynie ilościowym zwiększeniem złożoności, lecz wiąże się także z jakościową zmianą w funkcjonowaniu poznawczym. Dla przykładu poziom drugi umożliwia aktywną ocenę sytuacji oraz rozpoznanie, że istnieje wiele jej możliwych interpretacji, a także wybór jednej z nich jako najbardziej adekwatnej. Oznacza to, że podmiot nie tylko potrafi ocenić, co się dzieje, ale także wygenerować najbardziej stosowną reakcję – nie w sposób automatyczny, lecz kontekstowo i elastycznie. Przykładem może być zachowanie raka amerykańskiego, którego agresja jest zmienna osobniczo – zależna od odwagi jednostki i modyfikowana przez wcześniejsze doświadczenia (Schetz, 2011). Już na tym

poziomie ujawnia się istotna cecha wyższych form poznania społecznego – zdolność do uczenia się, przewidywania i dostosowywania reakcji w oparciu o złożoną ocenę sytuacji, a nie wyłącznie proste reakcje na bezpośrednie bodźce.

Oczywiście koncepcja Bermudeza mierzy się z poważną krytyką, dotyczącą co najmniej dwóch etapów proponowanego przez niego schematu rozwoju zdolności do przypisywania stanów mentalnych innym (Lurz, 2011). Między innymi jego twierdzenie głoszące, że przewidywanie zachowań innych na podstawie ich metareprezentacji epistemicznych wymaga nie tylko uwzględnienia profilu psychologicznego, ale też zaplanowania co jednostka powinna zrobić, gdy ten wykona przewidziane przez nią zachowanie. To zaś, jak twierdzi Bermudez, możliwe jest tylko dzięki zdolności do generowania zdań warunkowych, do których jednostka ma świadomy dostęp (Bermudez, 2003, 2009). Jak można zauważyć, w przytaczanych twierdzeniach jest kilka obszarów, które niezgodne są z doniesieniami empirycznymi. Niestety analiza ich wszystkich wymagałabym osobnego opracowania, dlatego postanowiłam skoncentrować się na sugestiach wskazujących na to, że zdolności z zakresu *mindreadingu* pojawiają się znacznie szybciej niż zdolności językowe umożliwiające ich precyzyjne opisywanie⁹⁴. Badania pokazują, że zdolność do odwoływania się do formy logicznej argumentu w celu odróżnienia poprawnych wnioskowań od tych fałszywych obserwujemy u dzieci dopiero w 4-6 roku życia (Gopnik, Graf, 1988). Co więcej, przed tym okresem, nie obserwuje się także zdolności do oceny siły dowodowej własnej argumentacji (Gopnik, Graf, 1988). Oznacza to, że dzieci do 4-6 roku życia nie wykazują świadomości rozumowań, które przeprowadzają. Natomiast, jak już wielokrotnie wskazywałam w niniejszej dysertacji, niektóre zdolności z zakresu *mindreadingu* widoczne są znacznie wcześniej⁹⁵.

4.5. Podsumowanie

Rozdział czwarty prezentuje kluczowe współczesne koncepcje teorii umysłu, stanowią podstawę analiz zdolności poznawczych zwierząt innych niż ludzie. Omówione zostały dwa główne nurty wyjaśniające mechanizmy interpretacji stanów

⁹⁴ Po więcej zapraszam do zapoznania się z Lurz, R. W. (2011). *Mindreading animals: The debate over what animals know about other minds*. Boston Review.
<https://doi.org/10.7551/mitpress/9780262016056.001.0001>.

⁹⁵ Dla przykładu detektor kierunku patrzenia rozwija się w okresie 9-14 miesięcy życia człowieka.

mentalnych przez zwierzęta: podejście symulacyjne, odwołujące się do mechanizmów empatii i automatycznego współodczuwania, oraz podejście metareprezentacyjne, zakładające obecność złożonych procesów tworzenia i operowania reprezentacjami stanów mentalnych innych podmiotów. Model symulacyjny trafnie wyjaśnia szybkie, afektywne reakcje społeczne obserwowane u wielu gatunków, jednak okazuje się niewystarczający w kontekście bardziej złożonych form poznania społecznego – takich jak rozumienie fałszywych przekonań czy ukrytych intencji. Z kolei, modele metareprezentacyjne umożliwiają analizę tych bardziej zaawansowanych procesów, wymagających refleksji, dystansu poznawczego oraz operowania hierarchicznie zorganizowanymi reprezentacjami mentalnymi. W rozdziale przytoczono również dane empiryczne sugerujące, że niektóre naczelnie oraz ptaki wykazujące zachowania, które trudno wytłumaczyć bez odwołania do stanów mentalnych, co podważa dominujące przekonanie o wyłącznej przynależności tego typu kompetencji do człowieka.

Istotnym elementem rozdziału jest analiza modelu Simona Barona-Cohana, który ujmuje zdolność do *mindreadingu* jako wynik interakcji wielu wyspecjalizowanych komponentów poznawczych, takich jak detektor intencjonalności (ID) czy detektor kierunku patrzenia (EED), wskazując na ich automatyczny, wrodzony i wielomodalny charakter. Koncepcja ta stanowi istotny punkt odniesienia w dalszych rozważaniach nad ontogenezą teorii umysłu i jej podstawowymi mechanizmami.

Rozdział ukazuje, że adekwatne ujęcie teorii umysłu wymaga integracji zróżnicowanych podejść teoretycznych, które pozwalają uchwycić zarówno automatyczne, afektywne reakcje społeczne, jak i złożone, refleksyjne procesy poznawcze. Wskazano również na konieczność krytycznego namysłu nad zasadnością utrzymywania wyraźnej granicy demarkacyjnej między człowiekiem a innymi gatunkami, szczególnie w świetle wyników badań nad zdolnościami poznawczymi zwierząt. Punktem wyjścia dla dalszych analiz, podejmowanych w kolejnych częściach pracy, jest przyjęcie modeli metareprezentacyjnych jako niezbędnych narzędzi w badaniach nad bardziej złożonymi formami poznania społecznego.

Rozdział V: Tendencyjność w badaniu teorii umysłu

5.1. Teoria umysłu poza gatunkiem ludzkim: argumenty i dowody

Jak już wielokrotnie zaznaczałam, interpretacje badań procesów poznawczych zwierząt innych niż ludzie, stanowią narracje wspierające określone postawy – od antropomorfizmu po antroponegacjonizm – jakie ludzie przyjmują wobec zwierząt pozaludzkich. Jak wskazywałam w części 3.2. *Dwie perspektywy interpretowania eksperymentów z udziałem zwierząt*, argumenty przemawiające za istnieniem teorii umysłu poza gatunkiem ludzkim konstruowane są w ramach perspektywy mentalistycznej, osadzonej w micie o *jedności człowieka ze światem i innymi zwierzętami*. Jest to struktura poznawcza, która porządkuje nasz obraz świata, w tym sposób interpretowania wyników badań z udziałem zwierząt. Tego rodzaju wyjaśnianie świata leży u podstaw darwinowskiej teorii ewolucji i znajduje swój wyraz w zasadzie ciągłości ewolucyjnej.

Kluczowe jest nie tylko jasne określenie zakresu teorii umysłu – co pozwala odróżnić je od szerszego zbioru zdolności poznawczych, takich jak *mindreading* – lecz także precyzyjne rozgraniczenie między tą zdolnością a znacznie prostszym mechanizmem, nieodwołującym się do stanów mentalnych, jakim jest *smart behaviour reading*. Zwolennicy interpretacji behawioralnych twierdzą, że zwierzęta mogą skutecznie przewidywać działania innych osobników bez konieczności odwoływania się do stanów umysłowych, opierając się jedynie na regularnościach w zachowaniu. Stanowisko to odwołuje się do zasady Morgana, czyli reguły parsymonii stosowanej w etologii i psychologii porównawczej, zgodnie z którą danego zachowania nie należy wyjaśniać poprzez odwołanie do wyższych procesów psychicznych, jeśli da się je wyjaśnić poprzez prostsze mechanizmy.

W tej części rozważań koncentruję się jednak na interpretacjach mentalistycznych zakładających, że w niektórych przypadkach odwołanie się do reprezentacji mentalnych jest niezbędne dla wyjaśnienia obserwowanych interakcji społecznych. Przedstawię najważniejsze stanowiska naukowe, w których formułowana jest argumentacja na rzecz przypisywania stanów umysłowych ZINL, szczególnie w kontekście teorii umysłu. W kolejnym podrozdziale (5.2. *Perspektywa krytyczna w kwestii teorii umysłu u zwierząt innych niż ludzie: wybrane badania*) przeanalizuję z kolei stanowiska wspierające podejście behawioralne. Zdecydowałam się na taki

układ, ponieważ prezentacja perspektywy mentalistycznej jako pierwszej pozwala zarysować ewolucyjny i epistemiczny kontekst zmian w rozumieniu poznania społecznego, w szczególności *mindreadingu*.

Historyczne tło relacji człowiek—inne zwierzęta, wraz z towarzyszącym mu mitem o *jedności człowieka z naturą*, omówiłam w 3.1. *Historia badania umysłów u zwierząt pozaludzkich*. W innych częściach⁹⁶ argumentowałam, że reguła ciągłości ewolucyjnej stanowi podstawę argumentacji mentalistycznej. Co więcej, wbrew często podnoszonym, przez zwolenników skrajnie antroponegacjonistycznych interpretacji zarzutom, nurt ten nie ignoruje zasady Morgana, lecz proponuje jej mniej rygorystyczne zastosowanie (Shettleworth, 2010a). Podkreśla, że pewnych złożonych zjawisk nie da się zredukować do prostszych form, ponieważ są one fundamentalne dla danego poziomu funkcjonowania poznawczego. Przykładowo, model teorii umysłu zaproponowany przez Simona Barona-Cohena zakłada istnienie detektora intencjonalności (ID), mechanizmu automatycznie przypisującego agencyjność obserwowanym obiektom, który zauważalny jest już u niemowląt. W tym kontekście zarzut Shettleworth (2010b), że zwolennicy mentalizmu kierują się „pierwotnym antropomorfizmem”, wydaje się nie w pełni trafny. O ile bowiem przypisywanie stanów mentalnych może być błędne, to sam mechanizm przypisywania intencjonalności wydaje się na tyle pierwotny i automatyczny, że trudno zakładać jego wyłączone występowanie u ludzi. Można oczywiście twierdzić, że poza gatunkiem ludzkim przyjmuje on uproszczoną formę, niewymagającą tworzenia metareprezentacji. Pozostaje jednak pytanie, czy takie rozdzielenie *smart behaviour reading* i *mindreading* jest w każdym przypadku uzasadnione. Owszem, istnieje szerokie spektrum zachowań, które można wyjaśniać bez odwołania do stanów mentalnych, jednak we wszystkich definicjach przyjętych w niniejszej dysertacji teoria umysłu wymaga zdolności do tworzenia metareprezentacji epistemicznej – zagnieżdżonej i wielopoziomowej. Przypisywanie tej zdolności innym gatunkom nie musi oznaczać antropomorfizacji, lecz może wynikać z empirycznego uznania poznawczej ciągłości międzygatunkowej.

⁹⁶ Między innymi w rozdziale III: *Narracje naukowe konstytuujące porządek społeczny w kontekście relacji człowiek — inne zwierzęta*.

5.1.1. Przelomowe badania Premacka i Woodruffa

W artykule *Does the chimpanzee have a theory of mind?* z 1978 David Premack i Guy Woodruff przedstawili badania, które znacząco wpłynęły na panujący ówczesnie pogląd, o unikalności człowieka w kontekście wskazywanej zdolności. Przedstawione przez nich eksperymenty sugerowały, że szympanse są zdolne do przewidywania zachowania na podstawie stanów mentalnych innego przynajmniej, gdy są to cele, motywacje czy intencje. Czyli stosując zawężoną definicję teorii umysłu, wskazane badania, nie dały odpowiedzi na zadawane przez eksperymentatorów pytanie, jednak pokazały, że inne zwierzęta zdolne są do przewidywania zachowania innego na drodze innych mechanizmów. Nie oznacza to jednak, że wyniki uzyskane przez Premacka i Woodruffa (1978) przemawiają za twierdzeniami wskazującymi, że ZINL nie posiadają teorii umysłu. Wręcz przeciwnie, ukazują jak bardzo ich zdolność do *mindreadingu*, podobna jest do tej występującej u ludzi. Sami autorzy wskazują nawet, że uzyskane wyniki sugerują brak znaczących różnic międzygatunkowych.

Uczestniczką eksperymentów Premacka i Woodruffa (1978) była szympanśica Sara, która w tamtym czasie miała czternaście lat. Urodziła się w Afryce, natomiast do laboratorium trafiła, gdy miała rok. Przez lata pobytu w ośrodku nauczyła się uproszczonego języka wizualnego, co pozwalało jej na komunikację z ludźmi. Jej treningi odbywały się pięć dni w tygodniu przez dziesięć lat. Dotyczyły głównie rekonstrukcji rozmontowanych obiektów, a także rozumowania przyczynowego. Sara miała duże doświadczenie z telewizją komercyjną, jest to ważne, ponieważ jej zadaniem było oglądanie nagrań wideo. W pierwszym eksperymencie prezentowano jej 4 filmy, trwające około 30 sekund. Film przerywany był ok. 5 sekund przed końcem, czyli nie pokazywano jej ostatniego fragmentu przedstawiającego rozwiązanie sytuacji. Warto także zauważyć oczywiste podobieństwo do problemów, które prezentowano Sarze, do tych, które Wolfgang Köhler stawiał przed swoimi szympankami. Na nagraniu można było zobaczyć aktora (A), znajdującego się w klatce, takiej jak zwykle trzyma się zwierzęta biorące udział w badaniach. Nagranie wyraźnie wskazywało, że (A) chce zdobyć pożywienie, które jest dla niego niedostępne na jeden z czterech sposobów:

- a. Banany były przymocowane do sufitu, przez co znajdowały się poza zasięgiem (A).
- b. Banany znajdowały się poza klatką, na tyle daleko, że (A) nie może sięgnąć go ręką.
- c. Banany znajdują się poza klatką, jednak dostęp do nich zasłania duże pudło

wewnątrz klatki.

- d. Podobny do poprzedniego, z tą różnicą, że pudło wewnątrz obciążone jest blokami cementu.

Po zatrzymaniu nagrania trener pilnujący przebiegu badania wkładał do kartonowego pudła dwa zdjęcia oraz wychodził z pomieszczenia. To pozwalało na wyeliminowanie efektu eksperymentatora. Wspomniane fotografie wykonane zostały wcześniej (podczas tworzenia nagrań wideo). Przedstawiały one (A) zaangażowanego w rozwiązanie problemów prezentowanych na nagraniu. Na jednym zdjęciu widać było, jak (A) wchodzi na pudło, aby dostać się do bananów nad swoją głową, na innym (A) leżał na boku, próbując przysunąć banany prętem. Na trzecim zdjęciu aktor przesuwiał pudło, a na czwartym ściągał bloki cementu leżące na przeszkodzie. Zadaniem Sary było wyciągnięcie zdjęć z pudełka, wybranie jednego z dwóch oraz położeniu go na określonym miejscu koło telewizora. Następnie sygnalizowała wykonanie zadania poprzez dzwonienie dzwonkiem. Gdy trener wracał do pomieszczenia, nagradzał słownie Sarę za każdym razem, gdy dokonała dobrego wyboru oraz zwracał jej uwagę, gdy szympanśica się pomyliła. Jak raportują autorzy, do takiej pomyłki doszło tylko 3 razy, za każdym razem w przypadku nagrania z cementowymi blokami. Uważa się, że wskazane pomyłki zaszły przez różnicę w możliwościach siłowych szympanśów i ludzi. Dla Sary przesunięcie pudła, bez usuwania ciężkich elementów, nie stanowiłoby trudności, w przeciwieństwie dla ludzkiego aktora. W artykule znajdziemy 3 interpretacje takiego wyniku, z czego jedna odrzucana jest na samym początku, ze względu na małe prawdopodobieństwo jej słuszności (Premack, Woodruff, 1978, str. 516). Nie będę jej poświęcać zbyt wiele miejsca, odnosi się ona do asocjacionizmu, czyli krótko mówiąc dotyczy dopasowywania elementów fizycznych na zdjęciu do odpowiadających im rekwizytów znajdujących się na nagraniu. W takiej sytuacji nie potrzebne są złożone procesy *mindreadingu*, wystarczy mechanizm generalizacji, czyli podstawowy mechanizm kategoryzacji. Dwie pozostałe interpretacje uznane są za prawdopodobne oraz na podstawie przeprowadzonego eksperymentu nie można w pełni wykluczyć żadnej z nich. Pierwsza głosi, że Sara korzysta ze zdolności do czytania umysłów innych, przypisując (A) co najmniej dwa stany mentalne: cele i intencje⁹⁷.

⁹⁷ Oczywiście w tym miejscu należy zaznaczyć, że używam tutaj terminologii Premacka i Woodruffa (1978), co wymaga użycia stosowanego przez nich skrótu myślowego. W terminologii przyjętej w niniejszej pracy przypisanie tak prostych stanów, jak cele, intencje czy motywacje zachodzą w ramach detektora intencjonalności.

W takiej interpretacji szympanśica widzi, że musi rozwiązać problem, ponieważ zakłada, że (A) chce banana oraz próbuje go zdobyć. Druga interpretacja odnosi się do prostszego procesu leżącego u podstaw decyzji uczestniczki, a konkretnie empatii. W takiej sytuacji Sara nie tyle korzysta z mechanizmów *mindreadingu* (ID czy EED), odpowiadając na pytanie, *co robi aktor?* a używa empatii, aby odpowiedzieć na pytanie, *co ja bym zrobiła na miejscu aktora?* Autorzy zauważają, że w każdej z wymienionych sytuacji szympanśica dokonuje atrybucji stanów mentalnych, aby odpowiedzieć na pierwsze ze wskazanych pytań, Sara musiałaby wytworzyć reprezentację epistemiczną stosowną do stanu mentalnego (A), natomiast do odpowiedzi na drugie pytanie wystarczy, że jednostka rozpozna sytuację, w jakiej znajduje się aktor oraz odniesie ją do siebie, zastanawiając się, *co ja bym zrobiła?* Różnica jest fundamentalna, gdyż jeżeli Sara odpowiada na pytania przy użyciu empatii, należy uznać, że na podstawie wskazanych eksperymentów nie można orzec, czy szympansy posiadają zdolność do tworzenia teorii umysłu⁹⁸, co było celem eksperymentu.

Chcąc rozstrzygnąć, która z interpretacji — ta odnosząca się do empatii czy ta odnosząca się do teorii umysłu, jest bardziej prawdopodobna, autorzy przeprowadzili kolejny eksperyment. Założono, że jeżeli Sara korzysta z pierwszego z wymienionych mechanizmów, wpływ na jej decyzję będzie miała tożsamość aktora. Wykorzystali tym samym sympatię, ściśle związaną z empatią. Skonstruowano dwa kolejne eksperymenty, obydwa były bardzo podobne do poprzedniego, a jego uczestniczką ponownie była Sara. W tym wypadku celem eksperymentu było sprawdzenie, na jakie pytanie odpowiada zwierzę biorące udział w eksperymencie:

- a. Co zrobiłby aktor?
- b. Co powinien zrobić aktor?
- c. Jaką czynność chciałabym zobaczyć u aktora?

W każdym z dwóch eksperymentów Sara oglądała nagrania wideo, gdzie ludzki aktor:

- a. Próbował wydostać się z klatki (WK),
- b. Wchodził w interakcję z zepsutą grzałką, sygnalizując, że jest mu zimno, poprzez Przyciskanie ramion do piersi i dreszcze (IG),

⁹⁸ Wyjaśniając ten wniosek przy użyciu terminologii przyjętej w niniejszej dysertacji, badania Premacka i Woodfuffa (1978) nie potwierdzają, że szympansy, do przypisywania stanów mentalnych innym, korzystają ze zdolności czytania umysłów innych w perspektywie Baron-Cohena, czy też, stosują w tym celu inny mechanizm (kompensacyjny) z obszaru poznania społecznego np. empatię.

- c. Próbował grać na odłączonym od prądu fonografie (GF),
- d. Nie mógł umyć brudnej podłogi, ponieważ wąż, którym próbował to zrobić, był nieprawidłowo podłączony do kranu (OW).

Nagrania urywały się kilka sekund przed zakończeniem, a eksperymentator prezentował Sarze trzy fotografie:

- a. W przypadku problemu (WK), klucza: zwykłego, skręconego i zepsutego,
- b. W przypadku (GF) i (OW), kabla i węża: właściwie podłączony, nie podłączony, podłączony, ale przecięty,
- c. W przypadku (IG), kartkę papieru: niezapaloną, zapaloną oraz spaloną.

Różnicą między eksperymentami był dobór aktorów, w pierwszym tą rolę pełnił jej ulubiony trener (A1). Natomiast do drugiego wybrano innego (A2), Sara go знаła, jednak nie wykazywała wobec niego wyraźnych afektów. W przypadku nagrań z udziałem (A1) uczestniczka wskazywała znaczną większość poprawnych odpowiedzi. Pomyliła się tylko raz przy nagraniu (a), jednak sami autorzy twierdzą, że przyczyną tego błędu jest jakość zdjęć kluczy. Jednak w przypadku nagrań z (A2) tylko 2 na 8 odpowiedzi, było prawidłowych. Taki wynik oznacza, że zachowanie Sary było odpowiedzią na pytanie: *jaką czynność chciałbym zobaczyć u aktora?* Interesujące w kontekście niniejszej dysertacji zdaje się fakt, że szympanśica Sara posługiwała się heurystyką afektu opisaną we wcześniejszych rozdziałach.

Premarck i Woddruff podejmują kilka prób wyeliminowania preferencji dotyczących aktorów, jednak wyniki uzyskane z ich eksperymentów pozostały niejednoznaczne. Zatem nie rozstrzygają kwestii przekonań, w tym wiedzy, a tym bardziej teorii umysłu, gdyż wyniki odzwierciedlają oczekiwania Sary, nie zaś mechanizmy *mindreadingu*. Jednak opisany artykuł stanowił punkt wyjścia do dalszych rozważań na temat posiadania teorii umysłu u zwierząt innych niż ludzie. Autorzy sugerują, aby w przyszłych eksperymentach rolę aktorów pełniły inne szympansy. Sugestia ta jest istotna, ponieważ w opisywanych badaniach zadaniem uczestniczki było przypisanie stanów mentalnych przedstawicielowi innego gatunku. W przypadku, gdy aktorem będzie szympanś, sprawdzamy zdolność do przypisywania stanów mentalnych przedstawicielowi tego samego gatunku, a to jest zdolność, którą ludzkie dzieci nabywają około 4 roku życia.

W tym miejscu należy zaznaczyć, że omawiane badanie spotkało się z często słuszną krytyką metodologii przyjętej przez autorów. Przede wszystkim zgadzam się z zarzutem dotyczącym braku warunku kontrolnego, jednak pozostałe zarzuty mają

charakter *ad ignorantiam*, w związku z tym postanowiłam przeanalizować odpowiedź Sue Savage-Rumbaugh, Duane Rumbaugh i Sarah Boysen (1978) zamieszczoną w *Behavioral and Brain Sciences*. Zarówno Sue Savage-Rumbaugh oraz jej były mąż Duane, znani są z badań nad zdolnościami językowymi jak i poznawczymi bonobo, koncentrujących się na wykazywaniu zaawansowanych zdolności poznawczych naczelnych, sugerując, że mogą one posiadać pewne aspekty teorii umysłu (Savage-Rumbaugh, Rumbaugh, Boysen, 1980; Savage-Rumbaugh, Shanker, Taylor, 1998). W swoim komentarzu z 1978 roku prezentują szeroką krytykę, na której łamach, przyjmują stanowisko głoszące, że nie ma dowodu na to, że Sara postrzegała scenariusze przedstawione na nagraniach jako problem do rozwiązania (*problem-to-be-solve*) ani, że postrzegała swoje wybory jako rozwiązanie na przedstawiony w formie filmu problem. Argumentują, że Premack i Woodruff (1978) nie przedstawili dowodu, wskazującego, że szympanśica biorąca udział w ich eksperymencie rozumie czym są elementy przedstawionych jej problemów oraz samą naturę rozwiązania. Co więcej, wskazują, że nie ma dowodu na to, że gdyby Sara była postawiona przed problemem takim jak aktor nagrań, potrafiłaby go rozwiązać. Oczywiście nie jest wiadome, czy Sara miała doświadczenie z eksperymentami prowadzonymi w paradygmacie zapoczątkowanym przez Köhlera, jednak wątpię w trafność zastosowania argumentu *ad ignorantiam* w przypadku omawianych eksperymentów. W ramach pierwszej serii nagrań prezentowanych szympanśicy przedstawiono jej klasyczne problemy wykorzystywane w paradygmatach stosowanych z udziałem zwierząt innych niż ludzie. Natomiast liczne badania, wykazują, że nie tylko szympansy zdolne są do rozwiązania takich problemów (Köhler, 1949; Hopper, Kurtycz, Ross, Bonnie, 2015; Melis, Tomasello, 2019). Dlatego zdaje się raczej, że nie ma wiarygodnej przesłanki, przemawiającej za tezą, że Sara nie byłaby w stanie sama rozwiązać prezentowanych jej problemów. To sprawia, że uznaję za bardziej wiarygodne stanowisko, w ramach, którego szympanśica biorąca udział w badaniu Premacka i Woodruffa (1978), byłaby w stanie rozwiązać sama przynajmniej zdania prezentowane jej w pierwszej serii. Nie widzę jednak powodów do tego, aby wątpić w moje stanowisko także w kwestii innych zadań prezentowanych jej w omawianych eksperymentach. Jeśli chodzi o argument wysuwany przez Savage-Rumbaugh, Rumbaugh, Boysen (1978) głoszący, że Sara nie rozumie, że na nagraniach przedstawiany jest jej problem do rozwiązania, należy uznać, że dynamika przedstawionych jej wideo przekracza jej możliwości poznawcze. Argument ten również ma naturę *ad ignorantiam*, także w tym wypadku uważam go za

nietrafiony, gdyż należy wziąć pod uwagę, że szympanśica biorąca udział w badaniu miała wcześniejsze doświadczenie z telewizją, co raczej sugeruje, że przynajmniej na podstawowym poziomie rozumiała, na czym polega oglądanie nagrań wideo⁹⁹. A także, że przypisywanie agencyjności jest domeną procesów znacznie prostszych niż ToMM. Jednak autorzy *Sarah's problems with comprehension* zaznaczają — Premack i Woodruff (1978), nie udowodnili, że Sara przypisuje agencyjność aktorowi znajdującemu się w sytuacji problemowej co wskazuje na jej rozwiązanie poprzez wybór fotografii. W konsekwencji odwołują się do zasady Morgana, stawiając tezę, że lepiej przyjąć interpretacje odwołujące się do prostszych mechanizmów poznawczych, a zatem założyć, że Sara nie wykorzystywała złożonych procesów z zakresu poznania społecznego. Wskazują, że u podstaw jej zachowania leżą proste mechanizmy asocjacyjne, sugerując, że szympanśica mogła osiągnąć tak wysoką trafność odpowiedzi, kierując się regułą dopasowania na podstawie podobieństwa fizycznego elementów znajdujących się na fotografii z tymi obecnymi na nagraniu, dlatego gdy używają określenia „poprawne” w kontekście odpowiedzi Sary, umieszczają je w cudzysłowie.

Savage-Rumbaugh, Rumbaugh, Boysen (1978) zauważają, że w drugiej serii nagrań prezentowanej szympanśicy, jej zadaniem było wybranie obiektu umożliwiającego rozwiązanie problemu aktora. Oznacza to, że Sara nie mogła skorzystać z mechanizmu wskazanego wcześniej. Dlatego argumentują, że w tym wypadku skorzystała ze strategii dopasowania do próbki, z którą miała wcześniejsze doświadczenia (Premack, Woodruff, 1976). To oznacza, że szympanśica kojarzyła klucz z zamkiem, wąż z kranem oraz kabel z gniazdkiem. Jedynym zastanawiającym przypadkiem dla autorów omawianego artykułu był przykład (b). Nie było dowodu na to, że Sara wiedziała, jak działa grzałka, oraz że wymaga ona rozpalenia, by dawała ciepło. Dlatego nie można było założyć, że u podstaw jej zachowania leży asocjacja: palący się knot — urządzenie dające ciepło. Argumentują jednak, że Sara wie, że knot z płomieniem bardziej pasuje do grzałki niż klucze, krany czy gniazdka. Na potwierdzenie swojej tezy autorzy przytaczają wyniki własnego eksperymentu, w którym szympansy wykonywały zadania oparte na dopasowaniu do próbki. Ich

⁹⁹ Co prawda autorzy omawianego eksperymentu nie doprecyzowują, jaki charakter ma doświadczenie Sary, więc nie ma pewności, czy było ono wystarczające do zrozumienia prezentowanego przez nich badania.

zadaniem było łączenie przedmiotów w pary na podstawie skojarzeń. W celu wykonania zadania zwierzęta wybierały obiekt z grupy przedmiotów, w której znajdowały się takie, które były znane uczestnikom badania oraz takie, których nie znały. Takie działanie nie wymagało żadnych zdolności z zakresu poznania społecznego, w tym ToMM. Na podstawie wykonanego przez siebie eksperymentu, autorzy stwierdzają, że fakt, iż szympansy zdolne są do dopasowania do próbki, w zadaniu niewymagającym atrybucji stanów mentalnych, stanowi wystarczającą przesłankę dla twierdzenia, że szympanse biorące udział w eksperymencie Premacka i Woodruffa (1978), mogła korzystać z mechanizmu wskazanego przez autorów *Sarah's problems with comprehension*, a co za tym idzie, wniosek Premacka i Woodruffa (1978) jest nie do utrzymania.

W przedstawionej wyżej argumentacji ponownie brak dowodu na to, że Sara korzystała ze zdolności z zakresu poznania społecznego oraz możliwość wykorzystywania w tym celu prostszych mechanizmów, stanowi dla autorów wiarygodną przesłankę do stwierdzenia, że lepszą interpretacją wyników prezentowanych w omawianym przez nich artykule, jest odniesienie do prostszych mechanizmów, jakimi są reguły dopasowania do próbki. Jest to dość interesujące, ponieważ konsekwencją tej interpretacji jest odmówienie uczestnicze eksperymentu Premacka i Woodruffa (1978) najprostszego ze wskazanych przez Baron-Cohena mechanizmu, pozwalającego na identyfikowanie obiektów sytuacji jako agentów, czyli detektora intencjonalności. Dla przypomnienia ID jest ewolucyjnie starszy niż gatunek ludzki, za czym świadczy m.in. fakt, że już noworodki wykazują przejawy jego obecności. Co więcej, szympansy są gatunkiem, którego strategią przetrwania jest życie w grupie, co sprawia, że odmawianie im tak prostego mechanizmu zdaje się nie mieć podstaw teoretycznych. Jak wcześniej wskazałam, ID jest mechanizmem omylnym, na jego drodze zdarza się przypisywanie agencyjności nie-agansom, korekcja tych informacji następuje w późniejszych etapach *mindreadingu*. Dane płynące z detektora intencjonalności biorą udział także w przetwarzaniu innych kluczowych informacji jak np. kierunek ruchu, co pozwala określić czy istnieje możliwość zderzenia z obiektem oraz zaplanować potencjalny unik. Dlatego uważam, że odmawianie istnienia tego mechanizmu u zwierząt innych niż człowiek, wydaje się w świetle współczesnej wiedzy nie do utrzymania. I chociaż teoria Baron-Cohena powstała długo po artykule opublikowanym przez Savage-Rumbaugh, Rumbaugh, Boysen (1978), chcę podkreślić, że w czasach im współczesnych dostępna była literatura, dotycząca detektora

intencjonalności (Blest, 1957). Dlatego wydaje się, że Savage-Rumbaugh, Rumbaugh, Boysen (1978) ignorują pewne przesłanki niezgodne z powszechnie panującym przekonaniem, że tylko człowiek zdolny jest do teorii umysłu w sensie szerokim.

W ostatnim zestawie nagrań prezentowanych Sarze szympanśica uzyskiwała wysoką trafność odpowiedzi, gdy na nagraniu widziała swojego ulubionego trenera, natomiast w przypadku A2, jej trafność spadała. Premack i Woodruff (1978) interpretują taki wynik jako niezrozumienie eksperymentu, wskazując, że zwierze kierowało się preferencją, powodującą, że częściej przypisywała trafne zachowanie aktorowi, którego darzyła pozytywnym afektem. Savage-Rumbaugh, Rumbaugh i Boysen (1978) zgadzają się z autorami omawianego przez nich artykułu co do tego, że uczestniczka nie rozumiała stawianych przez Premacka i Woodruffa (1978) zadań, podważają jednak podstawę wskazanego niezrozumienia. Oprócz przytaczanych wcześniej zarzutów dotyczących tego, czy Sara rzeczywiście przypisywała agencyjność obserwowanemu na nagraniu aktorowi, wysuwają kolejne. Wskazują, że nie wiadomo jakie problemy były stawiane przed A2, ponieważ Premack i Woodruff (1978) nie udostępnili w artykule materiałów, prezentujących nagrania i fotografie wykorzystane w tym eksperymencie. A co za tym idzie, przyznają, że ciężko im powiedzieć co Sara widziała. Sugerują jednak, że w nowym eksperymencie problem prezentowany na nagraniu lub obiekty na fotografiach były inne niż w pozostałych seriach, dlatego szympanśica nie była w stanie wybrać odpowiedniej alternatywy. Uważam jednak, że bardziej prawdopodobna jest interpretacja Premacka i Woodruffa (1978). Zakładam, że Sara wyposażona w ID oraz w doświadczenie z telewizją rozumiała, że przedstawiony na nagraniu agent znajduje się w sytuacji problemowej oraz dostrzegała, że fotografie reprezentują rozwiązanie lub jego możliwość. Moje stanowisko ukształtowane jest nie tylko na podstawie koncepcji Lesliego (1984) i Baron-Cohena (2011), ale także przez późniejsze badania, które omówię w dalszej części niniejszego rozdziału.

Savage-Rumbaugh, Rumbaugh i Boysen (1978) odnoszą się także do zdań, w których Sara wskazywała na obiekty nieuszkodzone, a nie te zepsute. Ich zdaniem, uczestniczka nie rozumiała, że tylko nieuszkodzone przedmioty mogą rozwiązać problem aktora. Wskazują, że omawiana seria została zaprezentowana Sarze, już po tym, jak dokonała dopasowania fizycznego (w poprzedniej serii). Dlatego interpretują jej zachowanie jako wyuczone za pomocą prostego warunkowania, twierdząc, że ta wybrała po prostu znane jej fotografie, zamiast tych nowych. Co ciekawe autorzy wspominają, że aby uznać interpretację Premacka i Woodruffa (1978), szympansy

musiałyby odpowiedzieć poprawnie przy pierwszej próbie, ponieważ inaczej ich zachowanie należy tłumaczyć w ramach wybierania alternatyw, wskazanych w poprzednim „powtórzeniu”, jeśli została za nie nagrodzona. Co prawda Sara osiągnęła wysoką trafność już w pierwszych próbach, pomimo to autorzy omawianej krytyki przyjmują, że lepszym wyjaśnieniem jest odniesienie się do prostszych procesów poznawczych takich jak dopasowanie fizyczne i do próbki, nie odwołując się do bardziej złożonych procesów.

Chociaż nie podzielam interpretacji przyjmowanych przez autorów *Sara's problem with comprehension* (Savage-Rumbaugh, Rumbaugh i Boysen, 1978), to doceniam wkład opisywanego artykułu w rozwój technik badawczych z zakresu poznania społecznego. Należy bowiem zauważyć, że argumenty wysuwane przez Savage-Rumbaugh, Rumbaugh i Boysen (1978) nie są jednoznacznie fałszywe. Wręcz przeciwnie, prawdą jest, że w czasach współczesnych wskazanym autorom nie posiadano licznych zaplecza teoretycznego i empirycznego, wspierającego założenia Premacka i Woodruffa (1978). Co więcej, jak wskazałam wcześniej, brak grupy kontrolnej stanowi istotny błąd metodologiczny, a także nie udało im się wyeliminować heurystyki afektu, co sprawia, że na podstawie tych badań nie można jednoznacznie stwierdzić, że szympansy zdolne są do przypisywania innym prostych stanów wolicjonalnych.

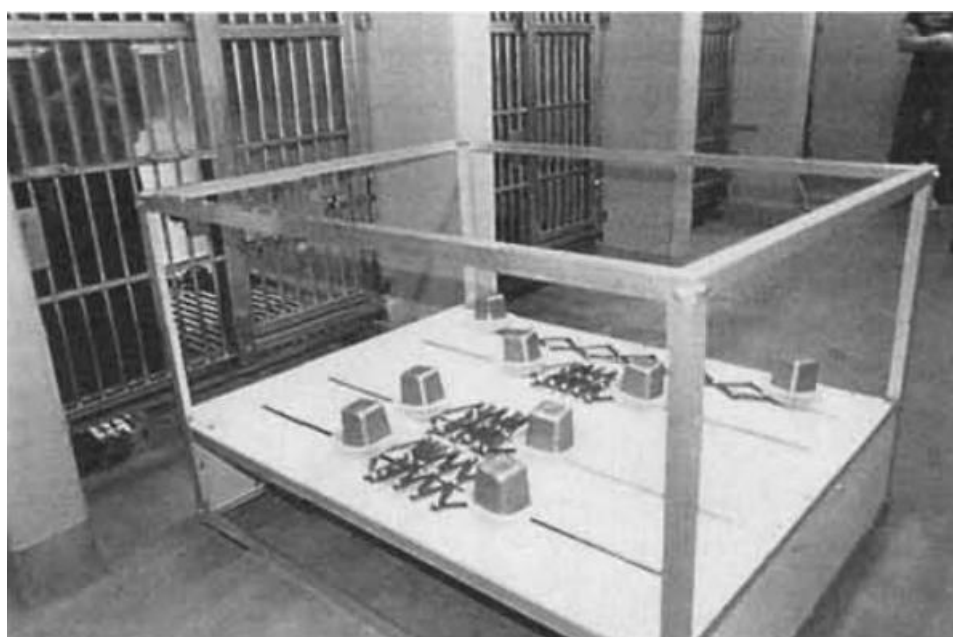
5.1.2. Badania Daniela Povinelliego i Kurta Nelsona

Nazwisko Povinelli może wydawać się zaskakujące w niniejszej części analizy, jednak zanim autor ten przyjął konsekwentnie sceptyczne stanowisko wobec zdolności zwierząt innych niż ludzie do teorii umysłu, opublikował klasyczny już artykuł przedstawiający pionierską metodologię opartą na wyborze obiektów wskazywanych przez eksperymentatora. Bezpośrednio przed eksperymentem uczestnicy brali udział w pilotażowym badaniu kooperacyjnym, w którym uczyli się wskazywać wabione kubki oraz reagować na sygnały przekazywane przez eksperymentatora. W badaniu brały udział cztery szympansy: Sheba (7 lat), Kermit (9 lat), Darrel (9,5 roku) oraz Sarza (28 lat). Ostatnia z uczestniczek brała wcześniej udział w eksperymentach Premacka i Woodruffa (1978), pozostali uczestnicy również mieli doświadczenie w badaniach dotyczących zdolności poznawczych. Wszystkie szympansy znały eksperymentatorów, których w badaniu było trzech: jeden pełnił rolę *wiedzącego* (wabienie aparatury), drugi *zgadującego* (wskazywanie losowej tacki), a trzeci

obsługiwał tekturową przesłonę, zakładaną na aparaturę w momencie umieszczania jedzenia. Istotnym elementem procedury była zamiana ról między *wiedzącym* a *zgadującym*, co miało na celu eliminację wpływu preferencji uczestników wobec konkretnych eksperymentatorów. Autorzy jednak nie podają informacji dotyczących relacji poszczególnych szympanśów z badaczami. Warto również zaznaczyć, że uczestnicy eksperymentu pozostawali w swoich klatkach (*home cages*), czyli wykonywali zadanie w dobrze znanym sobie miejscu.

W badaniu wykorzystano zmodyfikowaną wersję aparatury zastosowanej pierwszy raz w 1962 roku (Masona, Hollisa, 1962). Była to konstrukcja przypominająca stół, składająca się ze strony operatora oraz strony informatora (zdjęcie 9). Po stronie operatora znajdowały się cztery uchwyty umożliwiające poruszanie tackami. Na każdej tackie umieszczono kubek ustawiony do góry nogami, który przykrywał potencjalne pożywienie, dzięki takiemu zabiegowi nie było widać jaka, jest ich zawartość.

Zdjęcie 9. Aparatura wykorzystana w eksperymencie opisanym przez Povinelli’ego i Nelsona (1990)



[Źródło (Povinelli, Nelson, Boysen, 1990, str. 205)]

Eksperyment składał się z trzech faz, pierwsza z nich trwała 6 tygodni, podczas których uczestnicy przechodzili 10 prób każdego dnia. *Zgadujący* wychodził z pokoju, a podczas jego nieobecności *wiedzący* umieszczał kasek pod jednym z kubków. Następnie siadał w ustalonej odległości od aparatury, przyjmując neutralną pozycję

i kierując wzrok w stronę przedniej części klatki uczestnika. Po powrocie zgadującego obaj eksperymentatorzy pokazywali za pomocą gestu (Zdjęcie 10), gdzie znajduje się pożywienie, a zadaniem uczestników było poruszenie tacki wskazywanej przez *wiedzącego*. Zachowanie szympansów szybko się ustabilizowało, choć autorzy nie podają dokładnej liczby prób potrzebnych poszczególnym uczestnikom do konsekwentnego wskazywania właściwego eksperymentatora. Informują jednak, że kolejny etap w postaci wprowadzenia dodatkowego elementu różnicującego wprowadzono pod koniec trzeciego tygodnia, ale tylko dla Sheby, Sary i Darrela. Natomiast Kermitowi zaprezentowano go dopiero prawie tydzień później, co sugeruje, że miał największe problemy ze stawianym przed nim zadaniem. Wspomniany bodziec różnicujący stanowiły niebieskie czapki, które nosili *wiedzący* oraz osoba obsługująca przesłonę. Podczas zamiany ról eksperymentatorzy wymieniali między sobą czapkę na oczach uczestnika, aby ułatwić im identyfikację właściwego wskazania. Wprowadzenie bodźca różnicującego, mającego ułatwić zadanie sugeruje, że wyniki uczestników nie były w pełni satysfakcjonujące. Autorzy podkreślają jednak, że analiza wyników przed wprowadzeniem

Zdjęcie 10. Gest wskazywania wykorzystany w eksperymencie Povinelli’ego i Nelsona (1990).



[źródło: Povinelli, Nelson, Boysen, 1990, s.206]]

czapek ujawniła wyraźny wzorzec prawidłowego zachowania, zwłaszcza u Darrela i Sheby. Ponadto zauważono, że wyniki nie poprawiały się z tygodnia na tydzień, co stawia pod znakiem zapytania hipotezę o wyuczeniu się przez szympansy prostych reguł asocjacyjnych (Shettleworth, 2010, s. 442).

Druga faza eksperymentu miała na celu sprawdzenie czy szympansy rzeczywiście rozumiały zadanie oraz, czy ich wybory były uzależnione od obecności dodatkowych wskazówek. W tym celu usunięto niebieskie czapki, aby sprawdzić, czy uczestnicy nadal będą kierować się wskazaniem osoby, która miała dostęp do informacji o ukrytym pożywieniu. Wyniki pokazały, że eliminacja czapek nie wpłynęła znacząco na wybory szympanów, co sugeruje, że ich sukces nie był zależny od wskazówek różnicujących. Najslabiej w tej fazie wypadła Sarah – choć wciąż wykazywała preferencję wobec eksperymentatora — *wiedzącego*, odnotowano istotny spadek jej wyników w porównaniu do tych z poprzedniej fazy. Z kolei Darrel uzyskał wyniki nieco gorsze od Sheby i Kermita, jednak wciąż znacznie przekraczające granice losowości.

Trzecia faza eksperymentu była kluczowa, gdyż jej celem było sprawdzenie elastyczności zachowania uczestników oraz ich zdolności do przenoszenia zdobytej wiedzy na nową sytuację. W tej fazie *wiedzący* i *zgadujący* stali obok siebie, przy czym drugi z wymienionych, pozostawał w pomieszczeniu, ale w kluczowym momencie zasłaniał sobie głowę nieprzezroczystą papierową torbą. Tacki były wabione przez eksperymentatora obsługującego przesłonę. Ta modyfikacja pozwalała rozstrzygnąć między podejściem mentalistycznym a behawiorystycznym. Z perspektywy mentalistycznej szympansy powinny rozumieć, że widzenie prowadzi do wiedzy, natomiast podejście behawiorystyczne zakładało stosowanie prostych reguł, takich jak „wybieraj osobę, która była w pomieszczeniu” albo „wybierz osobę, która umieszczała jedzenie pod kubkiem”. Gdyby uczestnicy kierowali się którąś z wymienionych reguł, należałoby oczekiwać błędnych wyborów po zmianie procedury. Wyniki pokazały, że Darrel, Sheba i Kermit natychmiast poprawnie zastosowali wcześniejszą strategię, wykorzystując wskazania eksperymentatora, który widział proces ukrywania pożywienia, i ignorując wskazania osoby, która w kluczowym momencie zasłaniała twarz torbą. Co więcej, zaobserwowano, że uczestnicy często siadali przed wiedzącym jeszcze przed wykonaniem gestu wskazywania. Dodatkowo, w niektórych przypadkach, gdy przysłona wciąż zakrywała aparaturę, szympansy chwyciły za niewłaściwy uchwyt, ale po jej zdjęciu i dokładnym zauważeniu gestu wiedzącego, korygowały swoje

zachowanie. Wyniki te dostarczają wsparcia dla perspektywy mentalistycznej, sugerując, że szympansy mogły kierować się informacjami o dostępie do wiedzy, a nie jedynie obecnością w pomieszczeniu.

5.1.3. Badania Calla i Tomasello

W badaniach nad teorią umysłu u gatunków zwierzęcych innych niż człowiek, znaczącą rolę odgrywa Michael Tomasello. Przez wiele lat konstruował i interpretował eksperymenty w nurcie behawioralnym, wykazując przy tym sceptycyzm wobec możliwości przypisywania ZINL złożonych stanów mentalnych, jednak jego stanowisko, choć ostrożne, pozostawało wyważone. Na początku odmawiał gatunkom pozaludzkim wielu zdolności poznawczych uznawanych za typowo ludzkie, w tym zdolności do posiadania teorii umysłu (Tomasello, 2015). Jednak w miarę pojawiania się nowych doniesień empirycznych weryfikował własne stanowiska i wykazywał gotowość do ich korekty (Call, Tomasello, 2008; Tomasello, 2015). To podejście – łączące sceptycyzm z gotowością do zmiany zdania – sprawia, że historia jego zmieniającego się stanowiska wobec poszczególnych narracji dotyczących *mindreadingu* stanowi fascynujący przekrój, przedstawiający, jak rozwój w obszarze badań empirycznych, przyczynił się do odrzucenia wielu twierdzeń dotyczących wyjątkowości człowieka. Dlatego uważam, że historia ta zasługuje na osobne opracowanie. Choć analiza ta wykracza poza ramy teoretyczne niniejszej dysertacji, stanowi ciekawy punkt odniesienia dla dalszych kierunków badań.

W szczególności interesujące są zmiany widoczne w stanowisku Tomasella, zwłaszcza w kontekście wspieranych przez niego argumentów dotyczących teorii umysłu oraz — w nieco mniejszym stopniu — zdolności do odczytywania stanów mentalnych innych. Przez lata Tomasello i jego współpracownicy podkreślali, że teoria umysłu jest właściwa wyłącznie człowiekowi (Tomasello, 1999; Tomasello, Rakoczy, 2003). W nowszych pracach przyznaje się jednak, że u innych zwierząt — zwłaszcza naczelnych — można zaobserwować zachowania sugerujące zdolność do reprezentowania stanów mentalnych innych, choć nie osiągają one poziomu złożonych metareprezentacji drugiego rzędu (np. „on myśli, że ona wie, że...”); (Tomasello, Call, 2006; Tomasello, 2008; 2018).

W 2008 roku Call i Tomasello powracają do pytania postawionego przez Premacka i Woodruffa, prezentując artykuł, który stanowi zarówno podsumowanie

dotychczasowych ustaleń, jak i próbę precyzyjnego zdefiniowania, czym właściwie jest teoria umysłu. Autorzy przeciwstawiają się skrótowi myślowemu — o którym już wielokrotnie wspomniałam w swoich rozważaniach. Polega on na utożsamieniu teorii umysłu z *mindreadingiem*. Baron-Cohen (2009) identyfikuje źródło wskazanego problemu, w opisanym wyżej artykule Premacka i Woodruffa (1978). W tym celu Call i Tomasello (2011) zawężają zakres pojęciowy interesującej ich zdolności do reprezentowania przekonań innego – zwłaszcza wtedy, gdy przekonania te są fałszywe. Według tej definicji obserwator (O) może zostać uznany za podmiot posiadający teorię umysłu, jeśli potrafi przewidzieć zachowanie aktora (A) na podstawie przekonań (P), które – jak przypuszcza (O) – są niezgodne z rzeczywistością (Call, Tomasello, 2008; Mękarska, 2024). Tym samym wprowadzają rozróżnienie na teorię umysłu w sensie wąskim oraz szerokim, które już wielokrotnie przytaczałam w niniejszym opracowaniu. Tak ujęta teoria umysłu nie obejmuje więc całego spektrum *mindreadingu*, lecz koncentruje się na zdolności do przypisywania fałszywych przekonań. Zawężona definicja służy zwiększeniu precyzji metodologicznej i jednoznaczności interpretacyjnej.

Call i Tomasello (2008) opisują szereg eksperymentów, które miały na celu zbadanie, w jakim stopniu szympanse rozumieją stany mentalne innych osobników. Autorzy przytaczają zarówno wyniki eksperymentów inspirowanych badaniami Premacka i Woodruffa (1978), jak i obserwacje zachowań szympanów, które zdają się świadczyć o próbach manipulowania informacją – np. poprzez ukrywanie swojego zainteresowania pożywieniem przed rywalem. Takie zachowania można interpretować jako dowód na to, że szympanse przewidują zachowania innych osobników na podstawie ich przekonań – zwłaszcza tych związanych z wiedzą o świecie. Co ważne, eksperymenty zaprezentowane przez autorów *Does the chimpanzee have a theory of mind? 30 years later*, zostały skonstruowane w oparciu o paradygmat eksperymentalny wykorzystywany w badaniach z udziałem ludzkich dzieci. Zabieg ten zmniejsza ryzyko tendencyjnych interpretacji wyników, będących konsekwencją naturalnej preferencji grupy własnej. Jeżeli bowiem ZINL uzyskują wyniki porównywalne z ludzkimi dziećmi, a wyniki dzieci są interpretowane jako przejaw konkretnej zdolności poznawczej, wówczas nie ma podstaw, by odmawiać tej zdolności również ZINL – pod warunkiem, że procedura eksperymentalna została odpowiednio dostosowana do możliwości fizycznych i percepcyjnych uczestników (Call, Tomasello, 2008).

Kluczowymi badaniami prezentowanymi w artykule były te dotyczące

rozumienia celów i intencji, a także percepcji i wiedzy ¹⁰⁰ innych. Jednym z najważniejszych paradygmatów był tzw. „paradygmat konkurencyjny”, w którym szympanse musiały zdobyć pożywienie w obecności osobnika dominującego. Wyniki pokazały, że osobniki podporządkowane unikały jedzenia, które widział dominujący i wybierały to, które było poza jego zasięgiem wzroku – co sugeruje, że rozumiały, co inny osobnik może widzieć lub nie wiedzieć. Inne eksperymenty wykorzystywały zadania oparte na śledzeniu wzroku i gestów innych zwierząt. Szympanse były w stanie podążać za spojrzeniem człowieka lub innego szympansa, a także rozpoznawały, czy ktoś patrzy na obiekt, czy nie. W jednym z badań, gdy eksperymentator patrzył intensywnie w określone miejsce, szympanse najpierw podążały wzrokiem za jego spojrzeniem, a gdy niczego tam nie znajdowały, wracały spojrzeniem do eksperymentatora, jakby sprawdzały, co jest przedmiotem jego uwagi. Eksperymenty te pokazują, że szympanse wykazują złożone rozumienie stanów mentalnych innych, zwłaszcza w zakresie percepcji, wiedzy i celów. Nie wykryto zdolności do przetwarzania fałszywych przekonań, jednak autorzy zaznaczają, że w literaturze znaleźć możemy szereg opisów zachowań szympanów, które wymagają więcej niż tylko zrozumienia celów, intencji, motywacji i stanów percepcyjnych – np. taktyczne oszustwo.

Tomasello i Call (2008) argumentują także, że najbardziej miarodajne dane można uzyskać w warunkach rywalizacji o zasoby, zwłaszcza pożywienie – a nie w kontekście współpracy ¹⁰¹, który dominuje w wielu wcześniejszych badaniach.

¹⁰⁰ Należy jednak zaznaczyć, że taki rodzaj wiedzy przypisywany jest innemu na drodze informacji pochodzących z detektora kierunku parzenia oraz z uwspólniania uwagi.

¹⁰¹ W tym miejscu pragnę nadmienić, że Tomasello nadal podtrzymuje stanowisko, że zdolność do współdzielenia intencji (*shared intentionality*) jest unikalna dla ludzi i stanowi kluczową cechę odróżniającą człowieka od innych gatunków, w tym naczelnych. Uważa on, że współdzielenie intencji rozwija się u dzieci stopniowo w pierwszych miesiącach życia i obejmuje nie tylko rozumienie intencjonalnych działań innych, ale przede wszystkim motywację do dzielenia się emocjami, doświadczeniami i działaniami z innymi. Tomasello podkreśla, że podczas gdy naczelne rozumieją innych jako agentów celowo działających, to tylko ludzie angażują się w prawdziwą współpracę opartą na wspólnych celach i intencjach (Tomasello, Moll, 2010; Moore, Nichols, Mackey, 2020). Rozszerzeniem współdzielenia intencji jest kolejny etap rozwoju społecznego nazywany kolektywną intencjonalnością, który rozwija się ok. trzeciego roku życia, kiedy dzieci zaczynają myśleć w kategoriach my — oni i angażują się w normatywne zachowania społeczne. W późniejszym wieku (6–7 lat) dzieci stają się odpowiedzialnymi agentami, zdolnymi do argumentacji i rozumowania normatywnego.

Paradygmat konkurencyjny uznają za lepszy, gdyż stwarza sytuacje bardziej naturalne i motywujące, w których wyraźniej ujawniają się ewentualne zdolności do reprezentowania stanów umysłowych innych osobników. Choć argument Tomasella i Calla na rzecz paradygmatu kompetytywnego jest przekonujący, coraz więcej badań wskazuje, że również warunki kooperacyjne mogą dostarczać wiarygodnych i wartościowych danych na temat zdolności reprezentacyjnych zwierząt innych niż ludzie. Przykładowo, eksperymenty wykorzystujące zadania oparte na współpracy wykazały, że niektóre naczelne potrafią uwzględniać perspektywę partnera, dostosowując swoje działania do jego wiedzy, intencji czy ograniczeń percepcyjnych (Melis, Hare, Tomasello, 2006; Krupenye, Call, 2019). Co więcej, konteksty kooperacyjne odzwierciedlają inne istotne aspekty funkcjonowania społecznego, takie jak zdolność do dzielenia uwagi, wspólnego celu czy koordynacji działań, które również wymagają zaawansowanych kompetencji społeczno-poznawczych (Warneken, Tomasello, 2006; Carpenter, 2009). Z tego względu uważam, że niedocenianie tego rodzaju podejść, może prowadzić do zbyt wąskiego rozumienia zdolności reprezentacyjnych ZINL. Różnorodność paradygmatów – zarówno kompetytywnych, jak i kooperacyjnych – jest niezbędna dla pełniejszego uchwycenia złożoności tych zdolności oraz ich możliwych przejawów w zależności od kontekstu społecznego. Co ważne, autorzy podkreślają złożoność *mindreadingu* i wskazują, iż istotnym punktem zdawania sobie sprawy z tego, że inny ma swój własny świat wewnętrzny, jest przyjęcie, że ten świat różni się od *mojego*, czyli inny posiada także przekonania o świecie, niezgodne z własnymi. Dlatego uznali, że obserwator (O) posiada teorię umysłu, gdy zdolny jest do tworzenia przewidywań dotyczących zachowania aktora (A) na podstawie przekonań (P), co do których (O) przynajmniej podejrzewa, że są nieprawdziwe (Call, Tomasello, 2008; Mękarska, 2024).

5.1.4. Najnowsze dane empiryczne

W niniejszej części moich rozważań przedstawiam krótki przegląd najnowszych doniesień empirycznych dotyczących teorii umysłu u zwierząt innych niż ludzie. Jak wspominałam we wcześniejszych rozdziałach, interesująca mnie zdolność rozumiana jest jako tworzenie matareprezentacji przynajmniej [2] rzędu. Uznaje się ją za jeden z najbardziej złożonych mechanizmów poznawczych, a zarazem za cechę potencjalnie unikatową dla człowieka (*human-like*). Co istotne, w wielu przytaczanych niżej artykułów występuje skrót myślowy, który opisywałam już wielokrotnie we

wcześniejszych częściach niniejszej dysertacji, który utożsamia teorię umysłu z czytaniem umysłów innych. Dla przykładu w *Great apes use self-experience to anticipate an agent's action in a false-belief test* podkreślone zostaje, że autorzy używają terminu *teoria umysłu*, w sensie szerokim jako zdolność do przypisywania stanów mentalnych sobie i innym. Wskazują przy tym, że ciekawym aspektem interesującej ich zdolności jest przypisywanie fałszywych przekonań, co wymaga zdolności do tworzenia metareprezentacji epistemicznej [2] rzędu, która różni się od tego co sam obserwator wie. Taka decyzja podkreśla inspirację badaniami prowadzonymi przez Premacka i Woodruffa (1978), nie jest to bowiem jedyna zbieżność, tj. uczestnicy eksperymentów prowadzonych w 2019 roku tak, jak Ci w 1978, oglądali nagrania przedstawiające aktorów w różnych sytuacjach. Wyniki nowszych badań wykorzystują jednak metody niedostępne w roku 1978, korzystając z nieinwazyjnego *eye-trackingu* (Krupenye, Kano, Hirata, Call, Tomasello, 2016, 2017; Kano, Krupenye, Hirata, Call, 2017, Kanoa, Krupenye, Hirataa, Masaki Tomonaga, Call, 2019).

Zanim jednak przejdę do szczegółowej analizy eksperymentów zaprezentowanych przez zespół z Koyto, chcę omówić istotne problemy związane z paradygmatami, na których opierają się te eksperymenty — przede wszystkim, chodzi o spojrzenie wyprzedzające (*anticipatory looking*), czyli naturalną tendencję zwierząt do kierowania wzroku na miejsce, w którym spodziewają się, że agent podejmie działanie (Krupenye, Kano, Hirata, Call, Tomasello, 2017; Kanoa, Krupenye, Hirataa, Tomonaga, Call, 2019). W literaturze funkcjonują bowiem dwie konkurencyjne narracje wyjaśniające wyniki uzyskiwane we wskazanym paradygmacie używanym zarówno u dzieci ludzkich, jak i innych naczelnych. W ujęciu mentalistycznym uznaje się, że uczestnicy tych eksperymentów antycypują działania aktora, ponieważ przypisują mu określone stany umysłowe (np. fałszywe przekonania). W ujęciu behawioralnym to samo zachowanie tłumaczone jest przez odwołanie do prostych reguł asocjacyjnych — np. *jeśli ktoś coś widział, to wie, gdzie to jest*, bez konieczności przyjmowania, że uczestnik rozumie, iż agent może mieć przekonanie niezgodne z rzeczywistością (Perner, Ruffman, 2005). Czyli argumentują, że zamiast korzystać z mechanizmów *mindreadingu*, uczestnicy eksperymentów mogą wykorzystywać sprytne czytanie zachowania innego. Rozumiane jako repertuar wyuczonych, reguł warunkowych, których użycie nie wymaga odniesienia do przypisywania stanów mentalnych (Povinelli, Penn, 2007). Opiera się bowiem na

prostych skojarzeniach w rodzaju: *dominujący osobnik zazwyczaj wie, gdzie znajduje się jedzenie lub patrzący na jedzenie najczęściej później je zdobywa*. Perner podkreśla także znaczenie jednej z najczęściej przytaczanych reguł w kontekście paradygmatu fałszywych przekonań, głoszącą, że *agent będzie szukał obiektu, tam, gdzie go ostatnim razem widział* (Povinelli, Vonk, 2004; Perner, Ruffmana, 2005). Takie rozumowanie nie wymaga przypisywania fałszywych przekonań, wystarczy skojarzenie między percepcją a działaniem. Inne reguły skojarzeniowe postulowane w obszarze wskazanego mechanizmu sugerują, że uczestnik np., ludzkie dziecko, może rozumować: *jeśli ktoś coś widział, to wie, a jeśli nie widział, to nie wie* – ale to jeszcze nie oznacza, tego że rozumie, iż agent *myśli coś, co jest niezgodne z rzeczywistością* (Perner, Ruffman, 2005).

Niektóre pojęcia – takie jak *spojrzenie wyprzedzające* czy *detektor intencjonalności* – funkcjonują w literaturze w sposób znaczeniowo niejednoznaczny. Te same terminy bywają rozumiane inaczej w zależności od przyjmowanej perspektywy teoretycznej. W podejściu mentalistycznym uznaje się je za przejawy zdolności do przypisywania innym stanów mentalnych. Natomiast z perspektywy behawioralnej interpretowane są one jako wynik zastosowania prostych reguł skojarzeniowych lub warunkowych, niewymagających odniesienia do przekonań czy intencji. Argumentując, że tego rodzaju rozumowanie wystarcza do wyjaśnienia wyników, szczególnie w uproszczonych testach fałszywych przekonań – bez potrzeby postulowania istnienia zdolności metareprezentacyjnych. Zwracam jednak uwagę, że takie interpretacje pomijają aspekt elastyczności zachowań obserwowanych u uczestników eksperymentów, który trudno wyjaśnić na gruncie prostych schematów warunkowych¹⁰². Choć w artykułach empirycznych autorzy zazwyczaj doprecyzowują czy używają danego terminu w sensie *mindreadingu*, czy *smart behavior readingu*, to w tekstach przeglądowych i popularyzatorskich granice te często się zacierają. Powoduje to trudności w jednoznacznym odczytaniu znaczenia raportowanych wyników i może prowadzić do ich nadinterpretacji. Dlatego każdemu wyjaśnieniu danych uzyskanych na podstawie badań prowadzonych w paradygmatach odwołujących się do konwencji zjawiska spojrzenia wyprzedzającego, śledzenia kierunku patrzenia czy testu wiedzy i niewiedzy, powinna towarzyszyć epistemiczna ostrożność.

¹⁰² Temat ten rozwinę szerzej w części 5.2. *Perspektywa krytyczna w kwestii teorii umysłu u zwierząt innych niż ludzie: wybrane badania*.

Faktem jest, że brak narzędzi umożliwiających bezpośrednie badanie nieobserwowalnych procesów mentalnych istotnie ogranicza możliwości pomiaru interesującej mnie zdolności, jaką jest teoria umysłu, w sensie wąskim. Choć precyzyjne w pomiarze parametrów behawioralnych takich, jak kierunek i czas patrzenia, współczesne narzędzia badawcze wykorzystywane w psychologii rozwojowej i porównawczej nie dają możliwości bezpośredniego testowania metareprezentacji. Co ciekawe, widoczna jest wyraźna asymetria interpretacyjna wynikająca z przynależności gatunkowej uczestników. W przypadku ludzkich dzieci zakłada się, że zdolności metareprezentacyjne rozwijają się stopniowo, równolegle z mechanizmami sprytnego czytania zachowania innego, takimi jak detektor intencjonalności (ID) czy detektor kierunku patrzenia (EED), które wspólnie tworzą podstawę dla późniejszego, pełnego rozwoju zdolności z zakresu *mindreadingu* (Baron-Cohen, 1994; Baillargeon, Scott, He, 2010). Chociaż niektórzy badacze postulują, że wyniki badań z udziałem niemowląt i innych niewerbalnych zwierząt, również należy interpretować bez odwołania do stanów mentalnych, w tym sensie obecność *smart behavior reading*, występuje u nich, zanim pojawi się dojrzała zdolność metareprezentacyjna. Przyjęcie takiego założenia prowadzi do wniosku, że złożone formy metareprezentowania stanów umysłowych mogą wyłaniać się z prostszych mechanizmów, co w przypadku dzieci jest powszechnie akceptowane. Stanowisko to jest dobrze ugruntowane w szeroko przyjmowanych ramach teoretycznych np. w modelu Simon Baron-Cohena (2009) detektor intencjonalności rozumiany jako podstawowy mechanizm *mindreadingu*, pojawia się już u noworodków, co sugeruje, że stanowi podstawowe wyposażenie ludzkich systemów poznawczych, a co za tym idzie, prawdopodobnie możemy go także zaobserwować u innych zwierząt. Jednak zwolennicy narracji behawioralnych argumentują, że zdolności czytania sprytnego zachowania innych są szczytem ich możliwości poznawczych. Jak wielokrotnie już wskazywałam w niniejszej dysertacji, twierdzenia tego rodzaju oparte są na argumencie *ad ignorantiam*. Zakłada on, że skoro nie ma jednoznacznego dowodu na to, że ZINL zdolne są do *mindreadingu*, należy uznać, że u podstaw uzyskanych wyników eksperymentów leżą prostsze mechanizmy skojarzeniowe (czyli odwołanie do zasady Morgana). Na przykład szympanś może faworyzować wskazówki płynące od agenta, który patrzył na jedzenie, ponieważ nauczył się, że takie jednostki mają wyższe prawdopodobieństwo zdobycia pożywienia. Nie oznacza to jednak, że uczestnik badania przypisuje agentowi przekonanie. Czyli pomimo tego, że ZINL wykazują reakcje behawioralne podobne do ludzkich, te często

są interpretowane wyłącznie jako wynik działania prostych reguł asocjacyjnych, co odzwierciedla wspomniane wcześniej asymetryczne podejście do interpretacji danych eksperymentalnych. Jednocześnie jednak rośnie liczba badaczy, którzy postulują, że nawet w przypadku niemowląt reakcje obserwowane w klasycznych paradygmatach, takich jak patrzenie wyprzedzające czy naruszenie oczekiwań, nie muszą świadczyć o istnieniu metareprezentacji. Zamiast tego interpretowane są jako efekty działania skojarzeniowych reguł predykcyjnych, wykorzystujących dane percepcyjne i wcześniejsze doświadczenia. Takie stanowisko pozornie zmniejsza obserwowaną asymetrię interpretacyjną, jednak sprawia, że przeskok w rozwoju człowieka w porównaniu do innych gatunków, wydaje się jeszcze większy. Tego rodzaju luka interpretacyjna (między człowiekiem a innymi zwierzętami) czyni stanowisko behawioralne podatnym na zarzut niezgodności z darwinowską zasadą ciągłości ewolucyjnej, według której różnice międzygatunkowe mają przede wszystkim charakter ilościowy, a nie jakościowy. Jeśli bowiem mechanizmy predykcji zachowań są wspólne dla różnych gatunków, logicznym wydaje się założenie, że potencjał do ich dalszego rozwoju (w zdolności metareprezentacyjne) nie jest właściwy wyłącznie dla człowieka. Przyjmowanie istnienia tak wyraźnej granicy poznawczej oddzielającej ludzi od innych gatunków — zwłaszcza w oparciu o argument *ad ignorantiam* wykorzystywany w interpretacji doniesień empirycznych — może odzwierciedlać raczej kulturowe uprzedzenia niż rzeczywiste różnice funkcjonalne.

Coraz częściej w literaturze popiera się odejście od interpretowania wyników wskazanych paradygmatów jako bezpośrednich markerów *mind-readingu*, na rzecz traktowania ich jako wskaźniki funkcjonowania bardziej podstawowych mechanizmów predykcji zachowania. Zatem argument Baron-Cohena (2009) dotyczący twierdzenia, że rozwój detektora intencjonalności¹⁰³, obecny jest już u niemowląt, poddawany jest krytyce. Zwolennicy perspektywy behawioralnej uważają, że eksperymenty przytaczane przez Baron-Cohena (2009) nie są informacyjne w obszarze *mindreadingu*, a to co wskazany autor interpretuje jako mechanizmy właściwe czytaniu umysłów innych, wyjaśniają poprzez mechanizmy *smart behavior reading*. Niestety we wskazanym nurcie nie znajdziemy wiarygodnego wyjaśnienia, kiedy sprytne czytanie zachowania zamienia się w *mindreading*, jednak jej zwolennicy zapewniają, że skonstruowanie

¹⁰³ Rozumiany jako, mechanizm do przypisywania agencyjności innym oraz prostych stanów wolicjonalnych (np. intencje czy cele).

takich wyjaśnień będzie możliwe wraz z rozwojem narzędzi badawczych ze wskazanego obszaru. Zgadzam się z twierdzeniem, że paradygmaty spojrzenia wyprzedzającego, czy naruszenia oczekiwań testują proste zachowania behawioralne takie jak ID (*intentionality detector*) czy EED (*eye direction detector*), jednak uważam, że nie wyklucza to rozwoju bardziej złożonych zdolności – przeciwnie, może z nimi współwystępować i być ich filogenetycznym oraz ontogenetycznym fundamentem (Baron-Cohen, 1994). Zatem argumentacja funkcjonująca w zakresie nurtu mentalistycznego wydaje mi się bardziej prawdopodobna. Mechanizmy sprytnego czytania zachowania innego współlistnieją z początkowymi formami zdolności metareprezentacyjnych i mogą stanowić ich rozwojową podstawę (Baillargeon, Scott, He, 2010; Low, Apperly, Butterfill, Rakoczy, 2016). Dla przykładu, eksperymenty z udziałem niemowląt sugerują, że pewne formy rozumienia nie tylko intencji, ale i wiedzy mogą pojawiać się już we wczesnym okresie życia – około 15 miesięcy (Onishi, Baillargeon, 2005). Wyniki te opierają się na zastosowaniu zmodyfikowanych paradygmatów, takich jak naruszenie oczekiwań (*violation of expectation*) czy spojrzenie wyprzedzające, które bazują na analizie spontanicznych reakcji wzrokowych (Southgate, Senju, Csibra, 2007; Senju, Southgate, Snape, Leonard, Csibra, 2011). Naruszenie oczekiwań mierzone jest w sytuacjach eksperymentalnych, których dzieci biorące udział w badaniu, obserwują agenta wykonującego przewidywalne działania. Gdy agent nagle zachowuje się w sposób niezgodny z oczekiwaniami dziecka – np. kieruje się do miejsca, w którym nie powinien spodziewać się ukrytego obiektu – u uczestnika rejestrowany jest wzrost czasu patrzenia, interpretowany jako wskaźnik zaskoczenia. Reakcja ta sugeruje, że uczestnik posiadał pewne oczekiwania względem zachowania agenta, a więc przypisał mu intencje, cele czy nawet wiedzę (Onishi, Baillargeon, 2005). Interpretacja tych wyników jest jednak sporna, krytycy wskazują, że zachowania zaobserwowane w takich paradygmatach mogą być wyjaśnione za pomocą prostszych mechanizmów poznawczych — asocjacji, a nie poprzez właściwe tworzenie metareprezentacji przekonań innego (Heyes, 2014; Butterfill, Apperly, 2013).

Dyskusja dotycząca tego, czy nie-ludzkie naczelne również wykazują zdolności z zakresu *mindreadingu*, toczy się w nauce od dekad. W tym czasie zgromadzono wiele dowodów na to, że nie tylko człowiek potrafi przewidywać zachowania innych jednostek, rozumiejąc ich cele, percepcję czy wiedzę (Call, Tomasello, 2011). Pomimo to wciąż założenie, że zdolności metareprezentacyjne rozwijają się wraz z prostymi mechanizmami przypisywania agencyjności (ID) czy stanów percepcyjnych (EED)

innym, odrzucane jest w kontekście eksperymentów z udziałem zwierząt innych niż ludzie. Jednak jak już wielokrotnie wskazywałam w niniejszym opracowaniu, uważam, że takie rozumowanie narażone jest na tendencje wynikające z uprzedzeń. Dla lepszego zrozumienia prezentowanego przeze mnie stanowiska, w tym miejscu chcę zwrócić uwagę na inne problemy paradygmatów badań używanych w eksperymentach dotyczących czytania umysłów/sprytnego zachowania innego, z udziałem ludzkich dzieci. W literaturze coraz częściej podkreśla się ich ograniczoną trafność, niską powtarzalność i trudności w interpretacji wyników. Jednym z najbardziej znanych przykładów takiej krytyki, jest duża replikacja eksperymentu Southgate i współpracowników (2007), która nie potwierdziła wcześniejszych wyników sugerujących, że niemowlęta potrafią antycypować zachowanie innego w oparciu o jego fałszywe przekonanie (mierzone przez spojrzenie wyprzedzające, które jest także miarą detektora intencjonalności); (Dörrenberg, Rakoczy, Liszkowski, 2021). W warunkach testu dzieci w równym stopniu kierowały wzrok w miejsce odpowiadające zarówno temu, co same wiedziały, jak i fałszywemu przekonaniu agenta – co znacząco osłabia siłę dowodu na obecność metareprezentacji (Dörrenberg, Rakoczy, Liszkowski, 2021). Pojawiają się również wątpliwości dotyczące wpływu zmiennych kontekstowych (np. różnice w długości prób, kolejności prezentacji bodźców, atrakcyjności wizualnej elementów procedury eksperymentalnej) oraz obecność efektów ubocznych takich jak zmęczenie czy niestabilność motywacji uczestnika (Barone, Corradi, Gomila, 2019). Co więcej, istotnym zarzutem, który rzadko uwzględniany jest w interpretacjach omawianych badań, jest odwołanie się do znaczących różnic indywidualnych w wynikach uczestników. Nawet w tych samych warunkach niektóre dzieci konsekwentnie przewidują zachowanie aktora zgodnie z jego wiedzą, podczas gdy inne wykazują chaotyczny lub niespójny wzorec reakcji (Poulin-Dubois, Yott, 2018). Wyniki te sugerują, że *mindreading* nie rozwija się jednorodnie, tj. u wszystkich przedstawicieli gatunku w tym samym czasie. Raczej jest silnie zależne od czynników temperamentnych, poziomu rozwoju językowego, doświadczenia społecznego czy nawet bieżącego stanu emocjonalnego. W świetle powyższego należy uznać, że niewerbalne testy fałszywych przekonań i inne paradygmaty z zakresu badania czytania umysłów innych dostarczają ciekawych, choć trudnych do jednoznacznego zinterpretowania danych. Uważam, jednak, że ich ograniczenia metodologiczne nie powinny prowadzić do ich całkowitego odrzucenia, ale raczej do ostrożnego traktowania ich wyników. Zamiast pytać, czy niemowlęta lub inne naczelne posiadają

teorię umysłu, bardziej trafne może być badanie, w jakim zakresie i w jakich warunkach wykazują one kompetencje, będące potencjalnym fundamentem dla bardziej złożonych operacji metareprezentacyjnych. Mimo wspomnianych ograniczeń, testy fałszywych przekonań, również w wersji niewerbalnej, pozostają jednym z najważniejszych narzędzi umożliwiających badanie początkowych form zdolności z zakresu *mindreadingu*. Ich największą zaletą jest to, że pozwalają na wgląd w procesy poznawcze jednostek, które nie posługują się językiem – takich jak niemowlęta czy zwierzęta pozaludzkie. W przeciwieństwie do klasycznych, werbalnych wersji (np. zadania Maxi’ego), paradygmaty spojrzenia wyprzedzającego czy naruszenia oczekiwań umożliwiają badanie zdolności predykcyjnych bez konieczności aktywnej odpowiedzi ze strony uczestnika. Dzięki temu możliwe jest uchwycenie bardzo wczesnych etapów rozwoju wskazanych zdolności – zanim dziecko rozwinie pełne kompetencje językowe czy refleksyjne. Co istotne, wiele z tych badań wskazuje na to, że dzieci już w drugim roku życia potrafią przewidywać zachowania innych na podstawie ich intencji czy celów – co, niekoniecznie dowodzi obecności złożonych metareprezentacji, sugeruje tylko istnienie mechanizmów pośrednich, warunkujących ich rozwój. Warto również zauważyć, że niewerbalne paradygmaty badawcze mają istotną wartość heurystyczną – mogą służyć nie tylko testowaniu obecności określonych zdolności, ale także badaniu ich komponentów i warunków występowania. Analiza różnic indywidualnych, reakcji kontekstowych oraz zmienności między zadaniami pozwala lepiej zrozumieć, jakie mechanizmy poznawcze wspierają rozwój teorii umysłu i w jakich warunkach stają się one funkcjonalne. Zamiast traktować te testy jako jednoznaczny wskaźnik obecności lub braku określonych aspektów *mindreadingu*, warto postrzegać je jako narzędzia heurystyczne – pozwalające śledzić rozwój poznania społecznego, a także identyfikować pewne punkty rozwoju ontologicznego szczególnie ważne dla przejścia między prostym reagowaniem na zachowanie innego a złożonym reprezentowaniem jego stanów umysłowych. Co więcej, rozwijanie i doskonalenie tych metod – poprzez lepszą kontrolę zmiennych zakłócających, większą powtarzalność oraz bardziej precyzyjną interpretację wskaźników – można w przyszłości znacząco zwiększyć ich wartość praktyczną i teoretyczną.

Właśnie taki heurystyczny potencjał testów fałszywych przekonań został twórczo wykorzystany w badaniu opublikowanym w artykule *Great apes anticipate that other individuals will act according to false beliefs* (Krupenye, Kanto, Hirata, Call,

Tomasello, 2016). Badanie to zapisało się w historii jako jeden z pierwszych przykładów udanego zastosowania paradygmatu niewerbalnych testów fałszywych przekonań w badaniu teorii umysłu u nieludzkich naczelných. Zespół badawczy na uniwersytecie w Kyoto zaprezentował w nim badanie wykorzystujące metodę *eye-trackingu* (Krupenye, Kanto, Hirata, Call, Tomasello, 2016). Eksperyment ten nie tylko rozszerzył zasięg metod stosowanych wcześniej u ludzkich dzieci na inne zwierzęta, ale również rzucił wyzwanie dominującym dotąd interpretacjom behawioralnym, wskazując na możliwość istnienia bardziej złożonych form rozumienia stanów mentalnych u szympanów, bonobo i orangutanów, co otwiera także dyskusję na możliwości metareprezentacyjne, innych gatunków zwierząt (np. ptaków). Zanim opublikowano wskazany eksperyment paradygmat fałszywych przekonań, nie był często stosowany w badaniach z udziałem naczelných (Kaminski, Call, Tomasello, 2008; Martcorena, Ruiz, Mukerji, Goddu, Santos, 2011; Martin, Santos, 2014). W żadnym z nielicznych przypadków jego zastosowania, uczestnicy nie osiągnęli wyników zbliżonych do ludzkich dzieci. To natomiast wspierało behawioralne interpretacje eksperymentów głoszące, że zdolność ta wymaga komponentu językowego i jest właściwa wyłącznie człowiekowi. Bardziej ostrożne interpretacje wskazywały na problemy proceduralne, które znacząco utrudniają konstruowanie eksperymentów tak, aby zminimalizować możliwość wyjaśnienia wyników eksperymentu poprzez odniesienie do sprytnego zachowania innych.

Punktem wyjścia dla przełomowych badań zrealizowanych na Uniwersytecie w Kyoto było zatem pytanie, czy gdy zastosuje się odpowiednią procedurę badania, naczelne mogą uzyskać poziomy wskaźników rozumienia fałszywych przekonań podobne do uzyskiwanych z niewerbalnymi ludzkimi dziećmi. W swoim eksperymencie autorzy wykorzystali miarę spojrzenia wyprzedzającego, ponieważ wcześniejsze badania wykazały, że szympanasy potrafią generować wskazane zachowanie w sytuacjach, w których działania agentów są ukierunkowane na cel (ID); (Hare, Call, Tomasello, 2001, Kano, Call, 2014, Senju, Southgate, Snape, Leonard, Csibra, 2011). Jednak nie udało się wykazać, by były one zdolne do wnioskowania w oparciu o fałszywe przekonania – a więc zrozumienia, że działania innych podmiotów mogą być determinowane nie przez aktualny stan rzeczy, lecz przez subiektywne reprezentacje rzeczywistości, które mogą być niezgodne z faktami. Oczywiście, jak już podkreślałam wcześniej, należy przy tym pamiętać, że zredukowanie testów fałszywych przekonań do paradygmatów niewerbalnych, faktycznie prowadzi do pomiarów prostych zdolności

takich jak ID czy EED, nie mówiąc wiele o zdolności do tworzenia metareprezentacji. W tym sensie zgadzam się z żartobliwym stwierdzeniem Povinelliego i Vonka (2004), że do badania zdolności metareprezentacyjnych potrzebny jest mikroskop. Jednak, w przeciwieństwie do zwolenników perspektywy behawioralnej — uważam, że postulowane w przypadku ludzkich dzieci założenie o równoległym rozwoju prostych mechanizmów predykcyjnych i bardziej złożonych zdolności metareprezentacyjnych, powinno być stosowane także w odniesieniu do zwierząt innych niż ludzie. Nie widzę wiarygodnego powodu, dla którego mielibyśmy przyjmować asymetryczne standardy interpretacyjne, w których podobne zachowania i wskaźniki traktowane są jako dowód na różne poziomy zdolności poznawczych — wyłącznie ze względu na przynależność gatunkową badanych. Takie podejście wydaje się epistemologicznie niekonsekwentne i wzmacnia antropocentryczne uprzedzenia, takie jak antroponegacjonizm, które mogą prowadzić do systematycznego zaniżania możliwości poznawczych zwierząt pozaludzkich. Dlatego zgadzam się z poglądem, że choć niewerbalne paradygmaty wykorzystywane w badaniach nad teorią umysłu — takie jak testy wykorzystujące detekcję intencjonalności (ID), kierunku patrzenia (EED) czy uproszczone testy fałszywych przekonań — dostarczają cennych informacji na temat funkcjonowania podstawowych mechanizmów poznawczych, to ich ograniczenia metodologiczne sprawiają, że nadal nie jesteśmy w stanie jednoznacznie określić relacji między sprytnym czytaniem zachowania a rozumieniem stanów umysłowych innych. Niemniej jednak, badania z udziałem ludzkich dzieci sugerują, że te dwa zbiory procesów poznawczych nie są od siebie oddzielone, lecz współlistnieją i prawdopodobnie pozostają ze sobą w ścisłym związku rozwojowym. Jak wspomniałam wcześniej, w literaturze coraz częściej podnoszone są wątpliwości dotyczące trafności i powtarzalności niewerbalnych paradygmatów wykorzystywanych w badaniach dotyczących *mindreadingu* z udziałem ludzkich dzieci. Problemy te dotyczą m.in. zmiennych kontekstowych, różnic indywidualnych oraz niejednoznacznych reakcji uczestników, co skłania wielu badaczy do ostrożności w interpretacji wyników. Niemniej jednak warto zauważyć, że przed przełomowym eksperymentem zespołu z Uniwersytetu w Kyoto, uczestnikom pozaludzkim nigdy nie udało się osiągnąć wyników porównywalnych z dziećmi w zadaniach tego typu (Krupenye i współpracownicy, 2016). Dopiero we wskazanym artykule zaprezentowano sugestie, że szympansy, bonobo i orangutany są w stanie antycypować działania agenta na podstawie jego fałszywego przekonania, ponieważ wykazują zachowanie w postaci

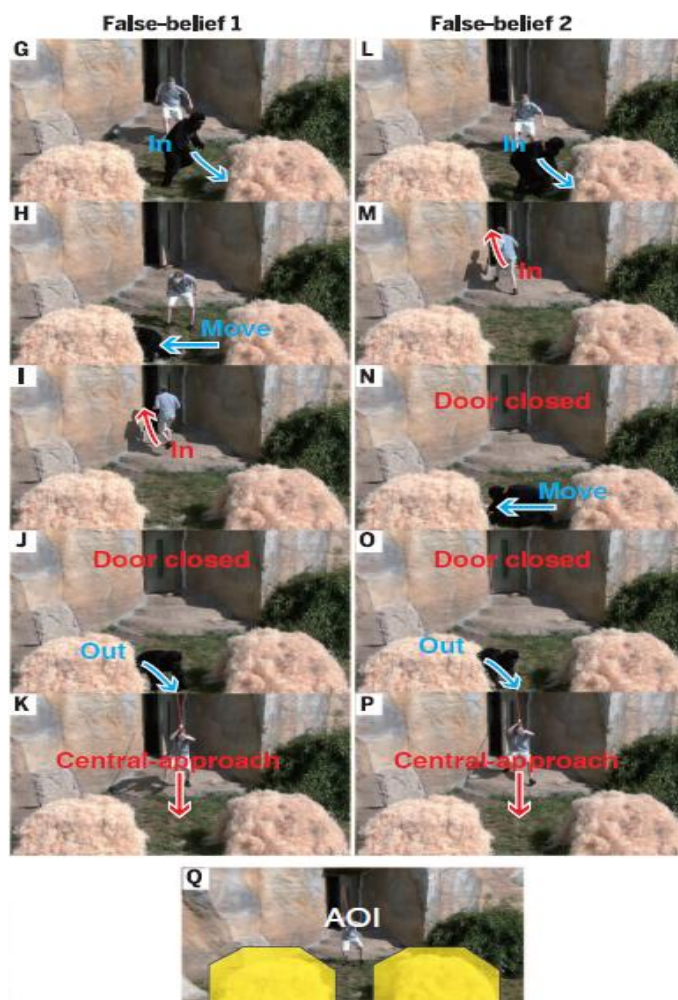
spojrzenia wyprzedzającego w odpowiednie miejsce. Wynik ten stanowi silne wyzwanie dla poglądu, że zdolność ta jest właściwa wyłącznie dla ludzi. Chociaż uzasadnione są pytania o ograniczenia metodologiczne używanych narzędzi oraz o to, w jakim stopniu rzeczywiście mierzą one metareprezentację. Eksperymenty Krupenye i współpracowników (2016, 2017, 2019) powinny skłaniać raczej do refleksji nad niedoskonałością samych procedur, niż do automatycznego redukowania zachowania naczelnych do prostych reguł behawioralnych. Obserwowana elastyczność zachowania, dostosowana do zmiennego kontekstu oraz możliwość antycypowania działań innych na podstawie niezgodnych z rzeczywistością przekonań — nawet jeśli częściowo oparte na mechanizmach sprytnego czytania zachowania — sugerują, że zdolności z zakresu poznania społecznego u ZINL są zdecydowanie bardziej złożone, niż dotąd zakładano. W moim przekonaniu, wnioski te powinny stanowić motywację do dalszego rozwoju metod eksperymentalnych oraz do większej epistemicznej wrażliwości, a nie do przedwczesnego zamykania dyskusji poprzez przyjmowanie najprostszych możliwych wyjaśnień, gdy istnieje tak wiele doniesień, sugerujących, że mechanizmy leżące u ich podstaw są bardziej złożone, niż sugerowałyby wcześniejsze badania.

Autorzy *Great apes anticipate that other individuals will act according to false beliefs* prezentują dwa eksperymenty, które oparte były o ten sam paradygmat (spojrzenia wyprzedzającego), ale wykorzystywały 2 różne scenariusze. W celu zminimalizowania wpływu złożonych, lecz wyuczonych reguł behawioralnych, uczestnicy konfrontowani byli z całkowicie nową dla nich sytuacją. To eliminuje zarzut, podniesiony w kontekście badań, które opisałam w części 5.1.1. *Przełomowe badania Premacka i Woodruffa*, m.in. o możliwość wyjaśnienia uzyskanych wyników poprzez odwołanie do dopasowania fizycznego (Savage-Rumbaugh, Rumbaugh, Boysen, 1978). W każdym z dwóch scenariuszy wykorzystano kompetytywną procedurę eksperymentalną, w której człowiek przebrany w strój małpy (A1), nazywany także *King Kongiem*, wykonuje pewne zachowania względem innego aktora (A2). Pierwszy scenariusz obejmował sytuację, w której (A2) szuka (A1), ukrywającego się w jednym z dwóch stogów siana (zdjęcie 11). Nagranie zawierające scenariusz drugi (zdjęcie 12), przedstawiało dwóch aktorów, jeden (A1) chciał zabrać kamień, jednak osiągnięcie tego celu uniemożliwiał mu (A2) — przebrany w strój *King Konga* człowiek. (A1) patrzył, jak (A2) chowa obiekt docelowy w jednej z 2 lokalizacji. Następnie przepędzony przez (A2) wychodzi z pomieszczenia. Pod jego nieobecność *King Kong* najpierw przenosi

kamień do drugiej lokalizacji¹⁰⁴, a następnie wychodzi, zabierając go ze sobą. Wtedy do pomieszczenia wraca (A1), podchodzi do miejsca, w którym może sięgnąć do obydwu lokalizacji. W obydwu przypadkach wyniki uzyskane przez nieludzkie naczelniki odpowiadały wynikom eksperymentów z udziałem ludzkich dzieci. Autorzy zaobserwowali, że uczestnicy wykazywali spojrzenie wyprzedzające w miejsce, gdzie (A1) ostatnio widział kamień, pomimo że uczestnik nie tylko widział obiekt po raz ostatni w innym miejscu niż (A1) oraz wie, że obiekt docelowy został ostatecznie zabrany przez (A2). Choć publikacja *Great apes anticipate that other individuals will act according to false beliefs* (Krupenye, Kanto, Hirata, Call, Tomasello, 2016) stanowiła przełom w zakresie narzędzi badawczych, wciąż nie eliminowała możliwości interpretacji uzyskanych wniosków w ramach prostej reguły skojarzeniowej *agent, będzie szukał obiektu tam, gdzie go ostatnim razem widział* (Povinelli, Vonk, 2004). W 2019 roku opublikowano jednak kolejny artykuł, w którym wykorzystano zmodyfikowaną wersję drugiego z opisanych scenariuszy zaprezentowanego w artykule wydanym w 2016 roku. Badania Premacka i Woodruffa (1978) nie były jedyną inspiracją, dla autorów *Great apes use self-experience to anticipate an agent's action in a false-belief test* (2019). Paradygmat zastosowany w eksperymencie zainspirowany został propozycją eksperymentalną, która początkowo dostosowana została do eksperymentów z udziałem ludzkich dzieci (Heyes, 1998; 2014; Maltzoff, Brooks, 2008). Autorka *Theory of mind in nonhuman primates* zaznacza, że rozwój w badaniach nad teorią umysłu wymaga wypracowania nowych procedur eksperymentalnych, które pomogą zdecydować o tym, jakie wyjaśnienia — behawioralne czy mentalistyczne — są bardziej adekwatne do uzyskanych wyników. Jak zaznacza, odkąd Premack i Woodruff (1978) opublikowali swój artykuł w środowisku naukowym, coraz częściej mówi się o tym, że ZINL nie tylko zdolne są do przypisywania prostych stanów wolicjonalnych, ale także do przypisania innym wiedzy. Zaproponowany przez nią paradygmat gogli (*goggles paradigm*), miał na celu rozróżnienie, czy uczestnicy rozumieją perspektywę wzrokową innych, czyli przypisują im mentalne stany percepcyjne, czy też opierają się na prostych regułach behawioralnych (np. *szukaj tam, gdzie agent ostatnio widział obiekt*). Potencjalne eksperymenty miały składać się z dwóch etapów. Pierwszy stanowił fazę nabywania własnego doświadczenia z goglami.

¹⁰⁴ Ten zabieg sprawia, że lokalizacja, w której uczestnik ostatnim razem widział kamień, jest inna niż ta, w której widział go (A1).

Zdjęcie 11. Pierwszy scenariusz eksperymentalny wykorzystany przez Krupenye i współpracowników (2016).



[Źródło (Krupenye, Kano, Hirata, Call, Tomasello, 2016, str. 111)]

Zdjęcie 12. Drugi scenariusz eksperymentalny wykorzystany przez Krupenye i współpracowników (2016).



[Źródło (Krupenye, Kano, Hirata, Call, Tomasello, 2016, str. 112)]

Uczestnicy mieli okazję zapoznać się dwoma rodzajami gogli, które z zewnątrz wyglądały identycznie, ale różniły się właściwościami widzenia (przezroczyste i nieprzezroczyste). Etap drugi stanowił fazę testową, w której uczestnicy obserwowali aktorów noszących gogle, ale nie wiedzieli, która osoba nosi te przezroczyste, a która nieprzezroczyste. Celem badania jest sprawdzenie czy uczestnik potrafi przewidzieć, kto widzi ukryty obiekt, używając nie tylko obserwacji zachowania innych, ale także własnego doświadczenia z goglami. Wykorzystanie własnego doświadczenia sugeruje przypisanie innemu odpowiedniego stanu mentalnego (widzenia lub niewidzenia). W ten sposób test redukuje możliwość, że uczestnicy po prostu reagują na obserwowane zachowania, a nie na reprezentacje mentalne. Należy jednak zaznaczyć, że procedura ta jest zorientowana na tworzenie metareprezentacji epistemicznej [1] rzędu, dotyczącej wiedzy. Kolejne badania po raz pierwszy wykorzystały procedurę zaproponowaną przez Heyes (1998), jednak uzupełnioną o spojrzenie wyprzedzające oparte na fałszywym przekonaniu (Maltzoff, Brooks, 2008; Senju i inn., 2008). Uczestnikami ich eksperymentów były 18- miesięczne ludzkie dzieci. W etapie pierwszym dzieci zapoznawały się z opaskami, jednak tym razem podzielono je na dwie grupy, w każdej z nich uczestnicy zapoznawali się z opaskami na oczy. Jedna grupa dzieci zapoznawała się z opaską nieprzezroczystą, druga z taką, która umożliwiała widzenie. Oczywiście, zgodnie z propozycją Heyes (1989) obydwie opaski wyglądały identycznie. Następnie obydwie grupy oglądały nagranie, na którym aktor widział, jak przedmiot jest ukrywany w konkretnym miejscu, a następnie zakładano mu opaskę. Wtedy usuwano obiekt docelowy, z określonej lokalizacji. Dzieci, które doświadczały opaski nieprzezroczystej traktowały aktora, tak, jakgdyby posiadał fałszywe przekonanie, wykonując spojrzenie wyprzedzające do miejsca, w którym aktor ostatnio widział obiekt. Nie wykryto takiego zachowania u dzieci, które doświadczały opaski przezroczystej. Omówiony paradygmat przetestowano także w eksperymentach z udziałem zwierząt innych niż ludzie. Wyniki nie ujawniły jednak istotnej statystycznie preferencji wykonywania spojrzenia wyprzedzającego (Penn, Povinelli, 2007; Karg, M. Schmelz, Call, Tomasello, 2015). Jednak, gdy zmodyfikowano omawianą procedurę poprzez zastosowanie paradygmatu kompetytywnego, w którym jeden aktor (A1), próbował kraść jedzenie innemu (A2), który znajdował się za osłoną, uczestnicy prawidłowo wykonywali spojrzenie wyprzedzające (Hare, Call, Agnetta, Tomasello, 2000; Melis, Call, Tomasello, 2006). Analogicznie do opisanych wyżej eksperymentów w etapie poprzedzającym eksperyment właściwy, uczestnicy mieli okazję z zapoznaniem się dwóch różnych

barier, jedną przezroczystą, drugą nieprzezroczystą. Okazuje się, że uczestnicy częściej przewidywali próby kradzieży, gdy (A2) znajdował się za drugą z wymienionych osłon, natomiast znacznie rzadziej, gdy osłona była przezroczysta. Wynik ten sugeruje, że naczelne przechodzą test z goglami, przynajmniej w pewnych kontekstach. Oczywiście, do uzyskania pozytywnych wyników konieczne jest zwiększenie motywacji uczestników do rozwiązania zadania (Hare, Call, Agnetta, Tomasello, 2000; Melis, Call, Tomasello, 2006).

Zatem kolejne eksperymenty zespołu z Kyoto (2019) zainspirowane rozwiązaniami wykorzystywanymi w paradygmatach z udziałem ludzkich dzieci, zaprojektowano w oparciu o schemat procedury, proponowany przez Heyes (1989). W pierwszym etapie uczestnicy zapoznawali się z właściwościami materiałów dwóch różnych barier wyglądających identycznie na nagraniach, prezentowanych uczestnikom w późniejszym etapie eksperymentu. Część uczestników doświadczała bariery nieprzezroczystej, natomiast druga grupa doświadczała tej przezroczystej. Etap drugi stanowił właściwą część eksperymentu, w której wykorzystano scenariusz, prezentowany po raz pierwszy w 2016 roku. Bardziej szczegółowy opis tego warunku opisałam przy okazji analizy artykułu *Great apes anticipate that other individuals will act according to false beliefs*. Scenariusz opisany jako drugi w 2016 roku różni się od tego zaprezentowanego w 2019 roku w jednym punkcie. Gdy (A2) zostaje przepłoszony przez *King Konga*, nie wychodzi z pomieszczenia, tylko chowa się za barierą z materiału. W dalszej części filmu kamień był przenoszony do drugiej lokalizacji, a następnie usuwany z pola widzenia uczestnika eksperymentu. Gdy (A1) opuszczał scenę prezentowaną na nagraniu, zabierając ze sobą obiekt docelowy, (A2) wychodził zza bariery i kierował się do punktu, w którym mógł sięgnąć do każdej z dwóch możliwych lokalizacji. Wyniki pokazały, że w zależności od tego, czy (A2) stał za osłoną przezroczystą (a więc mógł widzieć przeniesienie kamienia), czy nieprzezroczystą (nie mógł tego widzieć), naczelne biorące udział w badaniu, wykazywały spojrzenie wyprzedzające zgodne z tym, co (A2) mógł wiedzieć lub czego widzieć nie mógł. Interpretacja wyników sugeruje, że zwierzęta były w stanie przewidywać działania agenta na podstawie jego fałszywego przekonania, a nie tylko zapamiętywać ostatnie miejsce, w którym widział on obiekt.

Chociaż zaprezentowane eksperymenty nie dostarczają jednoznacznych dowodów na istnienie pełnoprawnej zdolności mindreadingu u naczelnych, to jednak ich wyniki trudno jest w pełni wyjaśnić bez odwołania się do reprezentacji stanów

mentalnych. Samo zastosowanie reguł asocjacyjnych czy prostych strategii predykcyjnych wydaje się niewystarczające, by opisać elastyczność i kontekstową adekwatność obserwowanych zachowań. Antycypowanie działań agenta w zależności od tego, co widział lub czego nie mógł widzieć, wskazuje na uwzględnienie jego indywidualnej perspektywy epistemicznej – a więc na coś więcej niż tylko skojarzenie percepcji z działaniem. Wyniki te, choć wymagające ostrożnej interpretacji, stanowią ważny argument za koniecznością ponownego przemyślenia granic poznania społecznego u zwierząt innych niż ludzie. Takie stanowisko wspierane jest nie tylko przez eksperymenty z udziałem nieludzkich naczelnych, prowadzone w laboratoriach, ale także przez doniesienia dotyczące innych gatunków zwierząt (np. psy czy kruki), także tych wykorzystujących obserwację w naturalnych warunkach. Psy, które ewoluowały w bliskim kontakcie z ludźmi, wykazują zdolności do rozumienia ludzkich wskazówek społecznych, takich jak gesty wskazywania, co sugeruje, że potrafią przypisywać innym wiedzę (Karan, 2023). Badania Bray i współpracowników (2021) wskazują, że już szczenięta wykazują wrodzone rozumienie ludzkich stanów umysłu, co może mieć podłoże genetyczne (Karan, 2023). Co więcej, ptaki z rodziny krukowatych, zwłaszcza kruki, wykazują zachowania świadczące o rozumieniu perspektywy innych osobników, np. w kontekście ukrywania i ochrony zapasów jedzenia przed rywalami, co jest interpretowane jako przejaw zdolności do tworzenia teorii umysłu (Karan, 2023). Ważnym aspektem badań nad *mindreadingiem* jest zdolność do refleksji nad własnymi stanami poznawczymi. Badania z udziałem delfinów czy makaków, sugerują, że zwierzęta te, potrafią monitorować i regulować swoje stany, co oznacza, że ich procesy poznawcze muszą być poddawane jakiejś formie kontroli, co jest bliskie ludzkiej refleksji (Donovan, 2009). Zebrane na przestrzeni wielu lat dane wskazują, że niektóre zwierzęta inne niż ludzie, zwłaszcza naczelne, ale także psy, ptaki i delfiny posiadają podstawowe zdolności do: rozróżniania wiedzy od niewiedzy u innych, przewidywania zachowań innych na podstawie ich przekonań (także tych fałszywych) oraz dostosowywania swoich reakcji do perspektywy percepcyjnej i poznawczej innych podmiotów. Chociaż istnieją wątpliwości dotyczące zakresu i mechanizmów tych zdolności, najnowsze badania, zwłaszcza te wykorzystujące metody *eye-trackingu*, otwierają nowe perspektywy w badaniu *mindreadingu* poza gatunkiem ludzkim i podważają tradycyjny pogląd o wyłączności tej zdolności dla ludzi (Kano, Krupenye, Hirata, Tomonaga, Call, 2019).

5.2. Perspektywa krytyczna w kwestii teorii umysłu u zwierząt innych niż ludzie: wybrane badania

W niniejszej części koncentruję się na koncepcjach behawioralnych, zakładających, że zachowanie zwierząt innych niż ludzie w kontekście badań (w tym obserwacji), można wyjaśnić bez odwołania do procesów wykorzystujących metareprezentacje. W konsekwencji zwolennicy wskazanych teorii przyjmują, że tylko człowiek zdolny jest do teorii umysłu, a nawet w skrajnych wersjach, odmawiają innym gatunkom prostszych mechanizmów z zakresu *mindreadingu*, jak przypisywanie stanów percepcyjnych (Povinelli, Eddy, 1996). Dominacja perspektywy behawioralnej w czasach współczesnych Premackowi i Woodruffowi (1978), sprawiła, że artykuł wzbudził znaczące kontrowersje wśród współczesnych im badaczy. W konsekwencji niemal natychmiast pojawiły się publikacje bezpośrednio odnoszące się do artykułu *Does the chimpanzee have a theory of mind?*, kwestionujące wyciągnięte przez autorów wnioski. Pełne uznanie dla koncepcji zaprezentowanej w tym, przełomowym tekście nastąpiło dopiero wraz z publikacją opisanego wyżej artykułu: *Does the chimpanzee have a theory of mind? 30 years later* (Call, Tomasello, 2008).

Oczywiście chcę przypomnieć, że faktycznie autorzy popełnili szereg błędów metodologicznych, które szczegółowo opisałam w części poświęconej *Does the chimpanzee have a theory of mind?* (Premack, Woodruff, 1996).

Krytyczna perspektywa na teorię umysłu poza gatunkiem ludzkim obecna jest przede wszystkim w nurcie behawioralnym. Narracje dominujące w tej perspektywie odwołują się do mitów o *wyjątkowości człowieka*, wskazana struktura poznawcza konstytuuje porządek społeczny, w którym żyją badacze konstruujący i utrzymujący określone interpretacje badań z udziałem ZINL. Jak wykazałam w części 3.1 *Historia badania umysłów zwierząt pozaludzkich*, mit o *wyjątkowości człowieka* ma długą tradycję kulturową i naukową, służącą nie tylko wyjaśnianiu, ale i legitymizowaniu określonego porządku społecznego. W późniejszych etapach m.in. za sprawą Nikoli Tinbergena i Konrada Lorenza, w Europie rozwinęła się etologia – kierunek badań nad zachowaniem zwierząt, który choć nie wnikał bezpośrednio w strukturę wewnętrznych stanów umysłowych, to jego podejście — zakorzenione w analizie adaptacyjnej — otworzyło drogę do konceptualizacji procesów wewnętrznych i tym samym pozwoliło na opisanie pewnych procesów, które kiedyś zamknięte były w tzw. „czarnej skrzynce” procesów poznawczych. Można więc powiedzieć, że etologia stworzyła ramy,

w których zaczęto traktować zachowanie nie tylko jako reakcję na bodziec, lecz jako potencjalny przejaw aktywności umysłowej.

Argumentacja przedstawiona przez zwolenników nurtu behawioralnego odwołuje się do szeroko przyjmowanej (w etologii, psychologii, filozofii i kognitywistyce) reguły parsymonii wyznaczonej przez Loyda Morgana. Interpretacje wyników niżej omawianych eksperymentów często opierają się na argumencie *ad ignorantiam*. Jak wskazywałam w części 2.5. *Teoria dysonansu poznawczego*, argument ten bywa uzasadniony, jednak w analizowanych przypadkach uważam jego zastosowanie za nietrafione (Walton, 1999b). Co więcej, interpretuję konsekwentne odmawianie zwierzętom pozaludzkim zdolności z zakresu *mindreadingu* za przejaw antroponegacjonizmu. Uznaję, że najczęstszą strategią argumentacyjną utrzymującą wskazane uprzedzenie jest przedstawienie możliwości istnienia alternatywnego, prostszego mechanizmu wyjaśniającego zachowanie zwierząt. Takie wyjaśnienie przedstawia mechanizm nieodwołujący się do przetwarzania metareprezentacji oraz przyjmuje formę mechanizmu skojarzeniowego. Następnie za pomocą argumentu *ad ignorantiam* utrzymuje się, że nie ma dowodu na to, że ZINL biorące udział w badaniu, wykorzystują bardziej złożony mechanizm. Z tak przeprowadzonego rozumowania, wyciąga się wniosek, głoszący, że tylko człowiek zdolny jest do *mind-readingu*, inne zwierzęta wykorzystują prostszy mechanizm określany *smart behaviour reading*. Warto również podkreślić, że badacze omówieni poniżej, konsekwentnie utrzymują pogląd, iż zdolność do przypisywania stanów mentalnych jest wyłącznie ludzką właściwością. Przykładem jest Povinelli, który nadal twierdzi, że szympansy nie posiadają pełnej teorii umysłu i nie przypisują innym istotom stanów percepcyjnych w sposób analogiczny do ludzi (Povinelli, 1996; Povinelli, Eddy, 1996; Povinelli, Vonk, 2003; Povinelli, Vonk, 2004). Jego badania koncentrują się na założeniu, że choć szympansy mogą przejawiać złożone zachowania społeczne, nie implikuje to zdolności do rozumienia perspektyw czy doświadczeń percepcyjnych innych jednostek. Każdy z niżej prezentowanych autorów opiera swoje interpretacje na Zasadzie Morgana, omówionej w rozdziale 3.1 *Historia badania umysłów zwierząt pozaludzkich*. Prawo Oszczędności zakłada, że jeśli dane zachowanie można wyjaśnić przy użyciu prostszych mechanizmów poznawczych, należy przyjąć taką interpretację. Jak już podkreślałam, zasada ta odgrywa kluczową rolę w metodologii badań nad poznaniem ZINL, jednak w pewnych przypadkach jej stosowanie wynika nie z rzetelnej analizy, lecz z postawy antroponegacjonistycznej. Dla przypomnienia polega ona na konsekwentnym

negowaniu istnienia procesów poznawczych analogicznych do ludzkich u innych zwierząt. Ponadto *argument ad ignorantiam* bywa angażowany w związku z efektem potwierdzenia, który wzmacnia treści składające się na poznawczy komponent uprzedzeń. Takie rozumowanie, choć pozornie neutralne metodologicznie, w rzeczywistości często reprodukuje założenia wynikające z mitu o *wyjatkowości człowieka* i niezdolności innych gatunków do przetwarzania złożonych stanów mentalnych (takich jak metareprezentacje). W efekcie badania, które mogłyby sugerować obecność form *mindreadingu* u ZINL, są systematycznie reinterpretowane w duchu parsymonii — nie jako potencjalne dowody metareprezentacyjnych zdolności, lecz jako przejawy prostszych mechanizmów behawioralnych. Tego rodzaju strategia interpretacyjna, oparta na domniemaniu, niewiedzy i preferowaniu wyjaśnień skojarzeniowych, może jednak skutkować przekroczeniem norm dobrej praktyki epistemicznej. To w konsekwencji może prowadzić do pewnego rodzaju ślepoty poznawczej wobec danych niepasujących do przyjętego modelu. Właśnie w tym kontekście warto przyrzeć się argumentacjom wybranych autorów, którzy odgrywają kluczową rolę w utrzymywaniu perspektywy behawioralnej w badaniach nad teorią umysłu. W dalszej części przedstawię kilka tekstów, które bezpośrednio odnoszą się do klasycznej pracy Premacka i Woodruffa (1978). Każdy z nich zawiera odwołania bibliograficzne do tego tekstu i stanowi element szerszej debaty dotyczącej możliwości przypisywania stanów mentalnych zwierzętom innym niż ludzie.

Szczególną uwagę poświęcę stanowiskom dwojga badaczy: Josepha Povinelliego oraz Sary Shettleworth. Pierwszy z wymienionych projektuje eksperymenty i interpretuje uzyskane dane. Natomiast druga sama nie prowadzi badań empirycznych, lecz aktywnie komentuje i analizuje aktualne doniesienia w obszarze *animal cognition*. Oboje konstruują swoje narracje w perspektywie behawioralnej. Prace Povinelliego koncentrują się głównie na szympanсах i dotyczą testowania hipotez związanych z rozumieniem przez nie takich pojęć jak „widzenie”, „wiedza” czy „zamiar”. Czyli można powiedzieć, że zajmuje się badaniem procesów poznawczych z zakresu *mindreadingu* u zwierząt pozaludzkich. Uważam, że Povinelli jest jednym z najbardziej konsekwentnych reprezentantów nurtu sceptycznego wobec możliwości istnienia teorii umysłu, czy nawet zdolności do przypisywania stanów percepcyjnych innemu. Podobną opinię mam o Shettleworth, jej zainteresowania nie koncentrują się wyłącznie na teorii umysłu, lecz obejmują szerokie spektrum zagadnień związanych z ewolucją i mechanizmami poznawczymi u zwierząt. W kontekście niniejszej

dysertacji szczególnie istotne są dwa jej teksty, w których prezentuje klasyczne argumenty charakterystyczne dla stanowiska behawioralnego i podważa możliwość istnienia zdolności do tworzenia metareprezentacji u zwierząt innych niż ludzie. Wybrana autorka prezentuje skrajnie antroponegacjonistyczną postawę, która w konsekwencji prowadzi do odrzucenia interpretacji wyników eksperymentów z udziałem ludzkich dzieci jako świadczących o obecności stanów mentalnych. Dotyczy to w szczególności badań prowadzonych w paradygmatach śledzenia kierunku patrzenia (*gaze following*) czy przypisywania agencyjności. Stanowisko to podważa informacyjną wartość metod wypracowanych przez dekady w badaniach nad teorią umysłu, poprzez twierdzenia, że zachowania rejestrowane w tych procedurach odzwierciedlają jedynie działanie prostych mechanizmów typu *smart behavior reading*, a nie właściwego *mindreadingu*.

Uważam, że tak skrajnie sceptyczne stanowisko przekracza granice dobrej praktyki epistemicznej. Przede wszystkim dostrzegam w nim złamanie reguły wrażliwości, co konsekwencji prowadzi do odrzucenia paradygmatów takich jak śledzenie kierunku patrzenia, detektor intencjonalności czy testy fałszywych przekonań, jako informacyjne w kontekście przypisywania stanów mentalnych innym. To prowadzi do odrzucenia dobrze ugruntowanych metodologicznie tradycji badawczych z zakresu psychologii rozwojowej.

5.2.1. Daniel Povinelli i Timothy Eddy

Kolejne eksperymenty Povinelli'ego, które warto przeanalizować, zaprezentowane zostały w książce *What Young Chimpanzees Know about Seeing* (Povinelli, Eddy, 1996). Dotyczą one zdolności określanej w literaturze jako *wizualne przyjmowanie perspektywy innego* (*visual perspective taking*). U jej podstaw leży prosty mechanizm, opisany przez Baron-Cohena (2003) jako detektor kierunku patrzenia (EED). Zatem, w tym miejscu pragnę zaznaczyć, że zdolność ta nie wymaga zaangażowania mechanizmu, który w pierwszym rozdziale opisałam jako *przyjmowanie perspektywy innego*. Aktualnie istnieją liczne dowody, że śledzenie kierunku patrzenia jest ewolucyjnie starsze niż człowiek (Baron-Cohen, 1995; Hare, Call, Tomasello, 2001; Emery, Clayton, 2001; Bugnyar, Heinrich, 2005; Kaminski, Call, Tomasello, 2008). Wyniki licznych eksperymentów sugerują, że nie tylko szympansy potrafią rozumieć, co inny osobnik widzi lub nie widzi oraz wykorzystać tę wiedzę w rywalizacji o pożywienie (Hare, Call, Tomasello, 2001; Kaminski, Call, Tomasello, 2008). Co

więcej, okazuje się, że kruki zmieniają swoje strategie ukrywania pożywienia w zależności od tego, czy były obserwowane przez innego osobnika, co sugeruje, że rozumieją, co widzi inny ptak (Bugnyar, Heinrich, 2005). Bardzo ciekawe badania z udziałem sójek pokazały, że ptaki te uwzględniają obecność innych osobników podczas ukrywania pożywienia, oraz że dostosowują swoje zachowanie w zależności od tego, kto ich obserwował (Emery, Clayton, 2001). Okazało się, że jeśli sójka A chowa jedzenie w obecności sójki B, to później (gdy ma okazję) przenosi je w inne miejsce, jakby chciała zapobiec jego kradzieży. Oczywiście, jeśli podczas ukrywania jedzenia nie wykryła innych ptaków, nie wykazuje takiego zachowania. Co więcej, zwierzęta te uwzględniają wcześniejsze doświadczenia obserwatora, a konkretnie, jeśli sójka B (obserwator) wcześniej kradł jedzenie innym ptakom, to sójka A (chowająca pożywienie) jest ostrożniejsza, czyli częściej przenosi ukryty pokarm. Natomiast jeżeli obserwujący ptak nie ma historii kradzieży, zwierzę rzadziej przenosi jedzenie. To sugeruje, że sójki szacują prawdopodobieństwo, że inny ptak je okradnie.

Ze względu na ilość eksperymentów zaprezentowanych w na łamach książki *What Young Chimpanzees Know about Seeing* (Povinelli, Eddy, 1996), nie zdecydowałam się na szczegółowy opis wszystkich. Wybrałam natomiast czternaście z nich, opisanych w rozdziałach: *III. Understanding Who Can See You: Preliminary Investigation* (str. 25-67) oraz *IV. Understanding Who Can See You: Further Investigation* (str. 67-106). Autorzy przyjęli dwie ramy teoretyczne, które stanowią przeciwległe bieguny spektrum: behawiorystyczną i mentalistyczną. Povinelli i Eddy (1996) wskazują, że są zainteresowani testowaniem wyjaśnień mentalistycznych, dlatego ich hipoteza zerowa ma charakter skrajnie behawiorystyczny, gdyż głosi, że zachowanie szympanśów powinno być interpretowane bez odwołania się do nieobserwowalnych stanów i procesów mentalnych. Badacze przeprowadzili w sumie 6 eksperymentów w ramach, których szympansy konfrontowane były dwoma eksperymentatorami (dalej jako (E1) i (E2)), posiadającymi różne informacje wizualne o sytuacji. Uczestnicy wpuszczani byli do pomieszczenia testowego, w którym znajdowała się szyba oddzielająca ich od eksperymentatorów. Oboje znajdowali się 1,5 m od szyby, a jeden z nich (E1) zawsze miał odsłonięte oczy, (E2) zakryte. Zadaniem szympanśów było wykonanie gestu *proszenia o jedzenie* do (E1) lub (E2). Wybrany akt komunikacji gesturalnej wykonywany jest w naturalnym środowisku przez szympansy w celu uzyskania jedzenia od innego osobnika (Goodall, 1986). Chcę jednak podkreślić, że chociaż nazwa „gest proszenia” wskazuje na stan mentalny (chcenia), autorzy

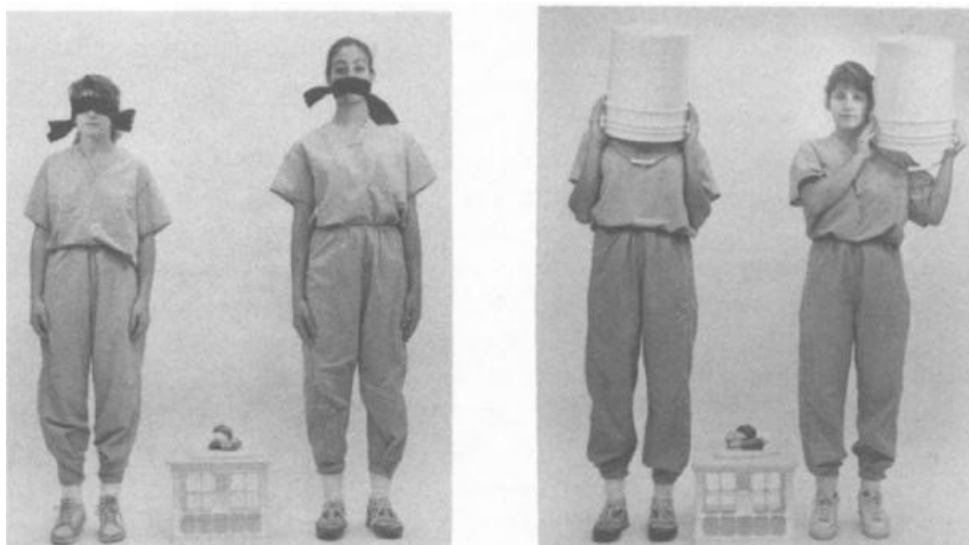
uważają, że takie zachowanie szympanсів nie wiąże się z posiadaniem stanu mentalnego. Na początku artykułu używają słowa „proszenie” w cudzysłowie, podkreślając w ten sposób, że ma ono jedynie charakter retoryczny. Autorzy *What Young Chimpanzees Know about Seeing* zakładają, że szympansy widzą różnicę między eksperymentatorami (wiedzą, że jeden ma zasłonięte oczy, a drugi nie) oraz przyjmują powszechne stanowisko, że gesty stanowią akty komunikacji intencjonalnej. Każda ze wskazanych wyżej perspektyw determinuje założenia dotyczące, tego jak uczestnicy badania powinni się zachować, zgodnie z przewidywaniami obydwu modeli teoretycznych. W ramach perspektywy mentalistycznej szympansy powinny wykorzystywać informację o tym, co widzi eksperymentator, czyli powinny wykonywać gest *proszenia o jedzenie* tylko w stosunku do tych eksperymentatorów, którzy ich widzą oraz ignorować (E2). Natomiast perspektywa behawioralna zakłada, że szympansy nie przypisują stanów percepcyjnych innym, dlatego jej przewidywania wskazują, że zwierzęta będą wykonywać gest *proszenia* losowo, do każdego z eksperymentatorów (bez wyraźnej preferencji). Oczywiście autorzy zaznaczają, że rama behawiorystyczna zakłada, że szympansy wykorzystują informację o oczach oraz spojrzeniu, mogą nawet wyuczyć się reguł dotyczących tego, do kogo warto gestykulować a do kogo nie.

Uczestnikami eksperymentów było pięć szympansic (Brandy, Jadine, Mindy, Megan, Kara) oraz jeden samiec (Apollo), wszystkie przyszły na świat w niewoli oraz każde z nich w swoim życiu było uczone gestu *proszenia o jedzenie*, oraz wykorzystywały go spontanicznie w komunikacji z trenerami. Omawiane eksperymenty wykorzystywały różne sposoby zasłaniania oczu, jednak przeprowadzone były w jednym paradygmacie. Warunek kontrolny nazywany *basile* służył do sprawdzenia czy szympansy są w ogóle zainteresowane sytuacją, w której muszą wybierać między dwoma eksperymentatorami, zamiast reagować na pojedynczą osobę. Autorzy chcieli także, aby zwierzęta podjęły decyzję między dwoma eksperymentatorami, w sytuacji niewymagającej rozumienia, kto co widzi (czyli takiej, która nie angażowała ich zdolności do wizualnego przyjmowania perspektywy innego). W pozostałych eksperymentach (E2) zasłaniał oczy rękami, opaską, wiadrem, dużym prostokątnym ekranem lub odwracał się tyłem do szympanсів biorących udział w badaniu (zdjęcie 13). Ciekawym przypadkiem był warunek patrzenia przez ramię, czyli sytuacja, w której obaj eksperymentatorzy byli odwrócenii tyłem, jednak (E1) miał twarz skierowaną w stronę zwierzęcia (zdjęcie 14). We wszystkich trzech warunkach, gdzie

(E2), miał zasłonięte oczy przez przedmiot, szympansy nie gestykulowały wybiórczo a losowo czyli do obydwu eksperymentatorów, bez wyraźniej preferencji. Taki sam wynik uzyskano w 1 z 3 eksperymentów, gdzie eksperymentator miał zasłonięte oczy w bardziej naturalny sposób np. rękami czy poprzez odwrócenie się tyłem. Oznacza to, że tylko w 2 z 6 eksperymentów szympansy wykazały natychmiastową niewyuczoną preferencję. Były to próby tył — przód i patrzenie przez ramię. Znaczna większość eksperymentów wykazała, że wykonywanie gestu do eksperymentatorów jest losowa. Dlatego autorzy uważają, że wyniki uzyskane ze wszystkich eksperymentów zgodne są z założoną przez nich hipotezą 0, czyli uznają, że najlepszą interpretacją zachowań zwierząt biorących udział w ich badaniach, stanowi odniesienie do perspektywy behawioralnej. Tym samym argumentując, że wyniki osiągnięte przez uczestników warunku patrzenia przez ramię, nie są przejawem wykorzystania informacji dotyczącej stanów percepcyjnych (E2) a reakcją na bodziec czołowy.

Utrzymują takie stanowisko nawet pomimo istotnych elementów procedury eksperymentalnej, które mogły znacząco wpłynąć na uzyskane wyniki. Podstawowym problemem wszystkich eksperymentów jest rozmieszczenie (E1) i (E2) w pokoju testowym. Po pierwsze stali oni 1,5 metra od szyby dzielącej ich od szympansa, co stanowi znaczącą odległość. Szczególnie biorąc pod uwagę, że ludzie są wyżsi od zwierząt biorących udział w eksperymencie, a co za tym idzie, ich oczy znajdowały się dużo niżej niż oczy ludzi. Możliwe, że taki dystans spowodował, że uczestnicy nie zwracali uwagi na twarz, tak intensywnie, jak zrobiliby to, gdyby eksperymentatorzy znajdowali się bliżej, a ich oczy byłyby na poziomie bardziej dostępnym dla szympanów. Co więcej, autorzy wskazują, że w warunkach tył — przód i *baseline*, gdy trener wpuszczał uczestników do pokoju testowego, ci patrzyli przez chwilę do góry, jednak nie ma pewności, że wystarczająco skupiali się na twarzach eksperymentatorów, przed dokonaniem wyboru. Ciekawe jest to, że we wskazanych eksperymentach szympansy osiągały wysoką trafność gestykulowania. Niestety autorzy *What Young Chimpanzees Know about Seeing* nie wspominają o spojrzeniu w górę w innych niż wymienione warunki. Nie oznacza to jednak, że uczestnicy nie patrzyli na twarze eksperymentatorów. Uważam, że zachowanie to mogło zostać np. przeoczone przez osoby dokonujące analizy nagrań badania. Należy pamiętać, że w czasie, gdy Powinelli i Eddy przeprowadzali swoje badania, nauka nie oferowała jeszcze tak zaawansowanych narzędzi do śledzenia patrzenia, jak *eye-tracker*.

Zdjęcie 13. Przedstawienie warunków eksperymentu z zasłoniętymi oczami



[Źródło (Povinelli, Eddy, 1996, str. 35)].

Zdjęcie 14. Warunek patrzenia przez ramię



[Źródło (Povinelli, Eddy, 1996, str. 50)].

Dlatego też nie podzielam niepewności autorów, dotyczącej zarejestrowania informacji percepcyjnej, przez zwierzęta biorące udział w eksperymentach. Współcześnie powszechnie wiadomo, że szympansy potrzebują znacznie mniej czasu do zapamiętania informacji wizualnej¹⁰⁵ (Biro, Matsuzawa, 1999; Kawai, Matsuzawa, 2000; Hayashi, Matsuzawa, 2003; Inoue, Matsuzawa, 2007; Tomonaga, 2007; Tomonaga, Imura, 2009). W eksperymencie przeprowadzonym przez Tetsuro Matsuzawę i jego zespół z Uniwersytetu w Kyoto młode szympansy (około 5-letnie) oraz studenci uniwersytetu zostali poddani testowi pamięci numerycznej. Na ekranie komputera pojawiała się siatka z dziewięcioma cyframi, które znikaly po krótkim czasie, a uczestnicy mieli za zadanie wskazać ich pierwotne położenie w odpowiedniej kolejności. Przy najkrótszym czasie ekspozycji (210 milisekund), młode szympansy osiągały 80% dokładności, podczas gdy wyniki ludzi spadały do około 40%. Autorzy sugerują, że szympansy wykorzystują tzw. pamięć fotograficzną (eidetyczną), pozwalającą im na błyskawiczne tworzenie i przywoływanie obrazów wizualnych. Zdolność ta może być adaptacyjna w naturalnym środowisku, umożliwiając szybkie reagowanie na bodźce, takie jak lokalizacja pożywienia czy unikanie zagrożeń. Wyniki te kwestionowały wcześniejsze przekonania o wyższości ludzi w zakresie wszystkich funkcji poznawczych.

Drugim nie mniej ważnym problemem omawianych eksperymentów jest fakt, że (E1) został poinstruowany, aby nie utrzymywać kontaktu wzrokowego z szympansamami. Co za tym idzie, nie kierował wzroku w stronę jedzenia, o które miał prosić uczestnik. Jak zauważają sami autorzy możliwe, że gdy szympansy patrzyły na twarze eksperymentatorów uznawały, że żaden z nich na nich nie patrzy (nawet (E1), który nie zasłaniał oczu). W takiej w sytuacji uczestnicy po prostu nie zdawali sobie sprawy, że podczas podejmowania decyzji powinni kierować się informacją o odsłoniętych oczach, czyli faktycznie nie wykorzystywały procesu przypisywania stanów percepcyjnych innemu.

¹⁰⁵ Szympanś osiągaający najwyższe wyniki w zapamiętywaniu informacji wizualnej nazywa się Ayumo oraz mieszka na terenie Uniwersytetu Kyoto, a konkretnie w instytucie badań prymatologicznych. Placówka ta przyjęła bardzo ciekawy paradygmat badań, zwierzęta dobrowolnie przychodzą do pomieszczenia, w których znajdują się komputery. Szympansy grają na nich w gry, a ich wyniki poddawane są różnorodnym analizom. Więcej szczegółów dotyczących wskazanych badań znaleźć można w popularnonaukowym filmie *The Cognitive Tradeoff Hypothesis* autorstwa popularnego na platformie YouTube kanału Vsauce.

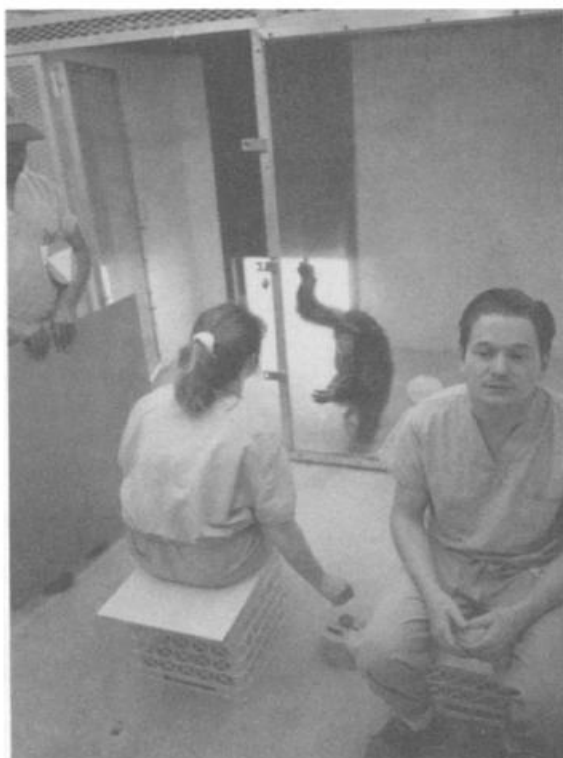
W kolejnym rozdziale książki Povinelliego i Eddy'ego (1996) autorzy prezentują eksperymenty 7-14, które zawierają korektę procedur eksperymentalnych, w założeniu, eliminującą problemy wcześniejszych badań. Uczestnikami były te same szympanse, które brały udział w 1-6, natomiast w warunku 9 dołączyła do nich Candy. Co ważne badacze zmienili dystans między eksperymentatorami a uczestnikami (z 1,5 m na 0,5 m) oraz wykorzystano skrzynki, na których siedzieli (E1) i (E2). Zmiana pozycji na siedzącą sprawiła, że oczy szympanów były na tym samym poziomie co eksperymentatorów. Większość prób przeprowadzono analogicznie do tych prezentowanych w III rozdziale książki, dlatego nie zdecydowałam się na umieszczenie szczegółowych opisów wszystkich eksperymentów zaprezentowanych w omawianym rozdziale. Chcę podkreślić, że (E1) poinstruowany był, aby wpatrywać się w miejsce, w którym szympan miał wykonać gest proszenia¹⁰⁶, czyli na wprost, nie zaś w oczy uczestnika. Eksperymentator wciąż nie łąpał kontaktu wzrokowego z szympanami ani tym bardziej nie zmieniał kierunku patrzenia np. na jedzenie. Jest to bardzo ciekawy zabieg, w przypadku badania detektora kierunku patrzenia, gdyż to właśnie ten mechanizm odpowiada za analizowaną przez Povinelliego i Eddy'ego zdolność (przypisywanie stanów percepcyjnych innym). Uważam to za kardynalny błąd procedury, przez który nie można traktować wyników tych eksperymentów jako informacyjne w zakresie EED.

Interesujące procedury wykorzystano w eksperymencie 12 gdyż w jego ramach, uczestnicy konfrontowani byli z nową sytuacją, gdzie faktycznie wykorzystano elementy śledzenia kierunku patrzenia. Pierwszy z wymienionych składał się z dwóch faz, każda z nich podzielona była na 4 sesje, które obejmowały sześć prób. W fazie 1 cztery z nich miały charakter kontrolny — dwóch eksperymentatorów siedziało naprzeciwko szympana, patrząc prosto przed siebie, nie wykonując przy tym żadnych dodatkowych ruchów głową ani oczami. Celem tego warunku było sprawdzenie, jak szympanse zachowują się w neutralnym kontekście, w którym żaden eksperymentator nie dawał sygnałów wzrokowych sugerujących, że powinny zwrócić uwagę na konkretny obiekt lub punkt w przestrzeni. Pozostałe próby wykorzystywały dwa

¹⁰⁶ Gest proszenia wykonywany był przez otwór w szybie, który autorzy nazywają „dziurą odpowiedzi” (*response hole*). W szybie oddzielającej ludzi od szympanów znajdowały się 2 takie otwory, każdy z nich odpowiadał jednemu z eksperymentatorów. Zwierzę dokonywało wyboru, kogo poprosić o jedzenie, aby dostać pożywienie, wyciągając rękę w stronę wybranego (E) przez odpowiadającą mu dziurę. Ta sama procedura była wykorzystywana we wcześniej opisanych eksperymentach 1-6.

warunki badawcze, jednym z nich jest opisany wcześniej tył — przód (zdjęcie 15), natomiast pozostałe wykorzystywały nowy paradygmat: skupiony — rozproszony (*attending—distracted*). W jego ramach (E1) pozostawał w neutralnej pozycji, patrząc prosto przed siebie (skupiony), natomiast (E2) unosił głowę i patrzył we wcześniej ustalony punkt na suficie (Zdjęcie 16). Takie ułożenie eksperymentatorów ułatwiało późniejszą analizę nagrań wideo, gdyż założono, że jeśli szympansy podążałyby za wzrokiem (E2), należało się spodziewać, że odwrócą głowę i spojrzą w ten sam punkt. Ciekawym zabiegiem zastosowanym przez autorów, było podzielenie uczestników na dwie grupy. Różnica między nimi opierała się na kolejności prezentowanych im warunków. Do pierwszej przydzielono Karę, Megan, Apollo i Candy, przechodzili oni najpierw paradygmat *attending-versus-distracted*, a następnie *back-versus-front*, natomiast Jadine, Brandy, Mindy były testowane w odwrotnej kolejności. Celem stworzenia dwóch grup, było zminimalizowanie wpływu kolejności prezentowanych zadań na ogólne wyniki eksperymentu. Chodzi o opisany w części 3.4.2. *Heurystyki i wynikające z nich tendencje w zachowaniu i podejmowaniu decyzji istotne w kontekście uprzedzeń międzygatunkowych*, efekt, w ramach którego kolejność prezentacji bodźców, w tym wypadku warunków eksperymentalnych wpływa na wyniki eksperymentu. W kontekście omawianych badań efekt miejsca w szeregu może prowadzić do zniekształcenia wyników, w postaci odmiennej reakcji uczestników w zależności od tego, czy dana manipulacja pojawiła się wcześniej, czy później.

Zdjęcie 15. Pozycje eksperymentatorów w warunku tył — przód.



[Źródło: Povinelli, Eddy, 1996, str. 69.]

Zdjęcie 16. Pozycje eksperymentatorów w warunku skupiony — rozproszony.



[Źródło (Povinelli, Eddy, 1996, str. 87)].

Wyniki pierwszej fazy wskazują, że uczestnicy podążali za wzrokiem (E2), na poziomie wyższym niż losowy, Povinelli i Eddy zaznaczają nawet, że uczestnicy zachowywali się tak, jakby śledzili wzrok eksperymentatora, co mogłoby sugerować, że rozumieją, że ktoś inny coś widzi. Jednak ich reakcje nie były konsekwentne, czyli nie wykazywały preferencji gestykulowania do (E1). Dlatego autorzy przyjmują, że ich reakcje nie były wystarczająco regularne, by można było uznać, że szympanse są zdolne do wizualnego przyjmowania perspektywy innego. Ciekawe jest dla mnie, że Povinelli i Eddy (1996) formułują swoją interpretację zachowań szympanów, na podstawie gestykulacji, chociaż biorąc pod uwagę cel, jakim jest przetestowanie zdolności wizualnego przyjmowania perspektywy innego, ich podstawowym punktem odniesienia, powinno być śledzenie kierunku patrzenia. Jednak chociaż autorzy zauważają, że uczestnicy śledzili kierunek patrzenia, czyli wykazywali zachowanie, które powinno stanowić podstawowe odniesienie w formułowaniu odpowiedzi na pytanie postawione w ich badaniach, zignorowali tę przesłankę na rzecz odniesienia się do losowego gestykulowania. Co ciekawe nie pojawia się pytanie: czy uczestnicy rozumieli zadanie? Jest to wyjątkowo ciekawe, ponieważ gest *proszenia* interpretuje się jako prośbę o jakiś obiekt np. jedzenie, natomiast w warunku *attending-versus-distracted* żaden z eksperymentatorów nie oferował zwierzęciu pożywienia, ani żadnego innego obiektu. Autorzy nie poświęcają tej kwestii szczególnej uwagi, jednak uważam, że brak elementu, o który szympan miałby prosić, wykonując zadanie, mógł prowadzić do pewnego niezrozumienia. Uznać należy, że gest proszenia jest sygnałem komunikacyjnym wyrażającym prośbę „daj mi x”, natomiast w opisanej procedurze x nie było obecne, co mogło wywołać dezorientację zwierzęcia. Autorzy jednak nie biorą tego pod rozwagę, zamiast tego sprawdzają, czy brak preferencji gestykulowania mógł wynikać z rozproszenia uczestników, spowodowanego nową sytuacją eksperymentalną, nie zaś faktycznym deficytem w omawianym zakresie.

Druga faza eksperymentu również składała się z czterech sesji, obejmujących sześć prób. Tym razem całe dwie pierwsze sesje miały charakter kontrolny, gdzie obaj eksperymentatorzy siedzieli patrząc na wprost, jeden z nich trzymał klocek, drugi pożywienie. Te 12 prób miało na celu przyzwyczajenie uczestników do sytuacji, w której oboje eksperymentatorów patrzyło na wprost. Dwie kolejne sesje miały charakter badawczy, wykorzystywały warunek (*baseline-plus-attending-versus-distracted*), w którym (E2) trzymający jedzenie, patrzył w górę, natomiast (E1)

niezmiennie kierował wzrok na otwór odpowiedzi, trzymając w ręce drewniany klocek (Zdjęcie 17).

Przewidywania obydwu ram interpretacyjnych były następujące: jeśli założenia behawiorystyczne były poprawne, to uczestnicy powinni wykazywać preferencje gestykulowania do (E2), co według autorów byłoby dowodem, na to, że szympanse ignorują kierunek patrzenia i wykonują gest do eksperymentatora, trzymającego jedzenie, natomiast jeśli szympanse wypadłyby w tych próbach losowo, czyli nie wykazywałyby żadnej preferencji, uznano by, że wyniki wspierają perspektywę mentalistyczną. Jest to interesująca konstrukcja proceduralna, gdyż zakładając, iż uczestnicy nie rozumieli postawionego przed nimi zadania, nie istnieje żadna przesłanka, dla której mieliby gestykulować, do (E1), który na nich patrzy oraz oferuje drewniany blok, zamiast pożywienia. Jednak autorzy przekonani są, że wyniki ich badań, wspierają perspektywę behawioralną, w ramach której wysoka trafność, jaką uzyskali uczestnicy w eksperymentach 10 i 11, nie jest konsekwencją ich zdolności do wizualnego przyjmowania perspektywy innego, a wyuczenia się reguły „gestykulować do osoby, której twarz jest widoczna” (*gesturing toward the experimenter whose eyes and/or face was visible*). Korzystanie ze wskazanej reguły, miało powodować dezorientację w sytuacji, gdy żaden eksperymentator nie miał zakrytej twarzy. Co ciekawe, wyniki drugiej fazy również wykazywały, że zwierzęta śledzą spojrzenie (E2), autorzy powołują się nawet na inne własne badania, które w tamtym czasie nie zostały jeszcze opublikowane, ale pokazują, że szympanse śledzą kierunek patrzenia, niezależnie od tego, czy człowiek kieruje w inną stronę całą głowę, czy tylko oczy (Povinelli, Eddy, 1996, str. 94).

Zdjęcie 17. Pozycje eksperymentatorów w warunku *baseline+attending-versus-distracted*



[Źródło (Povinelli, Eddy, 1996, str. 87)]

Jak twierdzą Povinelli i Eddy (1996), wyniki pokazują, że szympansy wyposażone są w system do śledzenia patrzenia, jednak jest on ewolucyjnie prostszy, a jego działanie nie jest związane ze zrozumieniem konsekwencji skupiania uwagi. Povinelli, Eddy (1996, str. 95) ponadto sugerują, że jeśli śledzenie spojrzenia może zachodzić bez zrozumienia stanów percepcyjnych innej istoty, to należałoby zrewidować badania nad niemowlętami ludzkimi, które traktują ten mechanizm jako dowód na zdolność do wizualnego przyjmowania perspektywy innego. W takim ujęciu należałoby uznać, że niemowlęta – podobnie jak szympansy – mogą podążać za wzrokiem innych bez rzeczywistego rozumienia ich perspektywy wzrokowej. Twierdzenie to wydaje się nadmiernie sceptyczne i nieuzasadnione w świetle dostępnych danych (Brooks, Meltzoff, 2005). Choć rzeczywiście istnieje możliwość, że zarówno szympansy, jak i ludzkie niemowlęta mogą podążać za kierunkiem patrzenia innej istoty bez głębszego rozumienia stanów mentalnych, to jednak interpretacja ta zdaje się opierać na nadmiernie redukcjonistycznym zastosowaniu zasady Morgana. Choć w interpretacji zachowań zwierząt należy preferować najprostsze możliwe wyjaśnienia, to jednak redukcjonizm ten nie powinien prowadzić do ignorowania bardziej złożonych mechanizmów, gdy istnieją na to empiryczne przesłanki. Upieranie się, że śledzenie

wzroku nie może być związane ze zrozumieniem percepcji innych, ponieważ można zaproponować prostsze wyjaśnienie, jest przykładem błędnego zastosowania argumentu *ad ignorantiam* – odrzucania bardziej złożonego wyjaśnienia wyłącznie z powodu braku całkowitej pewności co do jego poprawności. Uważam, że w ramach wyżej opisanych eksperymentów autorzy, nie przedstawiają dowodu na, brak rozumienia uwagi – jedynie stwierdzają, że ich badanie nie dostarcza dowodów na takie rozumienie. To jednak nie jest równoznaczne z dowodem na brak tej zdolności. Ich interpretacja ignoruje możliwość, że szympansy mogą mieć inną niż ludzie formę reagowania na uwagę eksperymentatora. Przede wszystkim założenie, że śledzenie spojrzenia może być wyłącznie prostą reakcją percepcyjną, nie uwzględnia szerokiego kontekstu funkcjonalnego, w jakim zachowanie to się pojawia. W przypadku ludzi stanowi ono jeden z fundamentów rozwoju *mindreadingu* – umożliwia koordynację uwagi, rozumienie intencji oraz uczestnictwo w interakcjach społecznych. Baron-Cohen (1995) sugerował, że śledzenie wzroku stanowi kluczowy etap w ewolucji zdolności mentalnych, co znajduje potwierdzenie w licznych badaniach nad niemowlętami (Butterworth, Jarrett, 1991; Brooks, Meltzoff, 2002; Senju, Csibra, 2008).

Uważam, że argumentacja Povinelli’ego i współpracowników (1996) narusza zasady dobrych praktyk epistemicznych opisanych w pierwszej części niniejszej dysertacji. Autorzy przyznają, że szympansy śledzą spojrzenie innych, jednak uznają, że konsekwencją tego mechanizmu nie jest rozumienie, co inny widzi. Argumentują, że szympansy biorące udział w badaniu nie wykazywały preferencji gestykulowania do odpowiedniego eksperymentatora. Takie podejście stanowi przykład selektywnego traktowania dowodów, co narusza zasadę unikania wad epistemicznych. Dodatkowo język używany przez autorów, gdy wyniki sugerują trafność perspektywy mentalistycznej, korzystają z określeń takich jak „najbardziej uderzające” (*the most striking*) co sugeruje, że taki wynik jest niezgodny z ich oczekiwaniami (Povinelli, Eddy, Hobson, Tomasello, 1996, str. 104). Co więcej, Povinelli i współpracownicy (1996), wprost stwierdzają, że ich badania pokazują, iż zachowania pozornie wskazujące na perspektywę mentalistyczną można wyjaśnić w kategoriach prostszych mechanizmów behawioralnych. Tego rodzaju sformułowanie sugeruje, naruszenie zasady wrażliwości epistemicznej – jeśli ich hipoteza byłaby błędna, ich metodologia i sposób interpretacji danych nie pozwalałyby im tego dostrzec. W konsekwencji wszelkie dowody przemawiające przeciwko behawioralnym interpretacjom zachowania

szympanów zostają uznane za niewłaściwe, co sprawia, że argumentacja Povinelliego nie spełnia kryterium śledzenia prawdy.

5.2.2. Stanowisko Sary Shettleworth

Sara Shettleworth jest jedną z kluczowych postaci w dziedzinie psychologii porównawczej i badań nad poznaniem zwierząt. Choć nie przeprowadza bezpośrednich badań empirycznych, także tych z zakresu poznania społecznego, jej wkład w tę dziedzinę polega na krytycznej analizie literatury oraz interpretacji wyników badań innych naukowców. W swoich publikacjach, takich jak *Cognition, Evolution, and Behavior* (2010a) czy *Clever animals and killjoy explanations in comparative psychology* (2010b) podkreśla znaczenie sceptycznych postaw wobec twierdzeń o istnieniu teorii umysłu u zwierząt innych niż ludzie. Wskazuje przy tym liczne trudności metodologiczne związane z testowaniem omawianej zdolności. Jednocześnie zaleca przyjmowanie ostrożnych postaw do stanowisk argumentujących w perspektywie mentalistycznej. Uważam, że postawa Shettleworth prezentowana w ramach dyskusji dotyczącej *mindreadingu*, cechuje się nadmiernym antroponegacjonizmem. W celu poparcia swojej tezy, chcę przeanalizować dwa wspomniane już teksty, czyli fragment książki oraz artykuł (Shettleworth 2010a, str. 441-455, 2010b). Teksty te świetnie się uzupełniają, prezentując nadmiernie antroponegacjonistyczne tendencje w interpretacji eksperymentów dotyczących zdolności zwierząt pozaludzkich do *mindreadingu*. Podręcznik *Cognition, Evolution, and Behavior* (Shettleworth 2010a, str. 441-455), zawiera szczegółowe komentarze autorskie odnoszące się do wielu kluczowych w omawianym obszarze badań, także tych, przeanalizowanych przeze mnie we wcześniejszych częściach (Premack, Woodruff, 1978; Povinelli, Nelson, 1990; Povinelli, Eddy, 1996). Natomiast *Clever animals and killjoy explanations in comparative psychology* (2010b) odczytuję jako informacyjny w kontekście afektywnego komponentu uprzedzeń autorki. Język używany przez Shettleworth ma charakter wręcz żartobliwy, a jej argumentacja jest znacznie mniej ostrożna pod względem merytorycznym. W jego tytule znajdują się dwa istotne odniesienia, do koncepcji, stanowiących pole krytyki Shettleworth. Po pierwsze, określenie interpretacji behawioralnych jako „zabijających zabawę” (*killjoy*), stanowi prześmiewcze nawiązanie do prac Denneta (1991), który określił w taki sposób podzielane przez nią argumenty. Dodatkowo angielskie słowo *clever* (tłum. sprytny) ma charakter pobłażliwy, a sama autorka zaznacza, że odnosi się ono do interpretacji mentalistycznych, przypisujących

cechy typowo ludzkie (*human-like*) innym zwierzętom. Zaznacza, że przypomina jej to czasy Darwina, który zaproponował sposób porównywania możliwości poznawczych ZINL z ludźmi. Autorka wskazuje, że sam Dawin był zwolennikiem anegdot, świadczących o tym, że inne zwierzęta rozwiązują problemy w sposób podobny do człowieka¹⁰⁷. Co jak pisałam w części 3.1 *Historia badania umysłów zwierząt pozaludzkich*, jest prawdą, świadczy za tym między innymi fakt, że na swojego następcę wybrał Ramonesa, prezentującego nadmiernie antropocentryczne postawy zniekształcające, analizy obserwacji. W jej argumentacji dostrzegam wyraźne przejawy antroponegacjonizmu rozumianego za de Waalem, jako tendencję do negowania *a priori* człowiekopodobnych cech u ZINL, któremu towarzyszy ignorowanie doniesień sugerujących trafność modeli mentalistycznych (komponent behawioralny); (de Wall, 2018, str. 39). Odnoszę także wrażenie, że podstawowym czynnikiem motywującym autorkę omawianego artykułu, jest obrona narracji (komponent afektywny) głoszącej, że poparcie dla antropomorficznych wyjaśnień zachowania zwierząt oparte jest na ograniczonym rozumieniu alternatyw w postaci prostych wyjaśnień oferowanych przez psychologię eksperymentalną i społeczną.

Jej krytyka opiera się na założeniu, że wiele zachowań interpretowanych jako dowód na zdolność do przypisywania innym stanów mentalnych można również wyjaśnić prostszymi mechanizmami poznawczymi, takimi jak warunkowanie, asocjacje czy wnioskowanie oparte na doświadczeniu. W konsekwencji autorka *Clever animals and killjoy explanations in comparative psychology* (2010b), przyjmuje skrajnie redukcjonistyczną postawę, opartą o kanon Morgana¹⁰⁸. W tym miejscu należy nadmienić, że autorka zauważa, że jest on „[...] powszechnie uważany za zakazujący niepopartego antropomorfizmu lub za formę brzytwy Ockhama¹⁰⁹”¹¹⁰ (Shettleworth,

¹⁰⁷ Argumentacja Shettleworth (2010b) zawarta w omawianym artykule odnosi do przypisywania zwierzętom innym niż ludzie zdolności powszechnie uznawanych za typowo ludzkie. Swoją uwagę autorka poświęca szczególnie zdolnością z zakresu poznania społecznego, w tym teorii umysłu, a także procesom uczenia się przez wgląd.

¹⁰⁸ Dla przypomnienia kanon Morgana głosi, że nie wolno tłumaczyć zachowania zwierząt za pomocą wyższych procesów psychicznych, jeśli można je przypisać procesom niższym.

¹⁰⁹ Brzytwa Okhama to powszechnie znana zasada funkcjonująca w paradygmacie skąpca poznawczego głosi, że podczas wyjaśniania zjawisk należy kierować się przede wszystkim prostotą. Zatem powinny składać się z jak najmniejszej liczby użytych w niej pojęć oraz przesłanek.

¹¹⁰ Tłumaczenie własne z [...] *commonly taken as forbidding unsupported anthropomorphism or as a form of Ockham's Razor; this precept states* (Shettleworth, 2010b, str. 477).

2010b, str. 477). Pierwsze z wymienionych definiuje, jako „przypisywanie cech ludzkich (*human-like*) innym zwierzętom, zwykle z sugestią, że odbywa się to bez solidnego uzasadnienia”¹¹¹ (Shettleworth, 2010b, str. 477). Zatem definicja ta różni się znacząco od tej przyjmowanej w niniejszej pracy. Przyjęte w części 3.1. *Historia badania umysłów zwierząt pozaludzkich*, wyznaczają szerszy zakres pojęciowy dla antropomorfizmu. Kiedy Shettleworth kładzie wyraźny nacisk na ryzyko związane z niepopartym antropomorfizmem, radząc przy tym unikać wyjaśnień zachowań zwierząt pozaludzkich, odnoszących się do procesów poznawczych nadmiernie przypominających ludzkie (*human-like*). Tym samym, w swoich analizach nawiązuje do stanowisk behawiorystycznych, podkreślając, że nawet skomplikowane reakcje zwierząt mogą wynikać z procesów uczenia się, a nie z rzeczywistego rozumienia stanów mentalnych innych istot. Ja natomiast przyjmuję, że antropomorfizm jest postawą o charakterze kulturowym i religijnym, której komponent poznawczy składa się ze stereotypów i mitów konstytuujących porządek społeczny. Oznacza to, że nie jest ona z natury niewłaściwa, oczywiście w historii badań dotyczących zdolności z zakresu społecznego u ZINL, znaleźć można przykłady bezzasadnego, przypisywania zwierzętom innym niż ludzie zdolności poznawczych specyficznych dla człowieka — w tym sensie, rozumiem określenie niepopartego antropomorfizmu (Schultz, Schultz, Andruszko, 2008). Tak samo uważam, że konsekwentne i bezzasadne odrzucenie wyjaśnień zachowania ZINL, odnoszących się do zdolności poznawczych i stanów mentalnych, nazwać należy niepopartym antroponegacjonizmem. Przyjmując aparat pojęciowy zaproponowany we wcześniejszych częściach niniejszego opracowania — przypisywanie ZINL zdolności do *mindreadingu*, opisać można, jako spektrum, którego przeciwległe bieguny stanowią antropomorfizm oraz antroponegacjonizm. Uporczywe odrzucanie argumentów pochodzących z któregośkolwiek bieguna, czy ignorowanie przesłanek niepasujących do konsekwentnie przyjmowanych narracji, interpretować należy, jako uprzedzenie pozytywne albo negatywne. Zatem ocena wiarygodności interpretacji eksperymentów z zakresu *mindreadingu* przez zwierzęta pozaludzkie narażona jest na wpływ tendencji wynikających z przyjmowania przez badaczy postaw antropomorficznych lub antroponegacjonistycznych. To cechuje się przypisywaniem większej wiarygodności narracją i przekonaniom (komponent poznawczy)

¹¹¹ Tłumaczenie własne z *Anthropomorphism: the attribution of human qualities to other animals, usually with the implication it is done without sound justification* (Shettleworth, 2010b, str. 477).

metalistycznym czy behawioralnym, a także potrzebie do obrony własnych przekonań (komponent afektywny). W efekcie jednostka przejawia tendencje (komponent behawioralny) do faworyzowania argumentów zgodnych z przyjmowaną przez siebie postawą. Dla przypomnienia chcę podkreślić, że mechanizm ten jest naturalną konsekwencją architektury poznawczej ludzkich systemów poznawczych, dlatego nie należy go rozumieć w klasycznych kategoriach dobry/zły. Dopóki badacze zdają sobie sprawę z tendencji, które mogą wpływać na kreowane przez nich narracje, mogą uważać na negatywny wpływ kategoryzacji społecznej. Problem pojawia się, gdy prezentowana postawa jest uprzedzeniem, prowadzącym do przyjmowania skrajnych perspektyw oraz uporczywego negowania zarówno doniesień empirycznych, jak i teoretycznych z przeciwnego bieguna spektrum (komponent poznawczy). Co często przyjmuje formę ignorowania lub wręcz nawet unikania argumentów niezgodnych z prezentowaną przez siebie postawą (komponent behawioralny), a także przejawia się silną potrzebą do obrony własnych przekonań (komponent afektywny).

Ważnym aspektem stanowiska Shettleworth jest zwrócenie uwagi na różnice międzygatunkowe w zakresie przetwarzania informacji, możliwości poznawczych oraz ich ekologii, w tym warunków życia. Ponadto proponuje podejście oparte na analizie adaptacyjnej, które kładzie nacisk na presję selekcyjną, uwzględniającą strategię przetrwania danego gatunku. Ja również popieram stanowisko, że właściwe interpretacje eksperymentów dotyczących zdolności zwierząt pozaludzkich do *mindreadingu*, takich jak teoria umysłu, powinna uwzględniać wskazane przez Shettleworth warunki. W tym miejscu niniejszego opracowania, szczególnie ważne są różnice terminologiczne. Prace analizowane przeze mnie mają już prawie 15 lat, chociaż w perspektywie stuleci omawianych w niniejszej dysertacji, może się to zdawać niewiele, jednak stanowi to znaczny przeskok czasowy dla omawianej dziedziny. Postaram się wyraźnie zaznaczyć punkty wspólne dla stanowiska własnego ze stanowiskiem znanej badaczki, a także, unaocznic zachodzące między nami rozbieżności.

Jednym z kluczowych argumentów Shettleworth jest twierdzenie, że badania nad teorią umysłu u zwierząt pozaludzkich nie uwzględniają alternatywnych wyjaśnień (Shettleworth, 2010a). Jest to ciekawy zarzut, ponieważ wszystkie eksperymenty, które opisałam w poprzedniej części tego rozdziału, uwzględniały różne interpretacje wyników oraz szacowały ich trafność. Premack i Woodruff (1978) wskazują na istotne ograniczenia zastosowanych przez siebie procedur, natomiast Call czy Tomasello

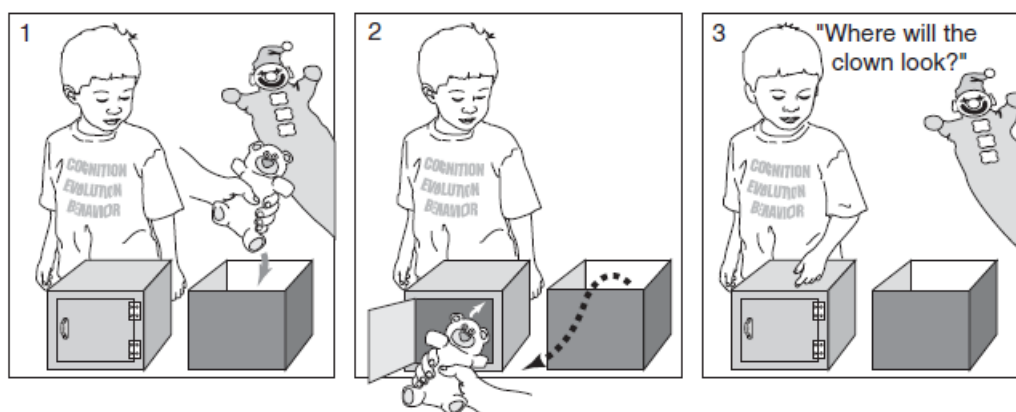
(2011) przez długi czas przekonywali, że tylko człowiek zdolny jest do przypisywania innym fałszywych przekonań. Natomiast Povinelli i Nelson (1990) przedstawiają kontrargumenty do stanowiska wyjaśniającego zachowania uczestników w kategoriach teorii uczenia się (Povinelli, Nelson, 1990, str. 207). Co więcej, zaznacza, że choć ludzie posiadają rozbudowaną teorię umysłu¹¹², u innych gatunków może ona przybierać bardziej ograniczone formy. Postulując, że niektóre zwierzęta mogą wykazywać ograniczone rozumienie intencji czy wiedzy innych osobników, ale niekoniecznie oznacza to, że posiadają zdolność do kompleksowego przypisywania przekonań, także tych fałszywych, jak ma to miejsce u ludzi. Deklaracja ta jest tak samo wiarygodna, jak ta Dunbara (2014). Pomimo wskazanego twierdzenia, w każdym z tekstów autorka utożsamia teorię umysłu z *mindreadingiem*, stosując przy tym skrót myślowy stawiający znak równości między omawianą zdolnością a czytaniem umysłów innych. Następnie autorka argumentuje, że ZINL nie są zdolne do przypisywania innym stanów mentalnych w formie metareprezentacji. Co sprawia, że jej deklaracje są w rzeczywistości puste. Stawia przy tym bardzo wysokie wymagania wobec eksperymentów mających potwierdzać istnienie teorii umysłu u zwierząt. Twierdzi m.in., że aby mówić o posiadaniu teorii umysłu, zwierzę powinno poprawnie reagować już od pierwszych prób w sytuacjach eksperymentalnych, co nie pozwala na zastosowanie prostych generalizacji z wcześniejszych doświadczeń. Jest to o tyle trudne, że w czasach, gdy omawiane teksty były publikowane, nauka nie dysponowała zaawansowanymi narzędziami, takimi jak *eye-tracking*. Dzięki nim można przeprowadzić eksperyment bez informowania uczestnika o jego zadaniu, co stanowiło dużą przeszkodę dla badań wykorzystujących detekcję kierunku patrzenia. Co więcej, autorka podkreśla konieczność projektowania eksperymentów w taki sposób, by umożliwiały one rozróżnienie pomiędzy konkurencyjnymi wyjaśnieniami zachowania, w tym szczególnie wykluczenie prostszych mechanizmów poznawczych. Takie założenie uważam, za praktycznie nieosiągalne ze względów metodologicznych. Po pierwsze celem procedury eksperymentalnej nie powinno być potwierdzanie jednej, z góry przyjętej interpretacji, lecz dążenie do najbardziej trafnego i zgodnego z regułą parsymonii wyjaśnienia danych empirycznych. Po drugie jak powszechnie wiadomo,

¹¹² Należy pamiętać, że wspomniana badaczka posługuje się pojęciem teoria umysłu w szerokim rozumieniu, czyli w jej pracach pod tym pojęciem kryje się zbiór zdolności nazywanych w niniejszej pracy czytaniem umysłów innych.

współczesne narzędzia naukowe nie umożliwiają metodologicznego oddzielenia zdolności z zakresu *mindreadingu* od sprytnego interpretowania zachowania. Przez dekady eksperymenty traktowane jako markery rozwoju ID, EED czy teorii umysłu, oceniane z perspektywy behawioralnej, stopniowo tracą swoją wiarygodność jako dowody na istnienie odpowiednich zdolności.

Zatem Shettleworth (2010a; 2010b) argumentuje za skrajnymi narracjami behawioralnymi, które postulują reinterpretację eksperymentów z zakresu *mindreadingu*, zaprzeczając ich informacyjności w zakresie przypisywania innym stanów mentalnych. W ten sposób odpira argument dotyczący luki poznawczej, twierdząc, że nie jest ona realna, a wytworzona przez przypisywanie ludzkim dzieciom zbyt wysokich możliwości (Shettleworth 2010a). Odwołuje się do testów fałszywych przekonań, a konkretnie paradygmatu z zabawkowym klaunem (zdjęcie 18), alogicznym do opisanego wcześniej¹¹³ testu z lalką Sally (zdjęcie 8).

Zdjęcie 18. Testy fałszywych przekonań w paradygmacie zabawkowego klauna.



[Źródło (Shettleworth, 2010a, s.441)]

Jako wyjaśnienie wyników pochodzących ze wskazanego paradygmatu proponuje mechanizm skojarzeniowy opisany po raz pierwszy przez Povinelliego i Vonka (2004). Zgodnie z ich propozycją wysokie wyniki zwierząt w testach prowadzonych w paradygmacie fałszywych przekonań, wynikają z zastosowania behawioralnej abstrakcji (*behavioral abstraction*). Jest to mechanizm z zakresu *smart behavior-reading*, który pozwala na lokalizację obiektów na podstawie wskazówek

¹¹³ Szczegółowe omówienie zagadnienia testów fałszywych przekonań znajduje się w części 4.3. *Test fałszywych przekonań*.

behawioralnych, za pomocą reguły *wybierz miejsce, w którym wychodzący widział przedmiot poraz ostatni*. Zwolennicy takiego wyjaśnienia postulują, że poprzez kombinację predyspozycji do zauważania i zapamiętywania, zwierzęta kategoryzują wskazówki behawioralne, jako istotne dla określonych systemów behawioralnych lub funkcjonalnie powiązanych reakcji. Korzystając z wyżej przytoczonego przykładu z sójkami — traktują one osobnika, który widział, jak chowają pokarm, jako konkurenta. W wyniku czego, gdy ten może zdobyć pożywienie, angażowane są specjalne dla gatunku zachowania obronne, które zmieniają się w zależności od przestrzeni (w ramach, której odbywa się sytuacja) lub statusu społecznego konkurenta. Uważam, że chociaż niektóre interpretacje mentalistyczne, przypisują ZINL z pozoru bardziej złożone zdolności, w rzeczywistości stanowią bardziej prawdopodobną, spełniającą zasadę parsymonii interpretację eksperymentów. Postulowanie, że zwierzęta posiadają zdolność rozumienia stanów mentalnych innych osobników — na przykład wiedzy, intencji czy przekonań — pozwala wyjaśnić ich elastyczne i szybkie reakcje bez konieczności zakładania ogromnych, złożonych zbiorów mechanizmów asocjacyjnych. Taki system asocjacyjny wymagałby znacznych zasobów poznawczych, szczególnie tych pamięciowych oraz byłby mniej adaptacyjny w zmiennych warunkach środowiskowych i społecznych. Natomiast zdolność do mentalistycznego rozumienia innych umożliwia zwierzętom natychmiastowe przewidywanie zachowań konkurentów i dostosowywanie własnych reakcji w sposób bardziej elastyczny i skuteczny. W tym świetle mentalistyczne wyjaśnienia nie tylko lepiej tłumaczą obserwowane zachowania, ale także spełniają zasadę oszczędności poznawczej. Pozwalają na bardziej zintegrowane i uniwersalne wyjaśnienie złożonych interakcji społecznych, bez konieczności mnożenia bytów (w postaci rozbudowanych systemów asocjacyjnych). Zdolność do rozumienia, co inni wiedzą lub czego chcą, stanowi więc efektywny mechanizm adaptacyjny, który ewolucyjnie mógł być preferowany ze względu na swoją elastyczność i ekonomię poznawczą. Oznacza to, że choć interpretacje oparte na „prostych” mechanizmach asocjacyjnych wydają się intuicyjne, to właśnie podejście mentalistyczne oferuje bardziej oszczędne i funkcjonalne wyjaśnienie złożonych zachowań zwierząt w kontekście rywalizacji o zasoby i innych interakcji społecznych. Co ciekawe, autorka zauważa, że badania świadczące o wyższych możliwościach poznawczych ZINL oparte są na skrupulatnych kryteriach behawioralnych, które służą do testowania mentalistycznych przewidywań (Clayton, Dickinson, 1998; Heyes, 2008). Jednocześnie uznaje, że ich celem nie jest przeciwstawienie się dobrze zdefiniowanym

alternatywom, a tendencyjne wpieranie antropomorficznych interpretacji (Shettleworth, 2010b, str. 478). Jak już wspominałam, język użyty *Clever animals and killjoy explanations* (2010b) odzwierciedla postawę, którą można odczytać jako uprzedzoną wobec interpretacji mentalistycznych. Dobrze obrazuje to sam początek tekstu, w którym autorka omawia zagadnienie złożoności wynikającej z prostoty. Przytacza tam przykład złożonego zachowania w postaci wręcz perfekcyjnie klimatyzowanego gniazda termitów (*perfectly air-conditioned termite nest*); (Shettleworth, 2010b, str. 477). Podsumowuje ten przykład żartobliwym stwierdzeniem, że chociaż może się to wydawać niewiarygodne, w naturze nie występują architekci termitów (*incredible as it seems, there is no termite architect*), co nie tylko trywializuje interpretacje mentalistyczne, ale sugeruje brak gotowości do ich uczciwego rozważenia.

Dla przypomnienia, autorka odpiera zarzuty dotyczące braku ciągłości ewolucyjnej poprzez wyjaśnienie, że ludzkim dzieciom przypisuje się znacznie wyższe procesy poznawcze, niż te badane w rzeczywistości. Uważa, że jakościowa różnica między człowiekiem a innymi zwierzętami ma źródło w ludzkich kompetencjach językowych. Nie oferuje jednak wyjaśnienia, jak mechanizm sprytnego zachowania przekształca się w *mindreading* oraz nie wskazuje ontogenetycznych markerów rozwoju złożonych procesów metareprezentacyjnych występujących u dorosłych ludzi. Oczywiście można się domyślić, że w proponowanej przez Shettleworth metodzie badania *mindreadingu*, dominują analizy języka metareprezentacyjnego, o którym pisałam w części 4.4. *Stopniowalność teorii umysłu w ujęciu Robina Dunbara i José Luisa Bermúdeza*. Uważam, że takie zawężenie perspektywy nie jest zadowalające z metodologicznego punktu widzenia, ponieważ teoretyczne przewidywania dotyczące złożonych zdolności metapoznawczych zdają się implikować wcześniejszy, rozpoznawalny rozwój tych kompetencji. Dlatego uznaję, że takie wyjaśnienie nie zapełnia luki widocznej między człowiekiem a innymi zwierzętami, wręcz przeciwnie, pogłębia omawiany problem. Przyjmując, że ludzie zdolni są do tworzenia ToMM rzędu [4], [6] czy [9] rzędu, natomiast szympansom, czyli przedstawicielom gatunku ewolucyjnie najbliższego człowiekowi, odmawia się zdolności do teorii umysłu rzędu drugiego, a nawet prostszych mechanizmów z zakresu poznania społecznego, jak EED, czy ID (Baron, Cohen, 2009, Dunbar, 2014).

Jak zaznacza Shettleworth (2010a; 2010b), żadne z istniejących danych nie wymaga teorii umysłu do wyjaśnienia, zachowań zwierząt w eksperymentach prowadzonych w paradygmacie fałszywych przekonań, ale też jej nie wyklucza.

Dlatego staje na stanowisku, że należy kierować się regułą parsymonii, tytuł jednej z części *Clever animals and killjoy explanations* (2010b) bezpośrednio odnosi się do Brzytwy Okhama, a jego treść poświęcona jest zaprezentowaniu przykładów złożoności wynikającej z prostoty. Powołuje się w nim na wspomniany wyżej przykład gniazda termitów, co wspiera narrację, że natura pełna jest przykładów złożoności niewymagającej złożonych procesów takich jak uwaga. Odwołuje się także do argumentu powszechnie przyjmowanego w ramach perspektywy mentalistycznej, głoszącego, że pewnych procesów nie da się sprowadzić do prostszych mechanizmów. Następnie formułuje kontrargument, zakładający, że brak możliwości wyjaśnienia niektórych zachowań zwierząt w ramach perspektywy behawioralnej, jest wynikiem braku wystarczającej wiedzy. Jest to o tyle ciekawe, że stanowi krytykę zastosowania przez jej zwolenników argumentu *ad ignorantiam*. W tym przypadku przyjmuje on formę: nie ma dowodu na to, że określone zdolności poznawcze, mogą zostać zredukowane do prostszych procesów mentalnych. Autorka wręcz zaznacza, że we wskazanych przypadkach brak dowodu, nie stanowi dowodu na brak, świadczy jedynie za tym, że jeszcze nie znaleziono właściwego wyjaśnienia. Jak wielokrotnie podkreślałam we wcześniejszych rozdziałach, ja także dostrzegam wady interpretacji opartych o brak wiedzy, jednak w tym wypadku uważam za fascynujący fakt, że w tym samym tekście, autorka sama powołuje się na *ad ignorantiam*, jako podstawę dla skrajnie antroponegacjonistycznych twierdzeń. Ze względu na brak możliwości jednoznacznego potwierdzenia zaangażowania procesów metareprezentacyjnych w zadania stawiane przez ZINL w paradygmatach charakterystycznych dla ID, EED czy ToMM, autorka odrzuca interpretacje wyników tych badań, odnoszące się do stanów mentalnych. Broni argumentu o ciągłości ewolucyjnej swojego stanowiska, wysuwając twierdzenie, że faktycznie niektóre aspekty *mindreadingu* mogą być wspólne dla człowieka i innych zwierząt. Natomiast inne jego komponenty mają charakter unikatowy dla ludzi. Przytacza twierdzenia Povinelliego, że ZINL śledzą kierunek patrzenia, ale nie rozumieją, że poprzez patrzenie nabywana jest wiedza. Powołuje się także na inne badania wskazujące, że przypisywanie intencji innym może odbywać się na drodze przewidywania zachowania innego, na podstawie wskazówek behawioralnych (np. kierunek patrzenia) bez rozumienia leżących u ich podstaw stanów (Tomasello, Call, Hare, 2003).

Przechodząc do kolejnego etapu krytycznej analizy stanowiska Shettleworth, pragnę podkreślić, że zgadzam się z deklaracją autorki, w ramach której osoby konstruujące, przeprowadzające oraz interpretujące wyniki, badań z udziałem ZINL, powinny pamiętać o własnych uprzedzeniach. Uważam jednak, że Shettleworth podobnie do Povinelli'ego faworyzuje interpretacje zgodne z własnymi przekonaniem. Czego przykładem jest powoływanie się autorki na przykłady testów dotyczących przypisywania prostych stanów wolicjonalnych innym, śledzenia kierunku patrzenia czy preferencyjnego wyboru wskazań eksperymentatora. Argumentuje, że zwierzęta mogą kierować się wyuczonymi schematami lub subtelnymi wskazówkami środowiskowymi, zamiast rzeczywiście rozumieć perspektywę drugiego osobnika, nawet wtedy, gdy wnioski samych autorów nie są w tej kwestii tak jednoznaczne (Premack, Woodruff, 1978; Povinelli, Nelson, 1990). Pierwszym, mniej rażącym przykładem jest jej analiza eksperymentów Premacka i Woodruffa (1978). Autorka zauważa, że elastyczność zachowania Sary faktycznie zgodna jest z przewidywaniami modeli mentalistycznych, jednak autorka zaznacza, że Sara była 10-letnią szympanśką z dużym doświadczeniem eksperymentalnym. A także miała okazję obserwować ludzi podczas wykonywania codziennych czynności. Na tej podstawie stwierdza, że Sara mogła w prosty sposób wybrać zdjęcie, które było uzupełnieniem dobrze jej znanych sytuacji. Przedstawia także kolejne alternatywne wyjaśnienie — wykonanie rozumienia „od siebie do innych” (*self-to-other*), przyjmującego postać: „kiedy ja patrzę na coś oczami, widzę x, zatem kiedy ktoś inny robi to samo, musi mieć takie same stany umysłu”. Stanowi to oczywiste odniesienie do teorii symulacyjnych, które jak wskazałam w części 4.1. *Współczesne filozoficzne koncepcje mindreadingu*, posiadają ograniczoną wartość eksplanacyjną w omawianym obszarze. Należy przy tym przyznać, że chociaż Premack i Woodruff konstruowali swoje eksperymenty w ramach teorii metareprezentacyjnej, współczesne doniesienia empiryczne wskazują, że oglądanie nagrań angażuje neurony lustrzane, które faktycznie mogłyby zaangażować procesy symulacyjne (Gallese, Goldman, 1998). Oczywiście przy założeniu, że Sara ma doświadczenie z udziałem w eksperymentach z obszaru rozwiązywania problemów w paradygmacie wyznaczonym przez Köhlera. Trzeba przyznać, że nie jest to scenariusz niemożliwy, jednak nie wydaje mi się on istotnie konkurencyjny dla analizy interpretacji zawartej w części 5.1.1. *Przełomowe badania Premacka i Woodruffa*. Dla przypomnienia pomimo błędów popełnionych w procedurze eksperymentalnej aktualnie powszechnie

uznaje się eksperymenty zaprezentowane w 1978 roku, za dobry punkt odniesienia do badań *mindreadingu* poza gatunkiem ludzkim.

Bardziej rażące tendencje w interpretacji widoczne są w analizie wyżej opisanego eksperymentu Povinelliego i Nelsona (1990). Autorka pisze:

Knower, and in fact, all four chimpanzees tested chose correctly some of the time. However, because the experiment went on for 300 trials, the animals had plenty of opportunity to learn a conditional discrimination: choose the person who was in the room when the container was baited. (Shettleworth, 2010a, str. 442)

Dla przypomnienia, jak wskazałam w części, w której opisywałam wskazany eksperyment, większość szympanсів, w większości prób odpowiadała poprawnie, a wyniki pokazały, że ich preferencja do wybierania wskazówki osoby, która widziała, gdzie umieszczona jest przynęta, była istotna statystycznie. Jak wskazywali sami autorzy szympanсы bardzo szybko nauczyły się wykorzystywać wskazówkę od właściwego eksperymentatora, a ich decyzje szybko osiągnęły wysoki poziom trafności, nawet w zmieniających się warunkach. Povinelli i Nelson (1990) podkreślili w swoim artykule, że pomimo ilości prób, uczestnicy szybko zaczęli wykazywać wyraźną preferencję wyboru osoby, która widziała. Osiągając pewien poziom trafności, wynik szympanсів nie poprawiał się z tygodnia na tydzień, co może sugerować, hipotezę przeciwną do stawianej przez Sarę Shettleworth (2010, str. 442). Co więcej, wskazują, że za wyjątkiem Kermita, nie zaobserwowano uczenia się w trakcie eksperymentu (Povinelli, Nelson, 1990, str. 207). W takich okolicznościach uważam, że określenie „przez pewien czas” (*some of the time*) użyte przez autorkę w przytoczonym wyżej cytacie, stanowi raczej przejaw tendencyjności przejawiającej się w figurze retorycznej, jaką jest litota, a nie rzetelnie świadczy o wynikach opisywanego przez Shettleworth artykułu.

W *Clever animals and killjoy explanations* (Shettleworth, 2010b) autorka zwraca uwagę na bardzo ciekawy wątek, uważa bowiem, że spadek popularności behawioralnych wyjaśnień zachowania zwierząt, wynika z mniejszej atrakcyjności narracyjnej popieranych przez nią interpretacji. Pisze nawet:

Entuzjazm, z jakim prasa popularna, a nawet naukowa, podchodzi do opowieści o sprytnych zwierzętach, sprzyja tej tendencji: psujące zabawę wyjaśnienia rzadziej trafiają na nagłówki niż historie o tym, że ośmiornice czy ptaki okazują się zaskakująco podobne do ludzi. (Shettleworth, 2010b, str.478).

Tym samym autorka sugeruje, że psychologia porównawcza częściej poszukuje „czegoś ludzkiego w zwierzętach” niż „czegoś zwierzęcego w ludziach”. Podzielam jej przekonanie, że taka asymetria epistemologiczna prawdopodobnie wynika z uprzedzeń. Uważam jednak, że zarzut ten powinien być przede wszystkim skierowany pod adresem osób zaangażowanych w szeroko pojęty przekaz medialny. Problem tendencyjnych doniesień – zwłaszcza w prasie popularnej i popularnonaukowej, a także w mediach społecznościowych – nie leży bowiem w gestii samych badaczy reprezentujących nurt mentalistyczny i nie może zostać przez nich jednoznacznie rozwiązany. Podzielam jednak jej obawy w zakresie wpływu takich przekazów na prezentowanie określonych postaw wobec ZINL, co szeroko opisałam w części 3.4.2. *Heurystyki i wynikające z nich tendence w zachowaniu i podejmowaniu decyzji istotne w kontekście uprzedzeń międzygatunkowych*.

Jednak Shettleworth rozwija odniesienie do wspomnianej asymetrii epistemicznej, w sposób, z którym trudno mi się z nią zgodzić. Argumentuje, że coraz więcej dowodów na znaczenie nieświadomych, automatycznych mechanizmów działania u ludzi podważa trafność mentalistycznych narracji nawet w kontekście ludzkiego poznania. Uważa zatem, że to co sama nazywa „odkrywaniem cech zwierzęcych u ludzi”, stoi w sprzeczności z utrzymywaniem, twierdzeń głoszących, że to co ludzkie (*human-like*) obecne jest także u innych gatunków. Według niej doniesienia wskazujące na prostotę mechanizmów przetwarzania informacji u człowieka powinny prowadzić do większego docenienia roli prostych mechanizmów u zwierząt pozaludzkich, co – jak pisze – umożliwiłoby rozwój „głębiej porównawczej psychologii poznawczej” (*a more deeply comparative cognitive science*); (Shettleworth, 2010b, s. 480). Uważam, że w istocie Shettleworth broni w ten sposób asymetrii epistemicznej, tyle że proponuje przesunięcie uwagi ku temu, co określa jako „odkrywanie cech zwierzęcych u ludzi”. W tym ujęciu niedostatecznie doceniona, a momentami wręcz marginalizowana zostaje wartość doniesień dotyczących „dostrzegania w innych zwierzętach tego, co ludzkie” – w tym złożonych, intencjonalnych procesów poznawczych. Uważam, że oba podejścia nie powinny być postrzegane jako wzajemnie wykluczające się, lecz jako komplementarne. Stanowią one

dwa bieguny tego samego epistemicznego kontinuum, a rzetelne badania nad poznaniem powinny uwzględniać zarówno możliwość istnienia ludzkich cech wśród innych gatunków, jak i zwierzęcych mechanizmów w obrębie ludzkiego umysłu. Tylko eksploracja obu tych kierunków – i napięcie między nimi – może doprowadzić do pełniejszego zrozumienia ewolucji i struktury procesów poznawczych.

Co najważniejsze, uważam, że sama rama pojęciowa, w której ujmuję się porównania międzygatunkowe – jako poszukiwanie „cech ludzkich u innych gatunków” lub „cech zwierzęcych u ludzi” – jest głęboko zakorzeniona w antropocentrycznym sposobie myślenia. Jak wielokrotnie podkreślałam w niniejszej dysertacji, ludzie są zwierzętami, a więc w naturalny sposób dzielą z innymi gatunkami wiele cech i mechanizmów. Próby wyodrębnienia tego, co „ludzkie” lub „zwierzęce”, mogą być heurystycznie użyteczne, ale zarazem zaciemniają rzeczywistość ontologiczną. Tego rodzaju rozróżnienia zakorzenione są w tradycji kartezjańskiego dualizmu oraz w humanistycznym dziedzictwie, które – jak pokazuje Donna Haraway (2008) – prowadzi do utrwalania porządku epistemicznego, w którym człowiek nieustannie zajmuje centralne miejsce poznawcze i moralne. W szczególności – sugerowanie, że pewne procesy są *niby* wspólne, albo że istnieje jakaś zasadnicza różnica, która dopiero musi zostać obalona lub potwierdzona, oznacza operowanie na granicy, która ma więcej wspólnego z tradycją filozoficzną niż z empiryczną wiedzą biologiczną. W efekcie nawet pytania badawcze sformułowane w ten sposób są niekiedy bardziej wyrazem naszej kulturowej potrzeby dystynkcji niż rzeczywistej potrzeby naukowego wyjaśnienia. Dlatego też uważam, że psychologia porównawcza powinna odejść od tego rodzaju binarnych podziałów i przyjąć bardziej ewolucyjnie spójną i metodologicznie neutralną perspektywę.

Zgadzam się z de Waal'em (2018), który wielokrotnie podkreśla, że to nie zwierzęta są „zbyt zwierzęce”, lecz raczej my jesteśmy „zbyt ludzcy” w swoich założeniach badawczych – co prowadzi do błędów zarówno typu I (fałszywe przypisanie) jak i typu II (fałszywe wykluczenie). Podobną myśl prezentuje Erika Fudge (2002), która zauważa, że „zwierzę” w kontekście badań naukowych często pełni funkcję epistemicznego kontrastu wobec „człowieka” – co sprawia, że różnice stają się bardziej istotne niż podobieństwa. Dlatego też uważam, że psychologia porównawcza – o ile ma ambicje rzeczywiście wyjaśniać mechanizmy poznawcze w świetle wspólnej ewolucyjnej historii – powinna odejść od binarnych podziałów typu „ludzkie versus zwierzęce” i przyjąć bardziej ewolucyjnie spójną oraz metodologicznie neutralną

perspektywę. Oznacza to nie tylko przeformułowanie hipotez badawczych, ale również rewizję samych kategorii, którymi się posługujemy.

5.3. Podsumowanie

Rozdział piąty stanowi krytyczną analizę interpretacyjnych napięć związanych z badaniem teorii umysłu u zwierząt innych niż ludzie. Zaprezentowałam w nim szczegółową analizę interpretacji eksperymentów – kluczowych w obszarze teorii umysłu, zarówno tych konstruowanych w ramach mentalistycznych, jak i behawioralnych. Wskazałam w jaki sposób przyjmowane przez badaczy ramy teoretyczne wpływają na interpretację wyników eksperymentalnych. Szczególne znaczenie przypisałam asymetrii w traktowaniu wyników badań z udziałem ludzi i innych zwierząt: podobne zachowania są często interpretowane odmiennie, w zależności od gatunku badanego podmiotu.

Na początku rozdziału przedstawiłam argumenty na rzecz interpretacji konstruowanych w perspektywie mentalistycznej, wskazując, że niektóre zachowania zwierząt trudno wyjaśnić bez odwołania do reprezentacji intencji, celów czy przekonań. Omówiłam klasyczne badania Premacka i Woodruffa (1978), stanowiące punkt wyjścia dla współczesnych analiz teorii umysłu u naczelnych, a także towarzyszące im kontrowersje metodologiczne. Przeanalizowałam również istotny artykuł autorstwa Povinelliego i Nelsona (1990), w którym autorzy sugerują, że szympansy mogą różnicować dostępność wiedzy u obserwatora, co stanowi przesłankę do wnioskowania, iż rozumieją one zależność między percepcją a stanem wiedzy. Wnioski te mogą być o tyle zaskakujące, że Daniel Povinelli —współautor badania — konsekwentnie projektuje swoje eksperymenty oraz interpretuje ich wyniki w ramach paradygmatu behawioralnego, zakładającego sceptycyzm wobec przypisywania zwierzętom stanów mentalnych. Co istotne, postulat ten odnosi również do badań z udziałem dzieci. Jednak w ich przypadku domniemywa się, że rozwój określonych zachowań – takich jak spojrzenia wyprzedzające – wiąże się z równoległym rozwojem zdolności do przypisywania stanów mentalnych. W przypadku zwierząt innych niż ludzie takie założenie nie jest przyjmowane, co prowadzi do interpretacyjnej asymetrii: identyczne zachowania u dzieci i u zwierząt oceniane są według odmiennych kryteriów teoretycznych. Szczegółowy opis stanowiska prezentowanego przez Daniela Povinelli znajduje się w części 5.2.1. *Daniel Povinelli i Timothy Eddy*. Dalsza część rozdziału

została poświęcona eksperymentom Calla i Tomasella, w których podjęto próbę precyzyjnego zdefiniowania teorii umysłu jako zdolności do przypisywania fałszywych przekonań. Zaprezentowano także współczesne badania wykorzystujące zaawansowane techniki, takie jak *eye-tracking*, które ujawniają potencjalne markery metareprezentacji u naczelnych. Zwrócono uwagę na metodologiczne i epistemologiczne wyzwania związane z interpretacją takich wyników, w szczególności na problem skrótu myślowego utożsamiającego *mindreading* z teorią umysłu.

W części 5.2. *Perspektywa krytyczna w kwestii teorii umysłu zwierząt innych niż ludzie: wybrane badania* podkreślam także, że interpretacje behawioralne, choć odwołujące się do zasady parsymonii, często opierają się na argumencie *ad ignorantiam*, zakładającym, że brak jednoznacznego dowodu na istnienie złożonych procesów poznawczych u zwierząt stanowi wystarczającą podstawę do ich odrzucenia. Tymczasem dowody empiryczne, choć niejednoznaczne, sugerują istnienie przynajmniej częściowych aspektów *mindreadingu* u niektórych gatunków, co kwestionuje ostre granice poznawcze dzielące człowieka od innych zwierząt i podważa tezy sprzeczne z darwinowską zasadą ciągłości ewolucyjnej. Zestawiając dane empiryczne z refleksją epistemologiczną, rozdział ukazuje, że rozróżnienie między *mindreadingiem* a *smart behaviour readingiem*, choć metodologicznie użyteczna, nie zawsze oddaje złożoność badanych zjawisk. Tym samym postuluję potrzebę większej ostrożności interpretacyjnej oraz otwartości na możliwość istnienia form pośrednich między prostymi mechanizmami skojarzeniowymi a pełnymi zdolnościami metareprezentacyjnymi.

Rozdział VI: niesprawiedliwość wynikająca z uprzedzeń względem zwierząt pozaludzkich

6.1. niesprawiedliwość związana z uprzedzeniami względem zwierząt pozaludzkich

Jednym z zainteresowań filozofii feministycznej są rozważania dotyczące etyki i polityki ludzkich praktyk epistemicznych. W swojej książce *Epistemic Injustice: Power and the Ethics of Knowing* (2007) Miranda Fricker argumentuje, że pewne grupy społeczne, takie jak kobiety, osoby kolorowe i inni członkowie społeczności marginalizowanych, często doświadczają niesprawiedliwości dotyczącej sfery wiedzy i informacji nazywanej epistemiczną. Źródłem tego zjawiska są uprzedzenia systemowe. Fricker twierdzi, że zrozumienie omawianej niesprawiedliwości i podjęcie działań w jej perspektywie, jest kluczowe dla stworzenia bardziej sprawiedliwego społeczeństwa. Na łamach swojej książki autorka wyróżnia dwa rodzaje niesprawiedliwości: testimonalną (*testimonial injustice*) i hermeneutyczną (*hermeneutical injustice*). Pierwsza z wymienionych ma miejsce, gdy wiarygodność jednostki jest podważana w określonym kontekście z powodu uprzedzeń. Dla przykładu uprzedzenia takie mogą przejawiać się tendencją do umniejszania lub braku uznania doświadczeń i potrzeb przedstawicieli grupy obcej (komponent behawioralny). Gdy ktoś doświadcza niesprawiedliwości testimonalnej, jego wypowiedzi nie są traktowane poważnie. Może to prowadzić do ignorowania lub niedowierzania osobie, nawet jeśli mówi prawdę i posiada wszelkie kompetencje do wygłaszania takich tez. Fricker przywołuje przykład, w którym kobieta, zgłasza molestowanie seksualne w miejscu pracy, może nie zostać potraktowana poważnie z powodu uprzedzeń płciowych lub założenia, że kobiety mogą przesadzać lub kłamać. Podobnie osoba o innym kolorze skóry może zostać zignorowana, gdy zgłasza doświadczenia związane z dyskryminacją lub rasizmem. Druga z wymienionych, to niesprawiedliwość hermeneutyczna, która pojawia się gdy jednostka doświadcza cierpienia na drodze krzywdy nie posiadającej nazwy w określonej kulturze/społeczności. Fricker zaznacza, że w takiej sytuacji powstaje luka hermeneutyczna, będąca źródłem niesprawiedliwości i niekorzyści na gruncie rozumienia własnego doświadczenia. Oznacza to, że niesprawiedliwość hermeneutyczna, odnosi się do sytuacji, w których marginalizowane grupy nie są w stanie w pełni wyrazić swoich doświadczeń lub perspektyw, z powodu braku zrozumienia społecznego lub zasobów językowych. Przykładem tego rodzaju

niesprawiedliwości jest sytuacja, w której kobieta cierpi z powodu nadużycia seksualnego, w czasie kiedy dane społeczeństwo nie wykształciło jeszcze pojęcia dla takiego rodzaju zachowania¹¹⁴.

Tak samo jak Fricker (2007) uważam, że „bycie podmiotem wiedzy” jest istotną wartością dla człowieka oraz dla przynajmniej niektórych innych żyjących istot¹¹⁵, a co za tym idzie, każda niesprawiedliwość epistemiczna jest krzywdząca. Wynika z niej cierpienie jednostki jako podmiotu wiedzy. Autorka *Epistemic Injustice: Power and the Ethics of Knowing* uważa, że krzywda wyrządzona przez każdą z wymienionych niesprawiedliwości, może być przyczyną np. zahamowania rozwoju osobistego. W niniejszej części rozważam opisywane przez Fricker (2007) zjawiska w kontekście uprzedzeń odnośnie do grupy jaką są zwierzęta pozaludzkie. Uważam także za istotne przedstawienie niesprawiedliwości jakiej mogą doświadczać osoby, które bronią praw ZINL lub przyjmują perspektywę mentalistyczną podczas interpretowania eksperymentów z udziałem zwierząt innych niż ludzie.

6.2. Podstawowe pojęcia w koncepcji niesprawiedliwości epistemicznej

Przechodząc do omówienia zagadnienia niesprawiedliwości, uważam za konieczne krótkie przedstawienie zjawiska władzy (*power*), które Fricker (2007) definiuje jako społecznie usytuowaną zdolność każdego podmiotu społecznego, służącą do kontrolowania działań innych. Jednak zanim przekieruję uwagę na władzę, chciałabym omówić kilka terminologicznych rozbieżności jakie pojawiają się między niniejszą dysertacją, a koncepcją Fricker (2007). Chodzi przede wszystkim o pojęcia

¹¹⁴ A co za tym idzie, podmiot takiej niesprawiedliwości nie jest w stanie pojąć swojego własnego doświadczenia.

¹¹⁵ Rozumiem przez to nie tylko odebranie głosu lub autorytetu poznawczego, ale także systematyczne pomijanie czy ignorowanie sygnałów, które jednostka wysyła w celu zakomunikowania swojego stanu. W tym sensie uważam, że każde stworzenie zdolne do przekazywania informacji o swoich potrzebach – choćby niejęzykowe – doświadcza dyskomfortu, gdy jego przekaz pozostaje niezauważony lub zlekceważony. Przykładem może być pies, którego opiekun nie reaguje na sygnały wskazujące na pragnienie, głód lub potrzebę wyjścia. Takie sytuacje stanowią moim zdaniem formę epistemicznej niesprawiedliwości, ponieważ podmiot zostaje pozbawiony sprawczości poznawczej w relacji. To natomiast uniemożliwia mu realizację najbardziej podstawowych potrzeb.

jakimi są uprzedzenia i stereotypy. Co więcej, autorka nie wyróżnia w swojej pracy kategorii mitów¹¹⁶.

W koncepcji przez nią rozwijanej stereotypy rozumiane są jako „szeroko przyjmowane asocjacje związane z grupą społeczną i przydzielanymi jej atrybutami. Skojarzenie to opiera się na generalizacji ujawniającej pewne, zazwyczaj epistemicznie złe inwestycje afektywne”¹¹⁷ (Fricker, 2007, str. 43). Takie zawężenie obszaru zainteresowań, pozwala na lepsze ujęcie stereotypów kulturowych, leżących u podstaw uprzedzeń systemowych. Natomiast uprzedzenia Fricker (2007) rozumie, jako sądy¹¹⁸ o charakterze pozytywnym lub negatywnym. Wykazują one pewne (zwykle epistemicznie) opory, wobec przyjmowania przekonań niezgodnych z istniejącymi już strukturami poznawczymi, ze względu jakąś afektywną inwestycję. Wydaje się, że zjawisko nazywane w omawianej książce uprzedzeniami, odpowiada mechanizmowi nazywanemu w niniejszym opracowaniu efektem potwierdzenia¹¹⁹. Należy pamiętać, że głównym obiektem zainteresowań autorki *Epistemic Injustice: Power and the Ethics of Knowing* jest systematyczna niesprawiedliwość epistemiczna, a przyjęcie wskazanych definicji, pomogło jej w lepszym zrozumieniu opisywanych przez nią pojęć.

W ramach koncepcji niesprawiedliwości epistemicznej uważa się, że gdy jednostka spostrzega nadawcę przyporządkowuje go do poznawczej kategorii społecznej, która odwołuje się (*invoking*) do stereotypów i mitów¹²⁰, np. dotyczących „wiedzy” o ludziach czy innych zwierzętach żyjących w świecie. Przydzielenie spostrzeganego obiektu społecznego do kategorii wiąże się z określeniem jego władzy tożsamościowej (*identity power*). Określenie tego parametru jest istotne dla procesów nadawania wiarygodności (*credibility*) interlokutorowi/interlokutorom sytuacji

¹¹⁶ Opisanych przeze mnie szczególnie w części 3.3.1. *Mity i inne narracje*.

¹¹⁷ Z angielskiego: „A widely held dispositive association between a social group and one or more attributes, where this association embodies a generalization that displays some (typically, epistemically bad affective investment)” (Fricker, 2007, s. 43).

¹¹⁸ Z angielskiego: „Prejudices are judgements, which may have a positive or negative valence, and which display some (typically, epistemically culpable) resistance to counter-evidence owing to some affective investment on the part of the subject.” (Fricker, 2007, s. 43).

¹¹⁹ Szczegółowe informacje na temat wskazanego mechanizmu znajdują się w części 2.5 *Teoria dysonansu poznawczego*.

¹²⁰ Sama autorka nie pisze o mitach, jednak mity jako struktury poznawcze legitymizujące strukturę poznawczą pełnią kluczową rolę w opisywanych procesach.

komunikacyjnej. Wiarygodność ta określa wartość „bycia podmiotem wiedzy” określonej jednostki społecznej w oczach postrzegającego. Władza tożsamościowa jest formą władzy społecznej (*social power*) (WS), stanowiącej konstrukt pozwalający członkom społeczności, wpływać na innych agensów¹²¹. WS zależna jest od dzielonych przez przedstawicieli grupy społeczno-wyobrażeniowych koncepcji, które wpływają na uczestników sytuacji komunikacyjnej uwikłanych w określone operacje z nią związane. Czyli władza tożsamościowa zależy od społecznych kategorii poznawczych, do których jednostka przyporządkowuje spostrzeganych agensów. Zarządzają one takimi konstruktami jak „co to znaczy być kobietą”, „mężczyzną” lub „starym” itp. Operacje tej władzy mogą mieć charakter czysto strukturalny, np. gdy jakaś grupa nie idzie na wybory, ponieważ jednostki do niej należące uważają, że „nie są osobami, które interesują się polityką”. Fricker (2007) uważa, że istnieje przynajmniej jeden rodzaj WS wymagający nie tylko praktycznej (*practical social co-ordination*), ale także wyobrażeniowej koordynacji społecznej (*imaginative social co-ordination*). Operacje takich rodzajów władzy zależą od tego, czy agenci sytuacji podzielają koncepcję społecznej tożsamości (*identity social*). W wyniku zaangażowania mechanizmu kategoryzacji, powstają także uprzedzenia, które Fricker nazywa tożsamościowymi (*identity prejudice*). Są to postawy przyjmowane w stosunku do przedstawicieli określonych grup społecznych. To właśnie uprzedzenia tożsamościowe są głównym źródłem niesprawiedliwości testimonialnej. Ze względu na fakt, że mechanizm kategoryzacji społecznej opisałam we wcześniejszych częściach¹²², nie uważam za konieczne ponownie omawianie jego roli w powstawaniu i utrzymywaniu postaw. Dyskryminacja wynikająca z uprzedzeń tożsamościowych, polega na przypisywaniu nadawcy niższej wiarygodności tylko ze względu na to, że jest on przedstawicielem określonej grupy. Autorka zaznacza, że każda niesprawiedliwość doświadczana przez jednostkę społeczną, polegająca na deprywacji z potrzeby „bycia podmiotem wiedzy”, ma charakter testimonialny. Fricker (2007) argumentuje, że opisywany rodzaj władzy jest ściśle związana z procesem wymiany wiedzy w ramach dyskursu społecznego. Podkreśla przy tym, że jego uczestnicy muszą spontanicznie szacować wiarygodność interlokutora. Robią to korzystając z heurystyk, stereotypów i jak uważam – mitów, gdyż to właśnie one legitymizują istniejącą strukturę społeczną (Fricker, 2007).

¹²¹ Agens rozumiany w ramach teorii socjologicznych jest elementem świata społecznego, który w określony sposób wpływa na sytuację społeczną/komunikacyjną.

¹²² Szczególnie w części 2.1. *Mechanizm kategoryzacji społecznej*.

Przykładem może być wspomniany wcześniej (w części 2.6. *Słodka reakcja*) mit króla kapłana, który powstał podczas rewolucji neolitycznej. Wraz z procesem osiedlania się ludzi, pojawiła się konieczność prowadzenia długich wojen o terytorium, co skutkowało koniecznością centralizacji władzy.

Władza tożsamościowa może być realizowana aktywnie albo pasywnie, jednak niezależnie od tego, zawsze jest wynikiem wyobraźniowej koordynacji społecznej. Przykładem tożsamościowej władzy aktywnej jest sytuacja, w której nadawca wykorzystuje swoją tożsamość np. płciową, do kontrolowania zachowania innego agensa. Natomiast pasywna realizacja władzy tożsamościowej polega na tym, że jednostka nie wyraża swojego zdania, ponieważ np. interlokutor jest innej płci, więc posiada przekonanie, że przy nim nie wypada. W celu lepszego wyjaśnienia dwóch wymienionych sposobów realizacji władzy, warto powołać się na przykład roli społecznej jaką, jest bycie nadzorcą parkingu. Władza związana z tym stanowiskiem może działać aktywnie, np. gdy osoba pełniąca określoną rolę wystawia mandat, za niepoprawne poruszanie się po parkingu, ale też pasywnie, gdy sama możliwość ukarania użytkowników parkingu, skłania ich do zachowania się w określony sposób. Czyli władza pozostaje w mocy nawet jeżeli nie jest w danym momencie realizowana. Tym samym, opisując pasywne działanie omawianego zjawiska, Fricker nie zgadza się z twierdzeniem Foucaulta, mówiącym, że władza istnieje tylko wtedy, gdy osadzona jest w działaniu. Dlatego mówi o władzy agenta (*agential power*) i władzy strukturalnej (*structurally power*). Pierwsza z wymienionych wykonywana (*exercised*) jest przez jednostkę, natomiast władza strukturalna realizowana jest bez żadnego zaangażowanego konkretnego podmiotu społecznego. Przykładem drugiego z wymienionych rodzajów, jest sytuacja, gdy pewna grupa ludzi pozostaje nieformalnie wykluczona społecznie, ponieważ z powodu pewnych złożonych czynników jej przedstawiciele mają tendencję do nie głosowania w wyborach. W takiej sytuacji żaden agens, czy działanie nie jest zaangażowane w wykluczanie ich z procesu demokratycznego. Jak zauważa Fricker — władza wpływająca na ich zachowanie jest tak bardzo rozproszona w systemie społecznym, że nie możemy przypisać jej podmiotu. Należy także pamiętać, że aby miała ona wpływ wymaga respektowania jej (przestrzegania zasad życia społecznego) przez innych uczestników sytuacji społecznej. Czyli wszystkie operacje władzy zależne są od kontekstu. Co więcej, władza ze swojej natury jest zjawiskiem strukturalnym, ponieważ zawsze zależy od społecznej koordynacji. Taki rodzaj władzy nazywamy diadycznym (*dyadic*), ponieważ uwzględnia podmiot władzy oraz jednostki jej

podległe. To zachowania tych drugich, pozwala na jej realizację. Da lepszego zrozumienia tego twierdzenia, posłużę się przykładem Wartenberga (1992), który opisuje omawiany mechanizm społeczny na przykładzie relacji nauczyciele — studenci/uczniowie. Utrzymanie systemu społecznego obowiązującego w kontekście wskazanej relacji, zależne jest m.in. od reguł bardziej lokalnych — panujących na wybranej uczelni, a narzędziami egzekfowania tego rodzaju władzy są np. oceny. Według Fricker (2007) utrzymanie takiej relacji wymaga przede wszystkim koordynacji akcji między odmiennymi grupami społecznymi (nauczyciele — studenci). Zarówno dla uczniów jak i dla nauczycieli, oceny muszą być w jakiś sposób ważne (*they take notice of grades*), gdyż inaczej nie wpływały by na zachowanie. Przykład ten pokazuje, że władza oparta jest na umowie społecznej, a właściwie, że układ społeczny jest konstytutywny dla tego rodzaju relacji. Co więcej, działania związane z władzą tożsamościową nie wymagają, aby jednostka je realizująca świadomie uznawała za prawdziwe stereotypy czy mity, które są w ich ramach wykorzystywane. Innymi słowy, podmiot WS nie musi podzielać przekonań czy wartości typowych dla poznawczego komponentu uprzedzeń, ponieważ kulturowe uprzedzenia systemowe często funkcjonują na poziomie niejawnych, automatycznych schematów.

6.3. niesprawiedliwość testymonialna

Jak wspomniałam we wstępie tego podrozdziału, niesprawiedliwość testymonialna wynika z nietrafnego szacunku wiarygodności nadawcy komunikatu. W swojej książce, Fricker (2007) wskazuje, że stereotypy¹²³ (komponent poznawczy) przetwarzane na drodze heurystyk leżą u podstaw wnioskowań, w ramach których jednostka dokonuje szacunku wiarygodności interlokutora, w codziennej wymianie testymonialnej. W wyniku takiego przetwarzania może dojść do dwóch rodzajów dysfunkcji w praktyce testymonialnej:

- a. Deficyt wiarygodności (*credibility deficit*): gdy wiarygodność nadawcy szacowana jest znacznie niżej, niż na to zasługuje,
- b. Nadmiar wiarygodności (*credibility excess*): gdy wiarygodność nadawcy szacowana jest znacznie wyżej, niż na to zasługuje.

Chociaż może się wydawać, że b. stanowi przywilej dla jednostki, zgadzam się z Fricker, że w niektórych sytuacjach ten rodzaj dysfunkcji testymonialnej może

¹²³ A także, jak sądzę mity.

stanować wadę (*disadvantage*). Za przykład służy jej sytuacja, w której dziecko dorastające w środowisku ciągle zawyżającym jego umiejętności (np. dziecko wysoko postawionego urzędnika państwowego), posiada fałszywy obraz siebie, co wpływa na nieumiejętność szacowania własnych możliwości poznawczych. Jednak przeważnie nadmiar wiarygodności niesie zysk dla nadawcy. Jak wskazałam w podrozdziale 2.4.. *Efekty oczekiwań*, zawyżanie możliwości poznawczych dziecka, może prowadzić do faktycznego rozwoju (na drodze efektu Rosenthala), co w przyszłości może przełożyć się na sukcesy zawodowe i prywatne. Z łatwością można wskazać także sytuacje, w których deficyt wiarygodności może być korzystny dla nadawcy. Fricker (2007) powołuje się na sytuację gdy członkowie Rady Cesarskiej nie decydują się na zamordowanie Klaudiusza, który miał zostać w przyszłości cesarzem Rzymu, ponieważ ten jękał się, co sprawiło, że przypisano mu niski poziom inteligencji. W rezultacie oszczędzono go, sądząc, że da się nim łatwo manipulować. Innym, bardziej współczesnym przykładem sytuacji, w której deficyt wiarygodności może być korzystny dla nadawcy, jest zdawanie egzaminu przez kobietę z przedmiotu społecznie uważanego za typowo męski. Zdarza się, że wykładowcy przypisują niższy poziom zdolności np. studentkom kierunków technicznych, w wyniku czego, ich oczekiwania względem egzaminowanych kobiet, są znacznie niższe, niż te które wykładowca żywi wobec męskiej części studentów. Oczywiście ta dysfunkcja testimonialna może być krzywdząca dla studentki zdającej egzamin. Zakładając, że podczas nauki do egzaminu, kobieta przyswoiła przynajmniej tyle samo materiału, co jej koledzy, zachowanie wykładowcy narusza jej naturalną potrzebę „bycia podmiotem wiedzy”.

Innym ważnym elementem koncepcji Fricker (2007) jest konieczność odróżnienia deficytu wiarygodności, którego źródłem są uprzedzenia, od zjawiska nazwanego niewinnym błędem (*innocent error*). Wynika on z posiadania fałszywych przekonań, dotyczących kwalifikacji, motywacji i doświadczeń interlokutora. W celu lepszego zrozumienia omawianego zjawiska, powołam się na następujący przykład Fricker: należy wyobrazić sobie sytuację komunikacyjną między etyczką i pracownikiem katedry medycznej. Przed spotkaniem pierwsza z wymienionych sprawdzała za pomocą wyszukiwarki Google swojego przyszłego interlokutora. W ramach krótkiego rozpoznania znalazła jego nazwisko na liście pracowników katedry medycznej pewnego uniwersytetu. Podczas rozmowy, gdy akademik powołuje się na dobrze znaną odbiorczynie literaturę oraz wystosowuje wobec niej prostą i trafną krytykę, ta przypisuje niższą wiarygodność jego argumentom, niż gdyby wiedziała, że

w rzeczywistości jej interlokutor jest z wykształcenia etykiem ze specjalizacją w naukach medycznych.

W tym miejscu warto zatrzymać się nad problemem różnic terminologicznych, jakie zachodzą pomiędzy przyjętym w niniejszym opracowaniu rozumieniem stereotypów a ujęciem zaproponowanym przez Fricker (2007). Autorka *Epistemic Injustice: Power and the Ethics of Knowing* definiuje stereotypy jako skojarzenia powszechnie obecne w systemach poznawczych ludzi. Są one zorganizowane wokół określonych grup społecznych oraz właściwości czy atrybutów przypisywanych członkom tych grup. Stereotyp w tym sensie funkcjonuje jako mentalny konstrukt, który – niezależnie od intencji – oddziałuje na przebieg interakcji społecznych, kształtując nie tylko sposób postrzegania innych, ale także strukturyzując epistemiczne relacje pomiędzy jednostkami. Tym, co szczególnie istotne w definicji Fricker, jest nacisk na zakorzenienie stereotypu w indywidualnym systemie poznawczym, przy jednoczesnym uznaniu jego społecznego wymiaru: stereotyp nie jest tylko jednostkowym przekonaniem, a bardziej formą poznawczego uogólnienia, którego treść i funkcja są ukształtowane kulturowo i społecznie. Można zatem powiedzieć, że autorka omawianej książki nazywa stereotypem zjawisko określane w niniejszej dysertacji jako narracja. W wyznaczanej przez nią perspektywie podkreślany jest wpływ, jaki stereotypy¹²⁴ – jako elementy poznawczo-afektywnych struktur – wywierają na epistemiczną pozycję jednostek, np. poprzez podważanie ich wiarygodności jako źródeł wiedzy (co skutkuje niesprawiedliwością testimonialną) lub ograniczanie zasobu dostępnych kategorii rozumienia własnych doświadczeń (co prowadzi do opisanie w następnym rozdziale niesprawiedliwości hermeneutycznej).

W niniejszej dysertacji przyjęto nieco odmienne, choć niekoniecznie sprzeczne, ujęcie stereotypów. Zgodnie z wyżej opisywanymi¹²⁵ definicjami, traktowane są one przeze mnie, jako kryteria przyporządkowywania obiektów – w tym agensów – do określonych kategorii społecznych. Innymi słowy, stereotyp pełni tu funkcję narzędzia poznawczego wykorzystywanego przez mechanizmy kategoryzacji społecznej. Ich główną wartością adaptacyjną jest uproszczenie przetwarzania informacji w sytuacjach społecznych: nowo spostrzegany agens zostaje automatycznie przypisany do określonej

¹²⁴ W tym kontekście używam terminu „stereotyp” w rozumieniu proponowanym przez Mirandę Fricker.

¹²⁵ Szczególnie w częściach 1.2.2. *Upředzenia, stereotypy, dyskryminacja*, 2.1. *Mechanizm kategoryzacji społecznej* oraz 3.3.2. *Stereotypy*.

kategorii. Właściwości przypisane agansom w tym procesie (tj. stereotypowe cechy), determinują heurystycznie przetwarzanie informacji, pozwalające przewidywać zachowanie lub intencje innych. Ujęcie to kładzie nacisk na operacyjny aspekt stereotypów – są one traktowane jako część dynamicznego procesu rozpoznawania i interpretowania świata społecznego. Mają one charakter funkcjonalny i są integralnym składnikiem codziennego funkcjonowania poznawczego, nawet jeśli w wielu przypadkach prowadzą do błędów atrybucyjnych, uprzedzeń lub niesprawiedliwości epistemicznych.

Zauważmy jednak, że różnica pomiędzy obiema definicjami ma charakter bardziej ontologiczny niż praktyczny. W podejściu Fricker stereotyp jest afektywnie nacechowaną, społecznie osadzoną narracją – czymś, co „nosimy w sobie” i co może oddziaływać niejawnie na nasze sądy. W niniejszym ujęciu stereotyp jest natomiast schematem poznawczym, mającym określoną funkcję adaptacyjną i operacyjną, która wyewoluowała, by efektywniej poruszać się w świecie społecznym. Innymi słowy, Fricker interesuje się bardziej epistemicznymi i etycznymi konsekwencjami posługiwania się stereotypami, podczas gdy w niniejszej pracy stereotyp traktowany jest przede wszystkim jako element struktury poznawczej, który można opisać i analizować przy pomocy narzędzi nauk poznawczych. Nie oznacza to jednak, że oba ujęcia są ze sobą niekompatybilne. Wręcz przeciwnie – można wskazać na kilka istotnych punktów wspólnych. Przede wszystkim, obie definicje podkreślają, że stereotypy są nierozzerwalnie związane z funkcjonowaniem systemów poznawczych i że są one ukierunkowane na określone grupy społeczne. Obie też uznają, że stereotypy – choć mogą być efektywne z perspektywy poznawczej – mogą prowadzić do niesprawiedliwych lub nietrafnych ocen. W obu podejściach dostrzega się także związek pomiędzy jednostkowym przetwarzaniem informacji a szerszymi strukturami społecznymi i kulturowymi. Można zatem powiedzieć, że gdy definicja przyjęta w niniejszej dysertacji koncentruje się bardziej na funkcji poznawczej i adaptacyjnej, ta przyjęta przez Fricker (2007) – na funkcji społecznej i etycznej. Oba jednak wskazują, że stereotyp jest pewną strukturą pośredniczącą pomiędzy poznaniem a społecznym działaniem, i jako taką, zasługującą na pogłębioną analizę w kontekście epistemicznej sprawiedliwości, psychologii społecznej oraz kognitywistyki. Tak samo, jak w koncepcji Fricker (2007) za przydzielane agansom, przyporządkowanym do określonej kategorii społecznej odpowiadających jej atrybutów odpowiedzialny jest mechanizm generalizacji. Innymi słowy niektóre stereotypy mogą mieć charakter

uniwersalny, a inne mogą być bardziej indywidualne. Drugie z wymienionych często są bardziej złożone np. aktywowane przez koniunkcję cech. Dodatkowo jak już wielokrotnie wspominałam uważam, że stereotypy stanowią poznawczy komponent uprzedzeń. Inaczej niż Fricker (2007), która definiuje uprzedzenia nie jako postawę, a sąd o charakterze pozytywnym lub negatywnym, który oporny jest na przekonania niezgodne z istniejącymi już strukturami poznawczymi. Jak wskazałam wcześniej taki opis zjawiska odpowiada omówionemu w części 2.5. *Teoria dysonansu poznawczego*. *poznawczego* efektowi potwierdzenia. Co ważne, Fricker (2007) zaznacza, że tendencja (komponent behawioralny) do nieuznawania argumentów niezgodnych z przekonaniami jednostki jest wynikiem pewnej inwestycji afektywnej.

Fricker (2007) zawęża pole swoich zainteresowań do systematycznej testimonialnej niesprawiedliwości, wynikającej z deficytu wiarygodności, będącego efektem prezentowanych przez odbiorcę uprzedzeń tożsamościowych. Systematyczne niesprawiedliwości są bardzo trwałe, ponieważ zależą od struktur poznawczych dotyczących koncepcji społecznej tożsamości. Stereotypy i mity będące komponentem poznawczym uprzedzeń systemowych leżą u podstaw dyskryminacji w każdej sferze życia jednostki: ekonomicznej, edukacyjnej, seksualnej, zawodowej i politycznej. Przyczynia się do tego fakt, że określone stereotypy mogą być systematycznie powiązane nie tylko z niesprawiedliwością tożsamościową. Takie zawężenie zakresu zainteresowań pozwala na wykluczenie uprzedzeń, powodujących deficyt wiarygodności, wywołujący niesprawiedliwość o charakterze lokalnym. Ten rodzaj niesprawiedliwości pozbawiony jest strukturalnego znaczenia społecznego. W celu lepszego wyjaśnienia omawianego zjawiska posłużę się przykładem wyobrazonego panelu redakcyjnego określonego czasopisma naukowego, który uprzedzony jest do określonej metody naukowej (MN). Ze względu na to, że członkowie komisji posiadają stereotypy, związane z przekonaniami, że MN jest wadliwym narzędziem (komponent poznawczy), nadają niskie noty wszystkim artykułom korzystającym ze wskazanego narzędzia (komponent behawioralny). W ramach praktyki testimonialnej odbiorca, czyli w tym przypadku Komisja Naukowa, nadaje niską wiarygodność analizom lub wynikom wykonanych/uzyskanych z udziałem konkretnego narzędzia. Oczywiście może to być krzywdzące dla autorów, gdyż może zaważyć na ich karierze, a nawet, w pewnym sensie hamuje rozwój naukowy, jednak wpływ tej niesprawiedliwości jest bardziej lokalny.

6.4. Upředzenia tożsamościowe a niesprawiedliwość testymonialna względem ZINL

Odpowiedź na pytanie *czy upředzenia tożsamościowe mogą być źródłem niesprawiedliwości testymonialnej u ZINL?* należy zacząć od postawienia dwóch pytań szczegółowych:

- a. Czy ZINL zdolne są do odczuwania niesprawiedliwości testymonialnej, która może się pojawić tylko u jednostki, której odmawia się statusu podmiotu wiedzy?
- b. Czy ZINL posiadają zdolności umożliwiające im przekazywanie informacji o sobie i świecie wewnątrz- lub międzygatunkowo?

Odpowiedź na a. warto zacząć od przełomowych prac dotyczących poczucia sprawiedliwości u ZINL. Zagadnienie to było wielokrotnie badane i obecnie powszechnie przyjmuje się, że niektóre gatunki przejawiają podstawowe formy wrażliwości na sprawiedliwość (Bekoff, Pierce, 2018). Przykładem może być klasyczna już książka *Chimpanzee Politics* (2000) autorstwa Fransa de Waala, w której autor pokazuje, że spory szympansov nie ograniczają się wyłącznie do zachowań opartych na agresji czy dominacji, ale obejmują także pojednanie i zacieśnianie więzi społecznych (np. przez całowanie). Dodatkowo badania prowadzone m.in. na kapucynkach pokazują, że te złoścą się, gdy nagrody rozdzielane są nieuczciwie lub gdy jedna z uczestniczek eksperymentu dokonuje egoistycznych wyborów, w wyniku czego, druga małpka otrzymuje mniej lubianą nagrodę (Brosnan, De Waal, 2003). Uważam, że skoro ZINL wykazują wrażliwość na równość i nierówność traktowania, można postawić pytanie, czy nie są także zdolne – przynajmniej potencjalnie – do wykształcenia bardziej złożonego poczucia sprawiedliwości, wrażliwego na to, czy dana jednostka ma potrzebę bycia postrzeganą jako kompetentna, skuteczna czy wiarygodna.

Kolejnym istotnym w omawianym obszarze pytaniem, jest: czy ZINL są podmiotami wiedzy? Klasyczna koncepcja wskazanego rodzaju przekonania, jest jedną z najstarszych perspektyw filozoficznych, znanych też jako teoria uzasadnionego prawdziwego przekonania (*justified true belief*). Odnosi się ona do wiedzy wyrażonej w sądach, czyli do tzw. „wiedzy, że” (Kluziński, 2015). Pierwszą pracą opisującą omawianą koncepcję, są platońskie dialogi — *Teajtet*, w których autor rozważa naturę wiedzy poprzez przedstawienie rozmowy Sokratesa z młodym Teaitetosem (Platon,

2005). W dialogu rozważane są różne perspektywy np. wiedzy jako postrzegania, czyli poglądu filozoficznego, w ramach którego, omawiany rodzaj przekonania jest związany ze zdolnością postrzegania rzeczywistości za pomocą zmysłów. Koncepcja ta koncentruje się na pozyskiwaniu informacji o rzeczywistości poprzez percepcję zmysłową oraz zaznacza, że wiedza polega na adekwatnym postrzeganiu rzeczywistości. Ostatecznie jednak uczestnicy dialogu uznają, że wiedza jest ściśle związana z prawdą, czyli, aby przekonanie mogło zostać uznane za wiedzę, musi być prawdziwe (gdzie prawdziwość rozumiana jest jako zgodność z rzeczywistością lub faktami), ponieważ ta musi być adekwatna do rzeczywistości. Sokrates podkreśla także, że samo posiadanie prawdziwej informacji nie jest wystarczające, gdyż przekonanie o prawdziwości twierdzenia jest subiektywnym aspektem wiedzy, dlatego podmiot, któremu ją przypisujemy, musi wierzyć, że określone twierdzenie jest prawdziwe. Co więcej, duża część dialogu skupia się na potrzebie uzasadnienia przekonań. Sokrates zadaje pytania, które mają na celu sprawdzenie, czy przekonania Teaitetosa są poparte odpowiednimi argumentami, co oznacza, że przekonanie, które jest wiedzą, powinno być oparte na solidnych dowodach lub racjonalnych przesłankach je wspierających. Jest to konieczne, aby odróżnić wiedzę od zgadywania czy przypadkowego trafu. Czyli w ramach teorii uzasadnionego prawdziwego przekonania uważa się, że przypisanie komuś wiedzy, może nastąpić wtedy i tylko wtedy, gdy spełnione zostaną trzy warunki:

- a. Podmiot jest przekonany, że p,
- b. Prawdą jest, że p,
- c. Posiada uzasadnienie.

Rewolucja w teorii wiedzy miała miejsce w 1963 roku, gdy Edmund Gettier opublikował krótki, lecz bardzo wpływowy artykuł pt. *Czy prawdziwe uzasadnione przekonanie jest wiedzą?* Przedstawiając dwa przykłady, argumentował, że wyżej wspomniane warunki nie są wystarczające do uznania przekonań za wiedzę. Jednym z klasycznych gettierowskich przypadków jest historia Smitha, który posiada dwóch znajomych — Jonesa i Browna. Na podstawie obserwacji, że Jones jeździ Fordem, Smith nabywa przekonanie, że ten jest właścicielem określonego samochodu. Konstruuje więc prawdopodobne twierdzenie, że „Jones posiada Forda, lub Brown jest w Barcelonie”, nie posiadając przy tym informacji o miejscu pobytu Browna. Smith nie wie jednak, że ford nie jest własnością Jonesa, gdyż ten niedawno sprzedał swój samochód, ale jeszcze nie przekazał go nowemu właścicielowi, oraz że Brown

rzeczywiście przebywa w Barcelonie. To sprawia, że przytoczone zdanie jest prawdziwe, ale z innych powodów niż myślał Smith, w wyniku czego, nie możemy uznać przytoczonego sądu za wiedzę. Przypadki Gettier'a doprowadziły do rozwoju teorii wiedzy oraz głosów mówiących, że konieczne jest rozwinięcie klasycznej koncepcji o dodatkowy warunek. Jego funkcją byłoby wyeliminowanie przypadków Gettier'a, jednak w praktyce podejście to także rodzi problemy. Istnieje duża trudność w ustaleniu, jak dokładnie powinien wyglądać dodatkowy warunek, który skutecznie eliminuje wszystkie przypadki Gettier'a, bez wprowadzania nowych problemów. Obejmują one niejasność definicji, arbitralność nowych warunków oraz trudności w praktycznym zastosowaniu.

Jedną z propozycji czwartego warunku jest teoria bezbłędnego uzasadnienia (*No False Lemmas*). Jej zwolennikiem jest Laurence Bonjour, który w swojej książce z 1985 roku *The Structure of Empirical Knowledge* argumentuje, że konieczne jest uwzględnienie braku fałszywych przekonań jako warunku w definicji wiedzy. Problem z tą propozycją polega na trudnościach w identyfikacji fałszywych przesłanek oraz określeniu metod ich weryfikacji. Inną propozycją jest reliabilizm, starający się uniknąć problemów, wynikających z przypadków Gettier'a, koncentrując się na jakości źródła wiedzy. Zwolennikiem tej koncepcji jest Alvin Goldman, który zaproponował (1979) pogląd, że wiedza jest prawdziwym przekonaniem powstałym na drodze niezawodnego procesu poznawczego, który generuje prawdziwe wnioski w sposób powtarzalny. Należy jednak pamiętać, że nawet niezawodne procesy mogą czasem prowadzić do prawdziwych przekonań w sposób przypadkowy, co nie spełnia intuicyjnych kryteriów wiedzy. Kolejną propozycją jest Teoria Bezpieczeństwa (*safety*), w ramach której przekonanie jest wiedzą wtedy i tylko wtedy, gdy w podobnych okolicznościach nadal byłoby prawdziwe. Zwolennikiem tej propozycji jest Timothy Williamson, który w swojej książce *Knowledge and Its Limits* (2000), zaproponował teorię bezpieczeństwa jako warunek konieczny do posiadania wiedzy. Według Williamsona przekonanie jest bezpieczne, jeśli w większości bliskich światów możliwych, w których agent formułuje to samo twierdzenie, jest prawdziwe. Propozycja, ta mierzy się z trudnościami polegającymi na określeniu, jak bliskie powinny być te światy i jak ocenić, czy przekonanie jest prawdziwe w nich wszystkich. Dodatkowo koncepcja ta prowadzi do skomplikowanych analiz modalnych, które są trudne do zastosowania w praktyce.

W związku z powyższym, dodanie czwartego warunku może prowadzić do nadmiernej komplikacji teorii wiedzy, co jest sprzeczne z pierwotną intencją stworzenia prostego, przejrzystego i praktycznego modelu. Brak wymienionych cech prowadzi do utraty użyteczności wyjaśnień wiedzy. Dlatego wielu filozofów odchodzi od klasycznych koncepcji, skłaniając się ku teoriom takim jak kontekstualizm, czy epistemologia cnót (Woleński, 2003). Centralnym założeniem pierwszego z wymienionych jest twierdzenie, że pojęcie „wiedzy” jest elastyczne, a warunki, które muszą zostać spełnione, aby można było uznać przekonanie za wiedzę — są wrażliwe na kontekst. Innymi słowy omawiana teoria sugeruje, że to, czy dane przekonanie kwalifikuje się jako wiedza, zależy od czynników kontekstowych. W codziennych sytuacjach standardy mogą być stosunkowo niskie np. jeśli ktoś twierdzi, że wie, gdzie zaparkował swój samochód i rzeczywiście znajduje go w wyznaczonym miejscu, uważamy to za wiedzę. Jednak w bardziej wymagających kontekstach, takich jak np. przewód sądowy, standardy zakwalifikowania przekonań jako wiedzy są znacznie bardziej rygorystyczne np. wymagają dodatkowych przesłanek przemawiających za tym, że świadek faktycznie coś wie. Kontekstualizm jest silnie związany z problem sceptycyzmu, a jego zwolennicy twierdzą, że w pewnym sensie sceptycyzm wynika z podniesienia standardów epistemicznych do niezwykle wysokiego poziomu, co w codziennym życiu nie jest konieczne. W ramach omawianego poglądu, w większości sytuacji życiowych stosujemy mniej surowe standardy, które pozwalają nam na praktyczne uznanie wielu przekonań za wiedzę, mimo że nie spełniają one najwyższych możliwych standardów. Jednym z głównych zwolenników kontekstualizmu jest Keith DeRose, który w swojej książce *The case for contextualism: Knowledge, skepticism, and context, vol. 1* (2011) argumentuje, że to samo przekonanie może być uznane za wiedzę w jednym kontekście, natomiast w innym np. bardziej wymagającym już nie. Pogląd ten podkreśla również rolę praktycznych konsekwencji w ustalaniu standardów wiedzy. Na przykład, jeśli od tego, czy ktoś wie, zależą ważne decyzje, standardy epistemiczne mogą być wyższe, ponieważ konsekwencje błędnego uznania czegoś za wiedzę są bardziej znaczące. Koncepcja ta spotkała się z szeroką krytyką ze strony invariantystów takich jak wspomniany wcześniej Timothy Williamson (2002), który twierdzi, że wiedza nie powinna być zależna od zmieniających się kontekstów. Inwariantyści uważają, że skuteczniejszym stanowiskiem jest podejście polegające na przyjęciu, że istnieje jeden uniwersalny standard wiedzy. Inne kontrargumenty wycelowane w tę koncepcję związane są z problem ścisłości definicji. Krytycy

twierdzą, że kontekstualizm nie dostarcza precyzyjnych kryteriów, które definiują, jakie standardy wiedzy obowiązują w danym kontekście. To prowadzi do niejasności i subiektywności w ocenie tego, co liczy się jako wiedza. Richard Feldman (1999) argumentuje, że kontekstualizm może prowadzić do problemów z jasnością i precyzją, ponieważ trudno jest ustalić, jakie konkretne standardy mają zastosowanie w różnych sytuacjach. Kolejnym zarzutem jest problem komunikacji i rozumienia. Kontekstualizm sugeruje, że różne osoby mogą mieć różne standardy wiedzy w różnych kontekstach, co może prowadzić do nieporozumień i trudności w komunikacji. Krytycy twierdzą, że jeśli standardy wiedzy mogą się zmieniać w zależności od kontekstu, to może to prowadzić do sytuacji, w której dwie osoby nie mogą się zgodzić, czy ktoś posiada wiedzę na podstawie tych samych dowodów. Ponadto, kontekstualizm spotyka się z zarzutem, że nie rozwiązuje w pełni problemu sceptycyzmu. Choć kontekstualiści argumentują, że zmiana standardów epistemicznych może pomóc radzić sobie z problemem sceptycyzmu, krytycy twierdzą, że Ci jedynie przesuwają problem, ale nie rozwiązują go. Sceptycy mogą nadal podważać wiedzę, podnosząc standardy do ekstremalnie wysokiego poziomu w dowolnym kontekście.

Inną alternatywą dla klasycznych koncepcji wiedzy jest epistemologia cnót, która koncentruje się na cechach charakteru i intelektualnych dyspozycjach poznającego podmiotu. Zamiast skupiać się wyłącznie na właściwościach poszczególnych przekonań, omawiane podejście podkreśla znaczenie tzw. cnót epistemicznych, czyli norm, reguł i własności poznawczych (Kaluziński, 2015). Główna idea omawianej koncepcji głosi, że wiedza nie jest jedynie produktem szczęśliwego trafu czy zewnętrznych okoliczności, ale wynikiem aktywnego, dążenia do prawdy przez podmiot poznający. Cnoty epistemiczne są postrzegane jako trwałe cechy indywidualne, które pomagają jednostce osiągać prawdziwe przekonania w sposób powtarzalny. Jednym z kluczowych zwolenników epistemologii cnót jest Ernest Sosa, który w swojej książce *Knowledge in Perspective: Selected Essays in Epistemology* (1991) wprowadził rozróżnienie między wiedzą „animalistyczną” (*animal knowledge*) a wiedzą „refleksyjną” (*reflective knowledge*). Pierwsza z wymienionych jest podstawową formą wiedzy, natomiast wiedza refleksyjna to bardziej złożony konstrukt, który wynika z autorefleksji i świadomego zastosowania cnót epistemicznych, co obejmuje zrozumienie i metakognitywną kontrolę procesów poznawczych. Chociaż pogląd ten wskazuje, że zwierzęta pozaludzkie posiadają pewien rodzaj wiedzy, podkreśla wyjątkowość człowieka, głosząc, że tylko człowiek posiada wiedzę refleksyjną.

Ciekawe stanowisko znajdziemy w pracach innej ważnej przedstawicielki opisywanego nurtu — Lindy Zagzebski, która w swojej pracy *Virtues of the mind: An inquiry into the nature of virtue and the ethical foundations of knowledge* (1996) argumentuje, że cnoty epistemiczne są analogiczne do cnót moralnych. Według niej zarówno cnoty moralne, jak i epistemiczne są dyspozycjami, które prowadzą do dobrych rezultatów – w przypadku drugich z wymienionych, do prawdziwych przekonań. Zagzebski podkreśla, że cnoty te nie tylko pomagają w osiąganiu wiedzy, ale także w rozwoju osobistym i intelektualnym.

Chociaż zwolennicy omawianej koncepcji twierdzą, że cnoty epistemiczne są związane ze zdolnością do refleksji oraz moralnym rozumowaniem, czyli elementami uznawanymi za specyficznie ludzkie, dostrzegamy u innych zwierząt np. u sokoła doskonały wzrok (O'Rourke, Hall, Pitlik, Fernández-Juricic, 2010), czy echolokację (Griffin, 1958) u nietoperzy, a te mogą być traktowane jako formy percepcyjnej niezawodności, co jest podstawowym elementem cnót epistemicznych. Dodatkowo np. pszczoły potrafią komunikować się nawzajem oraz nawigować na duże odległości, wykazując zdolności, które można uznać za formy epistemicznej cnoty związanej z orientacją w przestrzeni i komunikacją (Seeley, 1995). Co więcej, niektóre zwierzęta, jak np. małpy (Whiten, Horner, de Waal, 2005; Biro, Inoue-Nakamura, Tonooka, Yamakoshi, Sousa, Matsuzawa, 2003), delfiny (Herman, 2002) czy ptaki (Zentall, 2006), potrafią uczyć się poprzez obserwację i naśladownictwo, co może być uznane za zdolność poznawczą, która przybliża je do prawdziwych przekonań. Natomiast pytanie „czy zwierzęta pozaludzkie zdolne są do refleksji?” stanowi przedmiot intensywnych debat wśród naukowców i filozofów. Powszechnie refleksja i zdolność do introspekcji uważane są za cechy uniktowe dla człowieka, jednakże, badania nad kognitywnymi zdolnościami zwierząt dostarczają coraz więcej argumentów wskazujących, że pewne formy refleksji mogą występować nie tylko u ludzi (Robison-Greene, 2021). Istnieją sugestie, że samorozpoznanie, planowanie przyszłości, empatia, czy rozwiązywanie problemów są zdolnościami wymagającymi refleksji (Gallup, 1998, Povinelli, Vonk, 2003, Suddendorf, Corballis, 2007, de Waal, 2008). Griffin w swojej książce *Animal Minds: Beyond Cognition to Consciousness* (2001) argumentuje, że zwierzęta wykazują zdolności kognitywne, które mogą być uznane za refleksję, natomiast Kristin Andrews w *The Animal Mind: An Introduction to the Philosophy of Animal Cognition* (2020) omawia przesłanki świadczące za stanowiskiem, w ramach, którego ZINL mogą posiadać pewne formy refleksji i świadomego myślenia, choć te mogą różnić się od

zdolności występujących u człowieka. Wyżej wymienione argumenty sugerują, że stanowczość w odmawianiu zwierzętom pozaludzkim zdolności do refleksji może mieć źródło w rozumowaniu motywowanym, nie zaś racjonalnym.

Teoretycy praw zwierząt pozaludzkich od lat '70 XX wskazują, że ZINL są istotami czującymi oraz podobnymi do człowieka na wiele moralnie i etycznie istotnych sposobów. Dlatego współczesne ruchy sprawiedliwości społecznej działające na rzecz zmiany statusu moralnego w systemach politycznych i kulturowych, walczą o to, aby uwzględnić je w ludzkich systemach społecznych jako grupy polityczne, przynależące do dzielonych z człowiekiem wspólnot i terytoriów (Meijer, 2021). Meijer (2021) wskazuje także, że biorąc pod uwagę podmiotowość ZINL, ograniczenie się do rozważań na temat ich dobrostanu nie są wystarczające pod względem normatywnym i epistemicznym. To wiąże się także z odejściem od układu, w ramach którego to „my” — ludzie, mówimy w ich imieniu. Takie stanowisko wskazuje, że systemy polityczne i kulturowe powinny być oparte na komunikacji międzygatunkowej, a nie projekcji. Oczywiście zaznacza, że obecnie nie ma dowodu, na istnienie zdolności językowych u jakiegokolwiek gatunku innego niż człowiek. Jednocześnie podnosi, że odmawianie zwierzętom pozaludzkim jakichkolwiek kompetencji komunikacyjnych jest nie do utrzymania, oraz że źródłem takiego podejścia jest antropocentryczna i antroponegacionistyczna postawa, wspierana przez mity i stereotypy mówiące, że ludzie są najważniejszymi istotami we wszechświecie. W rezultacie, rzeczywistość interpretowana jest przez pryzmat ludzkiego sposobu postrzegania świata, a człowiek staje się punktem odniesienia stanowiącym ramy oceny innych zwierząt (Meijer, 2021). Aktualnie nazywanie „językiem” komunikacji ZINL, budzi kontrowersje, powodowane przez długą historię antropocentrycznego podejścia do jego rozumienia (Calarco, 2008). Dla przykładu w Teorii Sprawiedliwości kompetencje językowe są warunkiem koniecznym do bycia podmiotem politycznym (Rawls, 2009). W związku z tym, zgadzam się z autorką *Zwierzęta mówią*, że wymagana jest zmiana polegająca na wypracowaniu nowego rozumienia języka oraz uważam, że wspomniana książka stanowi dobry początek i solidny fundament dla bardziej obiektywnych teorii i definicji wskazanych zdolności. Pogląd, w ramach którego tylko ludzie są podmiotami politycznymi, ma długą historię, której fragmenty opisałam w części 3.1. *Historia badania umysłów zwierząt pozaludzkich*. We wspomnianej wtedy filozofii Arystotelesa, znajdziemy definicję człowieka jako zwierzęcia politycznego, jedynego posiadającego zdolności językowe i *logos*, czyli rozum pozwalający na odróżnienie dobra od zła

(Aristoteles, *Polityka*, 2003). Przypisanie własnemu gatunkowi takiego atrybutu stanowi linię demarkacyjną między ludźmi a innymi zwierzętami oraz jest punktem wyjścia dla rozróżnienia między ludzkim językiem a użyciem głosu przez ZINL. Traktowanie ludzkiego języka jako wzorzec i jedyny prawdziwy przykład języka, widoczne jest w badaniach dotyczących zdolności językowych u ZINL, które sprawdzano pod względem tego, na ile te są w stanie przyswoić ludzki język, lub go zrozumieć. Wyniki tych eksperymentów stanowią argumenty dla stanowiska umacniającego stereotypy i mity na temat braku badanych zdolności u ZINL. Przypisanie kompetencji językowych tylko człowiekowi związane jest z mitem, że ludzi i inne zwierzęta dzieli jasna granica, jednak przekonania te nie są ugruntowane w biologii, wręcz przeciwnie ((Darwin, 2009, Aaltola, 2013). Wiele stanowczych tez dotyczących charakteru różnicy w zdolnościach poznawczych ludzi i innych zwierząt, nie przetrwały próby czasu. Ostatnie badania z zakresu zdolności językowych u ZINL sugerują, że różnica ta ma charakter stopnia nie rodzaju (Slobodchikoff, Paseka, Verdolin 2009; Pika, Bugnyar, 2011; Berg, Delgado, Cortopassi, Beissinger, Bradbury, 2012; King, Janik 2013; Berthet, Surbeck, Townsend, 2025). Perspektywa przyjmująca, że różnica ta ma charakter rodzaju, zakorzeniona jest w praktykach kulturowych, będących częścią narracji, że życie ludzkie jest ważniejsze od życia innych zwierząt (Calarco, 2008).

Zgadzam się ze stanowiskiem, że zwierzęta pozaludzkie powinny być uwzględnione w istniejących systemach politycznych i kulturowych (Meijer, 2021). Oczywiście różne gatunki zwierząt mają różne potrzeby oraz inaczej postrzegają świat, co wpływa na formę i charakter nawiązywanych relacji i komunikacji wewnątrz- i międzygatunkowej. Zatem, uwzględnienie w systemach społecznych wszystkich ZINL nie może odbywać się poprzez przydzielenie ich do jednej kategorii społecznej, wymaga raczej traktowania różnych gatunków jako odrębnych grup będących częścią większych wspólnot społecznych. Przyjmuję, że zwierzęta pozaludzkie posiadają *Umwelty*, czyli są indywidualnymi istotami z własnymi perspektywami na otaczający je świat oraz potrafią komunikować własne potrzeby. Co więcej, nawiązują relacje z przedstawicielami zarówno własnego, jak i innych gatunków, czego potwierdzeniem są np. badania pokazujące, że kontakt wzrokowy opiekuna i jego towarzysza powoduje produkcję oksytocyny u obydwu uczestników sytuacji społecznej (Nagasawa, Mitsui, En, Ohtani, Ohta, Sakuma, Kikusu, 2015). Dlatego uważam, że zwierzęta pozaludzkie posiadają „głosy”, które powinny być wysłuchane, a także, ignorowanie ich może być

źródłem niesprawiedliwości testimonialnej. Mając jednak na względzie, że ludzie, którzy przypisują zdolności językowe ZINL, oskarżani są o antropomorfizm, pragnę zaznaczyć, że nawet jeżeli przyjmiemy założenie, że ZINL nie posiadają zdolności językowych, opartych na mechanizmach zbliżonych do człowieka, nie można odmówić im posiadania pewnych kompetencji komunikacyjnych, a co za tym idzie zdolności do wymiany informacji o swoich potrzebach i istotnych elementach świata je otaczającego (Donaldson, Kymlicka, 2013). Współczesne badania z zakresu etologii i biologii sugerują, że wiele gatunków zwierząt pozaludzkich wykształciło na drodze ewolucji systemy komunikacyjne, złożone bardziej niż myśłano do tej pory. Dla przykładu w komunikacji delfinów butlonosych (King, Janik 2013), papug (Berg, Delgado, Cortopassi, Beissinger, Bradbury, 2012) i słoni afrykańskich (Pardo, Fristrup, Lolchuragi, Poole, Granli, Moss, Wittemyer, 2024), odnajdujemy funkcję referencyjną — naukowcy zauważyli, że wymienione zwierzęta nadają sobie imiona. Natomiast Simone Pika, Thomas Bugnyar (2011) donoszą, że kruki produkują gesty referencyjne. System komunikacyjny słoni afrykańskich jest na tyle złożony, że te posiadają sygnały nie tylko na zagrożenia takie jak pszczoły, ale też korzystają z różnych sygnałów określających różne grupy ludzi. Inne badania pokazują, że nieświszczuki opisują dokładnie ludzi napotkanych na swej drodze innym osobnikom (Slobodchikoff, Paseka, Verdolin, 2009). Ich opisy zawierają szczegółowe informacje o prędkości, z jaką idą, kolorze ubrań, włosów, a nawet przedmiotach, które ze sobą niosą. Co więcej, wielu naukowców raportuje rekurencyjny charakter śpiewu różnych gatunków ptaków np. sikorki bogatki (Gentner, Fenn, Margoliash, Nusbaum, 2006). Warto przypomnieć także, że niektóre gatunki wyewoluowały do życia z człowiekiem jak np. psy czy koty, co sprawiło, że wykształciły sprawne kanały porozumiewania się z ludźmi. Stanowisko to wspierane jest także przez badania z zakresu zdolności do atrybucji innych stanów mentalnych takich jak emocje (Hare, Woods, 2013). Istnieje szeroki zakres badań dotyczących mimiki zwierząt, nazywany AnimalFACS. Kurs w tym zakresie, polegający na oglądaniu wyrazów mimicznych wybranych zwierząt (np. HumanFACS, DogFACS, MonkyFACS), wraz z opisami mięśni użytych do określonych ruchów, dostępny jest online, a jego ukończenie poświadczane jest certyfikatem.

Biorąc pod uwagę wspomniane wyżej doniesienia uważam, że nawet jeżeli ZINL nie posługują się językiem, zdolne są do generowania komunikatów dotyczących siebie i świata, a ignorowanie tych komunikatów w wyniku uprzedzeń tożsamościowych, będzie powodować niesprawiedliwość testimonialną. Czyli

niesprawiedliwość wobec zwierząt pozaludzkich polegającą na marginalizacji ich głosu, ignorowaniu ich komunikatów i doświadczeń m.in. w kontekście badań naukowych i ochrony środowiska. Historia traktowania ZINL jako podmioty wiedzy pokazuje, że te mogą skutecznie pomagać człowiekowi w wielu sferach jak obronność, kryminalistyka czy alpinistyka. Dla przykładu rzymianie preferowali jako zwierzęta domowe/obronne gęsi niż psy. Zwyczaj ten ma korzenie w 390 r. p.n.e. gdy na Rym napadli Galowie (Littlewood, 2006). Wykończeni mieszkańcy zabarykowali się w Kapitolu, gdzie przez kilka dni bronili swojej pozycji. Pewnej nocy napastnicy wspięli się na górę od strony najbardziej stromego zbocza, czego nikt się nie spodziewał. Śpiący rzymianie nie zorientowaliby się na czas, gdyby nie gęsi, które obudziły ich swoim gęganiem i wskazały stronę ataku. W wyniku czego dowódca *Marcus Manlius Capitolinus* pokierował swoje oddziały w odpowiednie miejsce, które odparły atak wroga. Niestety obrońcy, którzy zasnęli podczas pełnionej warty, obwinili psy, które zasnęły razem z nimi o brak czujności, w wyniku czego bardzo brutalnie je ukarano — jednego z nich powieszono. Co ciekawe zauważmy, że powieszenie psa jest oznaką silnej antropomorfizacji, gdyż jest to kara stosowana wobec ludzi. W wyniku opisanych wydarzeń wiele kultur traktuje gęsi jak zwierzęta obronne np. w Chinach, gdzie w niektórych prowincjach (np. Xinjiang), zwierzęta te rutynowo wspierają pracę policji, szczególnie nocą (Tadeusiewicz, 2020).

Innymi zwierzętami traktowanymi jako podmioty wiedzy, są psy tropiące wykorzystywane w policji. Zmysł węchu jest dla nich dominujący, co pozwala na sprawne szukanie narkotyków, zwłok zamordowanych lub zaginionych ludzi. Rasy psów tropiących wykorzystywane do szukania nielegalnych substancji, szkolone są w ramach tzw. *noseworkingu*, polegającego na kojarzeniu poszczególnych zapachów z nagrodą. Początkowo trening odbywa się z użyciem zwykłych przysmaków, następnie stopniowo wprowadza się intensywnie pachnące zamienniki syntetyczne, które z czasem pies zaczyna kojarzyć z nagrodą. Takie szkolenia stanowią okazję do świetnej zabawy dla czworonogów oraz odbywają się z udziałem ras, które posiadają silną naturalną potrzebę tropienia, jak np. owczarki niemieckie, lub pointery. Gdy przedstawiciele wymienionych ras nie zaspokajają wspomnianej potrzeby, niespożytkowana energia może prowadzić do niepożądanych zachowań, jak niszczenie przedmiotów domowych lub ogrodowych. Co ważne, aktualnie w artykułach dotyczących szkoleń psów policyjnych, podkreśla się wagę wzmocnienia pozytywnego, co oznacza, że odchodzi się od stosowania kar, szczególnie tych cielesnych,

a skojarzenia tworzone są na drodze nagród. Badania pokazują, że takie podejście sprawia, że zwierzęta chętniej podejmują współpracę i osiągają lepsze wyniki w szkoleniu (Sport-dog.pl, 2022). Szkolenie psów tropiących zapach zwłok jest podobne, gdyż zwierzęta uczą się kojarzyć odpowiednie zapachy z pozytywnymi bodźcami, takimi jak, smaczki czy zabawki. W lutym 2024 roku w Polsce żyło 17 psów szkolonych do tak specyficznej pomocy w praktyce kryminalistycznej¹²⁶, jednym z nich jest Suza, czyli pies rasy pointer, której ulubioną nagrodą jest zabawa piłeczką (pap.pl, 2024).

Ciekawym przykładem traktowania ZINL jako podmioty wiedzy jest wykorzystanie psów do szukania ludzi zaginionych w górach, np. przez uwięzienie pod śniegiem, szczególnie w wyniku lawiny, czyli tzw. psy lawinowe. Rasami wykorzystywanymi w tym celu są owczarki niemieckie, belgijskie, labradory i sznauclery. Takie psy powinny cechować się niską masą ciała, mieć pogodne przysposobienie i dużo chęci do zabawy. Selekcji osobników dokonuje się już pośród szczeniaków, sprawdzając ich ciekawość, lękliwość, ufność, predyspozycje do zachowań agresywnych oraz zdolność do samokontroli. W pierwszym etapie hodowca przynosi pieska do nowego dla niego pomieszczenia, gdzie w rogu stoi eksperymentator, czyli jego potencjalny przyszły opiekun. Piesek stawiany jest na środku pomieszczenia, a hodowca wychodzi bez słowa. Zainteresowanie nowym terenem i eksploracja nowego miejsca zdradza predyspozycje do przyszłego zawodu. W następnym etapie opiekun podchodzi coraz bliżej, próbując zainicjować interakcję, jeżeli pies ochoczo reaguje na przywoływania, chętnie przyjmuje smaczki oraz wyraża chęć do zabawy, eksperymentator przechodzi do etapu trzeciego, w którym wykazuje zachowanie agresywne, polegające na przewróceniu szczeniaka na plecy oraz przetrzymanie go siłą przez dłuższy czas w takiej pozycji. Takie działanie ma na celu sprawdzenie reakcji agresywnych, jeżeli pies będzie się wyrywał, gryzł i drapał, nie nadaje się do pracy w górach, gdyż te wykorzystywane są w sytuacjach, w których osobnik nie może bać się obcych ludzi, ani ich atakować. Dlatego tylko te psy, które spokojnie zniosą niekomfortową sytuację, dopuszczone będą do dalszych części szkolenia. Poziom ciekawości sprawdzany jest przez pokazanie szczeniakowi głośnej

¹²⁶ Dane te pochodzą z artykułu zawierającego rozmowę Małgorzaty Barskiej z komendy Miejskiej Policji w Zielonej Górze z dziennikarką. Link: <https://www.pap.pl/aktualnosci/pierwszy-wyzeł-w-polskiej-policji-bedzie-poszukiwal-ludzkich-zwlok-nasze-wideo>.

i kolorowej zabawki oraz obserwowanie jego zachowań eksploracyjnych. Jeżeli pies szybko przezwycięża lęk, wykazując wyraźne zainteresowanie zabawką, dopuszczony jest do dalszego etapu. W jego ramach człowiek wykonuje zachowanie polegające na zrzuceniu metalowego przedmiotu na podłogę, co wywołuje huk. Jeżeli pies nie radzi sobie z nagłym stresem i potrzebuje dużo czasu, aby wrócić do równowagi emocjonalnej, odrzucany jest jako potencjalny kandydat do pracy w TOPR i GOPR (Szczucińska, 2018). Następnie opiekun zabiera psa do domu, gdzie śpi z nim przez 3 doby na posłaniu swojego towarzysza, co jak się uważa, pozwala na wytworzenie relacji zbliżonej do dziecko — matka. Gdy pies jest starszy, przechodzi przynajmniej półtoraroczny trening w terenie. W tym celu ludzie zakopują tzw. pozorantów ok. 2-3 m pod śniegiem, a zadaniem psów jest znalezienie ich w jak najkrótszym czasie — podczas egzaminu czas ten nie może przekraczać 30 min. Co ciekawe psy szkolone na pierwszy stopień ratownika lawinowego, nagradzane są poprzez pozwolenie na wykopanie znalezionej osoby, kolejne etapy nie dopuszczają takiego zachowania.

Niestety pomimo coraz lepszych warunków zwierząt pracujących w różnych zawodach i deklaracji współpracy opartej na poszanowaniu wzajemnych interesów, myślę, że często współpraca ta postrzegana jest przez pryzmat uprzedzeń tożsamościowych, które wiążą się ze stereotypami i mitami mówiącymi, że życie człowieka jest cenniejsze od życia przedstawicieli innych gatunków zwierząt. To może prowadzić do wykorzystywania psów lawinowych do akcji ratowniczych, w których tracą życie pod śniegiem. Oczywiście wyszkolone psy tropiące są bardzo drogie, a ich opiekunowie dbają o ich dobrostan psychiczny i fizyczny. Jednak czytając artykuły o opisanych wyżej psich zawodach, trudno nie zauważyć powszechnej obecności przekonania, że jeżeli znajdzie taka potrzeba, zwierzę zobowiązane jest do ratowania życia człowieka, kosztem własnego istnienia.

Kultura, rozumiana jako zbiór wartości, norm, symboli i narracji, odgrywa kluczową rolę w kształtowaniu postaw i przekonań jednostek oraz całych społeczeństw. Jak wskazuje Hańderek (2020), kultura nie tylko odzwierciedla rzeczywistość, ale także pomaga ludziom interpretować i radzić sobie z jej złożonością, w tym z trudnymi i niejednoznacznymi aspektami życia. Poprzez różnorodne formy wyrazu — literaturę, film, animacje czy inne wytwory — kultura umożliwia kreatywne ukazywanie konsekwencji ludzkich działań, często w sposób metaforyczny lub symboliczny. Dzięki temu odbiorcy, w tym dzieci i młodzież, mogą rozwijać wrażliwość na subtelne przekazy i złożoność relacji społecznych, co sprzyja budowaniu krytycznego myślenia

i empatii. W dalszej części tekstu przyjrzymy się, jak ta funkcja kultury realizuje się na przykładzie popularnej animacji Psi Patrol, której analiza pozwala dostrzec ukryte warstwy znaczeniowe i stereotypy obecne w mediach skierowanych do najmłodszych widzów.

Takie narracje wzmacniane są w ludzkich społeczeństwach od najmłodszych lat np. przez wytwory kultury popularnej, których świetnym przykładem jest bardzo popularna w ostatnich latach bajka Psi Patrol¹²⁷. Jej ciekawą analizę znajdziemy w materiałach Matthew Patricka, znanego na platformie YouTube pod pseudonimem MatPat¹²⁸. Trzy filmy na jego kanale *The Film Theorists* poświęcone są wskazanej bajce. Oczywiście już teraz pragnę podkreślić, że interpretacja MatPata nie jest oficjalnie potwierdzona przez twórców animacji, a stanowi jedynie interpretację wskazanego dzieła. Przytoczony przeze mnie przykład nie ma charakteru naukowego, chcę jednak pokazać, że opisywane we wcześniejszych częściach stereotypy i mity są głęboko zakorzenione w kulturze, która kształtuje ludzi od najmłodszych lat. Uważam, też, że tworzenie interpretacji bajek dla dzieci, pokazujących niepokojące aspekty wytworów, które oglądaliśmy w młodości, kształtują pewną wrażliwość, na nieoczywiste przekazy popkultury. Dla mojej generacji były to animacje takie jak Ed, Ed i Eddy¹²⁹, Chojrak Tchórzliwy Pies¹³⁰ czy Pora na Przygodę¹³¹. Mam nadzieję, że

¹²⁷ Bajka wyprodukowana przez Spin Master Entertainment i Nickelodeon Productions w latach 2013-2023. Tutaj muszę zaznaczyć, że biorąc pod uwagę popularność opisywanej produkcji, może się okazać, że w najbliższej przyszłości powstaną kolejne filmy.

¹²⁸ Influencer zajmujący się analizą różnych sfer kultury. Do 2023, czyli do czasu sprzedania i rezygnacji z prowadzenia kanałów, posiadał cztery konta na platformie youtube.com: The Game Theorists, The Film Theorists, The Food Theorists i The Style Theorists. Jego kanały osiągają ponad 5 mln wyświetleń, co wskazuje na ich ogromną popularność.

¹²⁹ Bajka wyprodukowana przez Cartoon Network Studios w latach 1999-2009. Fabuła dotyczy 3 chłopców, nazywanych potocznie Edami, zamieszkujących zaułek fikcyjnej dzielnicy osadzonej w Stanach Zjednoczonych. Kreskówka liczy 12 bohaterów, w wieku 12 lat. Głównym celem Edów w każdym odcinku jest zdobycie pieniędzy, które mają zamiar wydać na swoje ulubione słodycze — łąmiszczęki. Próbują pozyskać środki poprzez wyłudzenie ich od pozostałych mieszkańców zaułka, na różne, często mało moralne sposoby. Teorie tworzone wokół tej bajki wskazują, że bohaterowie nie żyją, a wydarzenia przedstawione w animacji pokazują życie tych dzieci po śmierci (Cartoon Corruption, 2023).

¹³⁰ Bajka wyprodukowana przez Cartoon Network Studios w latach 1999-2002. Fabuła opowiada o małym różowym psie o imieniu Chojrak oraz jego opiekunach—Muriel i Eustachym Motykach. Relacje między bohaterami są wyraźnie nakreślone, widzimy, że o ile kobieta kocha Chojraka, to mężczyzna go

młodzi ludzie, którzy aktualnie wychowują się, pochłaniając godziny materiałów z Psim Patrolem, w dorosłości będą zauważać złożoność relacji i przekazów materiałów, które oglądali.

Animacja opowiada o 10-letnim chłopcu — Ryderze (nr. 1), który opiekuje się sześcioma pieskami, wyposażonymi w 6 rodzajów narzędzi odpowiadających poszczególnym funkcjom społecznym: Chase (nr.2) jest psem policyjnym, Marshall (nr. 3) strażakiem, Sky (nr. 4) awiatorem, Rocky (nr. 5) to pies śmieciarz, Rubble (6.) jest budowniczym i konstruktorem oraz Zuma (nr. 7) ratownikiem wodnym. W każdym odcinku drużyna przeprowadza misję ratunkową, w ramach której wybierane są pieski pełniące kluczową rolę w danej akcji. MatPat uważa, że istnieje wiele wskazówek pokazujących, że psiaki giną podczas niektórych akcji ratunkowych, dlatego Ryder zastępuje je innymi szczeniakami. Do tego wniosku youtuber dochodzi po wielogodzinnym seansie ze swoim 3-letnim synem i rozpoznaniu dotyczącym szczegółowych informacji o świecie przedstawionym w bajce. Pierwszym sygnałem, który wzbudził ciekawość autora materiałów, był odc. 9a, 1 sezonu. Przedstawia on tzw. *origin story*¹³² Rubble (nr.6), czyli historię przyłączenia tego pieska do Psiego Patrolu. Okazuje się, że ten dołączył do drużyny jako ostatni, gdy został uratowany przez zespół, przypadkowo pojechał z nimi na misję ratowania człowieka uwięzionego pod śniegiem, gdzie łopata Rydera zawiodła, w wyniku czego to właśnie Rubble wykopał tunel i uratował ofiarę. Po powrocie z akcji psiak otrzymał odznakę (charakterystyczny symbol związany z jego rolą), hełm oraz własny *pupmobil*, czyli pojazd — w tym wypadku koparkę. Na pierwszy rzut oka historia ta może nie wzbudzać podejrzeń, jednak istnieje kilka zastanawiających szczegółów. Po pierwsze na początku odcinka, gdy Ryder wzywa drużynę do misji, a Rubble nie jest jeszcze członkiem zespołu,

nienawidzi, ciągle nazywa go głupim i próbuje się go pozbyć. Rodzina mieszka na odludnej farmie, którą często odwiedzają stwory, próbujące wyrządzić krzywdę bohaterom na wiele różnych sposobów. Jednak Chojrak, choć tchórzliwy, przezwycięża swój strach aby ratować ukochaną Muriel. Współczesne analizy omawianej bajki wskazują, że ta jest metaforą dysfunkcyjnej i przemocowej rodziny, gdzie pies jest odpowiednikiem dziecka (Nguyen, M. 2016).

¹³¹ Bajka wyprodukowana przez Cartoon Network Studios, w latach 2010-2018, opowiada o magicznym psie Jacku i człowieku Finnie, który marzy aby zostać bohaterem. Fabuła rozgrywa się fikcyjnym świecie, który przybrał obecny kształt po tzw. „wojnie grzybów”. Ta i wiele innych wskazówek sugerują, że akcja animacji ma miejsce po wojnie nuklearnej.

¹³² *Origin story* — to historia bohatera lub bohaterów, pokazująca wydarzenia prowadzące do zostania przez nich protagonistą lub antagonistą.

możemy zauważyć, że na ekranie widnieje emblemat z symbolem, który w późniejszej części odcinka ma oznaczać rolę Rubble. Dodatkowo na panelu sterowania Rydera znajduje się przycisk odpowiadający za sprzęt budowniczego i konstruktora, czyli funkcji w przyszłości pełnionych przez nowego szczeniaka. Poza tym, gdy Rubble oficjalnie dołącza do zespołu, otrzymuje hełm i *pupmobil*, które noszą ślady wcześniejszego użytkowania. Ciekawym elementem jest też fakt, że koparka, którą otrzymuje Rubble, jest dla niego za duża, a sam chłopiec mówi, że piesak będzie musiał do niej dorosnąć. Jest to ciekawe, biorąc pod uwagę fakt, że wszystkie inne pieski dostały *pupmobile* idealnie dopasowane do swoich rozmiarów. Ostatecznym argumentem przemawiającym za hipotezą, że przed przedstawianym w odcinku bohaterem, był inny pies pełniący rolę budowniczego i konstruktora, jest numeracja pojazdów — jak wskazałam wcześniej, koparka Rubble jest oznaczona numerem 6, co plasuje się na przedostatnim miejscu, a nie ostatnim, na co wskazuje historia przedstawiona w odcinku, gdyż jak pisałam Rubbel dołącza do drużyny jako ostatni. Co więcej, w odcinku 6 1 sezonu Marshall bierze udział w konkursie na najszybszego psa strażaka, co implikuje, że istniały inne psy, które pełniły tę rolę. Wniosek ten wspierany jest przez przesłankę, polegającą na wspomnieniu, że obecny rekord został ustalony przez innego pieska. Każdy, kto nie oglądał Psiego Patrolu, mógłby zasugerować prostsze wyjaśnienie przedstawionych faktów, mianowicie stwierdzić, że w świecie kreowanym przez bajkę istnieje wiele różnych Psich Patroli. Jednak w pełnometrażowym filmie będącym częścią omawianego świata, widzimy sytuację, w której przejezdny kierowca tira ulega wypadkowi, w wyniku czego potrzebuje ratunku. Prosi świadka, aby ten zadzwonił na policję, a gdy dostaje odpowiedź, że w mieście Adventure Bay jest Psi Patrol zamiast wskazanego organu, kierowca jest zdziwiony, natomiast na widok Chase — gadającego psa policjanta, zaczyna panikować. To oznacza, że taka jednostka jest czymś wyjątkowym i mało znanym w tym świecie. Oczywiście wraz z upływem sezonów Psi Patrol zyskuje coraz to większą sławę. Dodatkowo należy zaznaczyć, że Ryder jest genialnym dzieckiem, które tworzy wszystkie urządzenia używane przez Psi Patrol. To istotne, ponieważ wydaje się, że poza narzędziami dla swoich podopiecznych, tworzy także broń, dla głównych przeciwników swojej drużyny, jak np. Ekipa Kotów (*Cat Crew*). To sugeruje, że Ryder nie tylko naraża życie swoich podopiecznych, ale także tworzy zagrożenia, w wyniku których te mogą stracić życie, MatPat sugeruje, że Ryder robi to wszystko w celu korzyści materialnej, ponieważ wraz z popularnością Psiego Patrolu, produkuje coraz

więcej merchu¹³³ i bierze udział w coraz poważniejszych akcjach, za które niewątpliwie otrzymuje sowitą opłatę. Natomiast na podstawie dwóch przedstawionych w bajce *origin story* możemy wnioskować, że Ryder wybiera do swojej drużyny bezdomne, biedne i przestraszone szczeniaki, za którymi nikt nie będzie tęsknił. Dla przykładu w historii Chase’a dowiadujemy się, że ten został porzucony przez swoich opiekunów przy drodze w mieście a w rezultacie, gdy przerażony próbował zrozumieć, co się dzieje, prawie został potrącony przez samochód. Na szczęście Ryder uratował szczeniaka i przygarnął go do siebie, za co Chase jest mu bardzo wdzięczny. Historia ta pokazuje, że Ryder przygarnia szczeniaki z ulicy, aby wcielić je do swojego zespołu i poprzez wdzięczność manipulować psiakami, aby wykonywały niebezpieczne zadania, które mogą doprowadzić do ich śmierci. Co więcej, sam wywołuje wiele zagrożeń, co pozwala mu na zyskanie większej popularności. Wydaje się zatem, że Ryder nie przejmuje się życiem swoich podopiecznych, ponieważ te są zastępowalne. Czyli życie zwierząt ratujących ludzi, jest mniej ważne niż istnienie tych drugich.

Bardziej rzeczywistym przykładem narażania życia zwierząt ratujących ludzi, są szczury wykonujące pracę sapera. Zwierzęta te szkolone są przez różne organizacje takie jak APOPO (*Anti-Persoonsmijnen Ontmijnende Product Ontwikkeling*), czyli belgijsko-tanzańską organizację pozarządową, która szkoli szczury do wyszukiwania min lądowych oraz wykrywania gruźlicy. Ich podopieczni nazywani są HeroRATS¹³⁴ oraz jeśli wierzyć danym podawanym na oficjalnej stronie organizacji, dzięki niskiej wadze i bardzo wysokim zdolnością percepcyjnym w zakresie węchu, wyczuwają substancje wybuchowe zawarte w minach, co sprawia, że ignorują pozostałe metalowe elementy przeszukiwanego terenu. To czyni je znacznie skuteczniejszymi od najbardziej specjalistycznych wykrywaczy metali. Natomiast szczury wykrywające gruźlice pracują w lokalnych laboratoriach przyszpitalnych, zajmujących się badaniami z zakresu gruźlicy. Próbką wskazana przez szczura jako podejrzana, przekazywana jest do dalszych testów w siedzibie APOPO, gdzie sprawdzana jest zgodnie z kryteriami wskazanymi przez WHO. Okazuje się, że praca szczurów zwiększa wskaźnik wykrywalności gruźlicy nawet o 40%. Najbardziej znanym osobnikiem szkolonym przez opisywaną organizację jest Magawa, czyli szczur, który otrzymał Złoty Medal PDSA, przyznawany przez brytyjską organizację za bohaterstwo i poświęcenie służbie

¹³³ Czyli tematyczne gadżety i produkty związane z określonym tematem. W przypadku Psiego Patrolu są to plecaki i figurki, które sprzedawane są także w rzeczywistości.

¹³⁴ Link do oficjalnej strony APOPO: <https://apopo.org/herorats/>.

w siłach zbrojnych oraz cywilnych służbach ratunkowych brytyjskiej Wspólnoty Narodów (natemat.pl, 2022). Jest to szczególne wyróżnienie, gdyż Magawa był pierwszym szczurem w historii, który został odznaczony wskazanym medalem. To przyniosło mu zasłużoną sławę, gdyż uznaje się, że w trakcie swojej pięcioletniej służby w Kambodży wykrył prawie 100 min przeciwpiechotnych i 40 niewybuchów. Faktem jest, że organizacje takie jak APOPO dbają o swoich podopiecznych, redukując przypadki śmierci, do absolutnego minimum. Gdy jednak dojdzie do nieszczęśliwych zdarzeń, w których zwierzę straci życie, przeprowadzane jest dochodzenie badające przyczyny takich wydarzeń.

Osobiście uważam, że współpraca człowieka i innych zwierząt w opisanych wyżej przykładach jest zjawiskiem pozytywnym. Chociaż niektóre przypadki budzą wiele kontrowersji, szczególnie w dyskusji społecznej dotyczącej wykorzystywania i dyskryminowania zwierząt. Jestem zwolenniczką współpracy międzygatunkowej, nawet w opisanych wyżej przypadkach psów lawinowych, czy szczurów saperów. Uważam jednak, że współpraca ta powinna odbywać się z poszanowaniem interesów zwierząt pomagających w działaniach na rzecz ludzi, a co za tym idzie, należy pamiętać, że współpraca ta powinna mieć charakter umowy społecznej.

Koncepcja umowy społecznej w filozofii politycznej została pierwszy raz opisana przez Jeana-Jacques'a Rousseau w pracach *Rozprawa o nierówności między ludźmi* (1954) oraz *O umowie społecznej* (2002). Idea ta odnosi się do teorii dotyczącej podstaw społeczeństwa i władzy politycznej oraz opiera się na założeniu, że społeczeństwo powstaje poprzez dobrowolny kontrakt między jednostkami, zgadzającymi się przekazać część swojej wolności i władzy jednostkowej na rzecz wspólnego dobra społecznego. Istota tej umowy polega na tym, że ludzie, żyjąc w stanie natury, muszą zrzec się pewnych swobód w celu stworzenia wspólnoty politycznej, która zapewniłaby im bezpieczeństwo i sprawiedliwość. Autor wymienionych prac wskazuje na 3 główne etapy powstawania umowy społecznej:

- a. Pojawienie się społeczeństwa: według Rousseau pierwotnie ludzie żyli w stanie natury, w którym każdy człowiek jest suwerenem, niezależnym od innych jednostek. Jednak w wyniku wzrostu populacji i ograniczonych zasobów, życie w stanie natury staje się niemożliwe lub niebezpieczne, co prowadzi do powstania społeczeństwa.

- b. Ustanowienie umowy społecznej: które polega na dobrowolnej decyzji ludzi o zawarciu umowy społecznej, w której Ci zrzekają się części swojej wolności i władzy na rzecz wspólnego dobra społecznego. Umowa ta tworzy społeczeństwo i ustanawia rządzących.
- c. Stworzenie władzy państwowej: umowa społeczna prowadzi do powstania władzy państwowej, która ma reprezentować wolę ogółu społeczeństwa i dbać o jego dobro. Władza powinna działać zgodnie z interesem ogółu i respektować prawa jednostki.

Zatem w koncepcji Rousseau, umowa społeczna ma charakter zbiorowego aktu suwerenności, w którym jednostki decydują się na współpracę i podporządkowanie się dla dobra wspólnego. Należy jednak pamiętać, że autor *O umowie społecznej* krytykował istniejące w jego czasach formy rządów, argumentując, że wielu współczesnych mu monarchów i arystokratów naruszało umowę społeczną, sprawując władzę w sposób sprzeczny z interesem ogółu społeczeństwa. Dlatego też, Rousseau popierał demokratyczne formy rządów, które byłyby bliższe woli ogółu i szanowałyby prawa jednostki. Sam Rousseau nie rozważał zagadnienia umowy społecznej w kontekście zwierząt pozaludzkich, uważam jednak, że jego koncepcja, powinna znaleźć zastosowanie w relacjach międzygatunkowych. W omawianej już wyżej pracy *Zwierzeta Mówią* (Mejier, 2021) znajdziemy solidną podstawę do rozważań dotyczących tego, jak taka umowa mogłaby wyglądać. Oczywiście moim celem nie jest skonstruowanie kompletnych zasad umowy społecznej, a przedstawienie pewnych sugestii oraz wskazanie, że wiele ZINL funkcjonuje w relacjach noszących wszelkie oznaki umowy społecznej. Pamiętając, że w koncepcji Rousseau, istota umowy polega na wspólnej zgodzie jednostek na pewne ograniczenia swojej wolności indywidualnej w celu stworzenia wspólnoty, która zapewnia bezpieczeństwo i dobrobyt wszystkim jej członkom. Uważam, że międzygatunkowa umowa społeczna powinna obejmować przynajmniej cztery wymienione niżej aspekty:

- a. Uznanie godności i praw zwierząt pozaludzkich: międzygatunkowa umowa społeczna musi uwzględniać godność i prawa wszystkich gatunków zawierających umowę. Ludzie powinni zobowiązać się do traktowania zwierząt z szacunkiem i dbałością, respektując ich prawo do życia i dobrostanu.
- b. Ograniczenie wykorzystywania zwierząt: w ramach umowy społecznej ludzie powinni zobowiązać się do ograniczenia wykorzystywania zwierząt do celów,

które mogą być źródłem cierpienia. Zobowiązanie to powinno pociągnąć za sobą restrykcje dotyczące hodowli przemysłowej, eksperymentów na zwierzętach oraz wykorzystywania zwierząt w rozrywce.

- c. Ochrona środowiska naturalnego: międzygatunkowa umowa społeczna powinna obejmować zobowiązanie do ochrony środowiska naturalnego i różnorodności biologicznej, które są istotne dla przetrwania wszystkich form życia.
- d. Współpraca w celu dobra wspólnego: międzygatunkowa umowa społeczna między ludźmi a innymi gatunkami zwierząt powinna opierać się na idei współpracy w celu dobra wspólnego, czyli stworzenia harmonijnej wspólnoty.

Uwzględniając wymienione aspekty międzygatunkowa umowa społeczna, która stanowi próbę zrozumienia potrzeb innych form życia, prowadziłyby nie tylko, do uznania ich praw, ale też do bardziej zrównoważonej i etycznej relacji ludzi ze światem przyrody.

Jak wskazałam wyżej, wiele przedstawicieli różnych gatunków zwierząt działa na rzecz wspólnego dobra angażując się w pomoc człowiekowi np. poprzez służbę w policji, wojsku czy w służbach ratunkowych. Natomiast człowiek łamie warunki międzygatunkowej umowy społecznej we wszystkich opisanych wcześniej punktach. Uważam, że źródłem dyskryminacji w każdym z wyżej wymienionych aspektów są uprzedzenia tożsamościowe, przejawiające się jako tendencja do ignorowania lub lekceważenia komunikacji zwierząt (zarówno wewnątrz-, jak i międzygatunkowej). Pierwsze z wymienionych wiąże się z niedocenieniem niejęzykowych form komunikacji, dla przykładu psy korzystają głównie z komunikacji posturalnej, dlatego gdy osobnik wyraża niepokój, ludzie mogą go zignorować lub niewłaściwie interpretować jego zachowanie, co może prowadzić do nietrafnych reakcji lub zaniedbań. Uważam, że taka ignorancja jest przejawem uprzedzeń tożsamościowych, w ramach których ludzie ignorują realne komunikaty, na rzecz własnych projekcji. Natomiast lekceważące traktowanie informacji pochodzących od pozaludzkich interlokutorów, polega na traktowaniu komunikatów od osobnika gatunku innego niż człowiek, jako mniej wiarygodnej niż te przekazywane przez ludzi. Na przykład, jeśli kot sygnalizuje swojemu opiekunowi, że coś jest nie tak z jego zdrowiem, opiekun może go potraktować niepoważnie, co będzie prowadzić do opóźnionej reakcji na problemy zdrowotne zwierzęcia. Przykładem łamania międzygatunkowej umowy społecznej jest dyskryminacja w kontekście badań naukowych, w których zwierzęta laboratoryjne stanowią modele do badań chorób i procesów biologicznych. Co wiąże się

z nie uwzględnianiem ich pełni doświadczeń oraz potrzeb, co może prowadzić do niesprawiedliwości testimonialnej. Dodatkowo marginalizacja ich głosu jest wyraźnie widoczna w kontekście ochrony środowiska, gdzie aktualnie perspektywa zwierząt często jest ignorowana i lekceważona na rzecz ludzkich interesów gospodarczych lub innych priorytetów.

6.5. Niesprawiedliwość hermeneutyczna

Rozważania na temat niesprawiedliwości hermeneutycznej warto zacząć od próby odpowiedzi na istotne pytanie, które jest analogiczne do pytania zadawanego w filozofii feministycznej od początku jej powstania. Tak, jak wtedy pytano i pyta się do teraz: *w jaki sposób społeczne relacje władzy krępowały i krępują kobiecą zdolność do zrozumienia ich własnych doświadczeń?* Tak warto zapytać *jak relacje władzy funkcjonujące w obszarze relacji człowiek — inne zwierzęta ograniczają zdolność do rozumienia własnych doświadczeń pozaludzkim agansom społecznym?* Niestety rozważania te nie mieszczą się w zakresie zainteresowań niniejszej dysertacji, jednak uważam, ten aspekt za warty rozwinięcia w innych opracowaniach. Chcę jednak zaznaczyć, że zaprezentowane przez Fricker (2007) ramy pojęciowe, można zastosować do rozważań nad innymi rodzajami uprzedzeń. Dla przykładu aparat pojęciowy, który dostrzegamy w proponowanej przez nią koncepcji, wykorzystywany jest także w rozważaniach filozoficznych i naukowych dotyczących rasizmu czy homofobii. Tak jak, w przypadku niesprawiedliwości testimonialnej, uważam, że zwierzęta pozaludzkie mogą być podmiotami niesprawiedliwości hermeneutycznej. Szczegółowe wyjaśnienie tego zagadnienia znajduje się w następnej części tego rozdziału. Przejdźmy natomiast teraz do omówienia podstawowych pojęć istotnych z punktu widzenia analizy problemów centralnych dla niniejszej dysertacji.

Niesprawiedliwość hermeneutyczna dotyczy sytuacji, w których brak odpowiednich struktur pojęciowych lub kategorii poznawczych uniemożliwia zrozumienie doświadczeń pewnych grup społecznych oraz prowadzi do dyskryminacji w postaci marginalizacji, lub zignorowania perspektyw odmiennych od własnych, w procesie interpretacji rzeczywistości społecznej. Brak odpowiednich pojęć we własnym (indywidualnym) zasobie hermeneutycznym, wywołuje emocje takie jak dezorientacja czy wyobcowanie (*isolated*). Luka ta znacząco utrudnia komunikację międzygrupową, co wynika z różnic w aparatach pojęciowych (braki w etykietach

werbalnych oraz rozbieżności w ich znaczeniu). Może być zlokalizowana w bardziej lokalnych zasobach hermeneutycznych, np. w stereotypach i mitach (komponent poznawczy) przyjmowanych przez daną społeczność lokalnie, m.in. w przekonaniach wywodzących się świata wytworzonego przez środowisko rodzinne, szkolne czy zawodowe jednostki. W zależności od tego, w jakim zakresie hermeneutycznych zasobów występuje luka. Struktury poznawcze jednostki zorganizowane są w określony sposób, a to determinuje, które przekonania uznane zostaną za fakty społeczne. To prowadzi do przyjęcia postaw (uprzedzenia) utrudniających interpretowanie doświadczeń przedstawicieli innych grup. Co ciekawe sama nierówność hermeneutyczna jest trudna do wykrycia, gdyż systemy poznawcze jednostki jej nie identyfikują, generują jednak emocje związane z brakiem zrozumienia, czyli wspomnianą dezorientację i poczucie wyobcowania. Chociaż autorka *Epistemic Injustice: Power and the Ethics of Knowing* nie wyjaśnia mechanizmu leżącego u podstaw nieuświadamiania sobie istnienia luki, to uważam, że dobrym wyjaśnieniem jest odwołanie się do mechanizmu konfabulacji¹³⁵. Innymi słowy, omawiany rodzaj niesprawiedliwości pojawia się, gdy w zasobach hermeneutycznych jednostki lub społeczeństwa, istnieje luka informacyjna, obejmująca określony zakres albo ogół doświadczeń przedstawicieli innych grup. Przejawia się ona w zauważalnych brakach etykiet werbalnych, określających znaczną część własnego doświadczenia społecznego u przedstawicieli grup dotkniętych takim ograniczeniem (*disadvantage*) poznawczym.

Dobrym przykładem obrazującym tego rodzaju zjawisko jest relacja pochodząca z lat '60, na którą powołuje się Fricker (2007, str. 150). Jest to wypowiedź kobiety – Wendy Stanford – która po urodzeniu dziecka zachorowała na depresję poporodową. Jednak w czasach, w których chorowała, zjawisko to nie było powszechnie znane, czyli w ogólnych zasobach hermeneutycznych społeczeństwa znajdowała się istotna luka, która skutkowała m.in. niediagnozowaniem tej przypadłości, wynikającym nie ze złych chęci lekarza, ale raczej jego braków w informacyjnych na temat, słabo poznanego zaburzenia. Brak kolektywnego zrozumienia dotyczącego depresji poporodowej

¹³⁵ Konfabulacja rozumiana jest jako „błędnie ugruntowane przekonanie, które jest skutkiem wypełniania luk w systemie poznawczym” (Reuter, Dziarnowska, 2012, str. 1). W ramach epistemicznej perspektywy na konfabulację, mechanizm ten dotyczy szerokiego zakresu różnych domen wiedzy, takich jak np. percepcja wzrokowa czy percepcja ciała. Za Reuter i Dziarnowską (2012) przyjmuję, że konfabulacja to różnego rodzaju tryby pracy umysłu w warunkach niepewności.

prowadził do doświadczania przez Wendy niesprawiedliwości hermeneutycznej, jednak nie identyfikowała jej w taki sposób, przyjmując stereotypy mówiące, że stan, w którym się znajduje, jest jej winą. Co więcej, kobieta wskazuje, że jej mąż, który również posiadał lukę w swoich zasobach hermeneutycznych, swoim zachowaniem potwierdzał jej przekonania. Dopiero gdy przyjaciółka Wendy — Esther zabrała ją na spotkanie na MIT, gdzie podczas 44-minutowej sesji w małych grupach zaczęto mówić o depresji poporodowej, kobieta zrozumiała co się z nią dzieje, a co za tym idzie, przestała się za to obwiniać. Jak twierdzi autorka relacji, był to dla niej przełomowy moment, gdyż pozwolił jej zrozumieć własną sytuację i pozbyć się poczucia wyobcowania.

Studiując prace feministyczne można zauważyć, że jednym z pierwszych etapów podnoszenia świadomości społecznej dotyczącej dyskryminacji kobiet, było *wypowiadanie się* (*speak-out*). Innymi słowy, metoda ta jest pierwszym etapem zapełniania luki hermeneutycznej znajdującej się w ogólnych zasobach społeczeństwa. Istotnym produktem *wypowiadania się*, jest nadanie krzywdzącemu doświadczeniu etykiety werbalnej. Rozpowszechnienie odpowiedniego słowa określającego niezrozumiałe wcześniej odczucia, jest ostatecznym etapem wytworzenia adekwatnego aparatu pojęciowego. *Wypowiadanie się* pozwala na rozszerzenie indywidualnych struktur poznawczych (pojęć), oraz uzupełnienie kolektywnych zasobów hermeneutycznych. Fricker (2007) uważa, że proces ten pozwala na przezwycięzenie tradycyjnych interpretacji rzeczywistości oraz wytworzenie nowych bardziej odpowiednich. Korzystając z terminologii wynikającej z ram terminologicznych przyjętych w niniejszej dysertacji — wytworzenie nowych interpretacji rzeczywistości, wiąże się z powstaniem nowych narracji wyjaśniających świat (mity) oraz zmianę w kryteriach przyporządkowywania obiektów do kategorii (stereotypy).

Fricker wskazuje także, że nie każde doświadczenie hermeneutycznej niekorzyści wiąże się z byciem podmiotem niesprawiedliwości jej odpowiadającej. Za przykład podaje człowieka, który cierpi na zaburzenia widoczne w jego behawiorze, w tym jego zachowaniu społecznym (Fricker, 2007). Jednak żyje w czasach, w których jego przypadłość nie jest jeszcze rozpoznana. Taka osoba będzie doświadczać hermeneutycznej niekorzyści, ale nie wynikającej z niesprawiedliwości hermeneutycznej. Różnica w przytoczonym przed chwilą przykładzie, a wcześniej opisanym przypadkiem, polega na tym, że Wendy Stanford doświadczała poznawczej niekorzyści, której przyczyną jest nierówny udział kobiet w tworzeniu kolektywnych zasobów hermeneutycznych. Czyli w przeciwieństwie do człowieka chorującego na

niezidentyfikowane jeszcze zaburzenie, sytuacja Wedny wynika z nieuczciwie rozłożonego wpływu społecznego na kolektywne formy społecznego zrozumienia. Ogólne zasoby hermeneutyczne powinny być odbiciem perspektyw różnych grup tworzących społeczeństwo, natomiast gdy istnieje nierówność w poszanowaniu istotnego obszaru społecznego doświadczenia, członkowie grup posiadających więcej społecznej władzy narzucają swoje doświadczenie pozostałym, twierdząc, że jest ono adekwatne dla wszystkich, tym samym marginalizując realne doświadczenia grup nieuprzywilejowanych. Prowadzi to do podporządkowania i wykluczenia z praktyk, które konstytuują znaczenie społeczne istotne dla grup dotkniętych taką ograniczeniem. Czyli deprywacja wynikająca z omawianego rodzaju niesprawiedliwości, związana jest z przynależnością do grupy bezsilnej (*powerless*) społecznie. W związku z powyższym omawiane zjawisko ma charakter czysto strukturalny, gdyż nie popełnia jej żaden konkretny agens, a wynika z luki hermeneutycznej w ogólnych lub bardziej lokalnych zasobach hermeneutycznych społeczeństwa, która jest źródłem asymetrycznej poznawczej niekorzyści.

Autorka *Epistemic Injustice: Power and the Ethics of Knowing* wskazuje, że omawiana niesprawiedliwość może działać zarówno systemowo, jak i bardziej jednorazowo. Druga z wymienionych polega na zasłonięciu (*obscure*) istotnego obszaru doświadczenia społecznego przed kolektywnym zrozumieniem z powodu marginalizacji hermeneutycznej. Jednak jak w przypadku niesprawiedliwości testimonialnej Fricker (2007) skupia się przede wszystkim na luce obecnej w ogólnych zasobach hermeneutycznych, powodującej niesprawiedliwość systemową, „polegającą na zasłonięciu znacznego pola własnego doświadczenia społecznego, przed kolektywnym zrozumieniem z powodu uprzedzenia tożsamościowego w zbiorowym zasobie hermeneutycznym” (Fricker, 2007, str. 156). Sama Fricker zaznacza, że hermeneutyczna niesprawiedliwość jest systemowa wtedy, gdy marginalizacja dotyczy sfer życia innych niż hermeneutyczna. W tym miejscu należy przypomnieć, że w pracy Fricker „uprzedzenie” rozumiane jest jako mechanizm potwierdzający, istniejące w umyśle jednostki struktury poznawcze, natomiast w niniejszej dysertacji wskazane pojęcie rozumie się jako postawa, a co za tym idzie, składa się na nią znacznie szerszy zbiór mechanizmów poznawczych, afektywnych i behawioralnych, kształtujących wrażliwość różnych społeczności na postrzeganie doświadczeń przedstawicieli grup

innych niż własna ¹³⁶ (Fricker, 2007, str. 148). Uważam, że oprócz władzy tożsamościowej istotną rolę w omawianym rodzaju niesprawiedliwości pełni władza materialna ¹³⁷. Sprawia, że marginalizacja hermeneutyczna przyczynia się do wykluczenia socjo-ekonomicznego, w wyniku którego osoby nią dotknięte nie podejmują pracy w określonych zawodach, będących newralgicznym aspektem uczestnictwa we współtworzeniu ogólnych zasobów hermeneutycznych społeczeństwa.

Kluczowa krzywda wynikająca z niesprawiedliwości hermeneutycznej polega na tym, że osoba jej doświadczająca usiłuje dokonać aktu mowy, którego treść — choć leży w jej interesie — nie może zostać właściwie zakomunikowana ze względu na braki w dostępnych zasobach hermeneutycznych. Ograniczenia te czynią przekaz niezrozumiałym lub „niewidzialnym” dla odbiorcy, uniemożliwiając tym samym efektywną komunikację w sprawie, która ma dla nadawcy fundamentalne znaczenie. W celu lepszego zrozumienia konsekwencji płynących z systemowej niesprawiedliwości hermeneutycznej warto powołać się na przykład Edmunda White’a, który w swojej książce *A Boy’s Own Story* (1982) prezentuje doświadczenia dorastania jako homoseksualny chłopiec w latach ‘50 w Stanach Zjednoczonych. Książka jest częściowo autobiograficzna i zajmuje się szeroką gamą tematów związanych z tożsamością, seksualnością i akceptacją siebie. White przedstawia emocjonalne i psychologiczne zmagania głównego bohatera z jego seksualnością, w tym konfrontacje z nietolerancją i niezrozumieniem ze strony społeczeństwa i rodziny. Na łamach wskazanej książki wyraźnie widać, jak brak odpowiednich narzędzi językowych do wyrażenia swojego doświadczenia oraz marginalizacja ze względu na systematyczne uprzedzenia tożsamościowe wpływają na przedstawiciela grupy dotkniętej niesprawiedliwością hermeneutyczną. Brak aparatu pojęciowego do opisanja własnej tożsamości, wynikający z postrzegania i braku akceptacji wobec homoseksualności prowadzi do deficytów ram społecznych, które pozwalałyby w pełni zrozumieć i zaakceptować swoją tożsamość. Co więcej, bohater zmagają się z niezrozumieniem i odrzuceniem ze strony rodziny i rówieśników. Społeczeństwo, w którym dorasta,

¹³⁶ Niektóre z nich opisałam w rozdziale II: *Wybrane mechanizmy poznawcze leżące u podstaw tendencyjności związanej z uprzedzeniami*.

¹³⁷ W celu zdefiniowania pojęcia władzy materialnej Fricker (2007, str. 148) powołuje się na filozofię Marksa, w ramach której, ten kto posiada władzę materialną ma wpływ na praktyki, które generują znaczenie społeczne.

charakteryzuje się silną stygmatyzacją homoseksualności. Odrzucenie potęguje jego cierpienie i utrudnia mu budowanie zdrowej tożsamości, a brak pozytywnych modeli i wsparcia społecznego sprawia, że bohater czuje się zmuszony do ukrywania swojej prawdziwej tożsamości. Sytuacja ta ma swoje psychologiczne konsekwencje, takie jak wymienione wcześniej poczucie izolacji czy wstydu, ale też pojawienie się wewnętrznych konfliktów, znacząco niskiej samooceny, lęku i depresji. Krzywdy te wynikają z niemożności wyrażenia i akceptacji siebie w kontekście społecznych norm i oczekiwań.

6.6. Uprzedzenia tożsamościowe a niesprawiedliwość hermeneutyczna względem ZINL

Tak jak w przypadku analizy problemu niesprawiedliwości testimonialnej, tak w obszarze należy w pierwszej kolejności udzielić odpowiedzi na pytania szczegółowe. W przypadku krzywdy hermeneutycznej, należy rozważyć pięć wstępnych pytań:

- a. Czy ZINL posiadają zdolność do samoświadomości i refleksji?
- b. Czy formy komunikacji ZINL pozwalają na wyrażania własnych doświadczeń i potrzeb?
- c. Czy zwierzęta pozaludzkie mogą być świadomymi podmiotami struktur interpretacyjnych, które są źródłem marginalizacji hermeneutycznej jednostki?
- d. Czy niesprawiedliwość hermeneutyczna może być rozszerzona na zwierzęta pozaludzkie, biorąc pod uwagę ich zdolność do odczuwania bólu, cierpienia i emocji?
- e. Czy ludzkie pojmowanie niesprawiedliwości i hermeneutyki nie ogranicza nas w uznawaniu zwierząt pozaludzkich za podmioty tego rodzaju niesprawiedliwości?

Odpowiedź na niektóre z zadanych wyżej pytań znaleźć można w poprzednich częściach niniejszej dysertacji. W podrozdziałach 3.1. *Historia badania umysłów zwierząt pozaludzkich* oraz 2.3. *Myślenie analogiczne* argumentowałam, że koncepcja *Umweltu* pozwala na redukcję wpływu ograniczającej perspektywy antropocentrycznej (e.). Natomiast w części. 6.4. *Czy uprzedzenia tożsamościowe mogą powodować niesprawiedliwość testimonialną wobec ZINL?* argumentowałam za Meijer (2021), że zwierzęta pozaludzkie posiadają systemy komunikacyjne (b.), pozwalające im na

generowanie aktów komunikacyjnych dotyczących ich samopoczucia, jednak antropocentryczna perspektywa (d.) powoduje ich ignorowanie lub lekceważenie. Uważam, że źródłem niesprawiedliwości hermeneutycznej w kontekście zwierząt pozaludzkich jest to, w jaki sposób ludzie interpretują ich zachowania i doświadczenia, a konkretnie, ludzka tendencja do antropomorfizacji, czyli przypisywania zwierzętom ludzkich cech i motywacji, może prowadzić do niewłaściwego zrozumienia ich zachowań i potrzeb. Co więcej, we wskazanej części wspominałam także, o stanowiskach argumentujących, że ZINL posiadają pewien poziom refleksji i samoświadomości (a.). W tym miejscu trzeba uwzględnić w powyższych rozważaniach, głos Robina Dunbara, który w swojej książce *Nowa historia ewolucji człowieka* (2014), argumentuje, że zwierzęta pozaludzkie zdolne są do tworzenia reprezentacji dotyczących ich stanów mentalnych (takie jak np. jestem głodny)¹³⁸.

Ostatnim pytaniem, na które należy udzielić odpowiedzi, jest c. W kognitywistyce i psychologii (Nęcka, 2013), poznawcze struktury interpretacyjne rozumiane są jako mechanizmy umysłowe, których funkcją jest przetwarzanie, organizowanie i interpretowanie informacji o świecie. Struktury te są kluczowe dla rozumienia świata, podejmowania decyzji i interakcji z otoczeniem. Jednym z podstawowych elementów tych struktur są schematy, czyli mentalne reprezentacje pomagające organizować i interpretować informacje. Mogą one dotyczyć różnych dziedzin, takich jak obiekty (np. schemat „człowiek”), sytuacje (np. schemat „park”) czy role społeczne (np. schemat „dozorca parkingu”). Specyficznymi rodzajami schematów są skrypty, które opisują sekwencje zdarzeń w określonych kontekstach. Na przykład, skrypt „wizyta w restauracji” może zawierać kroki od wejścia do restauracji, poprzez zamówienie jedzenia, aż po płacenie rachunku. Kolejnym elementem omawianych struktur poznawczych są ramy, które pomagają interpretować i klasyfikować informacje. Zawierają one pewne stałe i zmienne elementy, pozwalające na elastyczne adaptowanie się do różnych kontekstów. Ramy są używane do interpretacji doświadczeń i do przewidywania przyszłych zdarzeń. Innym istotnym elementem struktur interpretacyjnych są prototypy i opisane wyżej kategorie

¹³⁸ Szersze omówienie tego zagadnienia znajduje się w ostatnim rozdziale dysertacji, gdzie rozwijam stanowisko Dunbara i wyjaśniam szczegółowe różnice terminologiczne między jego koncepcją a przyjętym punktem widzenia w tym opracowaniu.

poznawcze. Pierwsze z wymienionych rozumiane są jako najbardziej typowi przedstawiciele danej kategorii. Na przykład, dla kategorii *ptaki* prototypem może być gołąb jako najbardziej typowy przedstawiciel. Najbardziej złożonymi strukturami interpretacyjnymi są modele mentalne, czyli wewnętrzne reprezentacje rzeczywistości, umożliwiające symulowanie różnych scenariuszy i przewidywanie skutków działań. Umożliwiają one zrozumienie złożonych systemów i zależności między ich elementami. Przyjmując teorię ewolucji, należy założyć, że poznawcze struktury „budujące” umysły ZINL są przynajmniej podobne, do ludzkich elementów umysłowych. Wydaje się mało prawdopodobne, że zwierzęta pozaludzkie nie posiadają żadnych z wyżej wymienionych struktur, gdyż ich umysły również przetwarzają informacje, które muszą zostać zorganizowane i interpretowane, dlatego nie widzę uzasadnienia do zastosowania w tym wypadku argumentu *ad ignorantiam*. Co więcej, uważam, że nawet jeżeli przyjmiemy, że wyżej wymienione struktury są unikatowe dla człowieka, a co za tym idzie, relacja analogii jest nieuzasadniona, należy przyjąć, że nasze systemy poznawcze wytworzyły struktury homologiczne do umysłów innych zwierząt. Biorąc pod uwagę argumenty przedstawione przeze mnie w części 6.4. *Uprzedzenia tożsamościowe a niesprawiedliwość testimonialna wobec ZINL* uważam, że twierdzenie to jest mało prawdopodobne, a co za tym idzie, uznaję, że korzystanie z analogii w tym przypadku jest jak najbardziej uzasadnione.

Zwierzęta pozaludzkie mogą być zatem podmiotami niesprawiedliwości hermeneutycznej, a jej przejawy mogą różnić się od siebie w zależności od uprzedzenia ją powodującego. Uważam jednak, że wszystkie wymienione przeze mnie uprzedzenia mają negatywne konsekwencje dla równowagi ekologicznej i sprawiedliwości wobec wszystkich istot żyjących. Pierwszą wartą wzmiankowania postawą istotną w kontekście opisywanej niesprawiedliwości, jest szowinizm gatunkowy, prowadzący do ignorowania, marginalizacji lub lekceważenia perspektyw, potrzeb i praw przedstawicieli gatunków innych niż ludzki. Wskazana postawa wiąże się ze stereotypami i mitami mówiącymi, że gatunek ludzki jest lepszy lub bardziej wartościowy niż wszystkie inne gatunki¹³⁹, co prowadzi do wyłącznego skupienia się na ludzkich doświadczeniach, potrzebach i interesach, ignorując lub marginalizując potrzeby innych gatunków. W wyniku tego omawiane uprzedzenie może prowadzić do

¹³⁹ Dla przypomnienia w niniejszej pracy szowinizm gatunkowy definiowany jest jako: uprzedzenie przejawiające się tendencyjnością do faworyzowania członków własnego gatunku kosztem innych.

traktowania zwierząt pozaludzkich wyłącznie w kategoriach narzędzi lub zasobów służących do zaspokajania ludzkich potrzeb, bez uwzględniania ich własnych potrzeb, interesów i praw. W wyniku takiego kategoryzowania przedstawiciele innych gatunków doświadczenia i cierpienie ZINL jest bagatelizowane, a w konsekwencji traktowane jest jako mniej istotne lub mniej ważne niż ludzkie. To z kolei może przyczyniać się do niesprawiedliwości hermeneutycznej poprzez ignorowanie perspektyw zwierząt i brak empatii wobec ich doświadczeń. Co więcej szowinizm gatunkowy może prowadzić do wykluczenia innych gatunków z moralnej wspólnoty, traktując je jako mniej wartościowe lub niezdolne do uczestniczenia w moralnym życiu. W konsekwencji stają się podmiotem niesprawiedliwości hermeneutycznej poprzez ignorowanie ich moralnych praw i potrzeb. Kolejnym istotnym uprzedzeniem opisywanym w ramach niniejszej dysertacji jest karnizm. Uważam, że w jego rezultacie, niesprawiedliwość hermeneutyczna przejawia się w nadmiernej koncentracji na pewnych gatunkach zwierząt, co prowadzi do ignorowania perspektyw i potrzeb innych gatunków oraz do nieproporcjonalnego traktowania różnych form życia. Krnizm charakteryzuje się stereotypami i mitami, które mogą przyczyniać się do romantyzowania pewnych gatunków, traktując je jako niewinne lub niemal boskie, podczas gdy inne są traktowane jako mniej wartościowe lub nawet szkodliwe. W wyniku czego przyjęta zostaje hierarchia wartości, w ramach której pewne gatunki mogą być uznane za bardziej wartościowe lub godne ochrony niż inne. Podobne przejawy niesprawiedliwości hermeneutycznej dostrzec można w kontekście bambizmu. Jak wskazałam wyżej, postawa ta wiąże się ze stereotypami i mitami mówiącymi, że niektórzy przedstawiciele innych gatunków są istotami niezdolnymi do życia bez pomocy człowieka, a zatem traktowanie ich oparte jest na antropomorficznych projekcjach, a nie na reakcji wynikającej z realnych potrzeb postrzeganych zwierząt.

Interesującym, choć niewyczerpująco przeze mnie omówionym wątkiem pozostaje zagadnienie podwójnych standardów moralnych, jakie stosujemy w odniesieniu do ludzi i ZINL. Colin McGinn (1999), krytykując stanowisko Petera Singera, zwraca uwagę na fakt, że nasza skłonność do moralnego wartościowania różni się w zależności od tego, czy dotyczy ludzi czy zwierząt pozaludzkich. Argumentuje on, że o ile ludzie – nawet ubodzy lub znajdujący się w trudnej sytuacji życiowej – posiadają potencjał do społecznego współdziałania i zaradności, o tyle zwierzęta są wobec człowieka całkowicie bezbronne. Dlatego, zdaniem McGinna (1999), moralny obowiązek ochrony powinien być silniej akcentowany właśnie w odniesieniu do

zwierząt. Choć rozumowanie to jest intrygujące, uważam, że wymaga ono znacznie szerszego opracowania, uwzględniającego kontekst społeczny i komunikacyjny. Badaczki takie jak Eva Meijer wskazują bowiem, że zwierzęta pozaludzkie nie są jedynie biernymi odbiorcami ludzkich działań, lecz aktywnie uczestniczą w relacjach społecznych, dostosowując się – na miarę swoich możliwości – do obowiązujących norm

i interakcyjnych reguł. Jeśli przyjmiemy, że ograniczona zdolność rozumienia tych zasad nie wyklucza zwierząt z moralnej wspólnoty, lecz raczej uzasadnia pewien rodzaj społeczno-moralnej wyrozumiałości (jak w przypadku dzieci), wówczas również stosunek do ludzi znajdujących się w trudnej sytuacji przestaje być tak jednoznaczny, jak sugeruje McGinn 1999). O ile zatem ten kierunek refleksji uznaję za ważny i inspirujący, to jednocześnie uważam, że wykracza on poza zakres obecnego rozdziału i zasługuje na osobne, pogłębione opracowanie w przyszłych badaniach.

6.7. Podsumowanie

Rozdział szósty stanowi próbę zastosowania koncepcji niesprawiedliwości epistemicznej zaproponowanej przez Mirandę Fricker, do analizy uprzedzeń względem zwierząt pozaludzkich oraz ich epistemicznego wykluczenia. W szczególności skoncentrowałam się na testimonialnym wymiarze tej niesprawiedliwości, rozumianym jako systematyczne podważanie wiarygodności komunikatów jednostek na podstawie stereotypów i uprzedzeń tożsamościowych. Pokazano, że analogiczne mechanizmy mogą dotyczyć również zwierząt – zarówno poprzez marginalizację ich zdolności komunikacyjnych, jak i poprzez ignorowanie ich „głosów” w kontekście społecznym, naukowym czy politycznym.

W części tej podkreślałam, że uprzedzenia względem ZINL są wspierane przez struktury władzy tożsamościowej, której działanie często ma charakter niejawny, automatyczny i głęboko zakorzeniony w narracjach kulturowych. Podałam także krytycznej analizie kluczowe założenia, leżące u podstaw odmowy epistemicznego statusu zwierzętom, m.in. poprzez rekonstrukcję klasycznych i współczesnych teorii wiedzy oraz ukazanie ich ograniczeń w kontekście międzygatunkowym. Wskazano, że choć ZINL mogą nie spełniać rygorystycznych standardów wiedzy refleksyjnej, coraz więcej danych empirycznych i filozoficznych stanowisk wspiera tezę o ich poznawczej kompetencji i zdolności do przekazywania istotnych informacji.

Co ciekawe, podmiotem międzygatunkowej niesprawiedliwości testimonialnej może być także człowiek, który reprezentujące perspektywę mentalistyczną – badacze i badaczki, aktywiści działający na rzecz zwierząt pozaludzkich. Argumentowałam, że także oni mogą doświadczać niesprawiedliwości epistemicznej ze względu na systemowe uprzedzenia obecne w społeczności akademickiej i kulturowej. Tym samym, rozdział szósty nie tylko pogłębia rozumienie uprzedzeń wobec zwierząt, lecz także stanowi próbę poszerzenia kategorii sprawiedliwości epistemicznej o obszary rzadko uwzględniane w głównym nurcie debaty filozoficznej.

Zakończenie:

Niniejsza dysertacja stanowi próbę połączenia analizy epistemologicznej, refleksji meta-poznawczej oraz wrażliwości etycznej w podejściu do badań nad teorią umysłu u zwierząt innych niż ludzie. Jej celem nie było rozstrzygnięcie sporu dotyczącego tego, czy zwierzęta pozaludzkie posiadają teorię umysłu, lecz ujawnienie i krytyczna analiza tych aspektów poznawczych, metodologicznych oraz kulturowych, które kształtują dominujące sposoby interpretowania danych empirycznych. W toku pracy ukazałam, że przyjęcie określonych ram teoretycznych — zwłaszcza napięcia między podejściem mentalistycznym a behawioralnym — wpływa nie tylko na interpretację wyników, ale także na sam sposób formułowania pytań badawczych i konstruowania procedur eksperymentalnych.

W szczególności podkreśliłam, że poznanie, emocje i zachowanie są ze sobą nierozzerwalnie powiązane. Dlatego próby analizy izolowanych mechanizmów każdego z komponentów postaw, bez uwzględnienia ich wymiaru społecznego, mogą prowadzić do uproszczeń i niepełnych wniosków. Choć nie zdecydowałam się na wyodrębnienie osobnej sekcji poświęconej komponentowi afektywnemu, jego rola została wielokrotnie zaakcentowana, m.in. poprzez analizę uprzedzeń, motywacji autowaloryzacyjnych i narracji kulturowych. Mam nadzieję, że zawarte tu analizy będą stanowić punkt wyjścia do pogłębionych badań poświęconych właśnie emocjonalnym komponentom postaw wobec zwierząt pozaludzkich oraz roli afektu w procesach interpretacyjnych.

Jednym z kluczowych wniosków wynikających z przedstawionych rozważań jest potrzeba porzucenia binarnych opozycji ludzkie — zwierzęce na rzecz perspektyw bardziej ewolucyjnie spójnych i metodologicznie neutralnych. Tylko takie podejście pozwala na uczciwe i wrażliwe epistemicznie badanie poznania innych gatunków — odwołania do antroponegacjonistycznych mitów o *wyjątkowości człowieka*. W analizowanym kontekście dostrzegam także kilka możliwych kierunków dla dalszych badań. Po pierwsze, jak już wspominałam, mam nadzieję, że dysertacja ta stanie się punktem wyjścia do pogłębionej analizy afektywnych aspektów postaw wobec zwierząt pozaludzkich, ze szczególnym uwzględnieniem ich roli w kształtowaniu narracji naukowych i społecznych. Co więcej, uważam, że wskazanie jest skonstruowanie zestawienia omówionych perspektyw z innymi możliwymi ujęciami (np. enaktywizmem czy podejściem ekologiczno-dynamicznym), w celu stworzenia bardziej inkluzywnych ram badawczych. Co więcej dostrzegam konieczność krytycznej refleksji nad terminami używanymi w zakresie *mindreadingu* oraz sprytnego

zachowania innego, a także rekonstrukcję pojęcia teorii umysłu, które wciąż stosowane jest w dwóch różnych znaczeniach (szerokie i wąskie), często nawet w ramach jednej pracy (artykułu czy książki). Wyznaczone zakresy pojęć powinny uwzględniać różnorodność definicji i zastosowań określonych zdolności w kontekście międzygatunkowym.

Zrozumienie umysłów innych podmiotów — zarówno ludzkich, jak i pozaludzkich — wymaga nie tylko zastosowania narzędzi empirycznych, lecz również epistemicznej pokory, filozoficznej refleksji oraz etycznej wrażliwości. Przedstawione w niniejszej pracy analizy nie dostarczają ostatecznych rozstrzygnięć, lecz mają na celu krytyczne podważenie utrwalonych założeń oraz stworzenie przestrzeni dla nowych pytań i perspektyw badawczych. W kontekście współczesnych przemian, które redefiniują granice między *nami* a *innymi*, konieczne jest formułowanie narracji, które nie tylko opisują rzeczywistość, lecz również ją współtworzą w sposób odpowiedzialny, krytyczny i sprawiedliwy.

Bibliografia:

1. Aaltola, E. (2013). Empathy, intersubjectivity, and animal philosophy. *Environmental Philosophy*, 10(2), s. 75-96, <https://doi.org/10.5840/envirophil201310215>.
2. Abbott, I., Abbott, L. K., Grant, P. R. (1975). Seed selection and handling ability of four species of Darwin's finches. *The Condor*, 77(3), s. 332–335, <https://doi.org/10.2307/1366230>.
3. Aboud, F. E. (1988). *Children and prejudice*. Oxford: Basil Blackwell.
4. Adair, J. G., Epstein, J. S. (1968). Verbal cues in the mediation of experimenter bias. *Psychological Reports*, 22(3_suppl), s. 1045-1053, <https://doi.org/10.2466/pr0.1968.22.3c.1045>.
5. Adamus, M. (2018). Ekonomia altruizmu—o racjonalności zachowań prospołecznych. *Diametros*, (57), s. 1-22, <https://doi.org/10.13153/diam.1233>.
6. Ainsworth, M. D. S., Blehar, M. C., Waters, E., Wall, S. N. (2015). *Patterns of attachment: A psychological study of the strange situation*. Psychology Press.
7. Allen, T. J., Sherman, J. W., Conrey, F. R., Stroessner, S. J. (2009). Stereotype strength and attentional bias: Preference for confirming versus disconfirming information depends on processing capacity. *Journal of experimental social psychology*, 45(5), s. 1081-1087, <https://doi.org/10.1016/j.jesp.2009.06.001>.
8. Alley, T. (1983). Infantile head shape as an elicitor of adult protection. *Merrill Palmer Quarterly*. 29(4), s. 411–427.
9. Anderson, N. H. (1971). Integration theory and attitude change. *Psychological Review*, 78(3), s. 171–206, <https://doi.org/10.1037/h0030834>.
10. Andrews, K. (2012). *Do apes read minds? Toward a new folk psychology*. MIT Press.
11. Andrews, K. (2020). *The animal mind: An introduction to the philosophy of animal cognition*. Routledge.
12. Apicella, C., Norenzayan, A., Henrich, J. (2020). Beyond WEIRD: A review of the last decade and a look ahead to the global laboratory of the future. *Evolution and Human Behavior*, 41(5), s. 319–329. <https://doi.org/10.1016/j.evolhumbehav.2020.07.015>.
13. Aragón, O. R., Clark, M. S., Dyer, R. L., Bargh, J. A. (2015). *Dimorphous Expressions of Positive Emotion: Displays of Both Care and Aggression in Response to Cute Stimuli*. *Psychological Science*, 26(3), s. 259–273, doi:10.1177/0956797614561044.

14. Aronson E., Travis, C. (2007). *Błądzą wszyscy, ale nie ja. Dlaczego usprawiedliwiamy głupie poglądy, szkodliwe decyzje i złe działania?*. Smak Słowa.
15. Aronson, E., Aronson, J., Radzicki, J. (2007). *Człowiek-istota społeczna*. Państwowe Wydawnictwo Naukowe, Warszawa.
16. Asch, S. (1946). Forming Impressions of Personality. *Journal of Abnormal and Social Psychology*, 41, s. 258-290. <https://doi.org/10.1037/h0055756>.
17. Ashmore, R. D. (1970). The problem of intergroup prejudice. [W:] B. E. Collins (red.), *Social psychology: Social influence, attitude change, group processes, and prejudice*. Addison-Wesley Publishing Company, s. 245–296.
18. Ashmore, R. D., Boca, F. K. (1979). Sex Stereotypes and Implicit Personality Theory: Toward a Cognitive-Social Psychological Conceptualization. *Sex Roles*, vol. 5, s. 219-248, <https://doi.org/10.1007/BF00287932> s. 219-248.
19. Ashmore, R., Del Boca, F. (1981). Conceptual approaches to stereotypes and stereotyping. [W:] D. Hamilton (Ed.), *Cognitive processes in stereotyping and intergroup behavior*. Erlbaum.
20. Aubenque, P. (2010). *Problematyka mądrości u Arystotelesa*. Wydawnictwo Uniwersytetu Jagiellońskiego. Kraków.
21. Bacon, F. (1857). *The works of Francis Bacon*. Longman.
22. Baillargeon, R., Scott, R. M., He, Z. (2010). False-belief understanding in infants. *Trends in Cognitive Sciences*, 14(3), <https://doi.org/10.1016/j.tics.2009.12.005>.
23. Baker, H. K., Nofsinger, J. R. (Red.) (2010). *Behavioral finance: investors, corporations, and markets*, John Wiley & Sons, vol. 6, <https://doi.org/10.1002/9781118258415>.
24. Bandura, A. (2007). *Teoria społecznego uczenia się*. Państwowe Wydawnictwo Naukowe, Warszawa.
25. Barash, D. P. (2018). *Through a glass brightly: Using science to see our species as we really are*. Oxford University Press.
26. Bargh, J. A., Pietromonaco, P. (1982). Automatic information processing and social perception: The influence of trait information presented outside of conscious awareness on impression formation. *Journal of Personality and Social Psychology*, 43(3), s. 437-449, <https://doi.org/10.1037/0022-3514.43.3.437>.

27. Baron-Cohen, S. (1991). Precursors to a theory of mind: Understanding attention in others. [W:] A. Whiten (red.), *Natural Theories of Mind: Evolution, Development, and Simulation of Everyday Mindreading*. Oxford: Basil Blackwel, s. 233–251.
28. Baron-Cohen, S. (1994). *Mindblindness: An essay on autism and theory of mind*. MIT Press.
29. Baron-Cohen, S. (1995). The eye direction detector (EDD) and the shared attention mechanism (SAM): Two cases for evolutionary psychology. [W:] C. Moore P. J. Dunham (red.) *Joint attention: Its origins and role in development*. Lawrence Erlbaum Associates, s. 41–59.
30. Baron-Cohen, S. (2009). Rozwój zdolności czytania innych umysłów: cztery etapy. [W:] A. Klawiter (red.), *Formy aktywności umysłu. Tom 2*, Wydawnictwo Naukowe PWN, s. 273–290.
31. Baron-Cohen, S., Campbell, R., Karmiloff-Smith, A., Grant, J., Walker, J. (1995). Are children with autism blind to the mentalistic significance of the eyes? *British Journal of Developmental Psychology*, 13(4), s. 379–398, <https://doi.org/10.1111/j.2044-835X.1995.tb00687.x>.
32. Baron-Cohen, S., Leslie, A. M., Frith, U. (1985). Does the autistic child have a “theory of mind”? *Cognition*, 21(1), s. 37–46, [https://doi.org/10.1016/0010-0277\(85\)90022-8](https://doi.org/10.1016/0010-0277(85)90022-8).
33. Barrett, J. L. (2011). Cognitive science of religion: Looking back, looking forward. *Journal for the Scientific Study of Religion*, 50(2), s. 229–239, <https://doi.org/10.1111/j.1468-5906.2011.01564.x>.
34. Barrett, L. F. (2017). The theory of constructed emotion: an active inference account of interoception and categorization. *Social cognitive and affective neuroscience*, 12(1), s. 1–23, <https://doi.org/10.1093/scan/nsw154>.
35. Bartał, I. B.-A., Decety, J., Mason, P. (2011). Empathy and pro-social behavior in rats. *Science*, 334(6061), s. 1427–1430, <https://doi.org/10.1126/science.1210789>.
36. Bartoszewicz, M. (2017). Wybrane aspekty poznawcze i emocjonalne socjotechniki mediów. *Studia Medioznawcze*, 69(2), s. 93–106.
37. Bator, J. (2007). Słodka kulturka. Nurt „kawaii” w kulturze japońskiej. *Kultura i Społeczeństwo*. [W:] Bator, W. (2013). Święte–przekłete–pozbawione duszy. O traktowaniu zwierząt w doktrynach i praktykach religii świata. *Ethos. Kwartalnik Instytutu Jana Pawła II KUL*, 26(2 (102), s. 103–116.
38. Beck, B. B. (1967). A study of problem solving by gibbons. *Behaviour*, s. 95–109, <https://doi.org/10.1163/156853967x00190>.

39. Bednarek, M. (2021). Kopciuszki dla młodzieży. O rewritingach motywu ATU 510A w literaturze young adult. *Dzieciństwo. Literatura i Kultura*, 3(2), s. 11-34, <https://doi.org/10.32798/dlk.840>.
40. Bekoff, M., Pierce, J. (2018). *Dzika sprawiedliwość. Moralne życie zwierząt*. Copernicus Center Press, Toruń.
41. Bennett, J. (1978). Some remarks about concepts. *Behavioral and Brain Sciences*, 1(4), s. 557–560, <https://doi.org/10.1017/S0140525X00076573>.
42. Bensalah, L., Caillies, S., Anduze, M. (2016). Links among cognitive empathy, theory of mind, and affective perspective taking by young children. *The Journal of genetic psychology*, 177(1), s. 17-31, <https://doi.org/10.1080/00221325.2015.1106438>.
43. Bentham, J. (1958). *Wprowadzenie do zasad moralności i prawodawstwa*. Wydawnictwo Naukowe PWN.
44. Berg, K. S., Delgado, S., Cortopassi, K. A., Beissinger, S. R., & Bradbury, J. W. (2012). Vertical transmission of learned signatures in a wild parrot. *Proceedings of the Royal Society B: Biological Sciences*, 279(1728), s. 585-591. <https://doi.org/10.1098/rspb.2011.0932>.
45. Bermúdez, J. L. (2003). *Thinking without words*. Oxford University Press.
46. Bermúdez, J. L. (2005). *Philosophy of psychology: A contemporary introduction*. Routledge.
47. Bermúdez, J. L. (2009). *Decision theory and rationality*. Oxford University Press.
48. Berthet, M., Surbeck, M., Townsend, S. W. (2025). Extensive compositionality in the vocal system of bonobos. *Science*, 388(6742), s. 104-108, <https://doi.org/10.1126/science.adv1170>.
49. Bettelheim, B. (1976). *The Uses of Enchantment: The Meaning and Importance of Fairy Tales*. Knopf, New York.
50. Bílek, J., Nedoma, J., Jirásek, M. (2018). Representativeness heuristics: A literature review of its impacts on the quality of decision-making. *Scientific papers of the University of Pardubice. Series D, Faculty of Economics and Administration*. 43/2018.
51. Biro, D., Inoue-Nakamura, N., Tonooka, R., Yamakoshi, G., Sousa, C., Matsuzawa, T. (2003). Cultural innovation and transmission of tool use in wild chimpanzees: Evidence from field experiments. *Animal Cognition*, 6(4), s. 213-223. <https://doi.org/10.1007/s10071-003-0183-x>.

52. Biro, D., Matsuzawa, T. (1999). Numerical ordering in a chimpanzee (*Pan troglodytes*): Planning, executing, and monitoring. *Journal of Comparative Psychology*, 113(2), s. 178–185. <https://doi.org/10.1037/0735-7036.113.2.178>.
53. Blanchard, F. A., Crandall, C. S., Brigham, J. C., Vaughn, L. A. (1994). Condemning and condoning racism: A social context approach to interracial settings. *Journal of Applied Psychology*, 79(6), s. 993–997, <https://doi.org/10.1037/0021-9010.79.6.993>
54. Blest, A.D. (1957). The Function of Eyespot Patterns in the Lepidoptera. *Behaviour*, vol. 11, 209-258.
55. Bodenhausen, G. V. (1988). Stereotypic biases in social decision making and memory: Testing process models of stereotype use. *Journal of Personality and Social Psychology*, 55(5), s.726–737, <https://doi.org/10.1037/0022-3514.55.5.726>.
56. BonJour, L. (1985). *The structure of empirical knowledge*. Harvard University Press, Cambridge.
57. Boorstin, D. (1983). *The Discoverers*, New York: Random House.
58. Borgi, M., Cirulli, F. (2013). Children's preferences for infantile features in dogs and cats. *Human-Animal Interaction Bulletin*, <https://doi.org/10.1079/hai.2013.0010>.
59. Borgi, M., Cogliati-Dezza, I., Brelsford, V., Meints, K., Cirulli, F. (2014). Baby schema in human and animal faces induces cuteness perception and gaze allocation in children. *Frontiers in psychology*, vol. 5, <https://doi.org/10.3389/fpsyg.2014.00411>.
60. Bousfield, W. A. (1953). The occurrence of clustering in the recall of randomly arranged associates. *Journal of General Psychology*, vol. 49, s. 229-240, <https://doi.org/10.1080/00221309.1953.9710088>.
61. Bousfield, W. A., Cohen, B. H. (1953). The effects of reinforcement on the occurrence of clustering in the recall of randomly arranged associates. *The Journal of Psychology*, 36(1), s. 67-81, <https://doi.org/10.1080/00223980.1953.9712878>.
62. Bowlby, J. (1969). *Attachment* (Vol. 1). [W:] *Attachment and loss*. New York: Basic Books.
63. Bradshaw, J. W. (2016). Sociality in cats: A comparative review. *Journal of veterinary behavior*, vol. 11, s. 113-124, <https://doi.org/10.1016/j.jveb.2015.09.004>.
64. Bray, E. E., Gnanadesikan, G. E., Horschler, D. J., Levy, K. M., Kennedy, B. S., Famula, T. R., MacLean, E. L. (2021). Early Emerging and Highly Heritable Sensitivity to Human Communication in Dogs. *Current Biology*, vol. 31, s. 3132 – 3136, <https://doi.org/10.1016/j.cub.2021.04.055>.

65. Brehm, J. W. (1956). Postdecision changes in the desirability of alternatives. *The Journal of Abnormal and Social Psychology*, 52(3). <https://doi.org/10.1037/h0041006>.
66. Brewer, M. B., Brown, R. J. (1998). Intergroup relations. [W:] Gilbert, D. T., Fiske, S. T., Lindzey G., *Handbook of social psychology*, Boston: s. 554 –594.
67. Broadhurst, P. L. (1966). Behavioral inheritance: Past and present. *Conditional reflex: a Pavlovian journal of research & therapy*, 1(1), s. 3-15.
68. Brooks, R., Meltzoff, A. N. (2005). *The development of gaze following and its relation to language*. *Developmental Science*, 8(6), s. 535–543. <https://doi.org/10.1111/j.1467-7687.2005.00445.x>
69. Brosch, T., Sander, D., Scherer, K. R. (2007). That baby caught my eye... attention capture by infant faces. *Emotion*, 7(3), s. 685–689. <https://doi.org/10.1037/1528-3542.7.3.685>.
70. Brosnan, S. F., De Waal, F. B. (2003). Monkeys reject unequal pay. *Nature*, 425(6955), s. 297-299. <https://doi.org/10.1038/nature01963>.
71. Bruner, J. S. (1983). *Child's talk: Learning to use language*. Norton, Nowy York.
72. Buckner, C. (2013). A property cluster theory of cognition. *Philosophical Psychology*, 28(3), s. 307–336. <https://doi.org/10.1080/09515089.2013.843274>.
73. Bugnyar, T., Heinrich, B. (2005). Ravens, *Corvus corax*, differentiate between knowledgeable and ignorant competitors. *Proceedings of the Royal Society B: Biological Sciences*, 272(1573), s. 1641–1646,
74. Burnham, D., Kitamura, C., Vollmer-Conna, U. (2002). What's new, pussycat? On talking to babies and animals. *Science*, vol. 296, <https://doi.org/10.1126/science.1069587>.
75. Burns, T. A. (2017). Empathy, Simulation, and Neuroscience: A Phenomenological Case against Simulation-Theory. *Phenomenology and Mind*, vol. 12, s. 208-216, https://doi.org/10.13128/Phe_Mi-21119.
76. Butterfill, S. A., Apperly, I. A. (2013). How to construct a minimal theory of mind. *Mind & Language*, 28(5), s. 606–637. <https://doi.org/10.1111/mila.12036>.
77. Butterworth, G., Jarrett, N. (1991). What minds have in common is space: Spatial mechanisms serving joint visual attention in infancy. *British Journal of Developmental Psychology*, 9(1), s. 55–72. <https://doi.org/10.1111/j.2044-835X.1991.tb00801.x>.

78. Calarco, M. (2008). *Zoographies: The question of the animal from Heidegger to Derrida*. Columbia University Press.
79. Call, J., Tomasello, M. (2008). Does the chimpanzee have a theory of mind? 30 years later. *Trends in Cognitive Sciences*, 12(5), s. 187–192,
80. Carpenter, M. (2009). Just how joint is joint action in infancy? *Topics in Cognitive Science*, 1(2), s. 380–392, <https://doi.org/10.1111/j.1756-8765.2009.01026.x>.
81. Carraher, D. W., Schliemann, A. D., Brizuela, B. M., Earnest, D. (2006). Arithmetic and algebra in early mathematics education. *Journal for Research in Mathematics education*, 37(2), s. 87-115, <https://doi.org/10.2307/30034843>.
82. Carroll, J. M., Russell, J. A. (1996). Do facial expressions signal specific emotions? Judging emotion from the face in context. *Journal of personality and social psychology*, 70(2), s. 205-218, <https://doi.org/10.1037/0022-3514.70.2.205>.
83. Cassidy, J. (1994). Emotion regulation: Influences of attachment relationships. *Monographs of the society for research in child development*, 59(2-3), s. 228-249.
84. Castano, E., Yzerbyt, V., Paladino, M. P., Sacchi, S. (2002). I belong, therefore, I exist: Ingroup identification, ingroup entitativity, and ingroup bias. *Personality and Social Psychology Bulletin*, 28(2), s.135-143, <https://doi.org/10.1177/014616720228200>.
85. Celani, G. (2002). Human beings, animals and inanimate objects: what do people with autism like? *Autism*, vol. 6, s. 93–102. doi: 10.1177/1362361302006001007.
86. Centola, D., Baronchelli, A. (2015). The spontaneous emergence of conventions: An experimental study of cultural evolution. *Proceedings of the National Academy of Sciences*, 112(7), s. 1989-1994, <https://doi.org/10.1073/pnas.1418838112>.
87. Chambers, C. (2017). *The seven deadly sins of psychology: A manifesto for reforming the culture of scientific practice*. Princeton University Press.
88. Chandler, M. J., Greenspan, S. (1972). Ersatz egocentrism: A reply to H. Borke. *Developmental Psychology*, 7(2), s. 104–106, <https://doi.org/10.1037/h0033145>.
89. Chrostowski, W. (2005). Status zwierząt w Biblii. *Forum Teologiczne*, vol. 6, s. 7-22.
90. Clark, A., Chalmers, D. J. (2010). The extended mind. [W:] *Life And Mind: Philosophical Issues In Biology And Psychology*. The MIT Press.

91. Clayton, N. S., Dally, J. M., Emery, N. J. (2007). Social cognition by food-caching corvids: The western scrub-jay as a natural psychologist. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 362(1480), s. 507–522, <https://doi.org/10.1098/rstb.2006.1992>.
92. Clayton, N. S., Dickinson, A. (1998). Episodic-like memory during cache recovery by scrub jays. *Nature*, 395(6699), s. 272–274, <https://doi.org/10.1038/26216>.
93. Cohen, I. B. (1985). *Revolution in science*. Harvard University Press.
94. Cole, M., Gay, J., Glick, J. A., Sharp, D. W. (1971). *The cultural context of learning and thinking*. Basic Books, New York.
95. Colom, F., Vieta, E., Martínez-Arán, A., Reinares, M., Benabarre, A., Gastó, C. (2000). Clinical factors associated with treatment noncompliance in euthymic bipolar patients. *The Journal of Clinical Psychiatry*, 61(8), s. 549–555. <https://doi.org/10.4088/JCP.v61n0802>.
96. Correll, J., Park, B., Judd, C. M., Wittenbrink, B. (2002). The police officer's dilemma: Using ethnicity to disambiguate potentially threatening individuals. *Journal of Personality and Social Psychology*, 83(6), s. 1314–1329, <https://doi.org/10.1037/0022-3514.83.6.1314>.
97. Cosmides, L., Tooby, J. (2013). Evolutionary psychology: New perspectives on cognition and motivation. *Annual review of psychology*, 64(1), s. 201–229, <https://doi.org/10.1146/annurev.psych.121208.131628>.
98. Costello, K., Hodson, G. (2010). *Exploring the roots of dehumanization: The role of animal–human similarity in promoting immigrant humanization*. Erratum.
99. Cowley, Stephen J. (2007). The cradle of language : making sense of bodily connexions. [W:] D. Moyal-Sharrock (red.), *Perspicuous presentations: essays on Wittgenstein's philosophy of psychology*. New York: Palgrave-Macmillan.
100. Crandall, C. S., Eshleman, A. (2003). A justification-suppression model of the expression and experience of prejudice. *Psychological bulletin*, 129(3), s. 414–446, <https://doi.org/10.1037/0033-2909.129.3.414>.
101. Crisp, R. J., Hewstone, M., Rubin, M. (2001). Does Multiple Categorization Reduce Intergroup Bias? *Personality and Social Psychology Bulletin*, 27(1), s. 76–89, <https://doi.org/10.1177/0146167201271007>.

102. Crowell-Davis, S. L., Curtis, T. M., Knowles, R. J. (2004). Social organization in the cat: a modern understanding. *Journal of feline medicine and surgery*, 6(1), s.19-28, <https://doi.org/10.1016/j.jfms.2003.09.013>.
103. Czeremski, M. (2021). *Mit w umyśle: Ewolucyjno-kognitywne podstawy form mitycznych*. Wydawnictwo UJ.
104. Darwin, K. (2009). *O powstawaniu gatunków drogą doboru naturalnego czyli o utrzymywaniu się doskonalszych ras w walce o byt*. Wydawnictwo Uniwersytetu Warszawskiego, Warszawa.
105. Dasser, V., Ulback, I., Premack, D. (1989). The perception of intention. *Science*, 243(4889), s. 365–367. <https://doi.org/10.1126/science.2911743>.
106. de Waal, F. B. (2018). *Bystre zwierzę: czy jesteśmy dość mądrzy, aby zrozumieć mądrość zwierząt?*. Copernicus Center Press.
107. de Waal, F. B. M. (2008). Putting the altruism back into altruism: The evolution of empathy. *Annual Review of Psychology*, 59, str. 279-300, <https://doi.org/10.1146/annurev.psych.59.103006.093625>.
108. de Waal, F. B. M., Preston, S. D. (2017). Mammalian empathy: Behavioural manifestations and neural basis. *Nature Reviews Neuroscience*, 18(8), s. 498–509, <https://doi.org/10.1038/nrn.2017.72>
109. Decety, J., Jackson, P. L. (2004). The functional architecture of human empathy. *Behavioral and Cognitive Neuroscience Reviews*, 3(2), s. 71–100, <https://doi.org/10.1177/1534582304267187>.
110. Deese, J. (1959). On the prediction of occurrence of particular verbal intrusions in immediate recall.. *Journal of Experimental Psychology*, 58(1), s. 17–22, <https://doi.org/10.1037/h0046671>.
111. DeLoache, J., Pickard, M., Lobue, V. (2011). How very young children think about animals, [W:] *How Animals Affect Us: Examining the Influences of Human Animal Interaction on Child Development and Human Health*, American Psychological Association, Washington, s. 85-99.
112. Dennett, D. (1978b). *Brainstorms: Philosophical Essays on Mind and Psychology*, Bradford Books Cambridge.
113. Dennett, D. C. (1978a). Beliefs about beliefs. *Behavioral and Brain Sciences*, 1(4), s. 568–570. <https://doi.org/10.1017/S0140525X00076664>.

114. Dennett, D. C. (1983). Intentional systems in cognitive ethology: The "Panglossian paradigm" defended. *Behavioral and Brain Sciences*, 6(3), 343–390, <https://doi.org/10.1017/S0140525X00016393>.
115. Dennett, D. C. (1991). *Consciousness explained*. Little, Brown and Company.
116. DeRose, K. (2011). *The case for contextualism: Knowledge, skepticism, and context*, Oxford University Press, Oxford, vol. 1.
117. Devine, P. G. (1989). Stereotypes and prejudice: Their automatic and controlled components. *Journal of Personality and Social Psychology*, 56(1), s. 5-18, <https://doi.org/10.1146/annurev.psych.121208.131628>.
118. Dörrenberg, S., Rakoczy, H., Liszkowski, U. (2021). Do 18-month-olds really attribute mental states to others? A critical test of the anticipatory-looking false belief paradigm. *Royal Society Open Science*, 8(3), <https://doi.org/10.1098/rsos.201958>.
119. Duckitt, J. H. (1992). *The social psychology of prejudice*. New York: Praeger.
120. Dunbar, R. (2013). *Nowa historia ewolucji człowieka*. Copernicus Center Press.
121. Dunn, J., Dale, N. (1984). I a daddy: 2-year-olds' collaboration in joint pretend with sibling and with mother. [W:] *Symbolic play*, Academic Press, s. 131-158.
122. Dudkowska, A. (2023) *Umysły proste, Jose Luis Bermudez i Donald Davidson o racjonalności istot nietuzinkowych*. Wydawnictwo Katolickiego Uniwersytetu Lubelskiego.
123. Ehrlinger, J., Gilovich, T., Ross, L. (2005). Peering into the bias blind spot: People's assessments of bias in themselves and others. *Personality and Social Psychology Bulletin*, 31(5), s. 680-692, <https://doi.org/10.1177/0146167204271570>.
124. Eisenberg, N., Fabes, R. A., Miller, P. A., Fultz, J., Shell, R., Mathy, R. M., Reno, R. R. (1989). Relation of sympathy and personal distress to prosocial behavior: A multimethod study. *Journal of Personality and Social Psychology*, 57(1), s. 55–66. <https://doi.org/10.1037/0022-3514.57.1.55>.
125. Eisenberg, N., Morris, A. S. (2001b). The origins and social significance of empathy-related responding. A review of empathy and moral development: implications for caring and justice by ML Hoffman. *Social Justice Research*, 14(1), s. 95-120, <https://doi.org/10.1023/A:101257980572>.

126. Eisenberg, N., Wentzel, M., Harris, J. D. (1998). The role of emotionality and regulation in empathy-related responding. *School Psychology Review*, 27(4), s. 506-521, <https://doi.org/10.1080/02796015.1998.12085934>.
127. Eisenberg, N., Zhou, Q., Koller, S. (2001a). Brazilian adolescents' prosocial moral judgment and behavior: Relations to sympathy, perspective taking, gender-role orientation, and demographic characteristics. *Child Development*, 72(2), s. 518-534. <https://doi.org/10.1111/1467-8624.00294>.
128. Ekman, P., Friesen, W. V. (1971). Constants across cultures in the face and emotion. *Journal of personality and social psychology*, 17(2), s. 124-129 <https://doi.org/10.1037/h0030377>.
129. Emery, N. J., Clayton, N. S. (2001). Effects of experience and social context on prospective caching strategies by scrub jays. *Nature*, 414(6862), s. 443-446, <https://doi.org/10.1038/35106560>.
130. Entman, R. M., Rojecki, A. (2001). *The Black Image in the White Mind: Media and Race in America*. University of Chicago Press, Chicago.
131. Farrington, B. (1951). Temporis partus masculus an untranslated writing of Francis Bacon. *Centaurus*, 1(3), <https://doi.org/10.1111/j.16000498.1951.tb00507.x>.
132. Farrington, B. (1964). *The Philosophy of Francis Bacon*. Liverpool University Press.
133. Feldman, R. (1999). Contextualism and skepticism. *Philosophical perspectives*, vol. 13, s. 91-114, <https://doi.org/10.1111/0029-4624.33.s13.4>.
134. Festinger, L. (1962). Cognitive dissonance. *Scientific American*, 207(4), s. 93-106.
135. Festinger, L. (2007). *Dysonans poznawczy*. Wydawnictwo Naukowe PWN, Warszawa.
136. Festinger, L. (1955, May). *The relation between cognition and action*. [Artykuł odczytany podczas Symposium on Cognition, Boulder, Colorado].
137. Fiske, S. T. (1994). Social cognition. [W:] A. Tesser (red.), *Advanced social psychology*, NY: McGraw-Hill, Nowy York , s. 149-193.
138. Fiske, S. T. (1998). Stereotyping, prejudice, and discrimination. [W:] D. T. Gilbert, S. T. Fiske, G. Lindzey (Eds.), *The handbook of social psychology*, McGraw-Hill, s. 357-411.
139. Flannery, K., Marcus, J. (2012). *The creation of inequality: how our prehistoric ancestors set the stage for monarchy, slavery, and empire*. Harvard University Press.
140. Floridi, L. (2010). *Information: A very short introduction*. Oxford University Press, vol. 225.

141. Floridi, L. (2011). *The philosophy of information*. Oxford University Press.
142. Fode, K. L. (1960). *The effect of non-visual and non-verbal interaction on experimental bias* (Dysertacja doktorska, University of North Dakota).
143. Foerder, P., Galloway, M., Barthel, T., Moore III, D. E., Reiss, D. (2011). Insightful problem solving in an Asian elephant. *PloS One*, 6(8), <https://doi.org/10.1371/journal.pone.0023251>.
144. Foley, R. (1996). An evolutionary and chronological framework for human social behaviour. [W:] *Proceedings of the British Academy*, Oxford University Press, vol. 88, s. 95–118.
145. Folkman, S., Moskowitz, J. T. (2000). Positive affect and the other side of coping. *American psychologist*, 55(6), s. 647–654, <https://doi.org/10.1037/0003-066X.55.6.647>.
146. Fredrickson, B. L., Levenson, R. W. (1998). Positive emotions speed recovery from cardiovascular sequelae of negative emotion. *Cognition & Emotion*, vol. 12, s. 191–220, <https://doi.org/10.1080/026999398379718>.
147. Fredrickson, B. L., Mancuso, R. A., Branigan, C., Tugade, M. M. (2000). The undoing effect of positive emotions. *Motivation and emotion*, vol. 24, str. 237-258, <https://doi.org/10.1023/A:1010796329158>.
148. Fricker, M. (2007). *Epistemic injustice: Power and the ethics of knowing*. Oxford University Press.
149. Fudge, E. (2002). *Animal*. Reaktion Books, Londyn.
150. Gallese, V., Goldman, A. (1998). Mirror neurons and the simulation theory of mind-reading. *Trends in cognitive sciences*, 2(12), s. 493-501, [https://doi.org/10.1016/S1364-6613\(98\)01262-5](https://doi.org/10.1016/S1364-6613(98)01262-5).
151. Gallistel, C. R., King, A. P. (2011). *Memory and the computational brain: Why cognitive science will transform neuroscience*. John Wiley & Sons.
152. Gallup, G. G. (1998). Self-awareness and the evolution of social intelligence. *Behavioral Processes*, 42(2-3), s. 239-247. [https://doi.org/10.1016/S0376-6357\(97\)00079-X](https://doi.org/10.1016/S0376-6357(97)00079-X).
153. Garcia-Marques, L., Mackie, D. M. (1999). The impact of stereotype-incongruent information on perceived group variability and stereotype change. *Journal of Personality and social Psychology*, 77(5). <https://doi.org/10.1037//0022-3514.77.5.979>

154. Gaunt, R., Leyens, J. P., Sindic, D. (2004). Motivated reasoning and the attribution of emotions to ingroup and outgroup. *International Review of Social Psychology*, vol. 17, s. 5–20.
155. Gentner, T. Q., Fenn, K. M., Margoliash, D., Nusbaum, H. C. (2006). Recursive syntactic pattern learning by songbirds. *Nature*, 440(7088), s. 1204-1207. <https://doi.org/10.1038/nature04675>.
156. Gerbner, G. (1998). Cultivation analysis: An overview. *Mass Communication and Society*, 1(3-4), s. 175-194, <https://doi.org/10.1080/15205436.1998.9677855>.
157. Gerbner, G., Gross, L., Morgan, M., Signorielli, N. (1986). Living with television: The dynamics of the cultivation process. [W:] J., Bryant, D. Zillmann (red.), *Perspectives on media effects*, Hillsdale, NJ: Lawrence Erlbaum Associates, s. 17-40.
158. Gettier, E. (1963). Is Justified True Belief Knowledge?. *Analysis*, vol. 23, s. 121–123. <https://doi.org/10.1093/analys/23.6.121>.
159. Gigerenzer, G., Brighton, H. (2009). Homo heuristics: Why biased minds make better inferences. *Topics in cognitive science*, 1(1), s. 107-143, <https://doi.org/10.1111/j.1756-8765.2008.01006.x>.
160. Gilbert, C. D., Li, W. (2013). Top-down influences on visual processing. *Nature Reviews Neuroscience*, 14(5), s. 350–363. <https://doi.org/10.1038/nrn3476>.
161. Gilovich, T. (1991). *How we know what isn't so: The fallibility of human reason in everyday life*. NY: Free Press, Nowy York.
162. Gintis, H. (2011). Gene–culture coevolution and the nature of human sociality. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 366(1566), s. 878-888, <https://doi.org/10.1098/rstb.2010.0310>.
163. Godden, D. R., Baddeley, A. D. (1975). Context-dependent memory in two natural environments: On land and underwater. *British Journal of Psychology*, 66(3), str. 325–331, <https://doi.org/10.1111/j.2044-8295.1975.tb01468.x>
164. Goldman, A. I. (1979). What is justified belief?. [W:] *Justification and knowledge: New studies in epistemology*, Dordrecht: Springer Netherlands, s. 1-23.
165. Goldman, A. I. (2006). *Simulating minds: The philosophy, psychology, and neuroscience of mindreading*. Oxford University Press.
166. Gould, S. J. (2010). *The panda's thumb: More reflections in natural history*. WW Norton & Company.

167. Grandgeorge, M., Bourreau, Y., Alavi, Z., Lemonnier, E., Tordjman, S., Deleau, M., i in. (2014). Interest towards human, animal and object in children with autism spectrum disorders: an ethological approach at home. *European Child&Adolescent*, 24(1), s. 83-93, *Psychiatry*, <https://doi.org/10.1007/s00787-014-0528-9>.
168. Green, M. C., Brock, T. C. (2000). The role of transportation in the persuasiveness of public narratives. *Journal of Personality and Social Psychology*, 79(5), s. 701-721, <https://doi.org/10.1037/0022-3514.79.5.701>.
169. Greenwald, A. G., Banaji, M. R. (1995). Implicit social cognition: Attitudes, self-esteem, and stereotypes. *Psychological Review*, 102(1), s. 4–27. <https://doi.org/10.1037/0033-295X.102.1.4>.
170. Grether, D. M. (1992). Testing Bayes rule and the representativeness heuristic: Some experimental evidence. *Journal of Economic Behavior & Organization*, 17(1), s. 31-57, [https://doi.org/10.1016/0167-2681\(92\)90078-P](https://doi.org/10.1016/0167-2681(92)90078-P).
171. Griffin, D. R. (1944). Echo orientation and the nature of the “target” in the bat, with comments on the value of the term “target” in sensory physiology. *Journal of Experimental Zoology*, 98(1), s. 25-46, <https://doi.org/10.1002/jez.1400980103>.
172. Griffin, D. R. (1958). *Listening in the dark: The acoustic orientation of bats and men*. Yale University Press, USA.
173. Griffin, D. R. (2001). *Animal minds: Beyond cognition to consciousness*. University of Chicago Press, Chicago.
174. Griggs, R. A., Cox, J. R. (1982). The elusive thematic-materials effect in Wason's selection task. *British journal of psychology*, 73(3), s. 407-420, <https://doi.org/10.1111/j.2044-8295.1982.tb01823.x>.
175. Gross, J. J. (2013). Emotion regulation: Taking stock and moving forward. *Emotion*, 13(3), s. 359–365. <https://doi.org/10.1037/a0032135>.
176. Gross, J. J., John, O. P., Richards, J. M. (2000). The dissociation of emotion expression from emotion experience: A personality perspective. *Personality and Social Psychology Bulletin*, 26(6), 712-726, <https://doi.org/10.1177/0146167200268006>.
177. Guinote, A. (2017). How power affects people: Activating, wanting, and goal seeking. *Annual review of psychology*, 68(1), s. 353-381,

- <https://doi.org/10.1146/annurev-psych-010416-044153>.
178. Guthrie, E. R., Horton, G. P. (1946). *Cats in a puzzle box*. Rinehart.
179. Hadwin, J., Perner, J. (1991). Pleased and surprised: Children's cognitive theory of emotion. *British Journal of Developmental Psychology*, 9(2), s. 215–234.
<https://doi.org/10.1111/j.2044-835X.1991.tb00872.x>
180. Haggag, K., Paci, G. (2014). Default tips. *American Economic Journal: Applied Economics*, 6(3), s. 1-19, <https://doi.org/10.1257/app.6.3.1>.
181. Haidt, J. (2012). The righteous mind: Why good people are divided by politics and religion. *New York Pantheon*.
182. Hainline, L. (1978). Developmental changes in visual scanning of faces and forms. *Journal of Experimental Child Psychology*, 25(1), s. 90–115, [https://doi.org/10.1016/0022-0965\(78\)90069-2](https://doi.org/10.1016/0022-0965(78)90069-2).
183. Haith, M. M., Bergman, T., Moore, M. J. (1977). Eye contact and face scanning in early infancy. *Science*, 198(4319), s. 853–855. <https://doi.org/10.1126/science.918670>.
184. Hare B., Call J., Agnetta B., Tomasello M. (2000). Chimpanzees know what conspecifics do and do not see. *Animal Behavior*, vol. 59, s. 771–785. <https://doi.org/10.1006/anbe.1999.1377>.
185. Hare, B., Call, J., Tomasello, M. (2001). Do chimpanzees know what conspecifics know? *Animal Behaviour*, 61(1), s. 139–151. <https://doi.org/10.1006/anbe.2000.1518>.
186. Hare, B., Call, J., Tomasello, M. (2001). Do chimpanzees know what conspecifics know? *Animal Behaviour*, 61(1), s. 139-151, <https://doi.org/10.1006/anbe.2000.1518>.
187. Hare, B., Woods, V. (2013). *The genius of dogs: How dogs are smarter than you think*. Penguin.
188. Harkness, S., Super, C. M. (1994). The developmental niche: A theoretical framework for understanding the embeddedness of culture in the development of the child. *International Journal of Behavioral Development*, 17(2), s. 199–232, <https://doi.org/10.1177/016502549401700204>.
189. Harman, G. (1978). Studying the chimpanzee's theory of mind. *Behavioral and Brain Sciences*, 1(4), s. 576–577, <https://doi.org/10.1017/S0140525X00076743>.
190. Harris, J. R. (1995). Where is the child's environment? A group socialization theory of development. *Psychological Review*, 102(3), s. 458–489.
<https://doi.org/10.1037/0033-295X.102.3.458>.

191. Haslam, N., Bain, P., Douge, L., Lee, M., Bastian, B. (2005). More human than you: attributing humanness to self and others. *Journal of personality and social psychology*, 89(6), <https://doi.org/10.1037/0022-3514.89.6.937>.
192. Hayashi, M., Matsuzawa, T. (2003). Cognitive development in object manipulation by infant chimpanzees. *Animal Cognition*, 6(4), s. 225–233, <https://doi.org/10.1007/s10071-003-0185-8>.
193. Heider, F., Simmel, M. (1944). An Experimental Study of Apparent Behavior. *The American Journal of Psychology*, 57, s. 243-259, [https://doi.org/10.2307/141695057\(2\)](https://doi.org/10.2307/141695057(2)), 243–259.
194. Hempel, C. G. (1991). Introduction to problem of taxonomy. [W:] J. Zubin (red.), *Field studies in the mental disorders*. New York: Grune & Stratton, s. 3–22.
195. Herman, L. M. (2002). Vocal, social, and self-imitation by bottlenosed dolphins. [W:] *Imitation in animals and artifacts*. The MIT Press, str. 63-108, <https://doi.org/10.7551/mitpress/3676.003.0004>
196. Herman, L. M. (2006). Intelligence and rational behaviour in the bottlenosed dolphin. [W:] S. Hurley M. Nudds (red.), *Rational animals?* Oxford University Press, s. 439 – 467, <https://doi.org/10.1093/acprof:oso/9780198528272.003.0020>.
197. Herzog, H., Foster, M. (2010). *Some we love, some we hate, some we eat*. Hyper.
198. Heyes, C. (2008) Beast Machines? Questions of animal consciousness. [W:] L. Weiskrantz, M. Davies, *Frontiers of consciousness*. Oxford University Press, Nowy York, s. 259-274.
199. Heyes, C. (2014). Submentalizing: I am not really reading your mind. *Perspectives on Psychological Science*, 9(2), s. 131–143. <https://doi.org/10.1177/1745691613518076>.
200. Heyes, C. M., Frith, C. D. (2014). The cultural evolution of mind reading. *Science*, 344(6190), <https://doi.org/10.1126/science.1243091>.
201. Hoffman, M. L. (1996). *Empathy and moral development. Implications for Caring and Justice*, Cambridge University Press, vol. 35, s. 157-162.
202. Holroyd, J., Puddifoot, K. (2019). Implicit bias and prejudice. [W:] *The Routledge Handbook of Social Epistemology*, Routledge, s. 313-326.
203. Holton, G. (1973). *Thematic Origins of Scientific Thought*, Harvard University Press.
204. Hood, L., Bloom, L., Brainerd, C. J. (1979). *What, when, and how about why: A longitudinal study of early expressions of causality*. Monographs of the society for research in child development, s. 1-47.

205. Hopper, L. M., Kurtycz, L. M., Ross, S. R., Bonnie, K. E. (2015). Captive chimpanzee foraging in a social setting: a test of problem solving, flexibility, and spatial discounting. *PeerJ*, <https://doi.org/10.7717/peerj.833>.
206. Horn, L., Huber, L., and Range, F. (2013). The importance of the secure base effect for domestic dogs - evidence from a manipulative problem-solving task. *PLoS ONE*. doi: 10.1371/journal.pone.0065296.
207. Hoshikawa, Y., Yamamoto, Y. (1997). Effects of Stroop color-word conflict test on the autonomic nervous system responses. *American journal of physiology-Heart and circulatory physiology*, 272(3), <https://doi.org/10.1152/ajpheart.1997.272.3.H1113>.
208. Huges, S. (2016). *American Monsters: Tabloid Media and the Satanic Panic. 1970–2000*, *Journal of American Studies*, 51(3), s. 691-719, <https://doi.org/10.1017/S0021875816001298>.
209. Hume, D. (2005). *Badania dotyczące zasad moralności*. Zielona Sowa.
210. Inoue, S., Matsuzawa, T. (2007). Working memory of numerals in chimpanzees. *Current Biology*, 17(23), <https://doi.org/10.1016/j.cub.2007.10.027>.
211. Jaynes, J. (1970). The problem of animate motion in the seventeenth century. *Journal of the History of Ideas*, 31(2), s. 219-234.
212. Johnson, C. N., Maratsos, M. P. (1977). Early comprehension of mental verbs: Think and know. *Child Development*, s. 1743-1747, <https://doi.org/10.2307/1128549>.
213. Johnson, C. Wellman, M. (1980). Children's developing understanding of mental verbs: Remember, know and guess. *Child Development*, vol. 51, s. 1095-1102, <https://doi.org/10.2307/1129549>.
214. Johnson, M. H., Dziurawiec, S., Ellis, H., Morton, J. (1991). Newborns' preferential tracking of face-like stimuli and its subsequent decline. *Cognition*, 40(1-2), s. 1-19, [https://doi.org/10.1016/0010-0277\(91\)90045-6](https://doi.org/10.1016/0010-0277(91)90045-6).
215. Jones, E. E., Kohler, R. (1958). The effects of plausibility on the learning of controversial statements. *The Journal of Abnormal and Social Psychology*, 57(3), s. 315–320, <https://doi.org/10.1037/h0049162>.
216. Joy, M. (2017). Karnizm: Dlaczego zjadanie zwierząt to kwestia sprawiedliwości społecznej [W:] B., Błońska, W., Gogłóza, W., Klaus, Woźniakowska-Fajst (red.) *Sprawiedliwość dla zwierząt*. Springer Publishing Company, Nowy York.
217. Joy, M. (2020). *Why we love dogs, eat pigs, and wear cows: An introduction to carnism*. Red Wheel.

218. Kagan, J. (2004). The uniquely human in human nature. *Daedalus*, 133(4), s. 77-88, <https://doi.org/10.1162/0011526042365609>.
219. Kahneman, D. (2012). *Pułapki myślenia: O myśleniu szybkim i wolnym*, Media Rodzina, Poznań.
220. Kahneman, D., Tversky, A. (1974). Judgment under uncertainty: Heuristics and biases. *Science*, 185(4157), s. 1124-1131, <https://doi.org/10.1126/science.185.4157.1124>.
221. Kaminski, J., Call, J., Tomasello, M. (2008). Chimpanzees know what others know, but not what they believe. *Cognition*, 109(2), s. 224–234, <https://doi.org/10.1016/j.cognition.2008.08.010>.
222. Kano, F., Call, J. (2014). Cross-species variation in looking for meaning: Eye-tracking great apes' and humans' anticipatory looks in a false-belief task. *Animal Cognition*, 17(6), s. 1245–1256. <https://doi.org/10.1007/s10071-014-0753-0>.
223. Kano, F., Krupenye, C., Hirata, S., Call, J. (2017). Eye tracking uncovered great apes' ability to anticipate that other individuals will act according to false beliefs. *Communicative & Integrative Biology*, 10(2), <https://doi.org/10.1080/19420889.2017.1299836>.
224. Kano, F., Krupenye, C., Hirata, S., Tomonaga, M., Call, J. (2019). Great apes use self-experience to anticipate an agent's action in a false-belief test. *Proceedings of the National Academy of Sciences*, 116(42), s. 20904-20909, <https://doi.org/10.1073/pnas.191009511>.
225. Kanouse, D. E. (1971). *Language, labeling, and attribution*. General Learning Press.
226. Kappas, A. (2011). Emotion and regulation are one!. *Emotion Review*, 3(1), s. 17-25, <https://doi.org/10.1177/1754073910380971>.
227. Kaproń-Charzyńska, I. (2012). Pojęcie mądrości w kontekście edukacji aksjologicznej. *Paedagogia Christiana*, 1(29), s. 109–128.
228. Karg, K., Schmelz, M., Call, J., Tomasello, M. (2015). The goggles experiment: Can chimpanzees use self-experience to infer what a competitor can see? *Animal Behaviour*, vol. 105, s. 211–221, <https://doi.org/10.1016/j.anbehav.2015.04.028>.
229. Kawai, N., Matsuzawa, T. (2000). Numerical memory span in a chimpanzee. *Nature*, 403(6765), s. 39–40, <https://doi.org/10.1038/47405>.

230. Kennedy, J. S. (1992). *The new anthropomorphism*. Cambridge University Press, Cambridge.
231. King, B. J. (2018). *Osobowość na talerzu: kim są zwierzęta, które zjadamy?*. Wydawnictwo WAB-Grupa Wydawnicza Foksal.
232. King, S. L., Janik, V. M. (2013). Bottlenose dolphins can use learned vocal labels to address each other. *Proceedings of the National Academy of Sciences*, 110(32), s. 13216-13221. <https://doi.org/10.1073/pnas.1304459110>.
233. Klayman, J. (1995). Varieties of confirmation bias. [W:] G. H. Bower (red.), *Psychology of Learning and Motivation*, s. 385–418, [https://doi.org/10.1016/s0079-7421\(08\)60315-1](https://doi.org/10.1016/s0079-7421(08)60315-1).
234. Knox, R. E., Inkster, J. A. (1968). Postdecision dissonance at post time. *Journal of personality and social psychology*, 8(4p1), <https://doi.org/10.1037/h0025528>.
235. Kohlberg, L. (1981). *Essays on Moral Development: Vol. I. The Philosophy of Moral Development*. Harper & Row, San Francisco.
236. Köhler, W. (1949). Problem Solving by Chimpanzees. [W:] W. Dennis (red.), *Readings in general psychology*,. Prentice-Hall, Inc, s. 306–314, <https://doi.org/10.1037/11352-044>.
237. Kokocińska, A., Kaleta, T. (2015). Behawioryzm i behavior-myśl filozoficzna i badania przyrodnicze. *Kosmos*, 64(2), s. 221-227.
238. Kossak, S. (2019). *O zwierzętach i przyrodzie*. Fundacja Sąsiedzi.
239. Kosslyn S.M., Rosenberg R.S., przekł. B. Majczyzna i in. (2006), *Psychologia: Mózg, człowiek, świat*. Wydawnictwo Znak, Kraków.
240. Kravchenko, A. V. (2020). Chapter five Language and Mind: a biocognitive view, [W:] *Memory, Mind and Language*, Cambridge Scholars Publishing.
241. Kruger, J., Dunning, D. (1999). Unskilled and unaware of it: how difficulties in recognizing one's own incompetence lead to inflated self-assessments. *Journal of personality and social psychology*, 77(6), doi: <https://doi.org/10.1037/0022-3514.77.6.1121>.
242. Krupenye, C., Call, J. (2019). Theory of mind in animals: Current and future directions. *Wiley Interdisciplinary Reviews: Cognitive Science*, 10(6), <https://doi.org/10.1002/wcs.1503>.

243. Krupenye, C., Kano, F., Hirata, S., Call, J., Tomasello, M. (2016). Great apes anticipate that other individuals will act according to false beliefs. *Science*, 354(6308), s. 110–114, <https://doi.org/10.1126/science.aaf8110>.
244. Krupenye, C., Kano, F., Hirata, S., Call, J., Tomasello, M. (2017). A test of the submentalizing hypothesis: Apes' performance in a false belief task inanimate control. *Communicative & Integrative Biology*, 10(4), <https://doi.org/10.1080/19420889.2017.1343771>.
245. Kunreuther, H., Heal, G. (2012). *Managing catastrophic risk*. National Bureau of Economic Research.
246. Kurek, A., Król, R. (2016). Krzywa uczenia w kontekście praktyki chirurgicznej. *Chirurgia Polska*, 18(1-2), s. 16-23.
247. Lakoff, G., Johnson, M. (2008). *Metaphors we live by*. University of Chicago press.
248. Lam, C., Kitamura, C. (2012). Mommy, speak clearly: Induced hearing loss shapes vowel hyperarticulation. *Developmental Science*, 15(2), s. 212–221, <https://doi.org/10.1111/j.1467-7687.2011.01118.x>.
249. Lamarck, J. B. (1994). *Philosophie zoologique*. Paris: Flammarion.
250. Lehrner, A., Yehuda, R. (2018). Cultural trauma and epigenetic inheritance. *Development and psychopathology*, 30(5), s. 1763-1777.
251. Leslie, A. M. (1993). *A theory of agency*. Rutgers University Center for Cognitive Science.
252. Leyens, J. P., Cortes, B. P., Demoulin, S., Dovidio, J. F., Fiske, S. T., Gaunt, R., i. in. (2003). Emotional prejudice, essentialism, and nationalism. *European Journal of Social Psychology*, vol. 33, s. 704 –717, <https://doi.org/10.1002/ejsp.170>.
253. Leyens, J. P., Paladino, P. M., Rodriguez-Torres, R., Vaes, J., Demoulin, S., Rodriguez Perez, A., Gaunt, R. (2000). The emotional side of prejudice: The attribution of secondary emotions to ingroups and outgroups. *Personality and Social Psychology Review*, 4(2), s. 186 –197, https://doi.org/10.1207/S15327957PSPR0402_0.
254. Lima, S. L. (1985). Maximizing feeding efficiency and minimizing time exposed to predators: A trade-off in the black-capped chickadee. *Oecologia*, 66(1), s. 60–67, <https://doi.org/10.1007/BF00378552>.

255. Little, A., (2012). Manipulation of infant-like traits affects perceived cuteness of infant, adult and cat faces. *Ethology*, vol. 118, s. 775–782, doi: 10.1111/j.1439-0310.2012.02068.x.
256. Littlewood, R. J. (2006). *A commentary on Ovid's Fasti, Book 6*. Oxford University Press, USA.
257. Liu, H., Du, S. (2016). Can an overconfident insider coexist with a representativeness heuristic insider?. *Economic Modelling*, vol. 54, s. 170-177, <https://doi.org/10.1016/j.econmod.2015.12.032>.
258. Lobue, V., Bloom-Pickard, M., Sherman, K., Axford, C., Deloache, J. S. (2012). Young children's interest in live animals. *British Journal of Developmental Psychology*, 31(1), s. 57-69, <https://doi.org/10.1111/j.2044-835X.2012.02078.x>.
259. Lockwood, R. (1985). Anthropomorphism is not a four-letter word. Advances in animal welfare science, *Humane Society of the United States*, 1985(86), Washington.
260. Loftus, E. F., Palmer, J. C. (1974). Reconstruction of automobile destruction: An example of the interaction between language and memory. *Journal of Verbal Learning & Verbal Behavior*, 13(5), s. 585–589. [https://doi.org/10.1016/S0022-5371\(74\)80011-3](https://doi.org/10.1016/S0022-5371(74)80011-3).
261. Lord, C. G., Ross, L., Lepper, M. R. (1979). Biased assimilation and attitude polarization: The effects of prior theories on subsequently considered evidence. *Journal of personality and social psychology*, 37(11), <https://doi.org/10.1037/0022-3514.37.11.2098>.
262. Lorenz, K. (1935). The companion in the bird's world. A contribution to the study of the origin of social relations in animals. *Journal of Ornithology*, vol. 83, s. 137–213.
263. Lorenz, K. (1943). Die angeborenen formen möglicher erfahrung. *Zeitschrift für Tierpsychologie*, 5(2), s. 235-409, <https://doi.org/10.1111/j.1439-0310.1943.tb00655.x>.
264. Lorenz, K. (1981). *The foundations of ethology*. Springer verlag.
265. Lurz, R. W. (2011). *Mindreading animals: The debate over what animals know about other minds*. Boston Review, <https://doi.org/10.7551/mitpress/9780262016056.001.0001>.
266. Lutts, R. H. (1992). The trouble with Bambi: Walt Disney's Bambi and the American vision of nature. *Forest and Conservation History*, 36(4), s. 160-171.

267. Łukasiewicz, N. (2010). *Prawa zwierząt w kontekście społeczeństwa obywatelskiego i przeobrażeń cywilizacji*. [Nieopublikowana dysertacja magisterska] Uniwersytet Łódzki.
268. Łukaszewski, W. (1999). *Stereotypy i uprzedzenia u dzieci. Geneza, diagnoza i modyfikacja*. Wydawnictwo Instytutu Psychologii PAN, . Warszawa.
269. Łyczkowska, K. Szarzyńska, K. (1981). *Mitologia Mezopotamii*. Wydawnictwa Artystyczne i Filmowe, Warszawa.
270. MacLeod, R. B. (1959). Cumulative Record. BF Skinner. Appleton-Century-Crofts, *Science*, 130(3366), Nowy York, s. 34-35.
271. Macnamara, J., Baker, E. and Olson, C. (1976) Four-year-olds' understanding of pretend, forget, and know: Evidence for propositional operations. *Child Development*, vol. 47, s. 62-70, <https://doi.org/10.2307/1128283>.
272. Maier, N. R. (1956). Frustration theory: Restatement and extension. *Psychological review*, 63(6), s. 370–388, <https://doi.org/10.1037/h0046970>.
273. Maier, N. R. F. (1963). *Frustration: The Study of Behavior Without a Goal*. University of Michigan Press.
274. Malone J. C. (1975). The “paradigms” of learning. *The Psychological Record*, 25(4), s. 479-489, <https://doi.org/10.1007/BF03394340>.
275. Margul, T. (1986). *Religia, a przestrzeń i krajobraz*. Uniwersytet Jagielloński, Kraków, s. 140-142.
276. Margul, T. (1996). *Zwierzę w kulcie i micie*. Wydawnictwo Uniwersytetu Marii.
277. Marr, D., Vaina, L. (1982). Representation and recognition of the movements of shapes. *Proceedings of the Royal Society of London. Series B. Biological Sciences*, 214(1197), s. 501-524, <https://doi.org/10.1098/rspb.1982.0024>.
278. Martcorena, D. C., Ruiz, A. M., Mukerji, C., Goddu, A., Santos, L. R. (2011). Monkeys represent others' knowledge but not their beliefs. *Developmental Science*, 14(6), s. 1406–1416, <https://doi.org/10.1111/j.1467-7687.2011.01085.x>.
279. Martin, A., Santos, L. R. (2014). The origins of belief representation: Monkeys fail to automatically represent belief states. *Cognition*, 130(3), s. 300–308. <https://doi.org/10.1016/j.cognition.2013.11.016>.
280. Marvin, R. S., Greenberg, M. T. and Mossier, D. G. (1976) The early development of conceptual perspective taking: Distinguishing among multiple perspectives. *Child Development*, vol. 47, <https://doi.org/10.2307/1128810>.

281. Mason, W. A., Hollis, J. H. (1962). Communication between young rhesus monkeys. *Animal Behaviour*, 10, s. 211-221, [https://doi.org/10.1016/0003-3472\(62\)90040-4](https://doi.org/10.1016/0003-3472(62)90040-4).
282. Mastro, D., Tropp, L. R. (2004). The effects of interracial contact, media exposure, and ingroup/outgroup identification on perceptions of interracial relationships. *Journal of Social and Personal Relationships*, 21(5), s. 709-730, <https://doi.org/10.1080/08824090409359974>.
283. Matlin, M. F. (1994). *Cognition*. Harcourt Brace Publishers.
284. Maturana, H. R., Varela, F. J. (1991). *Autopoiesis and cognition: The realization of the living*, Springer Science & Business Media.
285. Maturana, H., Varela, F. (1972). *Autopoeisis and Cognition*. Reidel Publishing Company, Boston.
286. Mayo, E. (1945). *The Social Problems of An Industrial Civilization*. Harvard University Graduate School of Business Administration, Boston.
287. Mayr, E. (1961). Cause and effect in biology: kinds of causes, predictability, and teleology are viewed by a practicing biologist. *Science*, 134(3489), s. 1501-1506, <https://doi.org/10.1126/science.134.3489.1501>.
288. McFarland, D. (1981). *The Oxford companion to animal behaviour*, University Press: Oxford, Oxford.
289. McGinn, C. (1999). Our Duties to Animals and the Poor. [W:] W. D. Jamieson (red.), *Singer and His Critics*. Malden, MA: Blackwell Publishers, s. 150–161.
290. Meijer, E. (2021). *Zwierzęta mówią. W stronę demokracji międzygatunkowej*. Wydawnictwo Drzazgi, Okoniny.
291. Messier, F., Crete, M. (1985). Moose–wolf dynamics and the natural regulation of moose populations. *Oecologia*, 65(4), s. 503–512,
292. Melis, A. P., Call, J., Tomasello, M. (2006). Chimpanzees (*Pan troglodytes*) conceal visual and auditory information from others. *Journal of Comparative Psychology*, 120(2), 154–162. <https://doi.org/10.1037/0735-7036.120.2.154>.
293. Melis, A. P., Hare, B., Tomasello, M. (2006). Engineering cooperation in chimpanzees: Tolerance constraints on cooperation. *Animal Behaviour*, 72(2), s. 275–286, <https://doi.org/10.1016/j.anbehav.2005.09.018>.
294. Melis, A. P., Tomasello, M. (2019). Chimpanzees (*Pan troglodytes*) coordinate by communicating in a collaborative problem-solving task. *Proceedings of the Royal Society B*. <https://doi.org/10.1098/rspb.2019.0408>.

295. Meltzoff, A. N., Brooks, R. (2008). Self-experience as a mechanism for learning about others: A training study in social cognition. *Developmental Psychology*, 44(5), s. 1257–1265. <https://doi.org/10.1037/a0012888>.
296. Merritt, A. C., Effron, D. A., Monin, B. (2010). Moral self-licensing: When being good frees us to be bad. *Social and personality psychology compass*, 4(5), s. 344–357, <https://doi.org/10.1111/j.1751-9004.2010.00263.x>.
297. Merton, R. K. (2005). Samospełniające się proroctwo [W:] P. Sztompka, M. Kucia (red.) *Socjologia. Lektury* Wydawnictwo Znak, Kraków.
298. Mezulis, A. H., Abramson, L. Y., Hyde, J. S., Hankin, B. L. (2004). Is there a universal positivity bias in attributions? A meta-analytic review of individual, developmental, and cultural differences in the self-serving attributional bias. *Psychological bulletin*, 130(5), <https://doi.org/10.1037/0033-2909.130.5.711>.
299. Mękarska, J. (2024). Błędy poznawcze związane z szowinizmem gatunkowym oraz odmawianie teorii umysłu zwierzętom innym niż ludzie. *Analiza i Egzystencja*, 67 (3), <https://doi.org/10.18276/aie.2024.67-05>.
300. Minsky, M. (2007). *The emotion machine: Commonsense thinking, artificial intelligence, and the future of the human mind*. Simon and Schuster.
301. Minsky, M. (2007). *The emotion machine: Commonsense thinking, artificial intelligence, and the future of the human mind*. Simon&Schuster, Nowy York.
302. Mirosławska, M., Kofta, M. (2007). Zjawisko infrahumanizowania „obcych”: wstępny test hipotezy generalizacji Ja. *Psychologia Społeczna*, 2(1).
303. Miscione, J. L., Marvin, R. S., O'Brien, R. G. and Greenberg, M. T. (1978) A developmental study of preschool children's understanding of the words 'know' and 'guess'. *Child Development*, vol. 49, s. 1107-1113, <https://doi.org/10.2307/1128750>.
304. Mitchell, R. W., Thompson, N. S., Miles, H. L. (1997). *Anthropomorphism, anecdotes, and animals*. Suny Press.
305. Mithen, S. J. (1998). *The prehistory of the mind: A search for the origins of art, religion and science*. Thames & Hudson, London.
306. Moeran, B., Skov, L. (2013). *Women, media and consumption in Japan*. Routledge.
307. Monin, B., Miller, D. T. (2001). Moral credentials and the expression of prejudice. *Journal of Personality and Social Psychology*, 81(1), s. 33–43, <https://doi.org/10.1037/0022-3514.81.1.33>.

308. Moore, B. R., Stuttard, S. (1979). Dr. Guthrie and *Felis domesticus* : or, Tripping over the cat. *Science*, 205(4410), s. 1031–1033, <https://doi.org/10.1126/science.572990>.
309. Morgan, C. L. (1894). *An introduction to comparative psychology*. Walter Scott Publishing Compamy, <https://doi.org/10.1037/11344-000>.
310. Morris, P. H., Doe, C., Godsell, E. (2008). Secondary emotions in non-primate species? Behavioural reports and subjective claims by animal owners. *Cognition and emotion*, 22(1), s. 3-20, <https://doi.org/10.1080/02699930701273716>.
311. Mossler, D. G., Marvin, R. S., Greenberg, M. T. (1976). Conceptual perspective taking in 2- to 6-year-old children. *Developmental Psychology*, 12(1), s. 85–86, <https://doi.org/10.1037/0012-1649.12.1.85>.
312. Myserud, I. (1999). Review of *Unto Others: The Evolution and Psychology of Unselfish Behavior*. [W:] E. Sober, D. S. Wilson. *Politics and the Life Sciences*, 18(1), s. 145–149.
313. Nagasawa, M., Mitsui, S., En, S., Ohtani, N., Ohta, M., Sakuma, Y., Kikusui, T. (2015). Oxytocin-gaze positive loop and the coevolution of human-dog bonds. *Science*, 348(6232), s. 333-336. <https://doi.org/10.1126/science.1261022>.
314. Nagel T. (1997). Jak to jest być nietoperzem? [W:] T. Nagel. *Pytania ostateczne*, Fundacja Aletheia, Warszawa, s. 204–207.
315. Nakamura, J., Csikszentmihalyi, M. (2009). Flow theory and research. Handbook of positive psychology. [W:] Snyder, C. R., Lopez, S. J., Edwards, L. M., Marques, S. C., *The Oxford handbook of positive psychology*. Oxford University Press, Nowy York.
316. Napiórkowski, M. (2019). *Turbopatriotyzm*. Wydawnictwo Czarne.
317. Napiórkowski, M. (2022). *Naprawić przyszłość: Dlaczego potrzebujemy lepszych opowieści, żeby uratować świat*. Wydawnictwo Literackie.
318. Neisser, U. (1976). *Cognition and reality: Principles and implications of cognitive psychology*. Holt & Company.
319. Nelson, T. D. (2003). *Psychologia uprzedzeń*. Gdańskie Wydawnictwo Psychologiczne.
320. Nęcka, E., Orzechowski, J., Szymura, B., Wichary, S. (2013). *Psychologia poznawcza*. Wydawnictwo Naukowe PWN, Warszawa.
321. Nichols, K. A., Champness, B. G. (1971). Eye gaze and the GSR. *Journal of Experimental Social Psychology*, 7(6), s. 623–626. [https://doi.org/10.1016/0022-1031\(71\)90031-5](https://doi.org/10.1016/0022-1031(71)90031-5).

322. Nickerson, R. S. (1998). Confirmation bias: A ubiquitous phenomenon in many guises. *Review of general psychology*, 2(2), s. 175-220, <https://doi.org/10.1037/1089-2680.2.2.175>.
323. Nigam, T., Schwiedrzik, C. M. (2024). Predictions enable top-down pattern separation in the macaque face-processing hierarchy. *Nature Communications*, vol. 15. <https://doi.org/10.1038/s41467-024-51543-y>.
324. Nittono, H., Fukushima, M., Yano, A., Moriya, H. (2012). The power of kawaii: Viewing cute images promotes a careful behavior and narrows attentional focus. *PLoS One*, 7(9), <https://doi.org/10.1371/journal.pone.0046362>.
325. Nowogrodzka, A. (2014). Poznanie społeczne—czynniki wpływające na adekwatne funkcjonowanie człowieka w społeczeństwie oraz zaburzenia wynikające z deficytów tej funkcji. *Neuropsychiatry & Neuropsychology/Neuropsychiatria*.
326. Nozick, R. (1981). *Philosophical explanations*. Harvard University Press
327. Nshom, E., Croucher, S. (Red.). (2024). *Research Handbook on Communication and Prejudice*. Edward Elgar Publishing.
328. Onishi, K. H., Baillargeon, R. (2005). Do 15-month-old infants understand false beliefs? *Science*, 308(5719), s. 255–258, <https://doi.org/10.1126/science.1107621>.
329. O'Rourke, C. T., Hall, M. I., Pitlik, T., Fernández-Juricic, E. (2010). Hawk eyes I: diurnal raptors differ in visual fields and degree of eye movement. *PLoS One*, 5(9), e12802. <https://doi.org/10.1371/journal.pone.0012802>.
330. Palagi, E., Nicotra, V., Cordoni, G. (2015). Rapid mimicry and emotional contagion in domestic dogs. *Royal Society open science*, 2(12).
331. Palmer, C. (2010). *Animal ethics in context*. Columbia University Press.
332. Papież Jan Paweł II. (1991). *Sollicitudo Rei Socialis* [Encyklika]. Sekcja 34. Warszawa: Wydawnictwo Pax.
333. Pardo, M. A., Fristrup, K., Lolchuragi, D. S., Poole, J. H., Granli, P., Moss, C., Wittemyer, G. (2024). African elephants address one another with individually specific name-like calls. *Nature Ecology & Evolution*, s. 1-12, <https://doi.org/10.1038/s41559-024-02420-w>.
334. Pawlow, I. P. (1923). New researches on conditioned reflexes. *Science*, 58(1506), s. 359-361.

335. Payne, B. K. (2001). Prejudice and perception: The role of automatic and controlled processes in misperceiving a weapon. *Journal of Personality and Social Psychology*, 81(2), s. 181–192. <https://doi.org/10.1037/0022-3514.81.2.181>.
336. Pearce, J. M., Hall, G. (1980). A model for Pavlovian learning: Variations in the effectiveness of conditioned but not of unconditioned stimuli. *Psychological Review*, 87(6), s. 532–552. <https://doi.org/10.1037/0033-295X.87.6.532>.
337. Pearce, J. M., Kaye, H., Hall, G. (1982). Predictive accuracy and stimulus associability: Development of a model for Pavlovian learning. [W:] M. L. Commons, R. J. Herrnstein, A. R. Wagner (red.), *Quantitative analyses of behavior*, MA: Ballinger, Cambridge, s. 241–255.
338. Pearce, J. M., Mackintosh, N. J. (2010). Two theories of attention: A review and a possible integration. [W:] C. J. Mitchell M. E. Le Pelley (red.), *Attention and associative learning: From brain to behaviour*, Oxford University Press, Oxford, s. 11–39.
339. Penn D. C., Povinelli D. J., (2007). On the lack of evidence that non-human animals possess anything remotely resembling a ‘theory of mind’. *Philosophical Transactions of the Royal Society of London B: Biological Sciences*, vol. 362, s. 731–744. <https://doi.org/10.1098/rstb.2006.2023>.
340. Perdue, C. W., Dovidio, J. F., Gurtman, M. B., Tyler, R. B. (1990). Us and them: Social categorization and the process of intergroup bias. *Journal of personality and social psychology*, 59(3), s. 475–486. <https://doi.org/10.1037/0022-3514.59.3.475>.
341. Perner, J. (1991). *Understanding the representational mind*. MIT Press.
342. Perner, J., Ruffman, T. (2005). Infants' insight into the mind: How deep? *Science*, 308(5719), s. 214–216. <https://doi.org/10.1126/science.1111656>.
343. Pervin, L. A., John, O. P. (2002). *Osobowość: Teoria i badania*. Wydawnictwo Uniwersytetu Jagiellońskiego, Kraków.
344. Pettigrew, T. F. (1998). Intergroup contact theory. *Annual review of psychology*, 49(1), s. 65-85. <https://doi.org/10.1146/annurev.psych.49.1.65>.
345. Pettigrew, T. F., Meertens, R. W. (1995). Subtle and blatant prejudice in Western Europe. *European journal of social psychology*, 25(1), s. 57-75, <https://doi.org/10.1002/ejsp.2420250106>.
346. Pika, S., Bugnyar, T. (2011). The use of referential gestures in ravens (*Corvus corax*) in the wild. *Nature communications*, 2(1), s. 560-573, <https://doi.org/10.1038/ncomms1567>.

347. Platon. (2007). *Państwo*. Wydawnictwo Naukowe PWN, Warszawa.
348. Platon. (2023). *Uczta*. Bellona, Warszawa.
349. Popper, K. R. (2015). *The logic of scientific discovery*. [W:] G. E. Moore, K. R. Popper, (red.), *Twentieth Century: Moore to Popper*. Central Works of Philosophy.
350. Poulin-Dubois, D., Yott, J. (2018). Probing the depth of infants' theory of mind: Disunity in performance across paradigms. *Developmental Science*, 21(4), <https://doi.org/10.1111/desc.12600>.
351. Povinelli, D. J. (1996). *Folk physics for apes: The chimpanzee's theory of how the world works*. Oxford University Press.
352. Povinelli, D. J., Vonk, J. (2003). Chimpanzee minds: Suspiciously human? *Trends in Cognitive Sciences*, s. 157-160. [https://doi.org/10.1016/S1364-6613\(03\)00053-6](https://doi.org/10.1016/S1364-6613(03)00053-6).
353. Povinelli, D. J., Eddy, T. J., Hobson, R. P., Tomasello, M. (1996). *What young chimpanzees know about seeing*. Monographs of the society for research in child development, <https://doi.org/10.2307/1166159>.
354. Povinelli, D. J., Vonk, J. (2004). We don't need a microscope to explore the chimpanzee's mind. *Mind & Language*, 19(1), s. 1–28. <https://doi.org/10.1111/j.0268-1064.2004.00214.x>.
355. Pratt, C., Bryant, P. (1990). Young children understand that looking leads to knowing (so long as they are looking into a single barrel). *Child Development*, 61(4), s. 973–982, <https://doi.org/10.2307/1130871>.
356. Pratto, F., Sidanius, J., Stallworth, L. M., Malle, B. F. (1994). Social dominance orientation: A personality variable predicting social and political attitudes. *Journal of Personality and Social Psychology*, 67(4), s. 741–763. <https://doi.org/10.1037/0022-3514.67.4.741>.
357. Premack, D., Woodruff, G. (1978). Does the chimpanzee have a theory of mind?. *Behavioral and brain sciences*, 1(4), s. 515-526, <https://doi.org/10.1017/S0140525X00076512>.
358. Preston, S. D., de Waal, F. B. M. (2002). Empathy: Its ultimate and proximate bases. *Journal of behavioral and brain science*, <https://doi.org/10.1017/s0140525x02000018>.
359. Prinz, J. (2007). *The Emotional Construction of Morals*, Oxford University Press.
360. Pritchard, D. (2005). *Epistemic luck*. Clarendon Press.

361. Pronin, E., Lin, D. Y., Ross, L. (2002). The bias blind spot: Perceptions of bias in self versus others. *Personality and Social Psychology Bulletin*, 28(3), s. 369-381, <https://doi.org/10.1177/0146167202286008>.
362. Prothmann, A., Ettrich, C., and Protthmann, S. (2009). Preference for, and responsiveness to people, dogs, and objects in children with autism. *Anthrozoös*, vol. 22, s. 161–171. doi: 10.2752/175303709X434185.
363. Puckett, N. N. (1991). *Popular Beliefs and Superstitions: A Compendium of American Folklore*. G. K. Hall. & Company. Boston.
364. Pullman, P. (2014). *Baśnie braci Grimm: dla dorosłych i młodzieży: bez cenzury*. Wydawnictwo Albatros.
365. Rawls, J. (2009). *Teoria sprawiedliwości*. Wydawnictwo Naukowe PWN, Warszawa.
366. Razran, G. (1957). Soviet psychology since 1950. *Science*, 126(3283), s. 1100-1107, <https://doi.org/10.1126/science.126.3283.1100>.
367. Razran, G. (1959). Pavlov the empiricist. *Science*, 130(3380), s. 916-917, <https://doi.org/10.1126/science.130.3380.916>.
368. Reuter, M., Dziarnowska, W. (2012). *Kognitywistyczne ujęcia konfabulacji: podejście pamięciowe vs. epistemiczne*. Studia Z Kognitywistyki I Filozofii Umysłu.
369. Richelle, M. (1993). *BF Skinner: a reappraisal*. Psychology Press.
370. Ricoeur, P. (2005). *Pamięć, historia, zapomnienie*. Universitas.
371. Robinson, R. (1971). Arguing from ignorance. *The Philosophical Quarterly*, 21(83), s. 97-108, <https://doi.org/10.2307/2218034>.
372. Roesch, M. R., Esber, G. R., Li, J., Daw, N. D., Schoenbaum, G. (2012). Surprise! Neural correlates of Pearce–Hall and Rescorla–Wagner coexist within the brain. *European Journal of Neuroscience*, 35(7), s. 1190–1200. <https://doi.org/10.1111/j.1460-9568.2011.07986.x>.
373. Rogers, E. M. (2003). *Diffusion of Innovations*. Free Press, Nowy York.
374. Rosenthal, R. (1981). Pavlov's mice, Pfungst's horse, and Pygmalion's PONS: Some models for the study of interpersonal expectancy effects. *Annals of the New York Academy of Sciences*, 364(1), s. 182-198, <https://doi.org/10.1111/j.1749-6632.1981.tb21315.x>.
375. Rosenthal, R., Fode, K. L. (1963). The effect of experimenter bias on the performance of the albino rat. *Behavioral Science*, 8(3), s. 183-189,

<https://doi.org/10.1002/bs.3830080302>.

376. Rosenthal, R., Fode, K. L., Friedman, C. J., Vikan, L. L. (1960). Subjects' perception of their experimenter under conditions of experimenter bias. *Perceptual and Motor Skills*, 11(3), s. 325-331, <https://doi.org/10.2466/pms.1960.11.3.407>.
377. Rosenthal, R., Jacobson, L. (1968). *Pygmalion in the classroom: Teacher expectation and pupils' intellectual development*. Holt, Rinehart & Winston.
378. Rosenthal, R., Rubin, D. B. (1978). Interpersonal expectancy effects: The first 345 studies. *Behavioral and Brain Sciences*, 1(3), s. 377-386.
379. Ross, L. (1977) *The intuitive psychologist and his shortcomings: Distortions in the attribution process*. [W:] L. Berkowitz, *Advances in experimental social psychology*. Academic Press, Nowy York, [https://doi.org/10.1016/S0065-2601\(08\)60357-3](https://doi.org/10.1016/S0065-2601(08)60357-3).
380. Ross, L., Lepper, M. R., & Hubbard, M. (1975). Perseverance in self-perception and social perception: Biased attributional processes in the debriefing paradigm. *Journal of Personality and Social Psychology*, 32(5), s. 880–892. <https://doi.org/10.1037/0022-3514.32.5.880>.
381. Rousseau J.J. (2002). *O umowie społecznej*. De Agostini, Ediciones Altaya Polska.
382. Rousseau, J.J. (1954). *Rozprawa o nierówności między ludźmi*. Wydawnictwo Naukowe PWN.
383. Ryczek, A. (2013). Krótka historia szowinizmu gatunkowego. *Polski Przegląd Nauk o Zdrowiu*, 2(35), s. 117-125.
384. Ryder, R. D. (1980). Szowinizm gatunkowy, czyli etyka wiwisekcji. *Etyka*, vol. 18, s. 39-47.
385. Safin, C. (2018) *Poza słowami. Co myślą i czują zwierzęta*. Wydawnictwo Krytyki Politycznej.
386. Saito, A., Shinozuka, K. (2013). Vocal recognition of owners by domestic cats (*Felis catus*). *Animal Cognition*, vol.16(4), s. 685–690. <https://doi.org/10.1007/s10071-013-0620-4>.
387. Saito, A., Takagi, S., Arahori, M., Fujita, K. (2019). Domestic cats (*Felis catus*) discriminate their names from other words. *Scientific Reports*, vol. 9, <https://doi.org/10.1038/s41598-019-40616-4>.
388. Samelson, F. (1978). From "race psychology" to "studies in prejudice": Some observations on the thematic reversal in social psychology. *Journal of the History of the*

- Behavioral Sciences*, s. 265-278, [https://doi.org/10.1002/1520-6696\(197807\)14:3%3C265::AID-JHBS2300140313%3E3.0.CO;2-P](https://doi.org/10.1002/1520-6696(197807)14:3%3C265::AID-JHBS2300140313%3E3.0.CO;2-P).
389. Samson, A. C., Gross, J. J. (2012). Humour as emotion regulation: The differential consequences of negative versus positive humour. *Cognition & emotion*, 26(2), s. 375-384, <https://doi.org/10.1080/02699931.2011.585069>.
 390. Savage-Rumbaugh, E. S., Rumbaugh, D. M., Boysen, S. (1978). Sarah's problems of comprehension. *Behavioral and Brain Sciences*, 1(4), s. 555-557, <https://doi.org/10.1017/s0140525x0007655x>.
 391. Savage-Rumbaugh, E. S., Rumbaugh, D. M., Boysen, S. (1980). Do Apes Use Language? One research group considers the evidence for representational ability in apes. *American Scientist*, 68(1), s. 49-61.
 392. Savage-Rumbaugh, E. S., Shanker, S., Taylor, T. J. (1998). *Apes, language, and the human mind*. Oxford University Press, Nowy York.
 393. Schacter, D. L. (2001). *Siedem grzechów pamięci*. Warszawa: Prószyński i S-ka.
 394. Schank, R. C., Abelson, R. P. (2013). *Scripts, plans, goals, and understanding: An inquiry into human knowledge structures*. Psychology press.
 395. Schimmack, U. (2001). Pleasure, displeasure, and mixed feelings: Are semantic opposites mutually exclusive?. *Cognition & Emotion*, 15(1), s. 81-97, <https://doi.org/10.1080/02699930126097>.
 396. Schetz, A. (2011). O tak zwanym problemie prostych umysłów. *Diametros*, vol. 30, s. 41-60.
 397. Schopenhauer, A. (2014). *Erystyka, czyli sztuka prowadzenia sporów*. Wydawnictwo SBM.
 398. Schultz, D. P., Schultz, S. E., Andruszko, R. (2008). *Historia współczesnej psychologii*. Wydawnictwo Uniwersytetu Jagiellońskiego.
 399. Schwarz, N. (1998). *Accessible content and accessibility experiences: The interplay of declarative and experiential information in judgment*. [W:] J. Metcalfe, (red.). *Metacognition: A Special Issue of Personality and Social Psychology Review*. <https://doi.org/10.4324/9781315799278>.
 400. Schwarz, N., Bless, H., Strack, F., Klumpp, G., Rittenauer-Schatka, H., Simons, A. (1991). Ease of retrieval as information: Another look at the availability heuristic. *Journal of Personality and Social Psychology*, 61(2), s. 195-202. <https://doi.org/10.1037/0022-3514.61.2.195>.

401. Sebeok, T. A., Rosenthal, R. E. (1981). *The Clever Hans phenomenon: Communication with horses, whales, apes, and people*. Annals of the New York Academy of Sciences.
402. Seeley, T. D. (1995). *The wisdom of the hive: The social physiology of honey bee colonies*. Harvard University Press, Cambridge.
403. Senju, A., Csibra, G. (2008). Gaze following in human infants depends on communicative signals. *Current Biology*, 18(9), s. 668–671.
<https://doi.org/10.1016/j.cub.2008.03.059>.
404. Senju, A., Southgate, V., Snape, C., Leonard, M., Csibra, G. (2011). Do 18-month-olds really attribute mental states to others? A critical test. *Psychological Science*, 22(7), s. 878–880, <https://doi.org/10.1177/0956797611411584>.
405. Serpell, J. (1999). *W towarzystwie zwierząt*. Państwowy Instytut Wydawniczy.
406. Serpell, J. A. (1996a). *In the company of animals*. Cambridge University Press. Cambridge.
407. Serpell, J. A. (2002). Anthropomorphism and Anthropomorphic Selection—Beyond the "Cute Response". *Society & Animals*, 10(4), s. 437–454,
<https://doi.org/10.1163/156853002320936926>.
408. Shelley, M. (2013) *Frankenstein*. Wydawnictwo Vesper.
409. Sherif, M. (1988). *The robbers cave experiment: Intergroup conflict and cooperation*. [Orig. pub. as *Intergroup conflict and group relations*]. Wesleyan University Press.
410. Sherman, G. D., Haidt, J., and Coan, J. A. (2009). Viewing cute images increases behavioral carefulness. *Emotion*, vol. 9, s. 282–286, <https://doi.org/10.1037/a0014904>.
411. Sherman, J. W., Stroessner, S. J., Conrey, F. R., Azam, O. A. (2005). Prejudice and stereotype maintenance processes: attention, attribution, and individuation. *Journal of personality and social psychology*, 89(4), s. 607–622, <https://doi.org/10.1037/0022-3514.89.4.607>.
412. Shettleworth, S. J. (2010a). *Cognition, evolution, and behavior*. Oxford University Press, Nowy York, s. 441–455.
413. Shettleworth, S. J. (2010b). Clever animals and killjoy explanations in comparative psychology. *Trends in Cognitive Sciences*, 14(11), s. 477–481, <https://doi.org/10.1016/j.tics.2010.07.002>.
414. Siegel, S. (2012). Cognitive penetrability and perceptual justification. *Noûs*, 46(2), s. 201–222.

415. Siegel, S. (2020). Bias and perception. In *An introduction to implicit bias*. Routledge, s. 99-115.
416. Siemaszko, A. (1993). *Granice tolerancji: O teoriach zachowań dewiacyjnych*. Państwowe Wydawnictwo Naukowe, Warszawa.
417. Sierra, M. (2009). *Depersonalization: A new look at a neglected syndrome*. Cambridge University Press.
418. Sikkens, T., Bosman, C. A., Olcese, U. (2019). The role of top-down modulation in shaping sensory processing across brain states: implications for consciousness. *Frontiers in systems neuroscience*, vol. 13.
419. Sikorska, J., Trojan, M. (2018). „My” i „one”. Szowinizm gatunkowy w kontekście badań psychologicznych nad relacjami międzygrupowymi. *Ethos. Kwartalnik Instytutu Jana Pawła II KUL*, 31(2), s. 161–179, <https://doi.org/10.12887/31-2018-2-122-10>.
420. Simon, H. A., Newell, A. (1971). Human problem solving: The state of the theory in 1970. *American Psychologist*, 26(2), s. 145–159. <https://doi.org/10.1037/h0030806>.
421. Simon, S. S., Tusch, E. S., Holcomb, P. J., Daffner, K. R. (2016). Increasing working memory load reduces processing of cross-modal task-irrelevant stimuli even after controlling for task difficulty and executive capacity. *Frontiers in human neuroscience*, vol. 10.
422. Simpson, G. E., Yinger, J. M. (1985). *Racial and cultural minorities: An analysis of prejudice and discrimination*. NY: Plenum Press, Nowy York.
423. Singer, P., Alichniewicz, A., Szczesna, A. (2018). *Wyzwolenie zwierząt*. Marginesy.
424. Singh, P., Minsky, M., Eslick, I. (2004). Computing commonsense. *BT Technology Journal*, 22(4), s. 201-210, <https://doi.org/10.1023/B:BTTJ.0000047599.89995.3c>.
425. Slobodchikoff, C. N., Paseka, A., Verdolin, J. L. (2009). Prairie dog alarm calls encode labels about predator colors. *Animal cognition*, 12, s. 435-439. <https://doi.org/10.1007/s10071-008-0203-y>.
426. Slovic, P., Finucane, M. L., Peters, E., MacGregor, D. G. (2007). The affect heuristic. *European journal of operational research*, 177(3), s. 1333-1352. <https://doi.org/10.1016/j.ejor.2005.04.006>.
427. Slovic, P., Kunreuther, H., White, G. F. (2016). Decision processes, rationality and adjustment to natural hazards. [W:] *The perception of risk*, Earthscan Publications, Routledge, s. 1-31.

428. Smolla M, Akçay E. (2023). Pathways to cultural adaptation: the coevolution of cumulative culture and social networks. *Evolutionary Human Sciences*. 5(26), <https://doi.org/10.1017/ehs.2023.21>.
429. Sodian, B., Frith, U. (1992). Deception and sabotage in autistic, retarded and normal children. *Journal of Child Psychology and Psychiatry*, 33(3), s. 591–605, <https://doi.org/10.1111/j.1469-7610.1992.tb00893.x>.
430. Sorge, R., Martin, L., Isbester, K., i in. (2014). Olfactory exposure to males, including men, causes stress and related analgesia in rodents. *Nat Methods*, vol. 11, s. 629–632, <https://doi.org/10.1038/nmeth.2935>.
431. Sosa, E. (1991). *Knowledge in perspective: Selected essays in epistemology*. Cambridge University Press, <https://doi.org/10.1017/CBO9780511625299>.
432. Southgate, V., Senju, A., Csibra, G. (2007). Action anticipation through attribution of false belief by 2-year-olds. *Psychological Science*, 18(7), s. 587–592, <https://doi.org/10.1111/j.1467-9280.2007.01944.x>.
433. Sperber, D. (2000). Metarepresentations in an evolutionary perspective. *Metarepresentations: A multidisciplinary perspective*, [W:] D. Sperber, *Metarepresentations: A multidisciplinary perspective*. Oxford Academic, Nowy York, s. 117–137, <https://doi.org/10.1093/oso/9780195141146.003.0005>.
434. Stanovich, K. E., West, R. F. (2000). Advancing the rationality debate. *Behavioral and brain sciences*, 23(5), str. 701–717, <https://doi.org/10.1017/S0140525X00623439>.
435. Stefaniak, A., Górską, P. (2015). Syndrom międzygrupowej wrogości w Polskim Sondażu Uprzedzeń 2, [W:] A. Stefaniak, M. Bilewicz, M. Winiewski (red.), *Uprzedzenia w Polsce*, Warszawa, s. 131–155.
436. Stein, B. S., Bransford, J. D. (1979). Constraints on effective elaboration: Effects of precision and subject generation. *Journal of Verbal Learning & Verbal Behavior*, 18(6), s. 769–777. [https://doi.org/10.1016/S0022-5371\(79\)90481-X](https://doi.org/10.1016/S0022-5371(79)90481-X).
437. Sternberg, R. J. (2013). *Efekty oczekiwań interpersonalnych: wybór tekstów*. Wydawnictwo Naukowe Scholar.
438. Stolarska, U., Kaciński, M. (2011). *Efekt placebo*. Wydawnictwo Naukowe PWN.
439. Stone, D. L., Łukaszewski, K. M., Krueger, D. C., Canedo, J. C. (2020). A model of factors thought to influence unfair discrimination against immigrants in organizations. *Diversity in Organizations, Special Series, Research in HRM, Information Age Publishing*, Charlotte, Nowy York, s. 331–61.

440. Strack, F., Martin, L. L., Schwarz, N. (1988). Priming and communication: Social determinants of information use in judgments of life satisfaction. *European journal of social psychology*, 18(5), s. 429-442, <https://doi.org/10.1002/ejsp.2420180505>.
441. Stroh, W. R. (2022). How the Implicit Bias, Representative Heuristic, and Anchoring Effect Embed Institutionalized Racism into the American Judicial System. *Horizon Projects*, <https://doi.org/10.47611/harp.229>.
442. Suddendorf, T., Corballis, M. C. (2007). The evolution of foresight: What is mental time travel, and is it unique to humans?. *Behavioral and Brain Sciences*, 30(3), s. 299-313. <https://doi.org/10.1017/S0140525X07001975>.
443. Sykes, G. M., Matza, D. (1957). Techniques of neutralization: A theory of delinquency. *American Sociological Review*, 22(6), s. 664-670.
444. Szubaka, T. (2004). Behawioryzm,(w:) *Słownik filozofii*. Zielona Sowa, Kraków.
445. Św. Tomasz z Akwinu. (1970). św, Suma teologiczna, 2-2, zag. 64, art. 1, tłum. Bednarski EW, *Katolicki Ośrodek Wydawniczy „Veritas”*, Londyn.
446. Świerk, K. (2017). Diabeł w sutannie, czyli ontologia pewnych bytów według Indian Matsigenka z peruwiańskiej Amazonii. *Etnografia. Praktyki, Teorie, Doświadczenia*, vol. 3, s. 119-136.
447. Takagi, S., Saito, A., Arahori, M., Chijiiwa, H., Koyasu, H., Nagasawa, M., Kikusui, T., Fujita K., Kuroshima, H. (2023). Author Correction: cats learn the names of their friend cats in their daily lives. *Scientific Reports*, 12996(2023), <https://doi.org/10.1038/s41598-022-11573-7>.
448. Tajfel, H. (1982). *Social identity and intergroup relations*. Cambridge University Press.
449. Tajfel, H. (1984). Intergroup relations, social myths and social justice in social psychology. [W:] H. Tajfel (red.). *The social dimension*. Cambridge University Press, Cambridge, s. 695-715.
450. Tajfel, H., Billig, M. G., Bundy, R. P., Flament, C. (1971). Social categorization and intergroup behaviour. *European journal of social psychology*, 1(2), s. 149-178, <https://doi.org/10.1002/ejsp.2420010202>.
451. Tajfel, H., Turner, J. C. (2000). An integrative theory of intergroup conflict. [W:] M. J. Hatch, M. Schultz (red.), *Organizational identity: A reader*. Oxford University Press, s. 56–65, <https://doi.org/10.1093/oso/9780199269464.003.0005>.
452. Teich, M., Porter, R. Cook, H.J. (1992). *The Scientific Revolution in National Context*, Cambridge University Press, s. 127–129.

453. Templeton, A. R. (2013). Biological races in humans. *Studies in History and Philosophy of Science Part C: Studies in History and Philosophy of Biological and Biomedical Sciences*, 44(3), s. 262–271, <https://doi.org/10.1016/j.shpsc.2013.04.010>.
454. Thaler R. H., Sunstein, C., R., (2023). *Impuls. Jak podejmować właściwe decyzje dotyczące zdrowia, dobrobytu i szczęścia*, Wydawnictwo Zys i s-ka, Poznań.
455. Thompson, E. P., Roman, R. J., Moskowitz, G. B., Chaiken, S., Bargh, J. A. (1994). Accuracy motivation attenuates covert priming: The systematic reprocessing of social information. *Journal of personality and Social Psychology*, 66(3), str. 474 - 489, <https://doi.org/10.1037/0022-3514.66.3.474>.
456. Thorndike, E. L. (1920). A constant error in psychological ratings. *Journal of applied psychology*, 4(1), s. 25-29. <https://doi.org/10.1037/h0071663>.
457. Tinbergen N. (1976). *Badania nad instynktem*. PWN, Warszawa.
458. Tinbergen, N. (1942). An objectivistic study of the innate behaviour of animals. [W:] *An objectivistic Study of the innate behaviour of animals*. Brill.
459. Tinbergen, N. (1963). On aims and methods of ethology. *Zeitschrift für Tierpsychologie*, 20(4),s . 410–433, <https://doi.org/10.1111/j.1439-0310.1963.tb01161.x>.
460. Tizard, B., Hughes, M. (1984). *Young Children Learning*. Londyn: Fontana.
461. Tolman, E. C. (1922). A new formula for behaviorism. *Psychological Review*, 29(1), <https://doi.org/10.1037/h0074428>.
462. Tomasello, M. (2015). *Historia naturalna ludzkiego myślenia*. Copernicus Center Press.
463. Tomasello, M., Call, J., Hare, B. (1999). Five primate species follow the visual gaze of conspecifics. *Animal Behaviour*, 55(4), s. 1063-1069, <https://doi.org/10.1006/anbe.1997.0636>.
464. Tomasello, M., Carpenter, M., Call, J., Behne, T., Moll, H. (2005). Understanding and sharing intentions: The origins of cultural cognition. *Behavioral and Brain Sciences*, 28(5), s. 675-691, <https://doi.org/10.1017/S0140525X05000129>.
465. Tomasello, M., Melis, A. P., Tennie, C., Wyman, E., Herrmann, E. (2012). Two key steps in the evolution of human cooperation: The interdependence hypothesis. *Current anthropology*, 53(6), s. 673-692, <https://doi.org/10.1086/668207>.
466. Tomonaga, M. (2007). Visual search for orientation of faces by a chimpanzee (*Pan troglodytes*): face-specific upright superiority and the role of facial configural properties. *Primates*, 48(1), s. 1–12, <https://doi.org/10.1007/s10329-006-0011-4>.

467. Tomonaga, M., Imura, T. (2009). Faces capture the visuospatial attention of chimpanzees (*Pan troglodytes*): evidence from a cueing experiment. *Frontiers in Zoology*, vol. 6, <https://doi.org/10.1186/1742-9994-6-14>.
468. Trevarthen, C. (1979). Communication and cooperation in early infancy: A description of primary intersubjectivity. [W:] W. M. Bullowa (red.), *Before speech: The beginning of human communication*, Cambridge University Press, s. 321–347.
469. Trojan, M. (2013). *Na tropie zwierzęcego umysłu*. Wydawnictwo Naukowe Scholar Sp. z oo, Warszawa.
470. Tsang, J. (2002). Moral rationalization and the integration of situational factors and psychological processes in immoral behavior. *Review of General Psychology*, vol. 6, s. 25-50, <https://doi.org/10.1037/1089-2680.6.1.25>.
471. Turner, J. C. (1987). *Rediscovering the social group: A self-categorization theory*. Blackwell.
472. Tversky, A., Kahneman, D. (1974). Judgment under uncertainty: Heuristics and biases. *Science*, 185(4157), s. 1124-1131.
473. Tversky, A., Kahneman, D. (1983). Extensional versus intuitive reasoning: The conjunction fallacy in probability judgment. *Psychological Review*, 90(4), s. 293–315. <https://doi.org/10.1037/0033-295X.90.4.293>.
474. Urbański, M. (2009). *Rozumowania abdukcyjne. Modele i procedury*. Wydawnictwo Naukowe UAM, Poznań.
475. Urcuioli, P.J. (2006). Responses and acquired equivalence classes, [W:] Wasserman, E.A., Zentall, T.R. (red.). *Comparative Cognition: Experimental Explorations of Animal Intelligence*, Oxford University Press, s. 405–421.
476. Vaes, J., Paladino, M. P., Castelli, L., Leyens, J. P., Giovanazzi, A. (2003). On the behavioral consequences of infrahumanization: the implicit role of uniquely human emotions in intergroup relations. *Journal of personality and social psychology*, 85(6), <https://doi.org/10.1037/0022-3514.85.6.1016>.
477. Van Eemeren, F. H., Grootendorst, R., Jacobs, C. S., Jackson, S. A. (2002). *Reconstructing argumentative discourse*. The University of Alabama Press.
478. Vitale Shreve, K. R., Udell, M. A. R. (2015). What's inside your cat's head? A review of cat (*Felis silvestris catus*) cognition research past, present and future. *Animal Cognition*, vol. 18, s.1195–1206. <https://doi.org/10.1007/s10071-015-0897-6>.

479. Von Uexküll, J. (2013). *A foray into the worlds of animals and humans: With a theory of meaning*. Minnesota Press.
480. Walton, D. (1999a). Profiles of dialogue for evaluating arguments from ignorance. *Argumentation*, 13(1), s. 53-71, <https://doi.org/10.1023/A:1007738812877>.
481. Walton, D. (1999b). The appeal to ignorance, or argumentum ad ignorantiam. *Argumentation*, 13(1), s. 367-377, <https://doi.org/10.1023/A:1007780012323>.
482. Walton, D. (2004). *Burden of proof, presumption and argumentation*. Cambridge University Press.
483. Waring, T. M., Wood, Z. T. (2021). Long-term gene–culture coevolution and the human evolutionary transition. *Proceedings of the Royal Society B*, 288(1952), <https://doi.org/10.1098/rspb.2021.0538>.
484. Warneken, F., Tomasello, M. (2006). Altruistic helping in human infants and young chimpanzees. *Science*, 311(5765), s. 1301–1303, <https://doi.org/10.1126/science.1121448>.
485. Wartenberg, T. E. (1992). Situated Social Power, [W:] T., Wartenberg. *Rethinking Power*, State University of New York Press, Albany, s. 79-101.
486. Wellman, H. M. (1990). *The child's theory of mind*. MIT Press.
487. Wellman, H. M., Cross, D., Watson, J. (2001). Meta-analysis of theory-of-mind development: The truth about false belief. *Child Development*, 72(3), s. 655-684, <https://doi.org/10.1111/1467-8624.00304>.
488. Westen, D., Blagov, P. S., Harenski, K., Kilts, C., Hamann, S. (2006). Neural bases of motivated reasoning: An fMRI study of emotional constraints on partisan political judgment in the 2004 US presidential election. *Journal of cognitive neuroscience*, 18(11), s. 1947-1958. <https://doi.org/10.1162/jocn.2006.18.11.1947>.
489. Whitehead, H., Laland, K. N., Rendell, L., Thorogood, R., Whiten, A. (2019). The reach of gene–culture coevolution in animals. *Nature communications*, 10(1). <https://doi.org/10.1038/s41467-019-10293-y>.
490. Whiten, A., Horner, V., de Waal, F. B. M. (2005). Conformity to cultural norms of tool use in chimpanzees. *Nature*, 437(7059), s. 737-740.
491. Wikenros, C., Gicquel, M., Zimmermann, B., Flagstad, Ø., Åkesson, M. (2021). Age at first reproduction in wolves: Different patterns of density dependence for females and

- males. *Proceedings of the Royal Society B: Biological Sciences*, 288(1948), <https://doi.org/10.1098/rspb.2021.0207>.
492. Williamson, T. (2002). *Knowledge and its Limits*. Oxford University Press, USA.
493. Willis-Moroe, J. (2019). *Odczarowanie, czyli dogmaty religii w świetle ustaleń nauki*. Ridero.
494. Wilson-Mendenhall, C. D. (2017). Constructing emotion through simulation. *Current Opinion in Psychology*, vol. 17, <https://doi.org/10.1016/j.copsyc.2017.07.015>.
495. Wimmer, H. (1983). Beliefs about beliefs: Representation and constraining function of wrong beliefs in young children's understanding of deception. *Cognition*, 13(1), s. 103–128. doi:10.1016/0010-0277(83)90004-5.
496. Wittgenstein, L. (1924). *Tractatus Logico-Philosophicus*. *Buletin of the American Mathematical Society*, vol. 30, str. 179-181.
497. Wittgenstein, L. (2019). *Dociekania filozoficzne*. Wydawnictwo Naukowe PWN.
498. Wojciszke, B. (2009). *Człowiek wśród ludzi. Zarys psychologii społecznej*. Wydawnictwo Naukowe Scholar.
499. Wojciszke, B. (2011). *Psychologia społeczna*. Wydawnictwo Naukowe Scholar.
500. Woleński, J. (2003). *Epistemologia. Tom III: Prawda i realizm*, Wydawnictwo Aureus, Kraków.
501. Woodruff, G., Premack, D. (1979). Intentional communication in the chimpanzee: The development of deception. *Cognition*, 7(4), s. 333-362, [https://doi.org/10.1016/0010-0277\(79\)90021-0](https://doi.org/10.1016/0010-0277(79)90021-0).
502. Wypler, W. (2013). Samousprawiedliwianie zachowania a rozluźnianie standardów moralnych. [W:] J. Lipińska-Lokś, G. Miłkowska i A. Napadło-Kuczera (red.), *Zewnętrzne i wewnętrzne aspekty bezpieczeństwa indywidualnego i zbiorowego*. Oficyna Wydawnicza UZ, Zielona Góra, str. 27–41.
503. Zagorin, P. (2001). Francis Bacon's concept of objectivity and the idols of the mind. *The British Journal for the History of Science*, 34(4), s. 379-393, <https://doi.org/10.1017/s0007087401004411>.
504. Zagzebski, L. T. (1996). *Virtues of the mind: An inquiry into the nature of virtue and the ethical foundations of knowledge*. Cambridge University Press.

505. Zaki, J., Hennigan, K., Weber, J., Ochsner, K. N. (2010). Social cognitive conflict resolution: Contributions of domain-general and domain-specific neural systems. *The Journal of Neuroscience*, 30, s. 8481-8488, <https://doi.org/10.1523/JNEUROSCI.0382-10.2010>.
506. Zebrowitz, L. A., Franklin Jr, R. G. (2014). The attractiveness halo effect and the babyface stereotype in older and younger adults: Similarities, own-age accentuation, and older adult positivity effects. *Experimental aging research*, 40(3), str. 375-393, doi: 10.1080/0361073X.2014.897151.
507. Zentall, T. R. (2006). Imitation: Definitions, evidence, and mechanisms. *Animal Cognition*, 9(4), s. 335-353. <https://doi.org/10.1007/s10071-006-0039-2>.
508. Zieliński, T. (1927). *Hellenizm a judaizm*, cz. 1-2, Wydawnictwo J. Morkowicza–Towarzystwo Wydawnicze w Warszawie.
509. Ziółkowski (2016). *Syndrom obłądzonej twierdzy jako mechanizm socjotechniczny*. [W:] F. Pierzchalski, J. Golinowski, (red.). (2016). *Socjotechnika lęku w polityce*. Wydawnictwo Uniwersytetu Kazimierza Wielkiego, Bydgoszcz.
510. Zirkle, C. (1958). Pavlov's beliefs. *Science*, 128(3337), s. 1476-1476, <https://doi.org/10.1126/science.128.3337.1476.b>.
511. Zoble, E. J., Lehman, R. S. (1969). Interaction of subject and experimenter expectancy effects in a tone length discrimination task. *Behavioral Science*, 14(5), s. 357-363, <https://doi.org/10.1002/bs.3830140503>.

Netografia:

1. Allport, G. W. (1954). *The nature of prejudice*. The Beacon Press, Boston: <https://ia601500.us.archive.org/17/items/in.ernet.dli.2015.188638/2015.188638.The-Nature-Of-Prejudice.pdf>, data dostępu: 14.06.2025.
2. AnimalFACS, <https://animalfacs.com/>, Data dostępu: 14.06.2025.
3. APOPO, <https://apopo.org/herorats/>, data dostępu: 14.06.2025.
4. Aronson, E., Mills, J. (1959). The effect of severity of initiation on liking for a group. *The Journal of Abnormal and Social Psychology*, 59(2): https://web.mit.edu/curhan/www/docs/Articles/15341_Readings/Motivation/Aronson_Mills_1959_The_effect_of_severity_of_initiation.pdf, data dostępu: 14.06.2025.
5. Arystoteles (2003) Polityka: <http://biblioteka.kijowski.pl/arystoteles/polityka.pdf>, data dostępu: 14.06.2025.
6. Bandura, A., Walters, R. H. (1977). *Social learning theory*, Englewood Cliffs, NJ: Prentice hall, 1, s. 141-154: https://www.asecib.ase.ro/mps/Bandura_SocialLearningTheory.pdf, data dostępu: 14.06.2025.
7. Bandura, A. (2001). Social cognitive theory of mass communication. *Media Psychology*, 3(3), s. 265-299: http://cogweb.ucla.edu/crp/Media/Bandura_01.pdf, data dostępu: 14.06.2025.
8. Baron-Cohen, S. (2000). Theory of mind and autism: A fifteen-year review. *Understanding Other Minds: Perspectives from Developmental Cognitive Neuroscience, International Review of Research in Mental Retardation* 2, s. 3-20: http://docs.autismresearchcentre.com/papers/2000_SBC_Theory-of-mind-and-autism-review.pdf, data dostępu: 14.06.2025.
9. Barone, P., Corradi, G., Gomila, A. (2019). Re-thinking false-belief attribution: a comparative-developmental perspective. *Frontiers in Psychology*, vol. 10, <https://doi.org/10.3389/fpsyg.2019.02356>: <https://www.frontiersin.org/journals/psychology/articles/10.3389/fpsyg.2019.02356/full>, data dostępu: 14.06.2025.
10. Berny, J., Leśniewski, M. A., Górski, P. (2012). Motywacja w systemie zarządzania zasobami ludzkimi. Analiza teoretyczna problemu: <http://core.ac.uk/download/pdf/160237348.pdf>, data dostępu: 14.06.2025.

11. Borgi, M., Cogliati-Dezza, I., Brelsford, V., Meints, K., Cirulli, F. (2014). Baby schema in human and animal faces induces cuteness perception and gaze allocation in children. *Frontiers in psychology*, vol. 5:
<https://www.frontiersin.org/journals/psychology/articles/10.3389/fpsyg.2014.00411/full>, data dostępu: 14.06.2025.
12. Boyd, R. (1985). Evolutionary Psychology. *Britannica*,
<https://www.britannica.com/science/evolutionary-psychology>, data dostępu: 14.06.2025
13. Burdzik, T. (2015). Kawaii. Estetyka made in Japan. *Kultura Popularna*, 44(02), s. 156-165, <https://www.ceeol.com/search/article-detail?id=472707>, data dostępu: 14.06.2025.
14. Cartoon Corruption. (2023). You're Not Going To Believe What Courage The Cowardly: Dog Was REALLY About. [Film], (YouTube):
https://www.youtube.com/watch?v=9P9S42o_UY4, data dostępu: 14.06.2025.
15. Cartoon Corruption (2023). *This Horrific Theory Reveals The REAL Reason The Ed's Never Left The Cul-de-Sac*. [Film], (YouTube):
<https://www.youtube.com/watch?v=AqZ-1qqted8>, data dostępu: 14.06.2025.
16. Chomsky, N. (2008). Trzy elementy projektujące język. *Prace Językoznawcze*, vol. 10, s. 241-273:
https://bazhum.muzhp.pl/media/files/Prace_Językoznawcze/Prace_Językoznawcze-r2008-t10/Prace_Językoznawcze-r2008-t10-s241-273/Prace_Językoznawcze-r2008-t10-s241-273.pdf, data dostępu: 14.06.2025.
17. Clark, A. (1998). *Being there: Putting brain, body, and world together again*. MIT Press: <https://tiny.pl/dp1pj>, data dostępu: 14.06.2025.
18. Cullari, S., Kusché, C. A. (2025). Ego defense mechanisms. EBSCO Research Starters: <https://www.ebsco.com/research-starters/psychology/ego-defense-mechanisms>, data dostępu: 14.06.2025.
19. Damasio, A. (1994). Descartes' error: Emotion, rationality and the human brain. *Putnam*, New York:
https://ahandfulofleaves.wordpress.com/wp-content/uploads/2013/07/descartes-error_antonio-damasio.pdf, data dostępu: 14.06.2025.
20. Daszkiewicz, W. (2015). Ewolucjonistyczna teoria kultury w antropologii społecznej. *Zeszyty Naukowe KUL*, 58(3), s. 65–80:

- https://www.kul.pl/files/102/articles/2015_3/zn_kul_2015_3_a_daszkiewicz.pdf, data dostępu: 14.06.2025.
21. Donaldson, S., Kymlicka, W. (2013). A defense of animal citizens and sovereigns. *Law, Ethics and Philosophy*, s. 143-160:
<https://raco.cat/index.php/LEAP/article/view/294786/383319>, data dostępu: 14.06.2025.
 22. Dhont, K., Hodson, G., Loughnan, S., Amiot, C. E. (2019). Rethinking human-animal relations: The critical role of social psychology. *Group Processes & Intergroup Relations*, 22(6), s. 769-784:
https://www.pure.ed.ac.uk/ws/portalfiles/portal/113261736/Dhont_Hodson_Loughnan_Amiot_Intro_to_GPIR_SI_Animals_2019_FINAL.pdf, data dostępu: 14.06.2025.
 23. Donovan, P. (2009), Evidence Points to Conscious 'Metacognition' in Some Nonhuman Animals. Some animals may be able to reflect upon their states of mind. University of Buffalo:
<https://www.buffalo.edu/news/releases/2009/09/10440.html>, data dostępu: 14.06.2025.
 24. Eko-Unia. (2023). 10 000 podpisów przeciw elektrowni jądrowej w Choczewie złożone w GDOŚ, <https://eko-unia.org.pl/10-000-podpisow-przeciw-elektrowni-jadrowej-w-choczewie-zlozone-w-gdos/>, data dostępu: 14.06.2025.
 25. Encyklopedia Stanfordzka, (2003), Francis Bacon:
<https://plato.stanford.edu/entries/francis-bacon/#Ido>, data dostępu: 14.06.2025.
 26. Flavell, J. H., Green, F. L., Flavell, E. R. (1981). *The development of children's knowledge about visual perception*. Monographs of the Society for Research in Child Development, 46(5), vol. 192, s. 1–87. <https://doi.org/10.2307/1165985>:
https://bingschool.stanford.edu/sites/bingschool/files/1981_flavelletal2.pdf, data dostępu: 14.06.2025.
 27. Finkelhor, D. (2022). *The Internet, Youth Safety and the Problem of “Juvenoia” (Crimes against Children Research Center)*:
<https://www.unh.edu/ccrc/sites/default/files/media/2022-02/juvenoia-paper.pdf>, data dostępu: 14.06.2025.
 28. Foley R, Gamble C. (2009). The ecology of social transitions in human evolution. *Philosophical Transactions. The Royal Society Publishing*.
<https://doi.org/10.1098/rstb.2009.0136>:

- <https://royalsocietypublishing.org/doi/abs/10.1098/rstb.2009.0136>, data dostępu: 14.06.2025.
29. Glanzer, M., Cunitz, A. R. (1966). Two storage mechanisms in free recall. *Journal of verbal learning and verbal behavior*, 5(4), s. 351-360:
<https://meghnaspsychportfolio.wordpress.com/wp-content/uploads/2016/01/glanzer-cunitz-1966-serial-position-memory.pdf>, data dostępu: 14.06.2025.
30. Grabowska, B. (2011). Językowe aspekty szowinizmu gatunkowego. *Repozytorium.umk*:
<https://repozytorium.umk.pl/bitstream/handle/item/1006/B.%20Grabowska,%20J%C4%99zykowe%20aspekty%20szownizmu%20gatunkowego.pdf>, data dostępu: 14.06.2025.
31. Grabowska M. (2014) Potoczność w obrazie wiary w Boga na podstawie wypowiedzi współczesnych Polaków, *Słowo. Studia językoznawcze*, vol. 5:
https://www.ur.edu.pl/files/ur/import/private/77/IPID/czasopisma/slowo/slowo_5/05_grabowska.pdf, data dostępu: 14.06.2025.
32. Grabowska, B. (2014). Zmiany relacji człowiek–zwierzę, czyli cena postępu. *Kultura i wartości*, vol. 10, s. 105-120:
https://cejsh.icm.edu.pl/cejsh/element/bwmeta1.element.ojs-doi-10_17951_kw_2014_10_105/c/198-196.pdf, data dostępu: 14.06.2025.
33. Green, M. S. (2019). From signaling and expression to conversation and fiction. *Grazer philosophische studien*, vol. 96(3), s. 295-315:
<https://philpapers.org/archive/SGRFSa.pdf>, data dostępu: 14.06.2025.
34. Grochmal-Bach, B. (1997). Czynniki noetyczne w promowaniu zdrowia moralnego u dzieci niepełnosprawnych:
<https://rep.up.krakow.pl/xmlui/bitstream/handle/11716/8077/RND180--22--Czynniki-noetyczne--Grochmal-Bach.pdf?sequence=1>, data dostępu: 14.06.2025.
35. Grzelak, B., Stolarz, A., Gross, H., Sehn, J. Jak pachnie sprawca? Wartość diagnostyczna i praktyczne znaczenie ekspertyzy osmologicznej:
http://www.repozytorium.uni.wroc.pl/Content/58883/PDF/06_Malgorzata_Wisniewska.pdf, data dostępu: 14.06.2025.

36. Igniter media, (2010). *The Marshmallow Test | Igniter Media | Church Video*. [Film], (YouTube): https://www.youtube.com/watch?v=QX_oy9614HQ, data dostępu: 14.06.2025.
37. Kaluziński, B. (2015). Czy „Epistemologia Cnót” pozwoli nam rozwiązać problem Gettier’a?. *Internetowy Magazyn Filozoficzny Hybris*, (30), s. 158-174: https://dspace.uni.lodz.pl/xmlui/bitstream/handle/11089/20337/hybris30_Kaluzinski_158_174.pdf?sequence=1&isAllowed=y, data dostępu: 14.06.2025.
38. Karan, P. (2023). Understanding the Concept of Human Superiority: An Exploration of the Theory of Mind in Non-Human Species. *Advances in Anthropology*, 13(4), s. 295-304: <https://www.scirp.org/journal/paperinformation?paperid=129123>, data dostępu: 14.06.2025.
39. Krawczyk, T. (2022). Dysonans poznawczy czyli o mechanizmie ratującym ego: https://nowestandardy.pl/nauka/tomasz_krawczyk_dysonans_poznawczy.pdf, data dostępu: 14.06.2025.
40. Lamża, Ł. (2017) *Ewolucja naczelnych: od szczura do australopiteka*, Copernicus [Film], (YouTube): <https://www.youtube.com/watch?v=eCR7yIYvN8A>, data dostępu: 14.06.2025.
41. Linell, P. (2009a). *Rethinking language, mind, and world dialogically*. IAP: <https://tiny.pl/dp1nr>, data dostępu: 14.06.2025.
42. Linnell, P. (2009b). Building brand value through technical instruction-a case study in the use of a qualitative analysis method. Society for Technical Communication: <https://dora.dmu.ac.uk/server/api/core/bitstreams/84f0cfc0-7844-400c-8c05-4eb2e81a4bad/content>, data dostępu: 14.06.2025.
43. Low, J., Apperly, I. A., Butterfill, S., Rakoczy, H. (2016). Cognitive architecture of belief reasoning in children and adults: A primer on developmental research. *Philosophy Compass*, 11(7), s. 431–446, <https://doi.org/10.1111/phc3.12333>: https://pure-oai.bham.ac.uk/ws/portalfiles/portal/29362753/Low_et_al_cognitive_architecture_2016.pdf, data dostępu 14.06.2025.
44. Mafessoni, F., Lachmann, M. (2019). The complexity of understanding others as the evolutionary origin of empathy and emotional contagion. *Scientific reports*, 9(1): <https://www.nature.com/articles/s41598-019-41835-5>, data dostępu: 14.06.2025.

45. Main, P. (2023, November 30). Top-Down Processing and Bottom-Up Processing: <https://www.structural-learning.com/post/top-down-processing-and-bottom-up-processing>, data dostępu: 14.06.2025.
46. Marchewka, K. (2016). Wpływ błędów poznawczych na wolność wyboru pacjenta. *Filozofia w Praktyce*: <https://ruj.uj.edu.pl/server/api/core/bitstreams/65797b1a-5cea-4fd1-8b7e-e7143407ee62/content>, data dostępu: 14.06.2025.
47. Marczak, M. (2016) Efekt halo w psychologii–percepcja osób uprawiających sport i nieuprawiających sportu. *ResearchGate.net*: https://www.researchgate.net/profile/Milena-Marczak/publication/309313234_Efekt_halo_w_psychologii_-_percepcja_osob_uprawiajacych_sport_i_nieuprawiajacych_sportu/links/580925de08ae993dc0509fdb/Efekt-halo-w-psychologii-percepcja-osob-uprawiajacych-sport-i-nieuprawiajacych-sportu.pdf, data dostępu: 14.06.2025.
48. Marino, L. (2011). Cetaceans and primates: Convergence in intelligence and self-awareness. *Journal of Cosmology*, vol. 14, s. 1063-1079: https://www.wellbeingintlstudiesrepository.org/cgi/viewcontent.cgi?article=1029&context=acwp_asie, data dostępu: 16.04.2025.
49. Mäses, M., Wascher, C. A. F. (2022). Assessing cats' (*Felis catus*) sensitivity to human pointing gestures. *bioRxiv*, <https://doi.org/10.1101/2022.03.12.484069>: <https://www.biorxiv.org/content/10.1101/2022.03.12.484069v4.full>, data dostępu: 09.06.2025.
50. Moll, H., Nichols, R., Mackey, J. L. (2020). Intentionality hypothesis. Review of Philosophy and Psychology. <https://doi.org/10.1007/s13164-020-00489-3>: https://dornsife.usc.edu/henrike-moll/wp-content/uploads/sites/404/2023/11/Moll_Nichols_Mackey_2020.pdf, data dostępu: 07.06.2025.
51. Napiórkowski, M. (2018). *Mity polityczne wczoraj i dziś*. [Film], (YouTube): <https://www.youtube.com/watch?v=kTjc1E7z5DM&t=912s>, data dostępu: 03.11.2023.
52. NaTemat.pl. (2022). Nie żyje słynny szczur-saper. Magawa wytropił prawie 100 bomb w Kambodży i dostał medal: <https://www.youtube.com/watch?v=YejS1e56dKM>, data dostępu: 12.05.2024.

53. Napiórkowska, T. (2017). Wybrane aspekty strategii rozrodczej pajaków-poliandria. *Kosmos*, 66(3), s. 401-411:
<http://kosmos.icm.edu.pl/PDF/2017/401.pdf>, data dostępu: 06.04.2025.
54. Newell, A., Shaw, J. S., Simon, H. A. (1959). *Report on General Problem Solving. Carnegie Institute of Technology*:
https://iiif.library.cmu.edu/file/Newell_box00005_fld00314_doc0003/Newell_box00005_fld00314_doc0003.pdf, data dostępu: 31.05.2025.
55. Nguyen, M. (2016), Psychological Disorders in Courage the Cowardly Dog:
<https://prezi.com/wwbmbdfwbpm/psychological-disorders-in-courage-the-cowardly-dog/>, data dostępu: 01.06.2025.
56. Okarma H. (2012). Dlaczego wilki coraz częściej podchodzą do ludzi?:
https://zielona.interia.pl/przyroda/news-dlaczego-wilki-coraz-czesciej-podchodza-do-ludzi,nId,5094541#utm_source=paste&utm_medium=paste&utm_campaign=chrome, data dostępu: 29.09.2023.
57. Olszewski, B. (2015) Substancje psychoaktywne a ewolucja komunikacji symbolicznej w paleolicie i współczesności. [W]: Antropologia komunikacji. Od starożytności do współczesności, *Polska Akademia Nauk*, Wrocław, s. 159-172, <https://tiny.pl/dnxj5>, data dostępu: 24.07.2024.
58. Orosz, G., Krekó, P., Paskuj, B., Tóth-Király, I., Bóthe, B., Roland-Lévy, C. (2016). Changing conspiracy beliefs through rationality and ridiculing. *Frontiers in Psychology*, vol. 7:
<https://www.frontiersin.org/articles/10.3389/fpsyg.2016.01525/full?footnotes>, data dostępu: 20.06.2023.
59. OpenStax (2022). Introduction to Philosophy by OpenStax (Official Print Version, paperback version, B&W):
<https://openstax.org/books/introduction-philosophy/pages/2-1-the-brain-is-an-inference-machine>, data dostępu: 24.07.2024.
60. Pap.pl. (2024): Pierwszy wyżeł w polskiej policji. Będzie poszukiwał ludzkich zwłok:
<https://www.pap.pl/aktualnosc/pierwszy-wyzel-w-polskiej-policji-bedzie-poszukiwal-ludzkich-zwlok-nasze-wideo>, data dostępu: 19.06.2025.
61. Polański, G. (2013). Heurystyki wydawania sądów społecznych w mediach:
https://ktime.up.krakow.pl/symp2013/referaty_2013_10/polanski.pdf, data dostępu: 03.12.2024.

62. Polski Autyzm (2015). *Badanie teorii umysłu w autyzmie*:
<https://polskiauizym.pl/badanie-teorii-umyslu-w-autyzmie/>, data dostępu: 6.03.2025
63. Putnam H. (1975). The Meaning of 'Meaning':
<https://conservancy.umn.edu/items/7bfe66c3-bfb2-4be8-953a-58c997ce90bd>, data dostępu: 27.06.2024.
64. Robison-Greene, R. D. (2021). Human Identity, Animal Identity, and Reflective Endorsement. *Between the Species*, 24(1):
<https://digitalcommons.calpoly.edu/cgi/viewcontent.cgi?article=2223&context=bts>, data dostępu: 17.06.2024.
65. Robledo, J. P., Cross, I., Boada-Bayona, L., Demogeot, N. (2022). Back to basics: A re-evaluation of the relevance of imprinting in the genesis of Bowlby's attachment theory. *Frontiers in Psychology*:
<https://www.frontiersin.org/journals/psychology/articles/10.3389/fpsyg.2022.1033746/full>, data dostępu: 07.06.2025.
66. Rosch, E., Lloyd, B.B. (ed.). (1978). Principles of categorization. [W:] *Cognition and categorization*, Lawrence Erlbaum, s. 27-48,
https://commonweb.unifr.ch/artsdean/pub/gestens/f/as/files/4610/9778_083247.pdf, data dostępu: 09.01.2025.
67. Rousay, V. (2023). Top-Down Processing Examples in Psychology:
<https://www.simplypsychology.org/top-down-processing.html>, data dostępu: 06.04.2025.
68. Samani, R., Sharifian, F. (1997). Cross-language hierarchical spreading of activation. [W:] *Proceedings of the conference on language, cognition, and interpretation*. Isfahan: IAU Press, s. 11-23:
<http://web-archive.southampton.ac.uk/cogprints.org/1663/1/act2.htm>, data dostępu: 10.04.2025.
69. Schendan, H. E., Ganis, G. (2015). Top-down modulation of visual processing and knowledge after 250 ms supports object constancy of category decisions. *Frontiers in psychology*, vol. 6:
<https://www.frontiersin.org/journals/psychology/articles/10.3389/fpsyg.2015.01289/full>, data dostępu: 03.04.2025.

70. Schopenhauer, A. (2007). *Erystyka. Sztuka prowadzenia sporów*. Wolne lektury: <https://wolnelektury.pl/media/book/pdf/schopenhauer-sztuka-prowadzenia-sporow.pdf>, data dostępu: 01.06.2025.
71. Sevillano, V., Fiske, S. T. (2019). Stereotypes, emotions, and behaviors associated with animals: A causal test of the stereotype content model and BIAS map. *Group Processes & Intergroup Relations*, 22(6), s. 879-900: https://repositorio.uam.es/bitstream/handle/10486/693816/stereotypes_sevillano_GPIR_2019_ps.pdf?sequence=1, data dostępu: 31.05.2025.
72. Super, C. M., Harkness, S. (2010). *Culture and early childhood development: Implications for policy and programs*. *Humanity & Social Sciences Journal*, 5(1): https://www.researchgate.net/publication/286809793_Culture_and_Early_Childhood_Development_Implications_for_policy_and_programs, data dostępu: 04.05.2025.
73. Szymanek, K. *Argument z podobieństwa*. Wydawnictwo Uniwersytetu Śląskiego, Katowice: https://rebus.us.edu.pl/bitstream/20.500.12128/476/5Szymanek_Argument_z_Podobieństwa.pdf, data dostępu: 29.09.2023.
74. Sport-dog.pl. (2022) <https://sport-dog.pl/pl/i/Jak-szkolone-sa-psy-do-wykrywania-narkotykow/75>, data dostępu: 10.05.2024.
75. Stjernberg, F. (n.d.). *Theory of mind*. Internet Encyclopedia of Philosophy. <https://iep.utm.edu/theomind/>, data dostępu: 03.05.2025.
76. Szczucińska, P. (2018). Psy lawinowe: Cztery łapy na ratunek w śniegu. *National Geographic*, <https://www.national-geographic.pl/artukul/cztery-lapy-na-ratunek>, data dostępu: 10.05.2024.
77. Śliwak, J. (1999). Altruizm w koncepcji socjobiologicznej. *Roczniki Psychologiczne*, 2(1), s. 111-123: http://cw.kul.lublin.pl/Content/21146/RPsych_1999_vol02_111-123_SLIWAK.PDF, data dostępu: 06.04.2025.
78. Tadeusiewicz R. (2020). *Gęsi zamiast psów policyjnych?*, *Gazeta Krakowska*, <https://plus.gazetakrakowska.pl/gesi-zamiast-psow-policyjnych/ar/c15-15101172>, data dostępu: 29.07.2020.
79. Tomasello, M., Moll, H. (2010). The gap is social: Human shared intentionality and culture. [W:] P. M. Kappeler J. B. Silk (red.), *Mind the gap: Tracing the origins of*

- human universals*, Springer, s. 331–349, https://doi.org/10.1007/978-3-642-02725-3_16: https://dornsife.usc.edu/henrike-moll/wp-content/uploads/sites/404/2023/11/Tomasello_Moll_2010.pdf, data dostępu: 07.06.2025.
80. Tulving, E., Thomson, D. M. (1973). Encoding specificity and retrieval processes in episodic memory. *Psychological Review*, 80(5), s. 352-373: <https://alicekim.ca/9.ESP73.pdf>, data dostępu: 31.05.2025.
81. Vaughan, G.M. (1988). The psychology of intergroup discrimination. *New Zealand Journal of Psychology*, vol. 17, s. 1-14: <https://www.psychology.org.nz/journal-archive/NZJP-Vol171-1988-1-Vaughan1.pdf>, data dostępu: 07.01.2024.
82. Wagemans, J. (2003). Conceptualizing fallacies: The informal logic and pragma-dialectical approaches to the argumentum ad ignorantiam. [W:] *Proceedings of the Fifth Conference of the International Society for the Study of Argumentation. Proceedings of the fifth conference of the International Society for the Study of Argumentation*, Amsterdam, s. 1049-1051; https://pure.uva.nl/ws/files/1460924/91514_Wagemans_2003_Conceptualizing_fallacies.pdf, data dostępu: 14.03.2024.
83. Wardle, S. G. (2023, June 13). What it means if you can see faces in objects. [Film]. (YouTube): https://www.youtube.com/watch?v=ADRES_FILMU, data dostępu: 16.04.2025.
84. Whiten, A. (1996). When does smart behaviour-reading become mind-reading? [W:] P. Carruthers P. K. Smith (Eds.), *Theories of theories of mind*, Cambridge University Press, s. 277-293, <https://web.media.mit.edu/~cynthiab/Readings/whiten-96.pdf>, data dostępu: 06.03.2025.
85. Wieczorek K.A. (2017) *Argumenty ad ignorantiam*, *Filozofuj!*, 3(15), s. 28–29: https://filozofuj.eu/wp-content/uploads/2022/05/fi015_v1.pdf, data dostępu: 14.03.2024.
86. Wiersze dla Juniora: https://wiersze.juniora.pl/brzechwa/brzechwa_z07.html, data dostępu: 25.07.2024.
87. Wikipedia, pokój Ames:

- https://upload.wikimedia.org/wikipedia/commons/thumb/3/37/Ames_room-pl.svg/330px-Ames_room-pl.svg.png, data dostępu: 25.07.2024.
88. Wilson, E. O. (1986). *Biophilia*. Harvard university press:
https://books.google.pl/books?hl=pl&lr=&id=gNAuEAAQBAJ&oi=fnd&pg=PP6&dq=Biophilia++-+Wilson,+E.O.,1984&ots=xwVqPwhxBR&sig=CZG-4k42ULeD8xRIV1apECvTXgE&redir_esc=y#v=onepage&q&f=false, data dostępu: 22.03.2024.
89. Woleński, J. (2015). Szowinizm gatunkowy, humanitaryzm i animalocentryzm. *Przegląd Filozoficzny. Nowa Seria*, 24(2 (94)):
https://pf.uw.edu.pl/images/NUMERY_PDF/094/2-152_Wolenski.pdf, data dostępu: 29.09.2023.
90. Woźniak, A. (2016). Dyskurs wegański i wegetariański a szowinizm gatunkowy w memach internetowych. *Zoophilologica. Polish Journal of Animal Studies*, vol. 2, s. 151-164, https://bazhum.muzhp.pl/media/texts/zoophilologica-polish-journal-of-animals-studies/2016-tom-2/zoophilologica_polish_journal_of_animals_studies-r2016-t2-s151-164.pdf, data dostępu: 01.06.2025.
91. Wróblewski, Z. (2016). Umysły zwierząt: między naiwnym antropomorfizmem a dogmatycznym antroponegacjonizmem. *Zoophilologica. Polish Journal of Animal Studies*, vol. 2, s. 83-95,
https://bazhum.muzhp.pl/media/texts/zoophilologica-polish-journal-of-animals-studies/2016-tom-2/zoophilologica_polish_journal_of_animals_studies-r2016-t2-s83-95.pdf, data dostępu: 01.06.2025.
92. Vsauce (2015, 1 listopada). *Juvenioia* [Film], (YouTube):
https://www.youtube.com/watch?v=LD0x7ho_IYc, data dostępu: 14.04.2025.
93. Vsauce. (2018, 5 grudnia). *The cognitive tradeoff hypothesis* [Film], (YouTube):
<https://www.youtube.com/watch?v=ktkjUjcZid0>, data dostępu: 09.02.2025.
94. Trojan, M. (05.11.2013). Na tropie zwierzęcego umysłu [cz.I], [Film], (YouTube):
<https://www.youtube.com/watch?v=ZU-jaNwMF30&t=1297s>, data dostępu: 30.08.2023.
95. Zakrzewska, D. Wzajemne uprzedzenia wśród uczniów szkół licealnych z Polski i Izraela¹. *Bliżej siebie. Edukacja i dialog w grupie niejednorodnej kulturowo*, vol. 77, s. 77-92,

https://www.zywymost.org.pl/wp-content/uploads/2016/07/Bli%C5%BCej-siebie_srodek_nr6.pdf, data dostępu: 01.06.2025.

96. Zielińska, I. (2004). Media, interes, panika moralna. Nowa kategoria socjologiczna i jej implikacje:

https://depot.ceon.pl/bitstream/handle/123456789/6060/Zielinsk_KIS%5b1%5d.pdf?sequence=1&isAllowed=y, data dostępu: 20.02.2024.

97. Zowczak, M. (1984). Mit bohaterski jako opowieść o granicach ludzkich możliwości. *Etnografia Polska*, 28(2), s. 243-267,

<https://zbiory.cyfrowaetnografia.pl/public/1093.pdf>, data dostępu: 18.06.2024.

Summary:

The aim of this work is to analyze the influence of prejudices — understood as attitudes — on the interpretation of research on the cognitive processes of non-human animals, with particular emphasis on theory of mind. I am interested in whether, and if so, to what extent, researcher's prejudices affect their assessment of non-human animal's ability to use theory of mind. Theory of mind, that is, the competence to read, predict, and evaluate the beliefs of other individuals, constitutes — in the approach adopted and substantiated in this dissertation — a manifestation of the so-called mind-reading, meaning the ability to access others' mental states. One of the theses that play an important role in my analysis is the claim that anthropocentric assumptions, adopted by many researchers, significantly distort the assessment of the mental abilities of non-human animals by, among other things, denying them the competence to create a theory of mind. The main research problem of the work concerns the question of to what extent the adopted interpretative frameworks, based on anthropocentric assumptions, distort the assessment of the mental capabilities of non-human animals?

The dissertation discusses two main interpretative perspectives functioning in the indicated problem area: mentalistic, which assumes the possibility of constructing explanations of non-human animal behavior by referring to mind-reading; and behavioral, postulating limiting explanations to simpler mechanisms, the ability of the so-called clever behavior reading, where an animal not having the ability to read other individual's mental states can efficiently or cleverly predict other people's behavior (hence the term clever behavior reading). The research method used to achieve these goals is primarily conceptual analysis. It contains a complex analysis of the epistemic and methodological consequences of adopting specific interpretative frameworks. Additionally, I also reconstruct the hidden assumptions influencing the assessment of non-human animal's mental capabilities, which I present and discuss in detail using meta-analytical methods.

The paper is situated in the interdisciplinary context of philosophy of mind, epistemology and cognitive psychology. It addresses a socially and scientifically important topic, revealing that interpretations of the results of experiments involving non-human animals are shaped by prejudices, that can lead to epistemic and moral injustice. The approach used is analytical and critical, but it does not aim to resolve the empirical question of whether non-human animals are capable of forming a theory of mind. Rather, it shows how theoretical assumptions resulting from prejudices shape both research practices and interpretative frameworks used in empirical sciences and in the construction of theoretical models in

philosophy of mind. Thus, it draws attention to the epistemological, ethical, and methodological challenges associated with studying the mind-reading capacities of others.

The conclusions drawn from this multidimensional and multilayered analysis indicate that biases — both explicit and implicit — influence the design of experiments, the selection of data, and the interpretation of results. They also emphasize the need for reflection on the language of science, theoretical frameworks, and cognitive categories used in research on non-human animals. The use of the concept of epistemic injustice opens a discussion on the responsible relationship between humans and other animals, which must be based on honest narratives about the cognitive abilities of non-human animals. The author recommends further research on the influence of affect on the interpretation of results and the development of more methodologically neutral approaches to the study of non-human minds.

OŚWIADCZENIE DOKTORANTA

Janina Mękarska

imię i nazwisko

Uniwersytet Szczeciński – Instytut Filozofii i Kognitywistyki

Szkoła Doktorska Uniwersytetu Szczecińskiego

Forma: stacjonarne

Oświadczam, że moja praca pt.:

Wpływ uprzedzeń na interpretacje badań procesów poznawczych zwierząt innych niż ludzie,
na tle debaty nad teorią umysłu:

- a. została napisana przeze mnie,
- b. nie narusza praw autorskich w rozumieniu ustawy z dnia 4 lutego 1994r. o prawie autorskim i prawach pokrewnych (Dz. U. Nr 24, poz. 83 z póź. zm.) oraz dóbr osobistych chronionych prawem,
- c. nie zawiera danych i informacji , które uzyskałem w sposób niedozwolony,
- d. nie była podstawą nadania tytułu naukowego ani lub zawodowego ani mnie ani innej osobie.

Ponadto oświadczam, że treść pracy przedstawionej przeze mnie do obrony, zawarta na przekazywanym nośniku elektronicznym jest identyczna z jej wersją drukowaną.

Szczecin, dn.

.....

podpis

OŚWIADCZENIE

Wyrażam zgodę na udostępnienie mojej pracy pt. „Wpływ uprzedzeń na interpretacje badań procesów poznawczych zwierząt innych niż ludzie, na tle debaty nad teorią umysłu”.

..... dnia.....

miejsowość

.....

podpis