

Analiza & Egzystencja

czasopismo filozoficzne

AiE!V

67

2024

Uniwersytet Szczeciński

Rada Naukowa – Arianna Betti (Vrije Universiteit Amsterdam, Netherlands), Piotr Bołtuć (University of Illinois, Springfield, USA), Arkadiusz Chrudzimski (Uniwersytet Jagielloński, Kraków), Józef Dębowski (Uniwersytet Warmińsko-Mazurski, Olsztyn), Adam Drozdek (Duchesne University, USA), Piotr Gutowski (Katolicki Uniwersytet Lubelski Jana Pawła II, Lublin), Stanisław Judycki (Uniwersytet Gdański), Andrzej Klawiter (Uniwersytet im. Adama Mickiewicza, Poznań), Dariusz Łukasiewicz (Uniwersytet Kazimierza Wielkiego, Bydgoszcz), Josef Moural (Usti nad Labem, Czech Republic), Zbysław Muszyński (Uniwersytet Marii Curie-Skłodowskiej, Lublin), Robert Piłat (Uniwersytet Kardynała Stefana Wyszyńskiego, Warszawa), Roger Pouivet (Université de Lorraine, France), Alexander Pruss (Baylor University, USA), Andrzej Przyłębski (Uniwersytet im. Adama Mickiewicza, Poznań), Włodek Rabinowicz (Lund University, Sweden), John Skorupski (University of St Andrews, UK), Marek Szulakiewicz (Uniwersytet im. Mikołaja Kopernika, Toruń), Jerzy Szymura (Uniwersytet Jagielloński, Kraków), Monika Walczak (Katolicki Uniwersytet Lubelski Jana Pawła II, Lublin), Urszula Żeglén (Uniwersytet im. Mikołaja Kopernika, Toruń)

Lista recenzentów dostępna jest na stronie: wnus.usz.edu.pl/aie

Redaktor naczelna – Renata Ziemińska (Uniwersytet Szczeciński)

Komitet redakcyjny – Mirosław Rutkowski (Uniwersytet Szczeciński),
Adriana Schetz (Uniwersytet Szczeciński), Tadeusz Szubka (Uniwersytet Szczeciński)

Sekretarz redakcji – Katarzyna Ciarcińska

Adres Redakcji – Instytut Filozofii i Kognitywistyki Uniwersytetu Szczecińskiego
71-017 Szczecin, ul. Krakowska 71, e-mail: egzystencja@usz.edu.pl

Wersja papierowa jest wersją podstawową

Elektroniczna wersja czasopisma jest dostępna na stronie: wnus.usz.edu.pl/aie

Czasopismo jest indeksowane w bazach: SCOPUS, DOAJ (Directory of Open Access Journals), ERIH PLUS (European Reference Index for the Humanities), CEJSH (The Central European Journal of Social Sciences and Humanities), CEEOL (The Central and Eastern European Online Library) oraz BazHum

Dofinansowane ze środków Ministerstwa Edukacji i Nauki na podstawie umowy nr RCN/SP/0352/2021/1 z dnia 20.10.2022; kwota środków finansowych stanowiących pomoc przyznaną w ramach programu „Rozwój czasopism naukowych” 40 890,00 zł

Redakcja językowa – Kamil Cekiera / e-DYTOR

Korekta – Paulina Kaczyńska-Domagalska

Skład komputerowy – Iga Bańkowska

Projekt okładki – Michał Starczewski

© Copyright by Uniwersytet Szczeciński, Szczecin 2024

ISSN (print) 1734-9923 | ISSN (online) 2300-7621

WYDAWNICTWO NAUKOWE UNIwersYTETU SZCZECIŃSKIEGO

Wydanie I. Ark. wyd. 8,5. Ark. druk. 9,3. Format A5. Nakład 36 egz.

Spis treści

| | |
|--|-----|
| Tomasz Walczyk, <i>Werystyczna epistemologia społeczna i problemy komunikacji sieciowej</i> | 5 |
| Mateusz Penczek, <i>Autonomia osobowa z perspektywy relacyjnej teorii autonomii Catriony Mackenzie</i> | 27 |
| Malwina Rolka, <i>Filozofia homo romanticus w zwierciadle teorii literackiej Friedricha Schlegla</i> | 55 |
| Adriana Schetz, <i>Geneza zachowań instynktownych</i> | 77 |
| Janina Mękaraska, <i>Błędy poznawcze związane z szowinizmem gatunkowym oraz odmawianie teorii umysłu zwierzętom innym niż ludzie</i> | 109 |
| Piotr Biłgorajski, <i>Wyobrażeniowy opór</i> | 127 |

Contents

| | |
|--|-----|
| Tomasz Walczyk, <i>Veritistic social epistemology and network communication problems</i> | 5 |
| Mateusz Penczek, <i>Personal autonomy according to Catriona Mackenzie's relational theory of autonomy</i> | 27 |
| Malwina Rolka, <i>Philosophy of homo romanticus in the mirror of Friedrich Schlegel's literary theory</i> | 55 |
| Adriana Schetz, <i>Genesis of Instinctive Behavior</i> | 77 |
| Janina Mękarska, <i>Cognitive biases associated with speciesism and the denial of theory of thought to non-human animals</i> | 109 |
| Piotr Biłgorajski, <i>The imaginative resistance</i> | 127 |



„Analiza i Egzystencja” 67 (2024), 5–26
ISSN (print): 1734-9923
ISSN (online): 2300-7621
DOI: 10.18276/aie.2024.67-01

ARTYKUŁY

TOMASZ WALCZYK

Uniwersytet Marii Curie-Skłodowskiej w Lublinie

ORCID: 0000-0002-4858-7336

email: tomasz.walczyk@mail.umcs.pl

Werystyczna epistemologia społeczna i problemy komunikacji sieciowej

Słowa kluczowe: epistemologia społeczna, werytyzm, prawda, rzetelność, dezinformacja, niesprawiedliwość epistemiczna, komora echa

Keywords: social epistemology, veritism, truth, reliability, disinformation, epistemic injustice, echo chamber

Veritistic social epistemology and network communication problems

Abstract

The project of veristic social epistemology is based on the evaluation of social practices directed at acquiring knowledge and avoiding error. This research aims to analyse the social and technological dimensions of information processes. A number of practices in the network environment have a significant impact on the cognitive processes of individuals. They give rise to the acquisition of both true and false beliefs, ranging from reliable information practices to unreliable disinformation practices. The prevalence of phenomena such as echo chambers, fake news, clickbait or deepfakes indicates that the condition of the contemporary infosphere is under serious threat. Disinformation processes, reinforced by technological progress, prompt reflection on the reliability of social practices. In addition to the strictly veristic consequences, the importance of the phenomenon of epistemic injustice should also be pointed out.

Wprowadzenie

Obecne przeobrażenia w obrębie technologii i strategii komunikacyjnych znajdują szeroki wydzźwięk w analizie wiedzy. Problematyce tej poświęca się wiele uwagi w ramach werystycznej epistemologii społecznej (dalej: WES), badając rozmaite zjawiska informacyjne pod kątem ich związku z wiedzotwórczymi praktykami społecznymi, a także, budzącymi niepokój, powszechnymi praktykami dezinformującymi. Projektowane i tworzone systemy komunikacyjne są coraz bardziej zautomatyzowane, wzrasta ich przepustowość informacyjna, co wiąże się dodatkowo z trudnościami z rozpoznawaniem prawdziwości publikowanych treści. Sporą część zadań w przestrzeni informacyjnej przejmują generatywne sztuczne inteligencje, jak choćby często przywoływany w debacie publicznej Chat GPT, należący do klasy dużych modeli językowych (ang. *large language models*). Wzrasta przy tym ryzyko powszechnej manipulacji z powodu samej istoty zapośredniczonej komunikacji masowej opartej na najnowszych technologiach¹. Jednym z istotnych wyzwań jest kwestia (względnej) anonimowości internautów. Brak odpowiedzialności za publikowane treści może stanowić zachętę do rozprzestrzeniania fałszywych informacji. Presja wywierana jest także odgórnie, ze strony instytucji zarządzających przepływem informacji. Dzieje się tak z uwagi na możliwość pozyskiwania szczegółowych danych o użytkownikach sieci oraz tworzenia spersonalizowanego strumienia informacji, opartego na funkcjonowaniu algorytmów profilujących. Konsekwencją rejestrowania aktywności sieciowej jest filtrowanie informacji, co istotnie zawęża zakres informacji docierającej do odbiorcy. Towarzyszące temu zjawiska to bańki epistemiczne (np. bańki filtrujące, ang. *filter bubbles*) oraz komory echa (ang. *echo chambers*), w obrębie których następuje pomijanie, a nawet odrzucanie i dyskredytowanie wielu istotnych (w tym rzetelnych) źródeł informacji (Nguyen, 2020). Ułatwia to rozpowszechnianie dezinformacji, co jest jednocześnie przyczyną polaryzacji poglądów i trudności komunikacyjnych. Paletę zagrożeń poszerza możliwość generowania spreparowanych treści wizualnych, co odnosi się do zjawiska o nazwie *deepfake*.

¹ Wskazany problem został szczegółowo omówiony w dyskusji na temat teleepistemologii (Goldberg, 2001), choć postęp w zakresie stosowanych technologii wymaga aktualizacji prowadzonych wówczas badań.

Środowisko sieciowe jest także platformą łatwego rozprzestrzeniania się zachowań i treści opartych na uprzedzeniach i dyskredytujących część uczestników sfery publicznej, co wiąże się z problemem niesprawiedliwości epistemicznej (Fricker, 2009; Frost-Arnold, 2014, 2023).

Środowisko, na gruncie którego wyrastają społeczne praktyki pozyskiwania informacji, jeszcze nigdy wcześniej nie było tak złożoną przestrzenią wielopoziomowych relacji. Refleksja na temat ludzkiego poznania nie może być zogniskowana jedynie na analizie kryteriów wiedzy stosowanych w obrębie jednostkowego podmiotu poznającego. Zgodnie z koncepcją Alвина Goldmana, jednego z najważniejszych przedstawicieli epistemologii społecznej, należy uwzględnić także działalność szeroko rozumianych systemów poznawczych, takich jak zespoły badawcze, składy sędziowskie, komitety czy korporacje. „Zakładając, że takie grupy posiadają przekonania, epistemologia społeczna zadaje pytanie, w jaki sposób te grupowe przekonania są uzasadnione, racjonalne lub konstytuują wiedzę” (Goldman, 2020, s. 11). Przemiany społeczne zachodzą rzecz jasna z udziałem zaawansowanych technologii. Badać należy zatem także przemiany w zakresie technologii telekomunikacyjnych, które w dynamiczny sposób dokonują gruntownych modyfikacji w obrębie funkcjonowania grup społecznych oraz instytucji. Wraz ze wzrostem zaawansowania stosowanych narzędzi, zmienia się także przestrzeń stosowanych praktyk poznawczych. Obok społecznego wymiaru wiedzy, istotny i nierozzerwalnie z nim związany jest wymiar technologiczny.

Jakie są konsekwencje tego mariażu, który nieustannie reżyseruje spektakl ludzkich interakcji i kierunki jego rozwoju? Dobrym punktem wyjścia analizy złożoności współczesnego społeczeństwa informacyjnego są szeroko zakrojone badania prowadzone w obrębie werystycznej epistemologii społecznej. Pozwalają one uchwycić problem wiedzy i dezinformacji w szerokim kontekście społeczno-technologicznym, wspomagając ocenę praktyk poznawczych pod kątem ich roli w nabywaniu prawdziwych lub fałszywych przekonań. Badania te zawierają niezwykle istotną lekcję dotyczącą metod neutralizowania bądź unikania wpływu praktyk dezinformujących. Mogą zatem być potrzebnym wsparciem w wymagających próbach nawigowania po burzliwej współczesnej infosferze², tak silnie naznaczonej wpływem różnorodnych, często zdradliwych prądów.

² Alvin Goldman stosuje pojęcie infosfery na określenie ogółu stanów informacyjnych. Mentalna infosfera odnosi się do ogółu przekonań, jakie żywią wszyscy ludzie

Werystyczna epistemologia społeczna i model praktyk społecznych

Ludzie z natury mają silną potrzebę zdobywania wiedzy i unikania fałszywych przekonań. Jej podstawą jest ciekawość, a także względy praktyczne, co jest integralnie wplecione w system wartościowania praktyk społecznych (Goldman, 2021a)³. Zdaniem Alvina Goldmana prototypem działań nakierowanych na zdobywanie wiedzy jest procedura zadawania pytań, które dotyczą interesujących zagadnień. Zainteresowanie konkretnym zagadnieniem można przypisać zarówno jednostkom, jak i grupom (np. składom sędziowskim). Projekt werystycznej epistemologii społecznej skupia się wokół ewaluacji społecznych praktyk pod kątem ich skuteczności w zdobywaniu wiedzy (Goldman, 2021a).

W jaki sposób wartość werystyczna łączy się z epistemologią społeczną? Nie ma nic społecznego w tym, że pojedynczy agent ma prawdziwe lub fałszywe przekonanie. Słusznie, ale szeroka gama praktyk i instytucji społecznych może mieć przyczynowy wpływ, często ogromny, na postawy jednostek, przechylając je albo w kierunku prawdziwych, albo w kierunku fałszywych przekonań. Duże sektory wymiany społecznej obejmują przekazywanie komunikatów – często zawierających informacje, dezinformacje, częściowe informacje. Praktyki komunikacyjne, które mają miejsce w tych sieciach społecznych mogą być badane z punktu widzenia ich wpływu na poziomy wartości werystycznych wielu jednostek.

(Goldman, 2009, s. 7)

Podstawę oceny werystycznej stanowi rzetelność tych praktyk, co przekłada się na wysoki współczynnik nabywania prawdziwych przekonań. Goldman wskazuje na wyraźną paralelę między społecznymi praktykami analizowanymi na gruncie epistemologii społecznej a problematyką rzetelności procesów poznawczych, opracowaną w ramach stanowiska zwanego

(Goldman, 2021, s. 274). Termin ten spopularyzował Alvin Toffler w książce *Trzecia fala*: „W każdej cywilizacji musi też istnieć »infosfera«, której zadaniem jest produkowanie i rozpowszechnianie informacji” (Toffler, 1985, s. 52).

³ Warto odnotować, że projekt epistemologii społecznej Goldmana skupiony jest wokół analizy procesów tworzenia wiedzy w „słabym” sensie, czyli wiedzy jako prawdziwego przekonania.

reliabilizmem (Goldman 1986, 2021b)⁴. Werystyczną epistemologię społeczną należy w sposób zdecydowany odróżnić od epistemologii społecznej opartej na kontestowaniu znaczenia prawdy korespondencyjnej, racjonalności i rzetelności. Reprezentantami tego stanowiska są Michel Foucault, Thomas Kuhn, Bruno Latour, Barry Barnes, David Bloor, Steven Shapin oraz Richard Rorty. Mimo wielu rozbieżności, Goldman sygnalizuje zaśluzoną rolę wskazanych badaczy w podkreślaniu społecznego charakteru poznania i wiedzy (Goldman, 2020, s. 11).

W ramach WES Goldman wyróżnia kilka wersji odnoszących się do odmiennych przejawów społecznych praktyk poznawczych (Goldman, 2020). Po pierwsze, możemy mówić o interpersonalnej WES, gdzie rozpatruje się sytuacje, w których indywidualny podmiot opiera się na zewnętrznych źródłach informacji, dokonując na ich podstawie rewizji własnych poglądów. Wiąże się to z pytaniem, w jakich warunkach zasadne jest zaufanie wobec twierdzeń postulowanych przez innych uczestników debaty. Po drugie, wskazuje kolektywną wersję WES, której przedmiotem badań są grupy i podmioty zbiorowe, takie jak jury, komitety, składy sędziowskie czy – obecnie szczególnie istotne – organizacje fact-checkingowe. „Kiedy epistemologowie starają się ocenić racjonalność i zasadność grupowych przekonań, pracują na (społeczno-) epistemologicznym terytorium” (Goldman, 2020, s. 12). Trzeci wariant WES jest instytucjonalny, skupiony wokół analizy wpływu szeroko rozumianych systemów (np. instytucji, sieci, rynku) na przepływ informacji wśród członków wspólnoty. W obrębie tej odmiany WES szerokich systemów nie traktuje się jako posiadaczy stanów doksastrycznych; istotę stanowi tu odgórne oddziaływanie systemu na przekonania jednostek i grup społecznych⁵.

⁴ Zgodnie z reliabilizmem Goldmana wiedzę należy przypisać osobie, która jest przekonana, że *p*; *p* jest prawdą; przekonanie, że *p* jest wynikiem rzetelnego procesu przekonaniotwórczego. W stosunku do klasycznej definicji wiedzy, zmianie ulega internalistycznie pojęty warunek uzasadnienia, który zastąpiony zostaje eksternalistycznym warunkiem rzetelności procesu (Goldman, 1986). Stopień rzetelności należy utożsamić z tendencją danego procesu do formułowania prawdziwych przekonań.

⁵ Dobrym przykładem są działania w sieci, opierające się na liczbie subskrypcji i polubień (tzw. lajków), co może wpływać pośrednio także na jakość publikowanych treści (m.in. ich prawdziwość, współczynnik błędów, wiarygodność). Wiąże się to wyraźnie ze zjawiskami bańki filtrującej, clickbaitu, a także fake newsów, które z perspektywy zasięgu i korzyści finansowych są często dużo bardziej skuteczne. Naczelną funkcją WES w tym zakresie powinna być analiza systemów transferujących informacje oraz

Proponowany model analizy werystycznej, uwzględnia stopień przekonania oraz wpływ zastosowanych praktyk na zmiany w obrębie stopnia przekonania⁶. Stosowany jest przy tym dwupoziomowy schemat wartościowania. Po pierwsze, stany takie jak wiedza, błąd i niewiedza posiadają fundamentalną wartość lub antywartość werystyczną (tzw. W-wartość). Po drugie, praktyki mają instrumentalną wartość werystyczną, co odnosi się wprost do wspierania lub utrudniania osiągnięcia fundamentalnej wartości werystycznej. Przez wzgląd na interesujące zagadnienie można mieć różne nastawienia do danego sądu: dawać mu wiarę czy też zdecydowanie go odrzucić lub zawiesić. „Jeżeli S daje wiarę zdaniu prawdziwemu, to W-wartość wynosi 1. Jeżeli odrzuca zdanie prawdziwe, to W-wartość wynosi 0. Jeżeli zawiesza sąd, to W-wartość wynosi 0,5. Pierwszy stan to wiedza, drugi błąd, a trzeci niewiedza” (Goldman, 2021, s. 158). Analiza werystyczna skupia się w prezentowanym ujęciu na zmianie W-wartości w czasie.

Załóżmy, że S był w czasie t_1 w pełni przekonany, że p (stopień przekonania o wartości 1) i jednocześnie p jest prawdą. Jeśli w obliczu wpływu, jaki wywarła na S dana praktyka społeczna, S w czasie t_2 zaczął odrzucać w całą stanowczością p (stopień przekonania o wartości 0), instrumentalna W-wartość tej praktyki jest ujemna i wynosi -1 . Zatem jeśli praktyka społeczna obniża fundamentalne W-wartości, jest to praktyka nieefektywna z perspektywy werytyzmu (a więc praktyka nierzetelna). Im większy wpływ takich praktyk, tym gorzej dla kondycji społeczeństwa informacyjnego. I odwrotnie, jeśli praktyka społeczna w określonym przedziale czasowym podwyższa fundamentalną W-wartość, jest to praktyka cenna z uwagi na cel werytyzmu, jakim jest poszukiwanie prawdy. Instrumentalna W-wartość takiej praktyki jest dodatnia i wpływa na przyrost prawdziwych przekonań. Należy uwzględnić ponadto dwa, zasadniczo odmienne, cele WES. Po pierwsze, istotne jest wskazanie i ewaluacja praktyk pod kątem ich zdolności do tropienia prawdy i wiedzy. Po drugie, badania powinny

metod minimalizacji wpływu praktyk dezinformujących czy treści tworzonych *ad hoc*, bez należytej kontroli.

⁶ Goldman utożsamia stopień przekonania z kategorią subiektywnego prawdopodobieństwa. Subiektywny stopień przekonania 0,65 w odniesieniu do danego zdania równy jest subiektywnemu prawdopodobieństwu 0,65, że zdanie to jest zgodne z prawdą. Stosuje przy tym model nieskończenie wielu stopni przekonania, gdzie konkretny stopień przekonania reprezentowany jest przez określony punkt przedziału od 0 do 1 (Goldman, 2021, s. 156).

skupić się na rozpoznaniu i ocenie praktyk wspierających unikanie błędów. „Niektóre praktyki epistemiczne pomagają nam odrzucić fałsz, ale robią to, pozostawiając nas w niewiedzy (tj. zawieszonym sądzie), co różni się od akceptacji prawdy” (Frost-Arnold, 2014, s. 67). Oczywiście sprawa ulega dodatkowej komplikacji, jeśli uwzględnimy stopień złożoności współczesnej infosfery, w której poszczególne praktyki tworzą zaawansowane systemy praktyk (np. praktyki realizowane w środowisku sieciowym, przy użyciu generatywnej SI)⁷.

Praktyki społeczne realizowane są w różnych aspektach interakcji społecznych, zarówno w wymiarze prywatnym, jak i publicznym. Prywatny odnosi się do praktyk komunikacyjnych, które opierają się na przekazywaniu świadectw innym jednostkom lub argumentacji na rzecz własnych poglądów w celu przekonania interlokutorów do swoich racji. Wymiar publiczny stanowią z kolei regulowane w wysokim stopniu praktyki publiczne, wdrażane w oparciu o ustalone uprzednio procedury (np. procesy sądowe). Na obu poziomach wymiany informacji ludzie komunikują się za pośrednictwem systemów elektronicznych, bez których nie można już sobie wyobrazić płynnie funkcjonującej infosfery. Ocena wiedzotwórczych praktyk społecznych to niewątpliwie trudne zadanie w warunkach oddziaływania tak złożonej i dynamicznej sieci. Z tego powodu jest to przedsięwzięcie wystawione na wiele problemów, jak choćby dotyczący wskazania W-lepszych praktyk w całym systemie sieciowego oddziaływania. Przykładowo, w jaki sposób skutecznie ocenić przydatność praktyki pozyskiwania informacji za pośrednictwem Chatu GPT? Zależy to od wielu czynników, m.in. z jakiej jego wersji korzysta użytkownik systemu, w jakim zakresie opanował on zdolność wydawania instrukcji czy też w jaki sposób w przyszłości rozwijać się będzie wskazana technologia. Mimo to kierunek, który obiera werytyzm, wydaje się słuszny; nie powinno się rezygnować z prób ewaluacji informacyjnych praktyk społecznych, gdyż mogłoby to skutkować chaosem na wielką skalę (biorąc pod uwagę dynamikę postępu w dziedzinie komunikacji

⁷ Goldman proponuje dwa tryby oceny praktyk: absolutny i porównawczy. Absolutny odnosi się do oceny każdej praktyki z osobna, porównawczy zaś ma na celu zestawienie różnych praktyk realizujących te same cele poznawcze, czyli ustalenie, która z nich jest „W-lepsza”. Metoda ewaluacji wykorzystuje średnią ważoną opartą na „historii trafień” konkretnych źródeł informacji (Goldman, 2021, s. 144–146). W odniesieniu do przekazu Goldman proponuje zastosowanie praktyki wnioskowania bayesowskiego przy trafnej ocenie warunkowych prawdopodobieństw.

sieciowej czy generatywnej SI)⁸. Cele towarzyszące tak rozumianej epistemologii społecznej mogą przyczyniać się do osłabienia wpływu praktyk dezinformacyjnych i manipulowania sferą publiczną⁹.

Internet jest przestrzenią ścierania się wielu wpływów, które zgodnie z przedstawionym modelem werystycznym przynoszą zgola odmienne rezultaty w odniesieniu do fundamentalnych W-wartości. Elektroniczne środowisko informacyjne jest nieocenionym zasobem wiedzy rozszerzonej (Pritchard, 2010, 2018)¹⁰. Oprócz istotnych korzyści, stanowi ono jednak wydajne narzędzie rozpowszechniania dezinformacji na masową skalę, co wyraźnie sygnalizują badacze inicjujący badania w ramach teleepistemologii (Dreyfus, 2001; Goldman, 2001). Potwierdzają to także współczesne

⁸ Na łamach tygodnika „The Economist” Yuval Noah Harari w pesymistycznym tonie komentuje konsekwencje zastosowania generatywnej SI. Ujmując to metaforycznie, stwierdza, że SI przyczynia się do „hakowania systemu operacyjnego ludzkiej cywilizacji”. Mając funkcję manipulowania i generowania treści językowych, a także dźwięków czy obrazów, SI zyskała praktycznie nieograniczony dostęp do fundamentu ludzkiej kultury, jakim jest język naturalny (Harari, 2023). Wysoką rangę problemu potwierdzają także szeroko zakrojone prace nad aktem regulującym stosowanie narzędzi SI, np. *The AI Act*, <https://artificialintelligenceact.eu/>.

⁹ Przykładem tak rozumianej ewaluacji może być analiza porównawcza publikacji serwisów informacyjnych, prowadzona pod kątem współczynnika prawdziwych i fałszywych informacji. Załóżmy, że nasza bezstronna analiza porównawcza pod kątem trafności źródeł informacji wykazuje, że odsetek trafności publikowanych treści w serwisie x wynosi 95%, y oraz z odpowiednio – 75% i 55%. W konsekwencji za najbardziej wiarygodne źródło informacji winien uchodzić serwis x. Nie zawsze tak się jednak dzieje, dlatego też Goldman wskazuje na konieczność uwzględnienia ujemnej wartości instrumentalnej danej praktyki. Zdarza się, że to właśnie serwis o niskim współczynniku trafności publikowanych treści, choć dużym ładunku emocjonalnym, stanowi dla kogoś główne źródło informacji. Dodatkowym problemem trybu porównawczego jest możliwość uwzględnienia właściwie nieograniczonej liczby praktyk, co przekłada się na zasadniczą trudność w analizie danych i ich ocenie werystycznej. Nie zmienia to jednak potrzeby monitorowania publikowanych treści pod kątem ich zgodności z faktami, w czym mogą być pomocne analizy statystyczne, badania ankietowe, wywiady z niezależnymi i bezstronnymi ekspertami itp. W trybie porównawczym wszystkie wspomniane praktyki tworzą system i, o ile są realizowane rzetelnie, zwiększają szansę na detekcję nośników dezinformacji.

¹⁰ Wiedza rozszerzona wylania się na gruncie wzajemnej interakcji między podmiotem, jego zdolnościami poznawczymi i dodatkowymi funkcjonalnościami ze strony artefaktów poznawczych, co wspólnie owocuje daleko idącymi transformacjami w obrębie procesów pamięciowych (np. notatnik, dysk zewnętrzny), percepcyjnych i uwagowych (np. technologia poszerzonej rzeczywistości, ang. *augmented reality*).

badania nad zjawiskiem postprawdy¹¹ (d'Ancona, 2018; McIntyre, 2019). Mimo wielu słusznych obaw dotyczących kondycji infosfery sieciowej, warto unikać pułapek zastawionych przez nadmierny technopesymizm, pamiętając jednocześnie o potrzebie umiarkowanego dystansu wobec technoentuzjazmu identyfikowanego jako transhumanizm. Należy skupić się na możliwych do wdrożenia rozwiązaniach, dostrzegając zarówno szanse, jak i zagrożenia dla sfery publicznej, powodowane niesłuchanym wzrostem dynamiki postępu komunikacyjnego. Istotne jest przy tym zachowanie w centrum zainteresowań podstawowego celu werytyzmu, jakim jest ewaluacja praktyk społecznych pod kątem ich rzetelności i trafności. Jest to zadanie trudne do wykonania, aczkolwiek realizacja tego ambitnego projektu jest jak najbardziej możliwa. W związku z tym należy przyjrzeć się zjawiskom sieciowym pod kątem ich oddziaływania na procesy tworzenia i pozyskiwania wiedzy. Jakie praktyki okażą się pomocne w tym przedsięwzięciu, a jakie mają negatywne konsekwencje na werystyczną kondycję współczesnej infosfery?

Problem dezinformacji w komunikacji sieciowej

Komunikacja między ludźmi stanowi podstawowe ogniwo w procesach rozpowszechniania i zdobywania wiedzy. Bez trudu dostrzegamy

jak wszechobecna jest komunikacja w naszym życiu; w efekcie łatwo jest przyznać, że jesteśmy stale zanurzeni w przepływach komunikacyjnych, które kształtują i determinują kluczowe aspekty naszego codziennego życia. Społeczeństwo, w którym żyjemy, od dawna określa się jako »społeczeństwo komunikacji«: społeczeństwo, w którym komunikacja nie tylko odgrywa kluczową rolę, ale także zasadniczo charakteryzuje, co dla jednostki oznacza bycie jego częścią. Jesteśmy członkami naszego społeczeństwa w zakresie, w jakim utrzymujemy z sobą stosunki komunikacyjne. Relacje te definiują nasze statusy społeczne, kształtują kulturę i wartości oraz determinują procesy gospodarcze i systemy produkcji.

(Turbanti, 2022, s. 5)

¹¹ Postprawda łatwo upowszechnia się w wyniku zapośredniczonej i masowej komunikacji. Pojęcie to odnosi się do obrazu rzeczywistości, w którym najważniejszą rolę w kształtowaniu opinii publicznej przypisuje się emocjom i osobistym przekonaniom, a nie odwołaniu do faktów.

Rozwój strategii komunikacyjnych jest koniecznym warunkiem przetrwania, co wiąże się z dynamicznie postępującym, ewolucyjnym procesem transmisji kulturowej (Tomasello, 2002). Istotną konsekwencję stanowi tu ekonomizacja procesów poznawczych, która jest wynikiem odpowiedniego wykorzystania zasobów wiedzy i zdolności innych członków wspólnoty komunikacyjnej oraz dostępnej technologii.

W szerokim sensie komunikacja to ogół biologicznych, psychicznych, społecznych i technicznych procesów przekazywania informacji. W sensie węższym komunikowanie odnosi się do przekazywania informacji między jednostkami. Przedmiotem dociekań prowadzonych w obrębie WES jest komunikacja w tym pierwszym sensie, co dotyczy szeroko rozumianych systemów przetwarzania informacji. W najszerszej wersji system taki należałoby utożsamić z globalnym systemem informacyjnym, w którym zachodzi integracja na poziomie jednostek, grup społecznych, instytucji i technologii, czego efektem są ich wzajemne relacje, procesy, procedury oraz bogaty zasób stosowanych praktyk społecznych¹².

Wraz z postępem technologicznym narzędzia komunikacyjne ulegają znaczącym zmianom, co wpływa na spektrum stosowanych praktyk komunikacyjnych oraz zakres ich zastosowań. Synchroniczne formy komunikacji, które wymagały natychmiastowego odczytania komunikatu (np. mowa, sygnały dymne), stopniowo tracą na znaczeniu na rzecz asynchronicznych technologii komunikacyjnych, które umożliwiają odbiór informacji w praktycznie dowolnym czasie i miejscu (warunkiem jest dostępność sieci internetowej). Proces ten, zapoczątkowany wynalezieniem pisma, zwieńczony jest możliwością przetwarzania i zapisywania ogromnej ilości danych w środowisku sieciowym (np. popularnych chmurach internetowych). To oczywiście niesie z sobą ryzyko wycieku i kradzieży danych, a także ich późniejszej manipulacji przez podmioty do tego nieuprawnione. Masowość i postępująca automatyzacja procesów informacyjnych dodatkowo potęgują to ryzyko. Jest to bezpośrednia przyczyna szerokiego

¹² Jako dobry model tak rozumianego systemu mógłby posłużyć model poznania rozproszonego (ang. *distributed cognition, d-cog*) zaproponowany przez Edwina Hutchinsa (Hutchins, 1995). W modelu tym poznanie należy pojmować jako efekt współpracy między zespołami ludzi i artefaktami. Jednostka z konieczności przenosi część zadań poznawczych na innych ludzi i narzędzia. Bez tego wsparcia wiele operacji poznawczych byłoby niemożliwych do wykonania. Podobna zależność zachodzi na ogólnym poziomie funkcjonowania społeczeństwa informacyjnego.

rozpowszechniania się różnorodnych fake newsów, z czym koresponduje zjawisko rosnącego wykładniczo ruchu internetowego (ang. *internet traffic*)¹³. Dynamiczny wzrost ilości przetwarzanych informacji, które należałoby poddać sprawdzeniu pod kątem ich trafności, prowadzi tym samym do trudności w ocenie rzetelności praktyk.

Dezinformacja jest zjawiskiem złożonym, przejawiającym się na wiele sposobów i na wielu poziomach (m.in. społecznym i technologicznym)¹⁴. Zdaniem Dona Fallisa jest rodzajem informacji rozumianej jako coś, co reprezentuje określoną część świata; coś, co ma zawartość semantyczną. Jest to także typ informacji, która jednocześnie wprowadza w błąd i czyni to w sposób nieprzypadkowy (Fallis, 2015). Dezinformację definiuje się także jako komunikat, co do którego nadawca wie, że jest fałszywy, a mimo to nadaje go z intencją wprowadzenia w błąd (Fetzer, 2004). W tym ujęciu problem stanowią oszustwa, które nie prowadzą do dezinformacji (chęć okłamania kogoś nie zawsze kończy się dezinformowaniem). Badania empiryczne wskazują ponadto, że procesy dezinformowania mają podłoże ewolucyjne i pełnią istotną funkcję adaptacyjną w środowisku. Systematyczne wysyłanie mylących sygnałów często nagradza nadawcę, pozwalając mu przetrwać w otoczeniu (Skrms, 2010). Popularnym przykładem tego typu zachowań jest wydawanie fałszywych okrzyków alarmowych, których efektem jest ucieczka niepożądanych osobników z obszaru, z którego okrzyk został wyemitowany. Prezentowane ujęcie dezinformacji jest jednak zbyt wąskie, z uwagi na to, że nie zawsze rezultatem dezinformowania jest korzyść nadawcy. Unifikując przedstawione ujęcia, ogólna definicja przybiera następującą formę: „dezinformacja jest myłącą informacją, pełniącą funkcję wprowadzania kogoś w błąd” (Fallis, 2015, s. 413)¹⁵.

¹³ Pojęcie to oznacza liczbę danych, które są przesyłane za pośrednictwem sieci komputerowej w określonym czasie. Co istotne, systematycznie wzrasta odsetek danych transferowanych przy użyciu smartfonów (*What Percentage of Internet Traffic...*).

¹⁴ Pojęcie dezinformacji silnie koresponduje ze zjawiskiem infodemii. Opisuje je Światowa Organizacja Zdrowia (WHO), uznając, że jest to „nadmierna ilość informacji na temat jakiegoś problemu, utrudniająca znalezienie rozwiązania. Może powodować wprowadzanie w błąd, dezinformację i powstawanie pogłosek podczas stanu zagrożenia zdrowia. Infodemia może utrudniać skuteczną reakcję w zakresie zdrowia publicznego oraz wywoływać dezorientację i brak zaufania wśród obywateli” (WHO, 2020).

¹⁵ Podana definicja wyklucza ze zbioru dezinformacji przypadki żartów, satyry, sarkastycznych komentarzy, a także przypadkowych kłamstw. Wyłącza także przypadki, w których intencja wprowadzenia w błąd nie prowadzi do dezinformowania,

Komunikacja sieciowa, z uwagi na swą specyfikę, może produkować wiele niezwyfikowanych treści, prowadzących w konsekwencji do wzrostu współczynnika fałszywych przekonań. Goldman, analizując blogi, podkreśla silną tendencję ludzi do szukania treści, które potwierdzają ich wcześniejsze informacje czy uprzedzenia, ignorując kontrargumenty (Goldman, 2008). Społeczności sieciowe są często „inkubatorami i wzmacniaczami niepotwierdzonych plotek i nierzetelnych opinii. [...] mogą one być ostoją marginalnych grup i ekstremistów chcących się dzielić swoimi poglądami z podobnymi sobie. Odcinając się od odmiennych zapatrywań, takie społeczności mogą umacniać się w swoich przekonaniach bez konfrontacji z przeciwnymi argumentami” (Goldman, 2021, s. 284–285). Jest to konsekwencja wąsko rozumianego, selektywnego wyboru, który powszechnie obserwuje się w przypadku zjawisk takich jak bańki epistemiczne i komory echa¹⁶.

Oba fenomeny różnią się jednak znacząco. Bańkę epistemiczną należy rozumieć jako społeczną strukturę epistemiczną, wewnątrz której niektóre z ważnych informacji zostają pominięte. Bańki epistemiczne mogą być tworzone bez intencji wprowadzania w błąd, w wyniku zwykłych procesów selekcjonowania informacji. Przykładowo, możemy poszukiwać szczególnie interesujących nas informacji, budując przy tym relacje społeczne na podstawie podobnych poglądów ekonomicznych, ekologicznych czy politycznych. Należy jednak pamiętać, że używając sieci społecznościowych jako źródła informacji, narzucamy sobie pewien filtr epistemiczny, który najczęściej pomija pluralizm poglądów, zawężając strumień dostarczanej podmiotowi informacji (Nguyen, 2020). Podobne zjawisko opisuje Eri Parisel na podstawie prostego eksperymentu wykorzystującego mechanizm funkcjonowania wyszukiwarek internetowych i mediów społecznościowych (Pariser, 2012).

np. gdy pewna osoba, chcąc wprowadzić w błąd detektywa, nieświadomie naprowadza go na właściwy trop. Mimo chęci oszustwa, komunikat nie pełni funkcji wprowadzania w błąd, a zatem nie jest dezinformacją.

¹⁶ Z opisywanym problemem korespondują badania nad błędami poznawczymi i heurystykami. Ludzie z natury mają skłonność ulegania błędom poznawczym, co dobitnie udokumentowały badania dotyczące heurystyk dostępności, zakotwiczenia oraz reprezentatywności (Tversky, Kahneman, 2012). Środowisko sieciowe, mimo wielu niewątpliwych zalet, może sprzyjać pojawianiu się i wzmacnianiu błędów poznawczych. Przykładowo, może ono potęgować tzw. efekt potwierdzenia, który odpowiada tendencji podmiotu do preferowania informacji, które wzmacniają jego stopień przekonania, często z pominięciem kwestii prawdziwości/fałszywości danej informacji.

Zwraca uwagę na wiele negatywnych skutków spowodowanych spersonalizowaniem i filtrowaniem informacji. Zadaniem baniek epistemicznych, w tym baniek filtrujących, jest personalizowanie treści, co w konsekwencji może prowadzić do blokowania przed użytkownikiem treści niepasujących do przetworzonego przez algorytmy profilu. Zdaniem Parisera algorytmy odpowiadają za budowanie unikalnego uniwersum informacyjnego, dostosowanego do każdego użytkownika sieci, opierając się na zebranych i magazynowanych danych na temat użytkowników. Możliwym efektem funkcjonowania algorytmów tego typu jest tworzenie zupełnie różnych, niewspółmiernych obrazów rzeczywistości. Rodzi to radykalizację postaw odbiorców, coraz silniej utwierdzających się w przekonaniu, że określony zasób informacji wyznacza jedyną słuszną wykładnię rzeczywistości.

W ten sposób docieramy do zjawiska komory echa, w obrębie którego zachodzi nie tylko filtrowanie informacji, ale również jej aktywne dyskredytowanie. Członkowie danej komory echa charakteryzują się nieufnością wobec tych spoza własnej przestrzeni informacyjnej. Komory echa funkcjonują poprzez izolowanie swoich uczestników od innych, niezależnych i zewnętrznych źródeł epistemicznych (Jamieson, Cappella, 2008). „Mechanizm ten wykazuje uderzające podobieństwo do niektórych opisów indoktrynacji w obrębie kultu. Dyskredytując osoby z zewnątrz, komory echa sprawiają, że ich członkowie są nadmiernie zależni od zatwierdzonych wewnątrz źródeł informacji. W bańkach epistemicznych inne głosy po prostu nie są słyszane; w komorach echa inne głosy są aktywnie podważane.” (Nguyen, 2020, s. 142).

Ze wskazanym problemem koresponduje problem niesprawiedliwości epistemicznej (ang. *epistemic injustice*)¹⁷, na który zwraca uwagę Miranda Fricker. Wyróżnia dwa wymiary problemu: niesprawiedliwość świadectwa (ang. *testimonial injustice*) i niesprawiedliwość hermeneutyczną (ang. *hermeneutical injustice*). „Niesprawiedliwość świadectwa pojawia się, gdy uprzedzenie powoduje, że słuchacz nadaje obniżony poziom wiarygodności słowu mówcy; niesprawiedliwość hermeneutyczna pojawia się na wcześniejszym etapie, gdy luka w zbiorowych zasobach interpretacyjnych

¹⁷ Obok problemu niesprawiedliwości epistemicznej, szeroko komentowanym zjawiskiem jest tzw. *implicit bias*, który odnosi się do wpływu nieuświadomionych uprzedzeń i stereotypów na podejmowane działania. „Ludzie mogą działać w oparciu o uprzedzenia i stereotypy, nie zamierzając tego robić” (Brownstein, 2019).

stawia kogoś w niesprawiedliwie niekorzystnym położeniu, gdy chodzi o nadanie sensu swojemu doświadczeniu społecznemu” (Fricker, 2009, s. 1). Przykładowo, niesprawiedliwość świadectwa objawia się w sytuacji, w której kobietę dyskredytuje się (np. umniejszając jej kompetencje) z powodu przesądu, że kobiety są z reguły nieracjonalne. Z kolei niesprawiedliwość hermeneutyczna może skutkować niezrozumieniem i krzywdą wobec osób, które doświadczają czegoś, co nie zostało w przestrzeni publicznej oficjalnie nazwane¹⁸. „Obie formy niesprawiedliwości epistemicznej są powiązane z pewnym rodzajem nierówności władzy. W przeciwieństwie do niesprawiedliwości świadectwa, niesprawiedliwość hermeneutyczna jest pojęciem strukturalnym i żaden podmiot nie utrwała niesprawiedliwości (nawet nieumyślnie). Jest to strukturalna wada wyobraźni społecznej” (Ziemińska, 2020, s. 53). Władza może radykalnie wzmacniać zjawisko niesprawiedliwości epistemicznej i jego wpływ na przestrzeń publiczną. Szczególną groźbę stanowi możliwość budowania systemu władzy i technologii w celu inwigilacji obywateli, w tym konkurentów politycznych, a w konsekwencji wywieranie na nich presji wynikającej z uzyskanej przewagi informacyjnej. Nie ma to nic wspólnego z budowaniem zaufania społecznego, opartego na swobodzie przepływu informacji i rzetelności źródeł. Zjawiska niesprawiedliwości epistemicznej mogą być z powodzeniem powielane, a nawet potężnie wzmacniane, w środowisku sieciowym. Sprzyja temu masowość i automatyzacja procesów komunikacyjnych, a także przeświadczenie bycia anonimowym, co może dodatkowo zachęcać do niestosownych zachowań w sieci¹⁹.

Dodatkowych trudności w werystycznej ewaluacji praktyk pozyskiwania informacji i wiedzy dostarczają nowe funkcje systemów opartych na generatywnej sztucznej inteligencji. Automatyzacja uzyskana przy pomocy dużych modeli językowych może spotęgować procesy powtarzania, rozwijania, umacniania czy rozsiewania fałszywych informacji, a nawet perswazji przy użyciu fake newsów, z uwagi na łatwość generowania i manipulowania treścią (Buchanan i in., 2021). Dotyczy to zarówno sfery tekstowej,

¹⁸ Jaskrawym przykładem tego typu luki był brak pojęcia molestowania seksualnego (do lat siedemdziesiątych XX wieku), co skutkowało wieloma problemami w klasyfikacji tego typu czynów jako krzywdzących.

¹⁹ Za przykład tego typu zachowań posłużyć może trollowanie (ang. *trolling*), określane jako akt „celowego prowokowania i/lub antagonizowania użytkowników w środowisku sieciowym, co często prowadzi do pożądanых, a czasem przewidywalnych rezultatów dla trolla” (Griffiths, 2014, s. 85).

jak i wizualnej, z uwagi na narzędzia wykorzystujące technologię *deepfake*²⁰. Podważa ona podstawowy standard uzasadniania twierdzeń z wykorzystaniem materiałów wizualnych, które traktowane są często jako ważne świadectwa w społecznych praktykach wymiany informacji (np. procesach sądowych czy debacie politycznej). Aby zneutralizować ten problem, niezbędne jest równoczesne rozwijanie specjalistycznej technologii przetwarzania wizualnego w celu rozwinięcia narzędzi rozpoznawania prawdziwych i fałszywych obrazów (Groh i in., 2021). Jednym z warunków *sine qua non* rzetelności praktyk społecznych funkcjonujących w obrębie filtrowania dezinformujących treści jest zatem równoległe rozwijanie technologii odpowiadających za detekcję fałszywych i zmanipulowanych treści.

Rola werytyzmu w zakresie sieciowych praktyk poznawczych

„Jak każda technologia sieć przystawia [...] zwierciadło ludzkości. Obok całego mnóstwa jej zalet dała również upust najgorszym ludzkim instynktom” (d’Ancona, 2018, s. 56). To zwierciadło, w którym dostrzegamy wzburzone oblicze współczesnego społeczeństwa informacyjnego, ukazuje złożony, w znacznej mierze niepokojący obraz. Składają się na niego dokumentowane i nagminnie występujące w przestrzeni publicznej zjawiska takie jak fake newsy, komory echa, clickbaity, *deepfake*’i, a także praktyki trollowania, kradzieży danych (np. *phishing*²¹), w ostateczności napięcia i eskalacje mogące sprzyjać globalnej wojnie informacyjnej. Tak prezentujący się obraz rzeczywistości prowadzi w efekcie do spadku zaufania społecznego, upadku autorytetów oraz polaryzacji poglądów i rozkwitu pseudonauki.

Teorie pseudonaukowe wyrastają na złożonej pożywce, w której odnaleźć można nie tylko niewiedzę i nieporozumienia merytoryczne, lecz także niepokój, lęk i nadzieję (głównie w przypadku medycyny alternatywnej), przesadną skłonność do doszukiwania się porządku (zwłaszcza

²⁰ Pojęcie to odnosi się do technologii umożliwiającej cyfrową obróbkę i manipulację wizerunku danej osoby, np. twarzy, ciała czy głosu. *Deepfake* to technologia wykorzystująca sieci neuronowe do tworzenia lub edytowania treści wideo albo obrazu w celu pokazania treści niezgodnej z wyjściowym materiałem źródłowym.

²¹ Metoda oszustwa, w której sprawca ma na celu podszycie się pod inną osobę lub instytucję z zamiarem wyłudzenia poufnych informacji. Należy zwrócić uwagę na możliwy rozwój metod phishingu w oparciu o technologię *deepfake*.

w przypadku teorii spiskowych), brak zaufania do instytucji publicznych i nauki oraz niezrozumienie tego, jak one funkcjonują (w zasadzie we wszystkich przypadkach).

(Lamża, 2020, s. 7)

Łatwość, z jaką najnowsze technologie komunikacyjne wspierają nierzetelne praktyki społeczne, spleta się z nadzieją, a zarazem realną możliwością zbudowania systemu rzetelnych praktyk, także opartych na wskazanych technologiach. Czy jako społeczeństwo gotowi jesteśmy do oczyszczenia zwierciadła i, w następstwie tego kroku, także krytycznego spojrzenia na własne oblicze? W jaki sposób w tym wymagającym przedsięwzięciu pomoc może werytyczna epistemologia społeczna? Czy cele stawiane przez werytyzm nie stoją w kontrze do dominujących trendów współczesnej infosfery, tak silnie naznaczonej powątpiewaniem w sensowność idei prawdy i rzetelności, co wzmacniane jest chwytliwymi hasłami postprawdy i „alternatywnych faktów”?

Werytyzm istotnie przeciwstawia się takiej wizji przestrzeni publicznej i w tym należy upatrywać jego fundamentalnej roli. Wskazując i analizując rzetelne praktyki społeczne, które wzmacniają efektywność procesów zdobywania wiedzy, jednocześnie jest szansą na uporządkowanie przestrzeni informacyjnej (m.in. poprzez projektowanie i stosowanie skutecznych metod rozpoznawania rzetelnych i nierzetelnych praktyk poznawczych²²). Tego typu funkcje skutecznie wdrażane są na poziomie organizacji, m.in. fact-checkingowych²³. Jako główny cel tego typu instytucji wskazuje się poprawę jakości debaty publicznej. Sposoby na osiągnięcie tego celu są różne: od weryfikacji komunikatów politycznych czy instytucjonalnych, po sprawdzanie przestrzeni publicznej pod kątem występowania popularnych zjawisk, jak fake newsy, clickbaity czy deepfake'i. Organizacje te

²² Praktyki takie mają niski współczynnik trafności i/lub są oparte na fundamencie nierzetelnych procesów poznawczych, np. myślenia życzeniowego. Mają także ujemną, instrumentalną wartość werytyczną.

²³ We wrześniu 2015 roku, w związku z inicjatywą Poynter Institute, została ufundowana Międzynarodowa Sieć Organizacji Fact-Checkingowych (*International Fact-Checking Network*, IFCN). W 2016 roku IFCN sformułowała kodeks zasad, które powinna spełniać każda organizacja sprawdzająca zgodność z faktami. Są to kolejno: bezstronność, transparentność, przejrzystość finansowania i stosowanej metodologii oraz zobowiązanie do stosownych korekt i aktualizacji raportów (<https://ifcncodeofprinciples.poynter.org>).

wykorzystują rzecz jasna technologie, które usprawniają procesy testowania autentyczności transferowanych informacji. Jedną z coraz częstszych prób tego typu, jest wykorzystanie dużych modeli językowych, jak Chat GPT, oraz próba odpowiedzi na pytanie o możliwość zastąpienia ludzi w procedurze sprawdzania trafności publikowanych treści²⁴.

Werytyzm dostarcza solidnego i zarazem przejrzystego modelu, który jest w stanie uchwycić złożoność i dynamikę społecznych praktyk poznawczych, w tym opartych na najnowszych technologiach. Opierając epistemiczne kontinuum na relacji między fundamentalnymi i instrumentalnymi W-wartościami, jednoczy w obrębie jednej struktury szeroką paletę zjawisk rozgrywających się w sferze publicznej, w której cele poznawcze jednostek czy grup społecznych uwikłane są w kontekst rywalizujących z sobą, różnorodnych praktyk instrumentalnych. Tego typu ewaluacja jest niezbędna w warunkach współczesnego środowiska informacyjnego. Różnorodne praktyki społeczne moderujące przepływ informacji winny być nieustannie monitorowane pod kątem ich wpływu na procesy zdobywania wiedzy. Cenniejsze z perspektywy werytyzmu są te praktyki, które z większą skutecznością prowadzą do osiągnięcia prawdziwych i trafnych wyników.

Jednakże cele przyświecające werytyzmowi, mimo swej niewątpliwej wagi, nie powinny być jedynymi aspiracjami. Społeczeństwo informacyjne powinno opierać się także na umiejętności rozpoznawania i eliminowania praktyk, które prowadzą do niesprawiedliwości epistemicznej (problemu wykluczenia i uprzedzeń). Epistemologia społeczna staje się więc w pewnej mierze także projektem etycznym, z uwagi na to, że dynamika środowiska sieciowego prowadzi do licznych konfliktów wartości. Jeśli pewna praktyka społeczna wytwarza wyższy wskaźnik trafności przekonań niż inne, ale jednocześnie wyklucza jednostki należące do mniejszości, to w takim systemie praktyk tkwi pewna istotna wada²⁵.

²⁴ Jak wykazują najnowsze eksperymenty, technologia generatywnej SI zawodzi na tym polu i nie jest jeszcze wystarczająco niezawodna i dokładna, aby móc zastąpić ludzi w tym wymagającym zadaniu (Abels, 2023).

²⁵ Jest to w wielu przypadkach poważny problem i częściowo słuszny wniosek. Są jednak praktyki, na co zwraca uwagę Goldman, które nie spełniają warunku równości szans w dostępie do wygłaszania opinii, nie będąc jednocześnie w żaden sposób wadliwymi (Goldman, 2021). Takie praktyki mogą z powodzeniem spełniać warunki rzetelności i sprawiedliwości epistemicznej. Jako prosty przykład posłużyć może serwis informacyjny

Rzetelność praktyk społecznych, obok skuteczności w tropieniu prawdy i unikaniu błędu, winna wspierać także procesy niwelowania niesprawiedliwości epistemicznej. Wiarygodność nadawcy komunikatu nie powinna być zależna od uprzedzeń i stereotypów uczestników debaty publicznej, gdyż taki stan w warunkach dynamicznej i złożonej infosfery sieciowej jest punktem wyjścia do dewaluacji zaufania społecznego i rozwoju praktyk destrukcyjnych przestrzeni publiczną. Obok rzetelności odpowiadającej wysokiemu wskaźnikowi trafności przekonań i eliminacji błędu, należy zatem wskazać na konieczność wdrażania zasad sprawiedliwości epistemicznej, co wiąże się z dystrybuowaniem dóbr epistemicznych w przejrzysty sposób, z poszanowaniem udziału innych (często diametralnie różniących się) członków wspólnoty komunikacyjnej²⁶.

Cele badawcze, jakie stawia przed sobą WES, należy umieścić w dynamicznym systemie i powinny być one w sposób elastyczny dostosowywane do określonego kontekstu. Dotyczy to na przykład problemu anonimowości. „Podmioty działające w złej wierze, ośmielone bezpieczeństwem anonimowości, mogą szerzyć uprzedzenia i mowę nienawiści, czyniąc społeczności internetowe wrogimi środowiskami dla uciskanych grup” (Frost-Arnold, 2014, s. 72). Jednak, co równie mocno podkreśla Karen Frost-Arnold, powszechnie dokumentuje się przypadki, w których sieciowa anonimowość jest jedyną możliwością bezpiecznego przekazania informacji. W wielu opresyjnych systemach za określone przekonania czeka nieuzasadniona i bezwzględna kara. W takich przypadkach względna anonimowość pomaga uniknąć niesprawiedliwego osądu. Praktyki oparte na anonimowości są zatem w pewnych sytuacjach korzystne także z werystycznej perspektywy, umożliwiając przekazanie relewantnego komunikatu z istotnym pominięciem represyjnych instytucji.

przygotowany przez grupę bezstronnych dziennikarzy i ekspertów. Oczywiście jest, że w ramach tak realizowanej praktyki możliwość spełnienia kryterium równości szans i równego dostępu do tworzenia publikowanej treści jest słusznie ograniczona.

²⁶ Koresponduje to wyraźnie z postulatami stawianymi przez Jürgena Habermasa, który argumentował, że epistemiczny i polityczny wzorzec społeczeństwa demokratycznego stanowi zjednoczona sfera publiczna. W tak zorganizowanej przestrzeni obywatele wspólnie podejmują poszczególne kwestie na równych prawach, odnosząc się do publicznego rozumu (Habermas, 2007).

Podsumowanie

Badania prowadzone w obrębie WES są cenne i niewątpliwie potrzebne do uporządkowania współczesnej debaty publicznej. Mogą odpowiedzieć na bardzo istotne pytania. Na jakie aspekty zwrócić uwagę, aby uchronić się przed trybami maszyny dezinformacyjnej? Jakie praktyki wykorzystać, aby zbudować skuteczny filtr poznawczy, chroniący przed rozbudowanym spektrum nierzetelnych procesów komunikacyjnych? Które spośród praktyk społecznych są skuteczne i właściwie stosowane w celu rozpoznawania prawdziwych informacji oraz unikania błędów? Są to możliwe zastosowania WES, co w warunkach dynamicznej komunikacji sieciowej należy stale wnikliwie rozwijać i aktualizować. Wciąż pojawiają się nowe technologie przetwarzania informacji, w związku z czym niezbędne jest podejmowanie stałych wysiłków w celu monitorowania systemów informacyjnych pod kątem trafności ich wyników (np. w przypadku popularnych w ostatnim czasie rozwiązań z zakresu generatywnej SI). Wdrażanie skutecznych procedur rozpoznawania rzetelnych i nierzetelnych praktyk poznawczych jest wymagane, aby niwelować skutki sieciowej dezinformacji, która z pełnym rozmachem kreśli niepokojący wizerunek współczesnej infosfery. Należy jednak pamiętać o dodatkowym wymogu, który powinien być przy tym spełniony, a jest nim nad wyraz istotne uwzględnienie postulatów związanych z ideą sprawiedliwości epistemicznej. Tylko w takich warunkach możliwe jest zbudowanie społeczeństwa opartego na zaufaniu społecznym, a więc takiego, w którym kwitnie właściwie rozumiana rzetelność.

Bibliografia

- Abels, G. (2023). Can ChatGPT fact-check? We tested. Pobrano z: <https://www.poynter.org/fact-checking/2023/chatgpt-ai-replace-fact-checking/> (29.05.2023).
- d'Ancona, M. (2018). *Postprawda*. Tłum. M. Sutowski. Warszawa: Wydawnictwo Krytyki Politycznej.
- Brownstein, M. (2019). Implicit Bias. W: E.N. Zalta (red.), *The Stanford Encyclopedia of Philosophy*. Pobrano z: <https://plato.stanford.edu/entries/implicit-bias/> (26.05.2023).

- Buchanan, B., Lohn, A., Musser, M., Sedova, K. (2021). Truth, Lies, and Automation. How Language Models Could Change Disinformation. Pobrano z: <https://cset.georgetown.edu/publication/truth-lies-and-automation/> (24.05.2023).
- Dreyfus, H.L. (2001). Telepistemology: Descartes's Last Stand. W: K. Goldberg (red.), *The Robot in the Garden. Telerobotics and Telepistemology in the Age of the Internet* (s. 48–63). Cambridge, MA: MIT Press.
- Fallis, D. (2015). What Is Disinformation?. *Library Trends*, 63 (3). DOI: 10.1353/lib.2015.0014.
- Fetzer, J.H. (2004). Disinformation: The use of false information. *Minds and Machines*, 2 (14). DOI: 10.1023/B:MIND.0000021683.28604.5b.
- Fricke, M. (2009). *Epistemic Injustice: Power and Ethics in Knowing*. Oxford: Oxford University Press.
- Frost-Arnold, K. (2014). Trustworthiness and Truth: The Epistemic Pitfalls of Internet Accountability. *Episteme*, 2 (1). DOI: 10.1017/epi.2013.43.
- Frost-Arnold, K. (2023). *Who Should We Be Online? A Social Epistemology for the Internet*. Oxford: Oxford University Press.
- Goldberg, K. (red.) (2001). *The Robot in the Garden. Telerobotics and Telepistemology in the Age of Internet*. Cambridge, MA: MIT Press.
- Goldman, A. (1986). *Epistemology and Cognition*. Cambridge, MA: Harvard University Press.
- Goldman, A. (2001). Telerobotic Knowledge: A Reliabilistic Approach. W: K. Goldberg (red.), *The Robot in the Garden. Telerobotics and Telepistemology in the Age of the Internet* (s. 126–142). Cambridge, MA: MIT Press.
- Goldman, A. (2008). Social epistemology of blogging. W: J. van den Hoven, J. Weckert (red.), *Information Technology and Moral Philosophy* (s. 111–122). New York: Cambridge University Press.
- Goldman, A. (2009). Social Epistemology: Theory and Applications. *Royal Institute of Philosophy Supplement*, 64. DOI: 10.1017/S1358246109000022.
- Goldman, A. (2020). The What, Why, and How of Social Epistemology. W: M. Fricke, i in. (red.), *The Routledge Handbook of Social Epistemology* (s. 10–20). New York: Routledge.
- Goldman, A. (2021a). *Wiedza a świat społeczny*. Tłum. A. Grobler. Warszawa: PWN.
- Goldman, A. (2021b). A Different Solution to the Generality Problem for Process Reliabilism. *Philosophical Topics*, 49 (2). DOI: 10.5840/philtopics202149217.

- Griffiths, M.D. (2014). Adolescent trolling in online environments: A brief overview. *Education and Health*, 32 (3).
- Groh, M., Epstein, Z., Fireston, C., Picard, R. (2022). Deepfake detection by human crowds, machines, and machine-informed crowds. *PNAS*, 119 (1). DOI: 10.1073/pnas.2110013119.
- Habermas, J. (2007). *Strukturalne przeobrażenia sfery publicznej*. Tłum. W. Lipnik, M. Łukasiewicz. Warszawa: PWN.
- Harari, Y.N. (2023). Yuval Noah Harari argues that AI has hacked the operating system of human civilization. *The Economist*. Pobrano z: <https://www.economist.com/by-invitation/2023/04/28/yuval-noah-harari-argues-that-ai-has-hacked-the-operating-system-of-human-civilisation> (20.05.2023).
- <https://ifcncodeofprinciples.poynter.org/> (27.05.2023).
- Hutchins, E. (1995). *Cognition in the Wild*. Cambridge, MA: MIT Press.
- Jamieson, K.H., Cappella, J.N. (2008). *Echo Chamber: Rush Limbaugh and the Conservative Media Establishment*. Oxford: Oxford University Press.
- Lamża, Ł. (2020). *Światy równoległe. Czego uczą nas płaskoziemcy, homeopaci, różdżkarze*. Wołowiec: Wydawnictwo Czarne.
- McIntyre, L. (2019). *Post-Truth*. Cambridge, MA: MIT Press.
- Nguyen, C.T. (2020). Echo chambers and epistemic bubbles. *Episteme*, 17 (2). DOI: 10.1017/epi.2018.32.
- Pariser, E. (2012). *The Filter Bubble: What The Internet Is Hiding From You*. London: Penguin Books.
- Pritchard, D. (2010). Cognitive Ability and the Extended Cognition Thesis. *Synthese*, 175 (1). DOI: 10.1007/s11229-010-9738-y.
- Pritchard, D. (2018). Extended Epistemology. W: J.A. Carter, A. Clark, J. Kallestrup, S.O. Palermos, D. Pritchard (red.), *Extended Epistemology* (s. 90–104). Oxford: Oxford University Press.
- Skyrms, B. (2010). *Signals: Evolution, Learning, and Information*. Oxford: Oxford University Press.
- Tomasello, M. (2002). *Kulturowe źródła ludzkiego poznawania*. Tłum. J. Rączaszek. Warszawa: PIW.
- Toffler, A. (1985). *Trzecia fala*. Tłum. E. Woydyło. Warszawa: PIW.

- Tversky, A., Kahneman, D. (2012). Osądy w warunkach niepewności: heurystyki i błędy poznawcze. W: D. Kahneman. *Pułapki myślenia*. Tłum. P. Szymczak (s. 559–580). Poznań: Media Rodzina.
- Turbanti, G. (2022). *Philosophy of Communication*. Cham: Palgrave Macmillan.
- What Percentage of Internet Traffic is Mobile?*. Pobrano z: <https://www.oberlo.com/statistics/mobile-internet-traffic> (27.05.2023).
- WHO (2020). *Coronavirus disease 2019 (COVID-19) Situation Report – 45*. Pobrano z: <https://apps.who.int/iris/handle/10665/331356> (22.05.2023).
- Ziemińska, R. (2020). The Epistemic Injustice Expressed in Normalizing Surgery on Children with Intersex Traits. *Diametros*, 17. DOI: 10.33392/diam.1478.

Nota o autorze

Tomasz Walczyk – doktor filozofii, adiunkt w Uniwersytecie Marii Curie-Skłodowskiej w Lublinie. Zainteresowania badawcze: epistemologia społeczna i rozszerzona, poznanie ucieleśnione, filozofia nowych technologii. Opublikował monografię *Teleepistemologia. Analiza rozszerzonych systemów poznawczych* (Kraków 2019).
Address for correspondence: Maria Curie-Skłodowska University, Faculty of Philosophy and Sociology, Institute of Philosophy, ul. Marii Curie-Skłodowskiej 4, 20-031 Lublin.

Cytowanie

Walczyk, T. (2024). Werystyczna epistemologia społeczna i problemy komunikacji sieciowej. *Analiza i Egzystencja*, 67 (3), 5–26. DOI: 10.18276/aie.2024.67-01.



MATEUSZ PENCZEK

Uniwersytet Śląski w Katowicach
ORCID: 0000-0002-9907-8283
email: mateusz.penczek@us.edu.pl

Autonomia osobowa z perspektywy relacyjnej teorii autonomii Catriona Mackenzie

Słowa kluczowe: autonomia, autonomia osobowa, autonomia relacyjna, relacyjne teorie autonomii, Catriona Mackenzie

Keywords: autonomy, personal autonomy, relational autonomy, relational theories of autonomy, Catriona Mackenzie

Personal autonomy according to Catriona Mackenzie's relational theory of autonomy

Abstract

The purpose of the article is to analyze the relational theory of personal autonomy proposed by Catriona Mackenzie. First, the theoretical context of the theory is shown, including key features of the relational approach to autonomy. Then the multidimensional analysis of the concept of autonomy proposed by Mackenzie and the theory of autonomy based on it are presented. According to the theory, seven conditions of autonomy should be distinguished. The conditions are grouped into three separate axes: self-determination, self-governance, and self-authorization. Moreover, the assessment of the level of autonomy of an individual should be contextual. Finally, an attempt is made to assess the value of the conception proposed by the Australian scholar. Among other things, it is argued that for the theory to be useful in assessing the autonomy of particular decisions it should be supplemented with at least one further condition: the coherence of a decision made by a given person with her practical identity.

Wprowadzenie

Wprowadzenie terminu „autonomia” do nowożytnej refleksji filozoficznej poświęconej jednostce i moralności przypisuje się Immanuelowi Kantowi, który zaczerpnął go z siedemnasto- i osiemnastowiecznej myśli politycznej (Schneewind, 1998, s. 483)¹. W ujęciu królewieckiego filozofa autonomia człowieka polega na tym, że „podlega on tylko swemu własnemu, a jednak powszechnemu prawodawstwu” (Kant, 1953, s. 67–68). W XIX i XX wieku krytyczna reakcja na Kantowskie rozumienie autonomii doprowadziła do pojawienia się pojęcia autonomii osobowej rozumianej jako kierowanie sobą w zgodzie z własną indywidualną i niepowtarzalną tożsamością osobową (Hinchman, 1996, s. 488–490)². Ujęcia autonomii osobowej proponowane we współczesnej literaturze filozoficznej są jednak tak bardzo zróżnicowane, że na pierwszy rzut oka może się wydawać, że jedynym łączącym je wątkiem jest to, iż chodzi o cechę osób i jest to cecha pożądana (Dworkin, 1988, s. 6; Taylor, 2009, s. 18–36). Mimo to niektórzy filozofowie przekonują, że można wskazać ich wspólny fundament. Postulują oni, aby odróżniać od siebie pojęcie autonomii oraz różne poświęcone jej koncepcje i teorie (Dworkin, 1988, s. 9; Mackenzie, 2014a, s. 15)³.

Wielu zwolenników tej dystynkcji uzupełnia ją o tezę, że mamy do czynienia z tylko jednym pojęciem autonomii. Do jego treści zalicza się zazwyczaj takie idee jak np. niezależność, decydowanie o sobie, autentyczność decyzji i działań, zdolność do kierowania swoim życiem ze swej własnej

¹ Termin ten pojawił się już w starożytnej Grecji, jednak wtedy odnosił się do posiadania lub tworzenia przez miasta-państwa swych własnych praw (Feinberg, 1986, s. 27–28).

² W filozofii współczesnej podejmuje się również takie zagadnienia jak autonomia moralna, autonomia konkretnych decyzji i działań, autonomia obywateli wobec politycznych instytucji państwa, a także autonomia jednostki względem pozapolitycznych instytucji społecznych, takich jak np. klasy społeczne, instytucje ekonomiczne (Beauchamp, Childress, 2009, s. 99–120; Dworkin, 1988, s. 10–12, 34–35, 100–120; Hill, 2013, s. 24–26; Schneewind, 2013, s. 146–168).

³ Treść pojęcia autonomii miałyby obejmować tylko podstawowe intuicje językowe na jej temat. Z kolei teorie i koncepcje autonomii miałyby w oparciu o te intuicje proponować konkretne jej warunki, poddawać je analizie, a także naświetlać związki pojęcia autonomii z innymi pojęciami, jak np. wolność, autentyczność, refleksja, odpowiedzialność, szacunek, uznanie (Dworkin, 1988, s. 9; Mackenzie, 2014a, s. 15; 2017, s. 523). Inspiracją jest tu wprowadzone przez Johna Rawlsa (1994, s. 15–16) analogiczne rozróżnienie dotyczące sprawiedliwości.

perspektywy, kontrola nad sobą, a także zdolność do budowania i realizacji własnej koncepcji wartościowego życia (Dworkin, 1988, s. 9–13; Mele, 2002, s. 529; Christman, 2005, s. 278; Anderson, Honneth, 2005, s. 137; Beauchamp, 2010, s. 79–80)⁴. Są też filozofowie, którzy skłaniają się ku tezie, że mamy tu raczej do czynienia z kilkoma różnymi pojęciami. Na przykład Joel Feinberg (1986, s. 27–51) zauważa, że termin „autonomia” zastosowany do osób ma aż cztery różne znaczenia: zdolność kierowania sobą, faktyczne kierowanie sobą, ideał charakterologiczny osoby kierującej sobą, a także prawo do kierowania sobą. Niektórzy badacze dodają, że w pewnych kontekstach termin ten oznacza także określony status społeczny przypisywany jednostce przez innych ludzi (Mackenzie, 2019, s. 147; 2021, s. 33). Zwolennicy tezy o jednym pojęciu autonomii mogą jednak przekonująco argumentować, że pierwsze ze wskazanych przez Feinberga znaczeń stanowi podstawę dla pozostałych (Taylor, 2009, s. 20–21)⁵.

Bez względu na interpretację, treść pojęcia (lub pojęć) autonomii osobowej jest bardzo ogólna i nie może dostarczyć precyzyjnych wskazówek przy budowaniu jej teorii (Beauchamp, 2010, s. 79–80; Dworkin, 1988, s. 9–13). W rezultacie teorie autonomii osobowej proponowane we współczesnej filozofii są ogromnie zróżnicowane⁶. Najbardziej chyba znany jest ich podział na proceduralne (zwane niekiedy neutralnymi pod względem treści) i treściwe (ang. *substantive*)⁷. Teorie autonomii dzielą się też na internalistyczne i eksternalistyczne. W tych pierwszych warunki autonomii

⁴ Niektórzy współcześni autorzy utożsamiają autonomię osobową z różnymi formami wolności jednostki, np. z wolnością pozytywną w ujęciu Isaiaha Berlina (zob. Stoljar, Voigt, 2021, s. 2) albo z wolnością woli w ujęciu Harry’ego Frankfurtza (zob. Christman, 2005, s. 279–280). Spotyka się to jednak z krytyką, wykazującą, że pojęcia autonomii i wolności są od siebie różne (Dworkin, 1988, s. 13–14; Taylor, 2009, s. 37–62).

⁵ James Stacey Taylor przekonuje z kolei, że nie istnieje jednoznaczne pojęcie autonomii, którego treść można po prostu odkryć. Należy raczej mówić o powszechnie przyjmowanych w filozofii współczesnej „konotacyjnych konturach” pojęcia autonomii, obejmujących bardzo szeroki wachlarz idei (Taylor, 2009, s. 1–3).

⁶ Ich kształt jest warunkowany m.in. przez ich założenia logiczne, teoretyczne, aksjologiczne i polityczne, a także cele, jakim służą. (Mackenzie, 2014a, s. 15).

⁷ Teorie proceduralne podają tylko formalne warunki autonomii, zaś teorie treściwe uwzględniają dodatkowo wymagania dotyczące wartości, treści przekonań, uczuć lub pragnień autonomicznej jednostki. W ramach teorii mocno treściwych wymaga się przyjmowania określonych wartości, tymczasem w teoriach słabo treściwych jedynie wyklucza się uznawanie niektórych wartości lub nie zezwala na obecność pewnych pragnień,

dotyczą wyłącznie cech umysłu jednostki, zaś w teoriach eksteralistycznych zalicza się do nich także wybrane czynniki „zewnętrzne” wobec niego, będące częścią otoczenia jednostki, np. jej relacje interpersonalne lub kontekst społeczno-kulturowy, w jakim żyje (Oshana, 1998, s. 83–86).

W ostatnich dekadach największe znaczenie ma jednak podział teorii autonomii na indywidualistyczne i relacyjne. W tych pierwszych: 1. warunki autonomii dotyczą wyłącznie cech umysłu jednostki; 2. cechy te ujmują się jako niezależne od takich czynników jak relacje interpersonalne jednostki lub społeczno-kulturowy kontekst jej życia. Tymczasem w relacyjnych teoriach autonomii niektóre z tego rodzaju czynników zalicza się do warunków autonomii lub przynajmniej przypisuje się im kluczowe znaczenie dla rozwoju i trwania cech umysłu jednostki, których dotyczą warunki autonomii (Christman, 2004, s. 143; Mackenzie, Stoljar, 2000, s. 21).

Podkreśla się, że wszystkie indywidualistyczne teorie autonomii osobowej mają charakter internalistyczny. Przykładowo taki charakter miały popularne w latach siedemdziesiątych XX wieku proceduralne ujęcia autonomii, akcentujące hierarchiczną strukturę ludzkiego umysłu, w których zazwyczaj wskazywano dwa jej warunki: zdolność człowieka do krytycznej refleksji oraz autentyczność jego przekonań, pragnień i decyzji (Christman, 2014, s. 373–374; Mackenzie, Stoljar, 2000, s. 13–26; Mackenzie, 2021, s. 34–35; Taylor, 2005, s. 1–24)⁸. Rozwijane w opozycji do nich relacyjne teorie autonomii są albo internalistyczne, albo eksteralistyczne. W tych pierwszych warunki autonomii dotyczą tylko określonych cech umysłu jednostki, ale jednocześnie podkreśla się, że cechy te istnieją dzięki sprzyjającemu im otoczeniu społecznemu i kulturowemu. Z kolei w drugich do warunków autonomii zalicza się wybrane elementy tego otoczenia (Mackenzie, 2021, s. 32–39).

Celem rozważań podjętych w niniejszym artykule jest krytyczna analiza i ocena kluczowych elementów relacyjnej i eksteralistycznej teorii autonomii osobowej zaproponowanej przez australijską filozof Catrionę Mackenzie⁹.

preferencji, postaw i ocen (Mackenzie, Stoljar, 2000, s. 13–21). Niekiedy też wymaga się odpowiedniego stosunku emocjonalnego do własnej osoby (Mackenzie, 2021, s. 36–37).

⁸ Prekursorem takiego ujęcia autonomii był Gerald Dworkin (1976, 1988). Do twórców hierarchicznego ujęcia autonomii często zalicza się też Harry’ego Frankfurta (1997), jednak jego hierarchiczna teoria dotyczyła wolności woli, a nie autonomii (Haworth, 1991, s. 129–131; Taylor, 2009, s. 37–40).

⁹ Poszczególne elementy tej teorii były sukcesywnie wypracowywane przez australijską badaczkę od lat dziewięćdziesiątych XX stulecia. Pełna jej wersja została przedstawiona

Aby możliwa była satysfakcjonująca realizacja tego zamierzenia, na początku wskazane zostały najważniejsze tezy relacyjnego podejścia do autonomii. Następnie przedstawiono zaproponowaną przez Mackenzie analizę treści pojęcia autonomii osobowej. W dalszej kolejności scharakteryzowano wyodrębnione przez nią wymiary i warunki autonomii, logiczne cechy pojęcia autonomii, a także sposób, w jaki jej zdaniem należy szacować poziom autonomii u konkretnych osób. Na koniec podjęta została próba wskazania mocnych i słabszych stron analizowanej teorii.

Najważniejsze założenia relacyjnego podejścia do autonomii osobowej

Relacyjne podejście do autonomii obecne jest w refleksji filozoficznej od co najmniej trzech dekad¹⁰. Ma swoich zwolenników przede wszystkim w nurcie feministycznym¹¹. Teorie autonomii proponowane w jego ramach wchodzą od kilku charakterystycznych założeń antropologicznych. Podkreśla się, że człowiek nie jest istotą całkowicie racjonalną oraz że nie ma pełnego wglądu we wszystkie motywy swoich decyzji i działań (Mackenzie, 2014a, s. 24). Akcentuje się również ucieleśnienie osoby ludzkiej. Zgodnie z tą ideą psychika człowieka jest głęboko zakorzeniona w jego ciele. Chodzi tu jednak nie tylko o to, że każdy proces psychiczny wymaga podłoża cielesnego, ale też o to, że ludzkie doświadczenie świata i siebie realizuje się poprzez ciało i dlatego jest przez nie warunkowane. Przykładowo to, jakim ciałem jest dana osoba (np. jak sprawnym, jakiego zdrowia, jakiej płci, rasy, urody) warunkuje to, czego ona doświadcza zarówno na poziomie biologicznym (np. wynikająca z płci biologicznej specyfika funkcjonowania cielesnego), jak i psychospołecznym (np. reakcje otoczenia społecznego na płeć,

w roku 2014 (Mackenzie, 2014) i od tej pory jest przez nią rozwijana w kolejnych tekstach (np. Mackenzie, 2015; 2017; 2019; 2021).

¹⁰ Jego początki miały miejsce w latach osiemdziesiątych XX stulecia. Za pierwsze próby relacyjnego ujęcia autonomii uznaje się niekiedy prace Jennifer Nedelsky z końca tej właśnie dekady i początku kolejnej (Benson, 2014, s. 87; Mackenzie, Stoljar, 2000, s. 4). Jak jednak słusznie zauważa anonimowy recenzent tej pracy, ważną inspiracją dla takiego sposobu myślenia są oczywiście prace Carol Gilligan z początku lat osiemdziesiątych i etyka troski.

¹¹ Warto dodać, że nie wszyscy sympatycy tego nurtu akceptująco odnoszą się do idei autonomii osobowej (Mackenzie, Stoljar, 2000, s. 3–12). Niektóre wątki feministycznej krytyki idei autonomii osobowej są podobne do krytyki tej idei przedstawionej przez zwolenników komunitaryzmu (Barclay, 2000, s. 52–68).

poziom sprawności, stan zdrowia, urodę). Budując samowiedzę, człowiek ujmuje swoje ciało jako bardzo ważną część siebie, doświadcza siebie poprzez swoje ciało, a to, co dzieje się z jego cielesnością, jest ważnym elementem narracji, jaką tworzy o sobie (Mackenzie, Scully, 2007, s. 342–344; Mackenzie, 2014a, s. 21; 2014b, s. 155–158).

Przyjmuje się też tezę o społecznej ontologii osoby, czyli „koncepcję osób, która kładzie nacisk na rolę ucieleśnionych praktyk społecznych (językowych i kulturowych), tożsamości grup społecznych, a także nieprzewidywalnych okoliczności historycznych w kształtowaniu się tożsamości jednostki” (Mackenzie, 2015, s. 56). Myśl tą niekiedy wyraża się pisząc, że osoba jest radykalnie usytuowana w określonym kontekście społeczno-kulturowo-historycznym. Ścisłej mówiąc, uznaje się, że kluczowe znaczenie mają tutaj dwa czynniki: relacje jednostki oraz społeczno-kulturowo-historyczny kontekst jej życia. Pierwsza z tych kategorii odnosi się do niepowtarzalnego układu aktualnych i minionych interakcji jednostki z innymi ludźmi. Druga obejmuje m.in. jej przynależność do grup społecznych o określonej strukturze, tożsamości kulturowej, historii, języku, tradycji, normach, instytucjach, praktykach, relacjach z innymi grupami społecznymi. Często podkreśla się, że do aktualnego kształtu tych czynników dodać należy ich formy minione, i to zarówno te bezpośrednio doświadczane przez daną jednostkę, jak i te, które kształtowały jej grupę społeczną przed jej narodzinami (Brisson, 2000, s. 284–285; Friedman, 2000, s. 40–41; Mackenzie, 2014a, s. 21)¹².

Na podstawie przedstawionych założeń tożsamość osobową jednostki uznaje się za funkcję metamorfoz jej cielesności oraz niepowtarzalnego układu czynników interpersonalnych, społecznych, historycznych i kulturowych tworzących środowisko jej życia. Oczywiście dostrzega się też, że konkretna jednostka może aktywnie kształtować siebie i swoje otoczenie, jednak nie uznaje się, aby była ona w stanie całkowicie kontrolować to, kim jest, co myśli o sobie, jakie decyzje podejmuje, a także to, jak żyje. Podkreśla się na dodatek, że ucieleśnienie człowieka oraz jego zależność od otoczenia społecznego przekładają się na jego podatność na cierpienia i krzywdę

¹² Metafizyczna wersja tezy o społecznej ontologii osoby głosi, że pewne określone relacje i czynniki społeczno-kulturowe są istotną częścią osoby. Jej wersja psychologiczna mówi zaś, że różnego rodzaju relacje i czynniki społeczno-kulturowe konstituują osobę, ale jednocześnie uznaje, że ich konkretne konfiguracje są przygodne (Christman, 2004, s. 144–145).

(Barclay, 2000, s. 52–68; Christman, 2004, s. 144–146; Friedman, 2000, s. 40–41; Mackenzie, 2014a, s. 21–23)¹³.

Także autonomię człowieka uznaje się za funkcję jego relacji interpersonalnych oraz jego otoczenia społeczno-kulturowego. Podkreśla się przy tym, że zazwyczaj są one bardzo złożone i nacechowane różnymi formami niesprawiedliwości niesprzyjającymi autonomii. Zauważa się też, że czynniki te mogą warunkować autonomię na dwa sposoby, co stanowi podstawę dla wyodrębniania dwóch odmian jej relacyjności: przyczynowej i konstytutywnej. Relacyjność autonomii ma charakter przyczynowy, jeśli przypisuje się im tylko rolę zewnętrznych bodźców przyczynowo kształtujących cechy umysłu jednostki, których dotyczą warunki autonomii. Takie rozumienie relacyjności nie implikuje przypisania autonomii natury społecznej. Relacyjność autonomii ma zaś charakter konstytutywny, gdy przyjmuje się, że określone relacje lub czynniki społeczno-kulturowe są kluczowymi składnikami autonomii lub cech umysłu będących warunkami autonomii. W takim wypadku autonomii przypisuje się naturę społeczną. Warto zaznaczyć, że obie wskazane formy relacyjności są często uwzględniane w ramach jednej teorii. Dzięki nim zwolennicy nurtu relacyjnego mogą podkreślać, że w ramach ich podejścia cechy umysłu jednostki, których dotyczą warunki autonomii, są interpretowane w sposób relacyjny (Mackenzie, Stoljar 2000, s. 22; Brison, 2000, s. 283; Mackenzie, 2014a, s. 15–23; Mackenzie, 2021, s. 36; Stoljar, 2011, s. 377; Stoljar, Voigt, 2021, s. 5)¹⁴.

¹³ Odróżnia się dwa rodzaje takiej podatności (ang. *vulnerability*). Pierwszy z nich to podatność na cierpienia i krzywdę wynikającą z ludzkiego ucieleśnienia i nierozdzielnie związanej z człowieczeństwem. Drugi zaś to podatność na cierpienia i krzywdę konkretnych osób lub grup wynikającą z ich szczególnej sytuacji, np. niższego statusu społecznego lub ekonomicznego, szczególnych potrzeb, słabszych zdolności, zwiększonej zależności od innych (Mackenzie, Rogers, Dodds, 2014, s. 4–6).

¹⁴ Na koniec warto dodać, że w ramach nurtu relacyjnego przyjmuje się też trzy ważne założenia etyczne. Pierwszym jest normatywny indywidualizm, czyli teza o aksjologicznym primacie jednostki nad grupami społecznymi i ciałami zbiorowymi. Drugie głosi, że każda osoba ludzka posiada niezbywalną i niestopniowalną godność, która sprawia, że niezależnie od okoliczności i stopnia autonomii zasługuje na szacunek i jest moralnie równa innym. Zgodnie z trzecim założeniem warunki społeczne ograniczające autonomię są niesprawiedliwe, a sprawiedliwe społeczeństwo powinno rozwijać społeczne, prawne i polityczne instytucje sprzyjające autonomii obywateli (Mackenzie, Stoljar, 2000; Mackenzie, 2014a, s. 20–23, s. 39; Mackenzie 2015, s. 56; Mackenzie, 2021, s. 40).

Pojęcie autonomii w ujęciu Mackenzie

Oryginalność podejścia zaproponowanego przez australijską badaczkę polega przede wszystkim na tym, że w ramach treści pojęcia autonomii wyróżnia ona trzy odrębne, choć powiązane z sobą osie: samostanowienie (ang. *self-determination*), kierowanie sobą (ang. *self-governance*) oraz upoważnienie siebie (ang. *self-authorization*). Na oś samostanowienia składają się dwa warunki autonomii: odpowiedni zakres wolności oraz dostęp do wartościowych możliwości. Oś kierowania sobą obejmuje kolejne dwa warunki: autentyczność tożsamości praktycznej oraz posiadanie umiejętności i kompetencji umożliwiających prowadzenie życia zgodnego z własną autentyczną tożsamością praktyczną. Z kolei oś upoważnienia siebie tworzą trzy warunki autonomii: odpowiedzialność za racje swojego postępowania, postawy związane z samooceną, a także uznanie społeczne rozumiane jako przypisanie jednostce przez jej otoczenie statusu autonomicznego podmiotu (Mackenzie 2014, s. 17–19; 2015, s. 55; 2017 s. 523; 2019, s. 148).

Broniąc takiego podejścia, Mackenzie zdecydowanie podkreśla, iż wskazane przez nią trzy osie nie są rozwinięciem jednej, bardziej podstawowej idei, którą można by uchwycić w jednolitym pojęciu autonomii. Odrzuca też myśl, że oddają one trzy różne pojęcia. Uzasadnia to podkreślając, że „pojęcie autonomii odgrywa ważną rolę w różnych społecznych i normatywnych praktykach” i dlatego należy zachować termin „autonomia”, a pojęcie, do którego on odsyła, uznać za wielowymiarowe (Mackenzie, 2014a, s. 39). Taki sposób rozumienia pojęcia autonomii osobowej jest podstawą dla jej teorii, w ramach której australijska badaczka przedstawia treść warunków autonomii uwzględnionych w ramach poszczególnych wymiarów, wyjaśnia logiczne cechy pojęcia autonomii, a także pokazuje, w jaki sposób należy dokonywać oceny poziomu autonomii u konkretnych osób.

Samostanowienie jako wymiar autonomii

Samostanowienie definiowane jest przez Mackenzie jako „posiadanie wolności i możliwości dokonywania oraz realizacji wyborów mających praktyczne znaczenie dla własnego życia, to znaczy wyborów dotyczących tego, czemu przypisywać wartość, kim być i co robić” (Mackenzie, 2014a, s. 25).

Warto podkreślić, że tak rozumiane samostanowienie odnosi się nie tyle do umiejętności jednostki, co raczej do jej statusu społecznego osoby autonomicznej, czyli uprawnionej do sprawowania pełnej kontroli nad swoim życiem (Mackenzie, 2019, s. 147–148). Status ten może przysługiwać jednostkom w różnym stopniu. Przypisanie go komuś w pełnym zakresie oznacza uznawanie tej osoby za równą innym, przyznanie jej pełnego spektrum swobód osobistych i politycznych, a także udzielenie dostępu do wszystkich ważnych dla niej opcji życiowych. W związku z tym oś samostanowienia obejmuje dwa warunki autonomii: odpowiedni poziom wolności oraz dostęp do możliwości. Warunki te nazywane są „zewnętrzny” lub „strukturalnymi”, ponieważ dotyczą wybranych aspektów społeczno-kulturowego kontekstu życia jednostki (Mackenzie, 2014a, s. 25).

Warunek wolności sprowadza się do wymagania, aby jednostce przysługiwały na odpowiednim poziomie podstawowe swobody osobiste i polityczne. Australijska badaczka zalicza do nich m.in. wolność myśli, wypowiedzi, stowarzyszania się, sumienia, religii, przynależności politycznej, ekspresji seksualnej, a także wolność od przymusu, manipulacji, wykorzystywania, przemocy (Mackenzie, 2014a, s. 25–27). Chodzi tutaj, najogólniej rzecz biorąc, o swobody wymieniane w dokumentach dotyczących praw człowieka (Mackenzie, 2015, s. 56–57). Warunek dostępu do możliwości sprowadza się z kolei do tego, aby jednostka miała wystarczający dostęp do ważnych dla niej możliwości, szans i opcji o charakterze osobistym, społecznym lub politycznym, takich jak np. edukacja, kariera zawodowa, założenie rodziny, opieka zdrowotna, opieka społeczna, dostęp do instytucji prawnych i politycznych, uczestnictwo w kulturze, aktywność sportowa, kulturotwórcza, religijna. O tym, które z tych możliwości i na jakim poziomie są ważne dla danej osoby współdecyduje oczywiście obrona przez nią koncepcja dobrego życia (Mackenzie, 2014a, s. 27–30)¹⁵.

¹⁵ Inspiracją dla wyodrębnienia tych warunków są relacyjne teorie autonomii o charakterze eksternalistycznym. Chodzi przede wszystkim o propozycje Mariny Oshany (1998, s. 93–94) i Susan J. Brison (2000, s. 285). Ujmują one autonomię jako konstytutywnie relacyjną (Stoljar, Voigt, 2021, s. 6). Analizując warunek dostępu do możliwości, Mackenzie (2014, s. 27–28) obszernie przywołuje także koncepcję autonomii Josepha Raza (1986). Inną inspiracją dla wyodrębnienia i teoretycznego ujęcia warunku dostępu do wartościowych możliwości są ponadto koncepcje sprawiedliwości Amartyi Sena, Marthy Nussbaum i Elisabeth Anderson (Mackenzie, 2014a, s. 27–30). Głębsza analiza tych związków wykracza jednak poza ramy niniejszego artykułu.

Mackenzie zauważa, że omawiany wymiar autonomii jest pojęciowo związany z wolnością (Mackenzie, 2017, s. 523), ale jednocześnie podkreśla, że nie można go sprowadzać wyłącznie do wolności negatywnej, rozumianej jako wolność jednostki od nieuzasadnionej ingerencji ze strony innych osób lub państwa. Zagwarantowanie komuś takiego samego zakresu podstawowych swobód jak innym nie oznaczałoby bowiem jeszcze tego, że byłby on w stanie korzystać z nich w takim samym stopniu jak oni. Dlatego oprócz wolności negatywnej samostanowienie jednostki wymaga też nieobecności różnych form dominacji ze strony innych (Mackenzie, 2019, s. 147–148), a więc nieobecności różnych form szeroko pojętej niesprawiedliwości (Mackenzie, 2014a, s. 22). Co więcej, aby jednostka mogła realnie korzystać z wielu przypisanych jej swobód, musi mieć też zagwarantowany dostęp do wartościowych opcji, takich jak np. edukacja czy zatrudnienie (Mackenzie, 2015, s. 56–57; 2017, s. 523–524; 2019, s. 147–148). Australijska badaczka zauważa też jednoczesne zachodzenie odwrotnej zależności: dla skutecznego korzystania z wartościowych możliwości niezbędna jest wolność od nieuzasadnionej ingerencji ze strony innych osób i instytucji. Podkreśla jednak, że w wielu sytuacjach oprócz wolności negatywnej niezbędne ku temu jest również aktywne wsparcie ze strony innych ludzi (Mackenzie, 2014a, s. 25). Oceniając stopień realizacji przez daną osobę obu wskazanych warunków autonomii, należy więc brać pod uwagę stopień, w jakim podlega ona różnym formom dominacji, a także w jakim może liczyć na wsparcie swojego otoczenia społecznego.

Kierowanie sobą jako wymiar autonomii

Kolejny wymiar autonomii wyróżniony w teorii Mackenzie to kierowanie sobą. Australijska badaczka zauważa, że „w głównym nurcie literatury o autonomii kierowanie sobą rozumie się zazwyczaj jako kwestię sprawowania wewnętrznej, podmiotowej kontroli nad swoją wolą oraz elementami swojej psychologii” (Mackenzie, 2019, s. 149). W jej zaś ujęciu kierowanie sobą to aktywne kształtowanie swojej autentycznej tożsamości praktycznej oraz podejmowanie zgodnych z nią decyzji i działań (Mackenzie, 2014a, s. 17–18, s. 31). Inaczej mówiąc, chodzi o „posiadanie umiejętności i zdolności koniecznych dla dokonywania wyborów i realizowania decyzji, które wyrażają własną, refleksyjnie tworzoną diachroniczną tożsamość praktyczną

lub są z nią zgodne” (Mackenzie, 2015, s. 58). Chodzi więc o wybory i decyzje, które są wyrazem własnych wartości i zobowiązań (Mackenzie, 2017, s. 524; Mackenzie, 2019, s. 149). Ten wymiar autonomii obejmuje dwa jej warunki „wewnętrzne” względem jednostki: posiadanie umiejętności i kompetencji niezbędnych dla kierowania sobą oraz autentyczność własnej tożsamości praktycznej¹⁶.

Pierwszy z tych warunków sprowadza się do wymagania, aby u jednostki obecne były na wystarczającym poziomie umiejętności i kompetencje psychologiczne kluczowe dla kierowania sobą. Przede wszystkim chodzi różnego rodzaju umiejętności i kompetencje poznawcze, w tym np. dokonywania krytycznej refleksji dotyczącej własnej tożsamości, emocji, decyzji i działań, a także niezbędne dla rozumienia informacji mających istotne znaczenie dla jednostki. Powinno się też wymagać obecności wybranych umiejętności i kompetencji wolicjonalnych (np. samokontrola, trwanie przy podjętej decyzji). Australijska badaczka podkreśla jednocześnie, że w wielu nierelacyjnych ujęciach autonomii zapomina się, że ogromne znaczenie dla kierowania sobą mają umiejętności i kompetencje emocjonalne, a także te związane z wyobraźnią (Mackenzie, 2019, s. 149). W ramach omawianego wymiaru autonomii powinno się więc uwzględniać umiejętności i kompetencje emocjonalne takie jak np. wrażliwość, umiejętność interpretacji cudzych i własnych emocji. Przeżywane emocje kształtują bowiem sposób, w jaki człowiek postrzega i interpretuje zarówno siebie, jak i świat. W związku z tym mają one niebagatelny wpływ na rezultaty, do jakich prowadzi krytyczna refleksja nad własną tożsamością praktyczną, a w konsekwencji także na ewentualne metamorfozy tej ostatniej (Mackenzie, 2002, s. 198–201, 203–204). Ważną rolę powinny odgrywać też umiejętności i kompetencje związane z wyobraźnią (np. wyobrażanie sobie różnych możliwych

¹⁶ Inspiracją dla uwzględnienia tych warunków są relacyjne koncepcje autonomii o charakterze internalistycznym i proceduralnym. Jedną z nich jest teoria Marilyn Friedman (2003, s. 4–5), zgodnie z którą autonomia polega na tym, że wybory i działania człowieka są realizacją tych jego potrzeb lub pragnień, które poddał krytycznej refleksji i zaaprobował jako zgodne ze swoją tożsamością. Mackenzie często powołuje się też na koncepcję Diany T. Meyers (1987, 2005), która ujmuje autonomię człowieka jako autentyczność jego „ja”, osiąganą dzięki umiejętności samopoznania, samookreślenia i decydowania o sobie. Ważną inspiracją jest także ujęcie autonomii rozwijane przez Johna Christmana (2004, 2009). Koncepcje zaliczające się do tej grupy mają charakter przyczynowo-relacyjny (Mackenzie, 2017, s. 518–519).

sposobów postępowania lub różnych wersji swojego życia)¹⁷. Warto również brać pod uwagę poziom umiejętności społecznych jednostki. Dialogiczne interakcje jednostki z innymi ludźmi mają bowiem istotny wpływ na rozwój i treść jej samowiedzy, a w związku z tym ich charakter będzie warunkował posiadany przez nią obraz własnej osoby (Mackenzie, 2014a, s. 32–33; 2015, s. 59; 2019, s. 149)¹⁸.

Warunek autentyczności tożsamości praktycznej odsyła do dwóch pojęć: tożsamości praktycznej i autentyczności. Tożsamość praktyczna jednostki to ten spośród różnych posiadanych przez nią obrazów samej siebie, który sprawia, że ceni ona siebie, uznaje swoje życie za warte przeżycia, a swoje działania za warte podjęcia (Korsgaard, 1996, s. 101). Jest to więc normatywna koncepcja własnej osoby, częściowo warunkowana przez czynniki biologiczne, psychologiczne, społeczne i kulturowe pozostające poza kontrolą jednostki, a częściowo konstruowana i przekształcana poprzez jej akty krytycznej refleksji (Mackenzie, 2008a, s. 10–11; Mackenzie, Walker, 2015, s. 379). W skład tożsamości praktycznej jednostki wchodzi takie elementy jak np. jej przekonania dotyczące siebie, wartości, cele, zobowiązania, kierunki zaangażowania (Mackenzie, 2014a, s. 17–18)¹⁹.

Autentyczność to z kolei cecha, którą przypisuje się różnym aspektom umysłu jednostki (np. pragnieniom, przekonaniom, tożsamości), a także jej działaniom. Stają się one autentyczne, kiedy dotycząca ich krytyczna refleksja jednostki prowadzi do tego, że zostają przez nią zaakceptowane (Dworkin, 1988, s. 15). Według innych ujęć, za autentyczne uznaje się te aspekty umysłu jednostki, które są charakterystyczne tylko dla niej i odróżniają ją od innych (Taylor, 2009, s. 30–31). Mackenzie przychyła się do pierwszej z tych definicji i analizowany tutaj warunek autentyczności tożsamości praktycznej sprowadza do tego, aby tożsamość praktyczna danej osoby była poddawana przez nią krytycznej refleksji i – w jej rezultacie – zaakceptowana jako „prawdziwie jej własna” (Mackenzie, 2014a, s. 31–32).

¹⁷ Obszerną i bardzo wnikliwą analizę wyobraźni zawiera tekst Mackenzie zatytułowany *Imagining Oneself Otherwise* (2000, s. 124–150). Głębsza analiza tej problematyki wykracza jednak poza ramy niniejszego artykułu.

¹⁸ Interesujące zestawienie różnego rodzaju umiejętności i kompetencji psychicznych ważnych dla kierowania sobą zawiera artykuł Diany Meyers (2014, s. 121).

¹⁹ W niektórych swoich pracach poświęconych problematyce tożsamości osobowej Mackenzie skłania się raczej ku narracyjnemu ujęciu tożsamości ze względu na jego diachroniczny charakter (Mackenzie, Walker, 2015, s. 380–381).

Niespełnianie tego warunku przez określony fragment tożsamości praktycznej danej osoby oznacza zaś, że przestał być już źródłem racji dla jej decyzji i postępowania.

Wyjaśniając znaczenie tego warunku autonomii, Mackenzie zauważa, że krytyczna refleksja nad różnymi aspektami własnej tożsamości praktycznej, a także wynikające stąd akty ich akceptacji bądź odrzucenia, stanowią ważne narzędzie, dzięki któremu człowiek przekształca siebie i dba o integrację swojej struktury motywacyjnej (Mackenzie, Stoljar, 2000, s. 13; Mackenzie, 2002, s. 189). Australijska badaczka precyzuje jednocześnie, że dla uznania przez daną osobę jakiegoś aspektu swojej tożsamości praktycznej za autentyczny nie jest konieczne pełne utożsamienie się z nim. Wystarczy, aby w wyniku krytycznej refleksji został on przez tę osobę uznany za coś, co nie jest jej obce. Podkreśla też, że badanie autentyczności danego aspektu tożsamości praktycznej wymaga wzięcia pod uwagę całego procesu jego kształtowania się oraz kontekstu społeczno-kulturowego, w jakim proces ten przebiegał (Mackenzie, 2015, s. 59–60)²⁰.

Mackenzie akcentuje potencjalnie negatywny wpływ opresyjnego kontekstu społecznego i kulturowego na zdolność jednostki do kierowania sobą. W jego rezultacie repertuar umiejętności i kompetencji jednostki ważnych dla kierowania sobą może być ograniczony, a te z nich, które zdołały się rozwinąć, mogą być zniekształcone lub osłabione (np. osłabiona może zostać zdolność krytycznej refleksji i umiejętność wyobrażania sobie innych wersji siebie). W przypadku długotrwałego funkcjonowania w opresyjnym otoczeniu dochodzi też do internalizacji niesprawiedliwych norm (np. dotyczących korzystania ze swobód, dostępu do możliwości, uznawania siebie za kogoś zdolnego do decydowania o sobie) i fałszywych przekonań na swój temat (np. krzywdzących stereotypów), a w tożsamości praktycznej jednostki pojawiają się takie elementy, które nie zostały w wystarczającym stopniu poddane przez nią krytycznej ocenie (Mackenzie, 2015, s. 59; 2017, s. 524; 2019, s. 149).

²⁰ Inspiracją dla tych uściśleń jest ujęcie autonomii autorstwa Johna Christmana, który sprowadza ją do posiadania kompetencji niezbędnych dla refleksji nad własnymi motywami, a także do autentyczności tych motywów. Drugi z warunków ujmuje zaś jako niewyobcowanie osoby wobec jej motywów będące rezultatem jej historycznie wrażliwej i adekwatnej autorefleksji, uwzględniającej jej diachroniczną praktyczną tożsamość oraz pozycję w świecie (Christman, 2009, s. 155).

Upoważnienie siebie jako wymiar autonomii

Ostatni z trzech wymiarów autonomii to stopień, w jakim jednostka uznaje siebie za upoważnioną do „sprawowania praktycznej kontroli nad swoim życiem, określania swych własnych racji działania, definiowania swoich wartości oraz kształtujących tożsamość praktycznych zobowiązań” (Mackenzie, 2014a, s. 18). Inaczej mówiąc, chodzi o stopień, w jakim dana osoba uznaje siebie za upoważnioną do tego, aby kierować swoim życiem, brać odpowiedzialność za swoje wartości, decyzje i działania, a także wyjaśniać innym swoje racje (Mackenzie, 2015, s. 61; 2017, s. 524; 2019, s. 149). Zaliczone tutaj zostały trzy warunki autonomii: odpowiedzialność przed sobą i innymi za racje swojego postępowania, odpowiednie postawy związane z samooceną, a także uznawanie jednostki przez innych ludzi za autonomiczną. Warunki pierwszy i drugi dotyczą cech jednostki, a ich realizacja łączy się z odpowiednim wartościowaniem własnej osoby. Z kolei warunek trzeci odnosi się do otoczenia społecznego jednostki i mówi o odpowiednim jej wartościowaniu przez to otoczenie.

Realizacja pierwszego z wymienionych warunków polega na gotowości jednostki do przedstawiania, wyjaśniania i obrony racji swojego postępowania, a także jej otwartości na ich ewentualne modyfikowanie pod wpływem racji przedstawianych przez innych. Jednostka powinna również przypisywać sobie prawo do oczekiwania od innych podobnej gotowości i otwartości (Mackenzie, 2014a, s. 18, 35)²¹. Chodzi więc o uznawanie siebie za podmiot w sensie moralnym (Mackenzie, 2015, s. 61) Australijska badaczka podkreśla, że warunek ten nie narzuca jednostce odpowiedzialności wobec konkretnych osób lub instytucji, a jej racje nie muszą ostatecznie znaleźć uznania w oczach innych. Zauważa też, że gotowość obrony swoich racji wymaga od jednostki poczucia, że sprawuje normatywną i epistemiczną kontrolę nad swoim życiem, czyli przekonania, że jeśli chodzi o jej życie

²¹ Inspiracją dla uwzględnienia warunku gotowości do odpowiadania za swoje racje są relacyjne ujęcia autonomii o charakterze internalistycznym i treściwym. Proponuje się w nich takie warunki jak umiejętność ujmowania normatywnych racji odnoszących się do własnego postępowania (Benson, 1991, s. 398, 401–402), a także gotowość do odpowiadania przed innymi za własne przekonania, wartości, cele, zasady i racje (Westlund, 2009, s. 34–35; Benson, 2005, s. 110–117).

i dążenia, to właśnie ona ostatecznie rozstrzyga co jest słuszne i co jest prawdą (Mackenzie, 2014a, s. 35–36)²².

Warunek postaw związanych z samooceną sprowadza się do przejawiania przez jednostkę odpowiedniego afektywnego nastawienia wobec siebie, w tym szacunku do siebie, zaufania do siebie i poczucia własnej wartości. Szacunek wobec siebie polega na uznawaniu swojego statusu moralnego za równy statusowi innych. Zaufanie do siebie oznacza z kolei ufność wobec własnych przekonań, emocji, ocen i działań, czyli np. przekonanie, że zazwyczaj, w normalnych warunkach, są one trafne, adekwatne i wiarygodne. Poczucie własnej wartości polega zaś na „uznawaniu swojego życia, swoich zobowiązań i swoich postaw za sensowne, wartościowe i interesujące” (Mackenzie, 2014a, s. 37)²³.

Realizacja trzeciego warunku polega z kolei na tym, że dana osoba jest uznawana przez innych ludzi i instytucje za kogoś, kto ma „pozycję społeczną autonomicznego podmiotu” (Mackenzie, 2014a, s. 35; 2015, s. 61), czyli za godny szacunku i decydujący o sobie podmiot działania (Mackenzie, 2014a, s. 37)²⁴. Chodzi więc o przypisywanie jednostce przez jej otoczenie społeczne autonomii rozumianej jako status²⁵. Warto podkreślić, że przyczyny braku realizacji tego wymogu często wykraczają poza kontekst najbliższych relacji jednostki (np. rodzinnych) i sięgają przekonań, norm i praktyk społecznych. Braki uznania są bowiem typowe np. dla stosunków społecznych cechujących się dominacją, a także nierównościami władzy,

²² Warto odnotować, że obrona swoich racji przez jednostkę zakłada korzystanie przez nią ze zdolności do podejmowania krytycznej refleksji nad sobą, a także niektórych innych zdolności i kompetencji uwzględnianych w ramach osi kierowania sobą.

²³ Wyjaśniając czym są szacunek i zaufanie do siebie, a także poczucie własnej wartości, Mackenzie powołuje się na eksternalistyczną i treściwą koncepcję autonomii Joela Andersona i Axela Honnetha (2005, s. 130–137). Podobne warunki są też jednak uwzględniane w relacyjnych ujęciach autonomii o charakterze internalistycznym i treściwym, np. uzasadnione zaufanie do siebie (McLeod, 2002, s. 103–131), a także zaufanie do siebie, poczucie kompetencji, poczucie własnej wartości, szacunek do siebie (Govier, 1993).

²⁴ Także i w tym przypadku inspiracją jest eksternalistyczna, treściwa i konstytutywnie relacyjna koncepcja autonomii Andersona i Honnetha, która wymaga, aby relacje społeczne, w jakich funkcjonuje jednostka, nacechowane były wzajemnym uznaniem, w tym m.in. szacunkiem dla autonomii i godności osób, a także poszanowaniem ich praw (Anderson, Honneth, 2005, s. 130–133, 137).

²⁵ Warto zauważyć, że przypisanie jednostce statusu osoby autonomicznej oznacza przypisanie jej przynajmniej podstawowego zakresu wolności i dostępu do możliwości.

statusu społecznego i ekonomicznego. Raz jeszcze ukazuje to wrażliwość autonomii na opresyjne konteksty społeczne i kulturowe (Mackenzie, 2015, s. 62; 2019, s. 149).

Mackenzie zwraca uwagę na silną zależność między warunkami autonomii składającymi się na oś upoważnienia siebie. Funkcjonowanie jednostki w ramach określonych „normatywnych struktur i praktyk uznania społecznego” wywiera wpływ na kształt jej postaw związanych z samooceną (Mackenzie, 2008b, s. 524–526; 2014, s. 37). One z kolei kształtują jej poczucie normatywnej i epistemicznej kontroli nad swoim życiem oraz gotowość do odpowiadania za racje swojego postępowania (Mackenzie, 2014a, s. 35–39; Mackenzie, 2021, s. 44). W związku z tym doświadczanie przez jednostkę odmowy uznania jej za autonomiczną może znacząco podważać jej szacunek i zaufanie do siebie, poczucie własnej wartości, a także gotowość do obrony swoich racji (Mackenzie, 2019, s. 149–150). Co więcej, nietrudno dostrzec, że długotrwałe doświadczanie tego rodzaju odmowy może mieć znaczący negatywny wpływ również na inne cechy jednostki ważne dla autonomii, to znaczy jej tożsamość praktyczną oraz poziom umiejętności i kompetencji ważnych dla kierowania sobą. Podsumowując, odmawianie jednostce autonomii rozumianej jako status może mieć negatywny wpływ na poziom jej autonomii rozumianej jako zdolność (Mackenzie, 2017, s. 523–525).

Logiczne cechy pojęcia autonomii i ocena jej poziomu

Bardzo oryginalny i nietypowy jest sposób, w jaki Mackenzie interpretuje logiczne cechy pojęcia autonomii. Jak zauważa, w filozofii zazwyczaj przyjmuje się, że pojęcie autonomii ma charakter jednolity (ang. *unitary*), to znaczy zakłada się, iż istnieje „jeden zbiór koniecznych i wystarczających warunków poprawnego stosowania tego pojęcia” (Mackenzie, 2014a, s. 16). Jej zdaniem takie podejście należy uznać za niewłaściwe. Po pierwsze, sprowadzanie autonomii do niewielkiej liczby niezależnych od kontekstu warunków koniecznych i wystarczających naraża takie ujęcie na kontrprzykłady. Po drugie, przy tego rodzaju podejściu szacowanie poziomu autonomii jest nieprecyzyjne. Jest tak dlatego, że zmiany kontekstu takiej oceny sprawiają, że niektóre warunki autonomii stają się ważniejsze od innych. Po trzecie, założenie, że pojęcie autonomii jest jednolite, jest po prostu błędne (Mackenzie, 2014a, s. 39–40; 2015, s. 55).

Autorka analizowanej koncepcji przyjmuje, że pojęcie autonomii ma charakter wielowymiarowy. Tę ostatnią cechę pojęcia definiuje się wyjaśniając, że „pojęcie F jest wielowymiarowe, jeśli to, czy i w jakim stopniu coś jest F, zależy od tego, jak to coś ma się pod względem wielu podstawowych wymiarów lub aspektów bycia F” (Hedden, Nebel, b.d.w., s. 1–2)²⁶. Propozycja australijskiej badaczki sprowadza się więc do tego, aby to, czy ktoś jest osobą autonomiczną, a także stopień, w jakim nią jest, zależało od tego, jak można go scharakteryzować pod względem trzech wskazanych przez nią wymiarów autonomii. Realizację każdego z uwzględnionych w ich ramach warunków autonomii uznaje jednocześnie za stopniowalną (Mackenzie, 2014a, s. 15–19, 40).

Co więcej, Mackenzie podkreśla też, że siedem wskazanych przez nią warunków autonomii osobowej nie tworzy zbioru warunków koniecznych i wystarczających. Zamiast tego postuluje, aby każdą ocenę autonomii uznawać za kontekstualną, to znaczy zależną od jej celu, a także od poziomu, na jakim się ją przeprowadza. Jak wyjaśnia, ocena autonomii może być dokonywana dla celu medycznego, osobistego, prawnego, politycznego i normatywnego, czyli etycznego. Jednocześnie można ją przeprowadzać na trzech różnych poziomach: całościowym, dziedzinowym (ang. *programmatic*) i lokalnym. Ocena całościowa obejmowałaby wszystkie dziedziny życia jednostki. Dziedzinowa skupiałaby się na wybranym obszarze jej funkcjonowania (np. praca, rodzina, przyjaźń, finanse). Z kolei lokalna ocena autonomii dotyczyłaby konkretnej decyzji (Mackenzie, 2014a, s. 19–20).

Za postulatem kontekstualności oceny autonomii kryje się pomysł, aby w zależności od jej celu i poziomu odpowiednio dookreślać, które jej warunki i wymiary uznaje się za ważniejsze, a które za mniej istotne, doprecyzowywać ich interpretację²⁷, a także przypisywać im minimalne wymagane progi ich realizacji. Po wprowadzeniu tego rodzaju uściśleń otrzymywalibyśmy obowiązujący tylko w konkretnym kontekście zestaw warunków, których spełnienie wystarczyłoby do przypisania minimalnej autonomii danej

²⁶ Szerzej rzecz ujmując, „przymiotnik F jest wielowymiarowy, jeśli to, czy jego forma pozytywna »x jest F« ma zastosowanie do obiektu, a także jeśli to, czy jego formy porównawcze, »x jest w takim samym stopniu F jak y« i »x jest bardziej F niż y«, mają zastosowanie do pary obiektów, zależy od tego, jak te obiekty mają się pod względem wielu podstawowych wymiarów bycia F” (D’Ambrosio, Hedden, 2023 s. 1).

²⁷ Przykładowo kontekst oceny może rozstrzygać o tym, jakie umiejętności i kompetencje mające znaczenie dla kierowania sobą należy uznać za najważniejsze.

osobie lub jej decyzji. (Mackenzie, 2014a, s. 15–20, 39–40; 2015, s. 54–56; 2017, s. 523; 2019, s. 148; 2021, s. 38–39, 44). Australijska badaczka proponuje też, aby – w zależności od kontekstu oceny autonomii – ewentualne deficyty w realizacji niektórych jej warunków mogły być do pewnego stopnia kompensowane przez wyższy poziom realizacji innych. Zdaje się przy tym przyjmować, że w wielu sytuacjach najtrudniejsze do skompensowania byłyby warunki autonomii związane z kierowaniem sobą (Mackenzie, 2014a, s. 15–20, 38–41). Podkreśla ponadto, że niedostateczne spełnianie warunków „zewnątrznych” (niski poziom wolności, ograniczony dostęp do możliwości czy też brak uznania) nie powinno stanowić jedynej podstawy dla odmowy przypisania komuś autonomii (Mackenzie, 2014a, s. 24, 38; 2017, s. 523). Jak można zauważyć, jedną z konsekwencji stosowania takiego podejścia byłoby to, że rezultat oceny poziomu autonomii u tego samego człowieka w danym punkcie jego życia mógłby się znacznie różnić w zależności od jej celu i poziomu (Mackenzie, 2014a, s. 19; 2015, s. 55).

Kluczowe cechy i mocne strony koncepcji Mackenzie

Australijska badaczka podkreśla, że w swojej teorii oddaje sprawiedliwość dwóm bardzo ważnym intuicjom dotyczącym autonomii obecnym w literaturze przedmiotu. Wedle pierwszej z nich autonomia jest statusem nadawanym człowiekowi przez jego otoczenie, co sprowadza się do uznawania go za osobę mającą prawo samodzielnie sprawować kontrolę nad swoim życiem. Druga z intuicji głosi zaś, że autonomia to zdolność człowieka do sprawowania kontroli nad swoim życiem, którą może on mieć nawet jeśli otoczenie odmawia przyznania mu statusu osoby autonomicznej (Mackenzie, 2015, s. 55; 2017, s. 523; 2019, s. 147). Znaczenie pierwszej intuicji mocno akcentują eksternalistyczne teorie autonomii, zaś drugiej – teorie internalistyczne (Mackenzie, 2021, s. 43). Ich połączenie dokonało się więc poprzez wzięcie pod uwagę inspiracji płynących z obu tych nurtów relacyjnej refleksji nad autonomią i uwzględnienie zarówno warunków autonomii odnoszących się do otoczenia jednostki (wolność, dostęp do możliwości, uznanie), jak i odnoszących się do cech jej umysłu²⁸.

²⁸ Inspiracje te zostały wskazane w przypisach 15, 16, 20–24.

Ponieważ teoria Mackenzie zawiera trzy warunki autonomii „zewnątrzne” wobec jednostki, należy ją zaklasyfikować jako eksternalistyczną. Obecność tych warunków sprawia też, że autonomia jest w niej ujęta jako relacyjna w sensie konstytutywnym. Taki charakter autonomii wynika również i z tego, że obecność niektórych kluczowych dla niej cech umysłu jednostki zakłada określony kontekst społeczny i normatywny. Przykładem może być gotowość do odpowiadania za racje swojego postępowania, która jest zawsze odpowiedzialnością wobec kogoś, a także zakłada otoczenie społeczne, w którym oczekuje się, że ludzie będą sobie wyjaśniali i uzasadniali swoje racje (Mackenzie, 2014a, s. 36).

Autonomia osobowa w ujęciu Mackenzie jest też jednocześnie – choć pod innym względem – relacyjna w sensie przyczynowym. Ten rodzaj relacyjności cechuje takie jej warunki, jak postawy związane z samooceną, a także wiele spośród zdolności i kompetencji ważnych dla kierowania sobą. Ich rozwój i trwanie u człowieka warunkowane są bowiem przyczynowo przez jego relacje interpersonalne i kontekst społeczno-kulturowy jego życia (Mackenzie, 2014a, s. 21–22, 36–38; Mackenzie, 2021, s. 38). Z powodu uwzględniania takich warunków autonomii, jak postawy związane z samooceną oraz gotowość do odpowiadania za swoje racje, teorię Mackenzie należy jednocześnie zaklasyfikować jako słabo treściwą.

Analizowana koncepcja spełnia sześć kryteriów satysfakcjonującej teorii autonomii zaproponowanych przez Dworkina (1988, s. 7–9). W przypadku pięciu z nich nie ma co do tego wątpliwości, bowiem teoria Mackenzie: jest logicznie spójna, przedstawia autonomię jako coś empirycznie możliwego, wyjaśnia jej wartość, pozostaje w zgodzie z typowymi filozoficznymi zastosowaniami pojęcia autonomii, a także z przyjmowanymi przez nas konkretnymi (pojęciowymi, normatywnymi, empirycznymi) sędziami dotyczącymi autonomii.

Pewien niepokój mogłoby co najwyżej budzić spełnienie przez nią kryterium ideologicznej neutralności, gdyż – jak się zauważa – realizacja takich warunków jak wolność i dostęp do możliwości wymaga otoczenia społecznego przyjmującego liberalny system wartości. Ich uwzględnianie w ramach teorii pociągałoby więc za sobą możliwość podważania autonomii osób, które co prawda są w pełni zdolne do kierowania sobą, ale muszą (lub chcą) żyć w nieliberalnym kontekście społecznym (Christman, 2004, s. 156–159). Koncepcja Mackenzie jest jednak odporna na taki zarzut, ponieważ nie traktuje normatywnie obciążonych „zewnątrznych” warunków

autonomii jako koniecznych, a kiedy bierze je pod uwagę, dopuszcza odpowiednie do kontekstu interpretowanie ich treści, przypisanie im mniejszej wagi i niezbyt wymagających progów minimalnych, a także kompensowanie niskiego poziomu ich realizacji (Mackenzie, 2014a, s. 40).

Kolejną zaletą teorii Mackenzie jest jej odporność na zarzuty stawiane koncepcjom definiującym autonomię jako identyfikowanie się z własną tożsamością praktyczną i wynikającymi z niej decyzjami (albo też jako brak poczucia wyobcowania wobec tej tożsamości i wynikających z niej decyzji). Jeden z nich głosi, iż koncepcje te nakazywałyby przypisać autonomię komuś, kto co prawda spełnia wskazany warunek, ale jednocześnie nie zdaje sobie sprawy, że jego tożsamość została ukształtowana drogą konsekwentnego poddawania go (przez innych ludzi lub instytucje) manipulacji (np. okłamywaniu). Tymczasem – jak się zauważa – takiego człowieka należałoby raczej uznać za paradygmatyczny przykład osoby o ograniczonej autonomii (Taylor, 2009, s. 1–17, 40–42).

Inny typowy zarzut wobec utożsamiania autonomii z proceduralnie rozumianą autentycznością zwraca uwagę, że jeśli tożsamość praktyczna człowieka ukształtowała się w opresyjnym kontekście społeczno-kulturowym wrogim wobec autonomii, to jej częścią są bezwiednie zinternalizowane normy i przekonania, nakazujące mu niepodejmowanie dążeń do samodzielnego decydowania o sobie. Człowiek taki będzie z dużym prawdopodobieństwem refleksyjnie akceptował swoją tożsamość jako „własną” i w pełni się z nią identyfikował, a mimo to jego codzienne funkcjonowanie będzie oczywiście naznaczone brakiem autonomii (Friedman, 2003, s. 19; Mackenzie, 2002, s. 190)²⁹.

Jak nietrudno zauważyć, teoria Mackenzie może bez trudu obronić się przed tego rodzaju krytyką, ponieważ w jej świetle akceptacja własnej tożsamości praktycznej nie jest warunkiem wystarczającym autonomii. Przywołanie zaś innych uwzględnionych w teorii jej warunków pozwoliłoby bez trudu wykluczyć z grona osób w pełni autonomicznych nieświadome

²⁹ Będzie tak nie tylko dlatego, że człowiek taki nie będzie nawet próbował autonomicznie decydować o sobie, ale również dlatego, że nie będzie rozwijał w sobie umiejętności i kompetencji ważnych dla autonomicznego kierowania sobą. Trzeba też podkreślić, że nie chodzi tutaj o przypadek kogoś, kto autonomicznie, w imię swoich wartości, wybiera brak autonomii (np. wybierając na jakiś czas całkowite posłuszeństwo wobec zwierzchników). Na koniec warto dodać, że przedstawiony w tekście zarzut jest niekiedy uznawany za argument na rzecz „treściwego” podejścia do autonomii.

ofiary manipulacji lub opresji społecznej. Przykładowo można uznać, że w obu przedstawionych właśnie sytuacjach naruszone są „zewnętrzne” warunki autonomii ujęte w teorii Mackenzie³⁰. Oprócz tego warto zauważyć, że w koncepcji Mackenzie autentyczność tożsamości praktycznej nie jest też warunkiem koniecznym autonomii, więc przy ocenianiu poziomu autonomii w niektórych kontekstach może po prostu zostać pominięta.

Łatwość, z jaką omawiana teoria radzi sobie z tymi zarzutami, pokazuje, jak dużą jej zaletą jest rezygnacja z ujmowania autonomii w kategoriach warunków koniecznych i wystarczających, uznanie jej za wielowymiarową, a także uzależnienie jej oceny od kontekstu. Mackenzie podkreśla, że dzięki tym zabiegom jej teoria umożliwia też precyzyjną diagnozę deficytów autonomii, co może następnie wspomagać proces projektowania odpowiednich interwencji i wsparcia dla osób nimi dotkniętych (2015, s. 63; 2017, s. 525; 2019, s. 148). Opis tych deficytów pozwala również identyfikować ich źródła w relacjach jednostki i w jej otoczeniu społeczno-kulturowym, dzięki czemu możliwa jest realizacja ważnej funkcji, jaką według Mackenzie powinna spełniać adekwatna koncepcja autonomii. Jest nią wyjaśnianie, w jaki sposób różne formy opresji społecznej mogą zagrażać autonomii jednostki (Mackenzie, 2015, s. 55–56; 2019, s. 147).

Uwagi krytyczne

Akcentując wartość omawianej teorii, nie można jednak pominąć kilku wątpliwości, jakie budzi. Australijska badaczka zdecydowała się uwzględnić autentyczność tożsamości praktycznej jako niekonieczny i niewystarczający warunek autonomii dopuszczalny w niektórych kontekstach. Zauważa przy tym, że krytyczna refleksja dotycząca tożsamości praktycznej powinna uwzględniać nie tylko stan aktualny, ale również historię i kontekst jej kształtowania się (Mackenzie, 2014a, s. 32). Podkreśla poza tym, że ocena autonomii powinna też brać pod uwagę to, jaki charakter miała socjalizacja danej

³⁰ Jeśli chodzi o pierwszy z przedstawionych zarzutów, można argumentować, że poddawanie kogoś konsekwentnej manipulacji oznacza, że narusza się jego swobody, a także kwestionuje jego status autonomicznego podmiotu. Jeśli chodzi o drugi zarzut, można zauważyć, że funkcjonowanie w otoczeniu nie sprzyjającym autonomii również oznacza, że jednostka ma ograniczony zakres wolności i dostępu do możliwości, a także nie jest uznawana za autonomiczny podmiot.

osoby, a także wymagać od niej zdolności i kompetencji, które ograniczałyby ewentualny wpływ socjalizacji opresyjnej (Mackenzie, Stoljar, 2000, s. 17). Dzięki temu zabezpiecza się przed zarzutami stawianymi wielu koncepcjom autonomii, że wybrane części tożsamości jednostki traktuje się w nich, jak gdyby były one jej prawdziwym „ja”, nieulegającym zmianom i wolnym od wpływu socjalizacji (Mackenzie, Stoljar, 2000, s. 14–15). Australijska badaczka świetnie zdaje też sobie sprawę z innych zarzutów dotyczących koncepcji uznających autentyczność za element autonomii (Mackenzie, 2014a, s. 31–32)³¹. Niestety ignoruje te z nich, które podważają sensowność brania pod uwagę warunku autentyczności w jakiegokolwiek formie.

Nietrudno zauważyć, że krytyczna refleksja jednostki, będąca podstawą oceny autentyczności danego składnika jej tożsamości praktycznej, musi opierać się na jakichś kryteriach³². Dostarczają ich te fragmenty owej tożsamości, które w danym momencie nie są przez jednostkę kwestionowane (Friedman, 2000, s. 41). Aby jednak taka refleksja mogła pełnić swoją rolę rozstrzygania o autentyczności, to kryteria, do których się przy tym odwołuje, również musiałyby być autentyczne. Inaczej mówiąc, także i one musiałyby zostać zaakceptowane dzięki krytycznej refleksji jako „własne”, co znowu rodzi pytanie o autentyczność kryteriów tej jeszcze bardziej podstawowej refleksji krytycznej. Pojawia się więc dylemat: albo zgodzimy się, że mamy tu do czynienia z regresem w nieskończoność, co oznaczałoby, że autentyczność nie jest możliwa, albo też przyjmiemy, że uzasadnienia autentyczności kolejnych fragmentów tożsamości praktycznej jednostki muszą się prędzej czy później odwołać do takich jej elementów, które ukształtowały się bez udziału refleksyjnej akceptacji (np. są wrodzone lub powstały wskutek niereflektowanego przez jednostkę wpływu jej otoczenia

³¹ Teoriom tym zarzucano też, że te elementy psychiki, dzięki którym mówi się o autonomii (autentyczności) innej jej części mogą powstawać na drodze manipulacji. Kolejny zarzut dodaje, że kiedy z kolei zapytamy o ich autonomię, powstaje regres w nieskończoność. Trzeci zarzut wskazuje, iż gdyby – próbując zapobiec wspomnianemu regresowi – po prostu uznać, że pewne części psychiki są źródłem autonomii innych jej części, to powstaje wątpliwość, jak źródłem autonomii może być coś, co samo nie jest autonomiczne (Taylor, 2005, s. 5–6). Zwracano też uwagę, że ignoruje się możliwość konfliktów między różnymi motywami jednostki oraz lekceważy się istnienie nieświadomych aspektów ludzkiej psychiki wymykających się refleksji (Mackenzie, Stoljar, 2000, s. 14–15).

³² Jest tak również wtedy, gdy refleksja ta bierze pod uwagę także historię i kontekst powstania badanych elementów tożsamości praktycznej.

społecznego i kulturowego). Jeśli wykluczy się arbitralne przypisanie im autentyczności, trzeba się zgodzić, że pozostają one już poza opozycją „tego, co autentyczne” i „tego, co nieautentyczne”. Także i w tym przypadku okazuje się więc, że autentyczność badanego aspektu tożsamości praktycznej nie jest ostatecznie możliwa. Jaki więc sens ma uwzględnianie w teorii warunku, którego rzeczywiste spełnienie nie jest możliwe?

Gdyby zaś obrońca warunku autentyczności powiedział, że jego prawdziwą funkcją jest tylko gwarantowanie wystarczającego stopnia integracji i spójności różnych aspektów tożsamości praktycznej jednostki, to w odpowiedzi należałoby zauważyć, że skuteczna realizacja tej funkcji również nie jest możliwa. Zgodnie z założeniami przyjmowanymi przez Mackenzie istnieją bowiem nieświadome aspekty psychiki człowieka, często skonfliktowane z jej częścią świadomą. Wymykają się one jego świadomej, krytycznej refleksji, ale mogą wpływać na jej rezultat, co czyni ją mało wiarygodną jako narzędzie oceny spójności i integracji tożsamości praktycznej.

Inna wątpliwość dotyczy tego, czy warunki autonomii wskazane przez Mackenzie wystarczą do oceniania autonomii konkretnych decyzji. W świetle jej teorii pełnia autonomii osobowej to posiadanie przez człowieka statusu społecznego jednostki autonomicznej, a także określonych cech i umiejętności. Podjęcie autonomicznej decyzji wymaga jednak czegoś więcej, a mianowicie faktycznego wykorzystania tych umiejętności. Z tego, że ktoś jest zdolny do decydowania o sobie w zgodzie z własną tożsamością praktyczną, wcale jeszcze nie wynika bowiem, że pewna jego konkretna decyzja będzie z nią zgodna (Faden, Beauchamp, 1986, s. 235–237). Niezbędne wydaje się więc uzupełnienie omawianej teorii o wymogi dotyczące autonomicznego podejmowania decyzji. Przede wszystkim chodziłoby o uzupełnienie osi kierowania sobą o warunek zgodności (lub niesprzeczności) konkretnych decyzji jednostki z jej tożsamością praktyczną. Nieobecność tego warunku w teorii Mackenzie jest tym dziwniejsza, że bezpośrednio nawiązuje do niego przedstawiona przez nią ogólna charakterystyka osi kierowania sobą (np. Mackenzie, 2014a, s. 31).

Na koniec warto też podkreślić, że w przypadku oceniania autonomii konkretnej decyzji duże znaczenie należałoby zawsze przypisywać temu, w jakim zakresie i stopniu *w momencie jej podejmowania* spełnione były trzy warunki autonomii dotyczące otoczenia. Od tego zależeć będzie bowiem odpowiedź na pytanie o obecność w tym momencie w otoczeniu jednostki czynników ograniczających autonomię jej decyzji. Typowo zalicza się

do nich: przymus, niektóre formy perswazji, a także manipulację, która obejmuje np. okłamywanie lub celowe zwodzenie dotyczące okoliczności ważnych dla podejmowanej decyzji (Bauchamp, Childress, 2009, s. 132–134). Niektórzy autorzy uznają wręcz, że nieuleganie w sposób nieświadomy manipulacji jest kluczowe dla autonomii decyzji (np. Taylor, 2009, s. 3–17).

Bibliografia

- Anderson, J., Honneth, A. (2005). Autonomy, Vulnerability, Recognition, and Justice. W: J. Christman, J. Anderson (red.), *Autonomy and the Challenges to Liberalism: New Essays* (s. 127–149). New York: Cambridge University Press.
- Barclay, L. (2000). Autonomy and the Social-Self. W: C. Mackenzie, N. Stoljar (red.), *Relational Autonomy: Feminist Perspectives on Autonomy, Agency, and the Social Self* (s. 52–71). Oxford: Oxford University Press.
- Beauchamp, T.L., Childress, J.F. (2009). *Principles of Biomedical Ethics*. Oxford: Oxford University Press.
- Beauchamp, T.L. (2010). *Standing on Principles. Collected Essays*. Oxford: Oxford University Press.
- Benson, P. (1991). Autonomy and Oppressive Socialization. *Social Theory and Practice*, 17 (3), 385–408. <https://doi.org/10.5840/soctheorpract199117319>.
- Benson, P. (2005). Taking Ownership. Authority and Voice in Autonomous Agency. W: J. Christman, J. Anderson (red.), *Autonomy and the Challenges to Liberalism: New Essays* (s. 101–126). New York: Cambridge University Press.
- Benson, P. (2014). Feminist Commitments and Relational Autonomy. W: A. Veltman, M. Piper (red.), *Autonomy, Oppression, and Gender* (s. 87–113). Oxford: Oxford University Press.
- Brison, S.J. (2000). Relational Autonomy and Freedom of Expression. W: C. Mackenzie, N. Stoljar (red.), *Relational Autonomy: Feminist Perspectives on Autonomy, Agency, and the Social Self* (s. 280–299). Oxford: Oxford University Press.
- Christman, J. (2004). Relational autonomy, liberal individualism, and the social constitution of selves. *Philosophical Studies*, 117 (1–2), 143–164.
- Christman, J. (2005). Procedural Autonomy and Liberal Legitimacy. W: J.S. Taylor (red.), *Personal Autonomy: New Essays on Personal Autonomy and its Role in Contemporary Moral Philosophy* (s. 277–298). New York: Cambridge University Press.

- Christman, J. (2009). *The Politics of Persons: Individual Autonomy and Socio-Historical Selves*. New York: Cambridge University Press.
- Christman, J. (2014). Relational Autonomy and the Social Dynamics of Paternalism. *Ethical Theory and Moral Practice*, 17 (3), 369–382. <https://doi.org/10.1007/s10677-013-9449-9>.
- D'Ambrosio, J., Hedden, B. (2023). Multidimensional adjectives. *Australasian Journal of Philosophy*. Pobrano z: <https://philarchive.org/archive/DAMMANv> (11.12.2023).
- Dworkin, G. (1976). Autonomy and Behavior Control. *Hastings Center Report*, 6, 23–28. <https://doi.org/10.2307/3560358>.
- Dworkin, G. (1988). *The Theory and Practice of Autonomy*. New York: Cambridge University Press.
- Faden, R.R., Beauchamp, T.L. (1986). *A History and Theory of Informed Consent*. Oxford: Oxford University Press.
- Feinberg, J. (1986). Harm to Self. The Moral Limits of The Criminal Law. Vol. 3. *Law and Philosophy*, 7 (1), 107–122.
- Friedman, M. (2000). Autonomy, Social Disruption, and Women. W: C. Mackenzie, N. Stoljar (red.), *Relational Autonomy: Feminist Perspectives on Autonomy, Agency, and the Social Self* (s. 35–51). Oxford: Oxford University Press.
- Friedman, M. (2003). *Autonomy, Gender, Politics*. Oxford: Oxford University Press.
- Govier, T. (1993). Self-Trust, Autonomy, and Self-Esteem. *Hypatia*, 8 (1), 99–120. <https://doi.org/10.1111/j.1527-2001.1993.tb00630.x>.
- Frankfurt, H.G. (1997). Wolność woli i pojęcie osoby. Tłum. J. Nowotniak. W: J. Hołówka (red.), *Filozofia moralności. Postanowienie i odpowiedzialność moralna* (s. 21–39). Warszawa: Aletheia.
- Haworth, L. (1991). Review: Dworkin on Autonomy. *Ethics*, 102 (1), 129–139.
- Hedden, B., Nebel, J.M. (b.d.w.). Multidimensional Concepts and Disparate Scale Types. Pobrano z: <https://philpapers.org/archive/HEDMCA.pdf> (3.12.2023).
- Hill, T.E. (2013). Kantian Autonomy and Contemporary Ideas of Autonomy. W: O. Sensen (red.), *Kant on Moral Autonomy* (s. 15–31). New York: Cambridge University Press.
- Hinchman, L. (1996). Autonomy, Individuality, and Self-Determination. W: J. Schmidt (red.) *What is Enlightenment? Eighteenth-Century Answers and Twentieth-Century Questions* (s. 488–516). Berkeley–Los Angeles–London: University of California Press.

- Kant, I. (1953). *Uzasadnienie metafizyki moralności*. Tłum. M. Wartenberg. Warszawa: Państwowe Wydawnictwo Naukowe.
- Korsgaard, Ch.M. (1996). *The Sources of Normativity*. New York: Cambridge University Press.
- Mackenzie, C. (2000). Imagining oneself otherwise. W: *Relational Autonomy: Feminist Perspectives on Autonomy, Agency, and the Social Self* (s. 124–150). Oxford University Press.
- Mackenzie, C. (2002). Critical reflection, self-knowledge, and the emotions. *Philosophical Explorations*, 5 (3), 186–206.
- Mackenzie, C., Scully, J.L. (2007). Moral imagination, disability, and embodiment. *Journal of applied philosophy*, 24 (4), 335–351.
- Mackenzie, C. (2008a). Introduction: Practical Identity and Narrative Agency. W: C. Mackenzie, K. Atkins (red.), *Practical Identity and Narrative Agency* (s. 1–28). New York: Routledge.
- Mackenzie, C. (2008b). Relational Autonomy, Normative Authority and Perfectionism. *Journal of Social Philosophy*, 39 (4), 512–533.
- Mackenzie, C. (2014a). Three Dimensions of Autonomy: A Relational Analysis. W: A. Veltman, M. Piper (red.), *Autonomy, Oppression, and Gender* (s. 15–41). Oxford: Oxford University Press.
- Mackenzie, C. (2014b). Embodied agents, narrative selves. *Philosophical Explorations*, 17 (2), 154–171. DOI: 10.1080/13869795.2014.886363.
- Mackenzie, C., Rogers, W., Dodds, S. (2014). Introduction: What is vulnerability and why does it matter for moral theory. W: C. Mackenzie, W. Rogers, S. Dodds (red.), *Vulnerability: New essays in ethics and feminist philosophy* (s. 1–29). Oxford: Oxford University Press.
- Mackenzie, C., Walker, M. (2015). Neurotechnologies, personal identity, and the ethics of authenticity. W: J. Clausen, N. Levy (red.), *Handbook of Neuroethics* (s. 373–392). Dordrecht: Springer.
- Mackenzie, C. (2015). Responding to the agency dilemma: Autonomy, adaptive preferences, and internalized oppression. W: M. Oshana (red.), *Personal autonomy and social oppression* (s. 48–67). New York–London: Routledge.
- Mackenzie, C. (2017). Feminist Conceptions of Autonomy. W: A. Garry, S.J. Khader, A. Stone (red.), *The Routledge Companion to Feminist Philosophy* (s. 515–527). New York–London: Routledge.

- Mackenzie, C. (2019). Feminist innovation in philosophy: Relational autonomy and social justice. *Women's Studies International Forum*, 72, 144–151. <https://doi.org/10.1016/j.wsif.2018.05.003>.
- Mackenzie, C. (2021). Relational Equality and the Debate Between Externalist and Internalist Theories of Relational Autonomy. W: N. Stoljar, K. Voigt (red.), *Autonomy and Equality: Relational Approaches* (s. 32–56). New York–London: Routledge.
- Mackenzie, C., Stoljar, N. (2000). Introduction. W: C. Mackenzie, N. Stoljar (red.), *Relational Autonomy: Feminist Perspectives on Autonomy, Agency, and the Social Self* (s. 3–31). Oxford: Oxford University Press.
- McLeod, C. (2002). *Self-Trust and Reproductive Autonomy*. Cambridge, MA: MIT Press.
- Meyers, D.T. (1987). Personal Autonomy and the Paradox of Feminine Socialization. *Journal of Philosophy*, 84 (11), 619–628. <https://doi.org/10.5840/jphil198784117>.
- Meyers, D.T. (2005). Decentralizing Autonomy. Five Faces of Selfhood. W: J. Christman, J. Anderson (red.), *Autonomy and the Challenges to Liberalism: New Essays* (s. 27–55). New York: Cambridge University Press.
- Meyers, D.T. (2014). The Feminist Debate over Values in Autonomy Theory. W: A. Veltman, M. Piper (red.), *Autonomy, Oppression, and Gender* (s. 114–140). Oxford: Oxford University Press.
- Mele, A.R. (2005). Autonomy, Self-Control, and Weakness of Will. W: R. Kane (red.), *The Oxford Handbook of Free Will* (s. 529–548). Oxford: Oxford University Press.
- Oshana, M.A.L. (1998). Personal Autonomy and Society. *Journal of Social Philosophy*, 29 (1), 81–102. <https://doi.org/10.1111/j.1467-9833.1998.tb00098.x>.
- Rawls, J. (1994). *Teoria sprawiedliwości*. Tłum. M. Panufnik, J. Pasek, A. Romaniuk. Warszawa: Wydawnictwo Naukowe PWN.
- Raz, J. (1986). *The Morality of Freedom*. Oxford: Clarendon Press.
- Schneewind, J.B. (1998). *The Invention of Autonomy. A History of Modern Moral Philosophy*. New York: Cambridge University Press.
- Schneewind, J.B. (2013). Autonomy after Kant. W: O. Sensen (red.), *Kant on Moral Autonomy* (s. 146–168). New York: Cambridge University Press.
- Stoljar, N. (2011). Informed Consent and Relational Conceptions of Autonomy. *Journal of Medicine and Philosophy*, 36 (4), 375–384. <https://doi.org/10.1093/jmp/jhr029>.

- Stoljar, N., Voigt, K. (2021). Introduction: A Relational Turn in Political Philosophy. W: N. Stoljar, K. Voigt (red.), *Autonomy and Equality: Relational Approaches* (s. 1–31). New York–London: Routledge.
- Taylor, J.S. (2005). Introduction. W: J.S. Taylor (red.), *Personal Autonomy: New Essays on Personal Autonomy and its Role in Contemporary Moral Philosophy* (s. 1–29). New York: Cambridge University Press.
- Taylor, J.S. (2009). *Practical Autonomy and Bioethics*. New York–London: Routledge.
- Westlund, A.C. (2009). Rethinking Relational Autonomy. *Hypatia*, 24 (4), 26–49. <https://doi.org/10.1111/j.1527-2001.2009.01056.x>.

Nota o autorze

Mateusz Penczek – doktor nauk humanistycznych w zakresie filozofii, adiunkt na Wydziale Nauk Społecznych Uniwersytetu Śląskiego. Zainteresowania badawcze: historia filozofii, filozofia i etyka współczesna, etyczne aspekty ochrony praw i podmiotowości osób z niepełnosprawnością, zagadnienia psychologiczne związane ze sprawiedliwością, opresją, niepełnosprawnością oraz pracą z osobami z niepełnosprawnością.

Address for correspondence: Wydział Nauk Społecznych UŚ, ul. Grażyńskiego 53, 40-126 Katowice.

Cytowanie

Penczek, M. (2024). Autonomia osobowa z perspektywy relacyjnej teorii autonomii Catriony Mackenzie. *Analiza i Egzystencja*, 67 (3), 27–54. DOI: 10.18276/aie.2024.67-02.



MALWINA ROLKA

Uniwersytet Śląski w Katowicach
ORCID: 0000-0001-6906-3550
email: malwina.rolka@us.edu.pl.

Filozofia homo romanticus
w zwierciadle teorii literackiej Friedricha Schlegla

Słowa kluczowe: Friedrich Schlegel, poezja romantyczna, refleksja, wolność, indywidualność, organicyzm

Keywords: Friedrich Schlegel, Romantic poetry, reflection, freedom, individuality, organicism

Philosophy of homo romanticus
in the mirror of Friedrich Schlegel's literary theory

Abstract

The article aims to analyse the philosophical content of selected aspects of the literary theory of Friedrich Schlegel – one of the initiators and main founders of the Jena Romantic school. Schlegel is primarily known as a brilliant literary critic who popularised the notion of Romantic poetry. Romantic poetry, however, is not only a fundamental concept of the new school of literary studies, but also an important element of Schlegel's philosophy of man, describing the model of modern humanity. In this paper, I assume that Schlegel's literary theory is a mirror that generates a reflection of Romantic man. In this perspective, all the main features of Romantic literature – self-consciousness, reflexivity, freedom, individuality, the striving for the infinite or organicism – in fact turn out to be attributes of its creator.

Wprowadzenie

W intelektualnej historii nowożytnej Europy niewiele jest terminów, które wzbudzały tak liczne spory i kontrowersje, jak słowo „romantyzm”. Wśród zarzutów podnoszonych przeciwko niemu wskazywano zwłaszcza na rozbieżności między zjawiskami występującymi pod tą nazwą w poszczególnych krajach europejskich, ale również na różnice dzielące reprezentantów tych samych tradycji narodowych. Już Wilhelm Dilthey postulował, aby „położyć zdecydowany kres nadużyciom, jakie od pięćdziesięciu lat z górą popełniane są w związku z tą nazwą” (Dilthey, 1982, s. 325), a Arthur Lovejoy upatrywał w niej „pomieszania terminologii i myśli, które od stulecia jest skandalem w historii i krytyce literackiej” (Lovejoy, 1955a, s. 234). Mimo wszystko możemy dość precyzyjnie określić miejsce i moment narodzin historycznego nurtu ideowego zwanego dziś romantyzmem, który ma swój początek w Niemczech u schyłku XVIII wieku, kiedy to młody Friedrich Schlegel – jeden z najbardziej błyskotliwych i oryginalnych krytyków literackich swoich czasów, zaliczany współcześnie do głównych przedstawicieli jenajskiej szkoły wczesnego romantyzmu – publikuje w drugim numerze czasopisma „Athenaeum” charakterystykę *romantische Poesie*. I choć samo to pojęcie stosował on już wcześniej, a ponadto było ono używane również w poprzednich wiekach – między innymi przez Ludovica Ariosto, Torquato Tasso, Jeana Chapelaina czy Thomasa Wartona (Wellek, 1949, s. 3) – to wydaje się, że dopiero wraz z publikacją Schlegla z 1798 roku „pokolenie przełomu wieków znalazło w nim trafny wyraz własnego rozumienia nowoczesności” (Jauss, 1999, s. 29), którego nie sposób jednak ograniczyć tylko do sfery teorii i praktyki literackiej. Poezja romantyczna w sensie, w jakim pojęcie to rozumiał Schlegel, rodzi się w orbicie wpływów myśli transcendentalnego idealizmu Johanna Gottlieba Fichtego, filozofii sztuki Friedricha Schillera oraz powieści Johanna Wolfganga Goethego *Lata nauki Wilhelma Meistra*, urastając tym samym do rangi nie tylko sztandarowego hasła nowej szkoły badań literackich, ale również kategorii antropologicznej opisującej model człowieka i jego stosunku do świata u progu nowoczesności. W niniejszym opracowaniu zakładam, że literacka teoria Schlegla jest swego rodzaju zwierciadłem, generującym odbicie *homo romanticus*, co umożliwia spojrzenie na ideę *romantische Poesie* z perspektywy antropologicznej, w ramach której jej główne cechy,

takie jak indywidualność, refleksyjność, wolność, dążenie do nieskończoności czy organicznej jedności, nie są niczym innym niż atrybutami poety, zmierzającego do rozpoznania w niej samego siebie.

Literackie źródła romantyzmu jenajskiego: *romantische Poesie* i *Roman*

Źródłem tego, co romantyczne Schlegel upatruje „w owej epoce rycerzy miłości i baśni, z której pochodzi samo to słowo i rzecz” (Schlegel, 2000, s. 172). W językach europejskich przymiotnik „romantyczny” pochodzi od słowa *romanice*, które w średniowiecznej łacinie oznaczało dosłownie „poezję tworzoną w mowie pospolitej” i odnosiło się do gatunku romanów rycerskich (Jauss, 1999, s. 30). Z uwagi na to, że „w odróżnieniu od wcześniejszego, XI-wiecznego eposu (*chanson de geste*), którego tematem są dzieje narodu i jego władców, romans rycerski obrazował prywatne życie bohaterów, historię ich przygód i doznań osobistych” (Żabski, 1990, s. 5), stanowi on prototyp powieści (niem. *Roman*). Jako nowoczesna forma romansu rycerskiego przedstawia „dzieje duszy, która wyruszyła w świat z zamiarem poznania siebie i szuka przygód, by się w ten sposób potwierdzić i dotrzeć do swej własnej istoty” (Lukács, 1968, s. 81). Schlegel nie pozostawia wątpliwości, że to właśnie *Roman* jest znakiem rozpoznawczym sztuki poetyckiej jego czasów, podobnie jak epopeja definiuje literacki świat starożytnych: „tak jak nasza poezja od powieści, tak grecka zaczęła się od eposu i w nim się potem rozplynęła” (Schlegel, 2000, s. 172). Etymologiczne pokrewieństwo *romantische Poesie* i *Roman* nie jest zatem przypadkowe, o czym przekonujemy się, zestawiając ich bardziej szczegółowe charakterystyki. Poezja romantyczna ma na celu „ponownie zjednoczyć wszystkie odrębne gatunki poezji oraz zbliżyć poezję do filozofii i retoryki. Chce ona i powinna także poezję i prozę, genialność i krytykę, poezję kunsztowną i poezję naturalną, już to mieszać, już to stapiać [...]” (Schlegel, 2009, s. 60–61). Odpowiada jej struktura *Roman* jako „przemieszania narracji, śpiewu i innych form” (Schlegel, 2000, s. 174), przypominającego znane z architektury „prawdziwe arabeski”, które, jak podkreśla niemiecki autor, stanowią „jedynie w naszej epoce naturalne wytwory romantyzmu” (Schlegel, 2000, s. 175). Za sprawą kluczowej w obu przypadkach kategorii „mieszaniny” (niem. *Mischung*) Schlegel „rozciągnął pojęcie powieści tak

dalece poza jej gatunkowe wyznaczniki, że niemal zbiega się ono z pojęciem poezji romantycznej” (Behler, 1993, s. 154), choć nie jest z nim tożsame. O ile bowiem „romantyczny rodzaj poetycki pozostaje jedynym, który jest czymś więcej niż rodzajem; który jest poniekąd samą sztuką poetycką” (Schlegel, 2009, s. 62), o tyle „powieść jest romantyczna tylko w pewnych granicach, jest powieścią stosowaną” (Schlegel, 2000, s. 173). Innymi słowy, we wczesnoromantycznej teorii estetycznej „sztuka jest continuum form, a powieść [...] zrozumiałym przejawem tego continuum” (Benjamin, s. 100).

Problem zachodzenia na siebie zakresów znaczeniowych pojęć *romantische Poesie* i *Roman* stał się następnie istotnym elementem sporu o genezę wczesnego romantyzmu. Rudolf Haym łączy pojęcie poezji romantycznej z „Athenaeum” z *Latami nauki Wilhelma Meistra*, które w jego opinii Schlegel miał uznać za uniwersalny wzór wszelkich późniejszych utworów tego typu oraz na tej podstawie ugruntować przekonanie o pierwszeństwie powieści wśród innych gatunków literackich (Haym, 1870, s. 250–255). Z kolei według Lovejoya to nie powieść Goethego, lecz esej Schillera *O poezji naiwnej i sentymentalnej* „był głównym instrumentem nawrócenia Schlegla na nową – romantyczną – wiarę estetyczną” (Lovejoy, 1955b, s. 218–219)¹. Zarazem jednak nie ulega wątpliwości, że „powieść *Wilhelm Meister* zostaje entuzjastycznie przyjęta i staje się wzorem dla niemal wszystkich romantycznych powieściopisarzy” (Emrich, 1978, s. 292), zapewniając „legitymizację dla ich przedsięwzięć, skoro wielki Goethe wskazał im drogę” (Paulin, 2005, s. 125). I choć Schlegel rzeczywiście nie nazywa *Wilhelma Meistra* dziełem romantycznym ani nie zestawia z tym przymiotnikiem nazwiska jego autora, to bez wątpienia dostrzega istotne zbieżności pomiędzy niektórymi cechami utworu a własną koncepcją *Roman*.

¹ Rozprawa Schillera powstaje w tym samym czasie, gdy Schlegel pisze tekst *Über das Studium der griechischen Poesie*, który ukazuje się dopiero w 1797 roku, na krótko po publikacji *O poezji naiwnej i sentymentalnej*. Oba pisma są tematycznie zbieżne i współcześnie powszechne jest przekonanie, że wpisują się w *querelle des anciens et des modernes* (spór „starożytników” z „nowożytnikami”), którego przedmiotem była „sprzeczność między wzorcowym charakterem antyku a postępowym charakterem usamodzielnionej nowoczesności” (Jauss, 1999, s. 62). Rozstrzygnięcia Schlegla i Schillera „aktualizują ponownie horyzont pytań i rozwiązań, wysuwanych sto lat wcześniej przez *querelle*” (Jauss, 1999, s. 69), w ramach którego pierwszy z nich lokuje się na pozycji „starożytników”, podczas gdy drugi przyjmuje stanowisko *modernes*. Pod wpływem lektury eseju Schillera Schlegel zmienia stronę w tym sporze, stając się rzecznikiem poezji nowoczesnej.

W odróżnieniu od Schillera, który uważał, że „forma »Meistra«, jak w ogóle wszelka forma powieściowa, jest absolutnie niepoetycka” (Schiller, 1974, s. 196), dla romantyków wraz z powieścią Goethego „na pierwszy plan wysunął się gatunek, który do tej pory nie był uznawany za medium poezji” (Behler, 1993, s. 168), a w którym odnaleźli oni najbardziej adekwatną formę wyrazu dla *romantische Poesie*. Równocześnie jenajczycy dążyli do spotęgowania cech strukturalnych *Wilhelma Meistra*: „podczas gdy Goethe bez skrupułów przerywał tok narracji dyskusjami (o Hamlecie) lub przeplatany wątkami, romantycy wprowadzili zasadę wielości form i stylów w ramach powieści” (Paulin, 2005, s. 126). Jej praktyczne zastosowanie Schlegel prezentuje w jedynej swej *Roman – Lucinde* (1799) – która paradoksalnie wśród pierwszych jej komentatorów z tej właśnie przyczyny zyskała opinię dzieła dyskredytującego cały gatunek powieści, utrzymującą się aż do XX wieku (Bernofsky, 2005, s. 90).

Romantyczna indywidualność w zwierciadle poetyckiej refleksji

Wbrew potocznemu przekonaniu o primacie uczucia nad rozumem istotą wczesnego romantyzmu nie są porywy namiętności, lecz „pasja myślenia oraz potrzeba – prawie abstrakcyjna, zgłaszana przez poezję – wyrażania się i spełniania poprzez refleksję” (Blanchot, 2013, s. 298). Pojęcie refleksji nieodmiennie wiąże myśl Schlegla z filozofią Immanuela Kanta i jego kontynuatorów, w tym przede wszystkim Fichtego, którego wykłady uniwersyteckie cieszyły się ogromną popularnością wśród młodych przedstawicieli kręgu jenajskiego. Głównym zadaniem, jakie stawia przed sobą Fichte, jest ustanowienie systemu Teorii Wiedzy, dzięki któremu „tak zwana dotąd filozofia byłaby nauką o nauce w ogóle” (Fichte, 1996, s. 25). Według Schlegla stanowi ona największe osiągnięcie myśli krytycznej, ponieważ odnosząc się do dwóch poziomów – z jednej strony do samego systemu wiedzy ludzkiej, a z drugiej do teorii tego systemu – „jest wszędzie filozofią i filozofią filozofii zarazem” (Schlegel, 2009, s. 98). Na pierwszym z tych poziomów mamy do czynienia z filozofującym podmiotem, który „jest bezspornie tylko Ja przedstawiającym”, natomiast na drugim z „Ja jako przedmiotem filozofowania” (Fichte, 1996, s. 69), czyli z pierwotną aktywnością ludzkiego ducha (absolutnym podmiotem, czystą świadomością), umożliwiającą dopiero wszelkie akty przedstawiania właściwe empirycznej

świadomości. Fichte odróżnia oba poziomy za pomocą niemieckich terminów *Darstellung* i *Vorstellung*, definiując „*Darstellung* jako aprioryczną aktywność świadomości, która polega na przedstawianiu empirycznego podmiotu jako *Vorstellung*” (Helfer, 1996, s. 67). Nieodzownym aktem ustanawiającym tę dwuwymiarową strukturę Teorii Wiedzy jako nauki i nauki o nauce jednocześnie jest refleksja, dzięki której „coś, co już samo w sobie jest formą – mianowicie konieczne działanie inteligencji – zostaje jako treść ujęte w nową formę, w formę wiedzy lub świadomości” (Fichte, 1996, s. 58). Innymi słowy, refleksja pozwala ująć filozofujący podmiot jako przedmiot filozofowania, w czym znajduje swój wyraz „nowa, epistemologiczna samoświadomość myślenia bezpośrednio teoriiotwórczego” (Siemek, 1977, s. 176). Właśnie te trzy podstawowe pojęcia Teorii Wiedzy – refleksję, przedstawienie i samoświadomość – Schlegel stosuje w charakterystyce swojej *romantische Poesie*. Zgodnie z definicją wyłożoną w „*Athenaeum*” poezja romantyczna „potrafi najnsadniej unosić się na skrzydłach refleksji poetyckiej w środku między przedstawionym a przedstawiającym, potęgując tę refleksję i powielając w nieskończonym szeregu luster” (Schlegel, 2009, s. 61). Schlegel nawiązuje tym samym do etymologii słowa refleksja, wywodzącego się od łacińskiego czasownika *reflectere* (odbijać), który opisuje optyczne zjawisko generowania obrazu przedmiotu na gładkiej powierzchni, pozostające w analogii do sposobu wytwarzania przedstawienia przez poznającą świadomość. Z kolei obraz powielany dzięki ustawionym naprzeciw siebie lustrom pozwala dostrzec w refleksji „proces operacji, które oprócz tego, że oznaczają aktywność zwierciadła reprodukującego przedmiot, warunkują odbicie zwierciadła, dzięki czemu widzi ono samo siebie” (Gasché, 1997, s. 16–17). Proces ten Schlegel nazywa również „pięknym samo-odzwierciedlaniem (*Selbstbespiegelung*)”, które w dziedzinie sztuki literackiej odpowiada transcendentalnemu przedstawianiu „wytwarzającego razem z wytworem” (Schlegel, 2009, s. 89). Refleksja pozwala zatem poecie oglądać samego siebie we własnym dziele literackim niczym w zwierciadle, w takim sensie jak w Teorii Wiedzy umożliwia rozpoznanie w przedstawieniu empirycznego podmiotu (wytworze) działania czystej świadomości (siły wytwarzającej)². Niemiecki autor dokonuje jednak tym samym dość

² Analogię tę lepiej oddaje oryginalny tekst fragmentu 116: „Und doch kann auch sie am Meissen zwischen dem Dargestellten und dem Darstellenden, frei von allem realen und idealen Interesse auf den Flügeln der poetischen Reflexion in der Mitte schweben,

dowolnej interpretacji słowa „transcendentalny”, które „ma w zestawieniu ze znaczeniem u Kanta sens wprost przeciwny, bo oznacza coś subiektywnego”, w związku z czym „Schlegel podmiot transcendentalny Fichtego pojmuje właśnie zbyt indywidualistycznie” (Łempicki, 1966, s. 156). W rezultacie nie mówi on już o Ja, ale o indywidualności, która stanowi centrum wczesnoromantycznej antropologii.

O fundamentalnym znaczeniu kategorii indywidualności w teorii Schlegla możemy wnioskować już z eseju *Über das Studium der griechischen Poesie*, gdzie podstawowym kryterium odróżniającym poezję nowoczesną od starożytnej okazuje się „przewaga tego, co charakterystyczne, indywidualne i interesujące” (Schlegel, 1979, s. 227). Jej przedmiotem jest unikalność jednostkowego doświadczenia: „odwołuje się ona do zróżnicowanego subiektywnego »zainteresowania« jednostek lub szczególnych typów umysłów”, a „za temat obiera sobie »das Charakteristische«, to znaczy indywidualną osobę lub неповtarzalną sytuację, a nie typ ogólny” (Lovejoy, 1955b, s. 207). Badanie typowych przejawów nowoczesnej poezji prowadzi Schlegla do wniosku, że jej celem „stała się oryginalna i interesująca indywidualność” (Schlegel, 1979, s. 244), co znajduje następnie swój wyraz w definicji *romantische Poesie*, która „potrafi tak się zatracić w przedstawianiu, że można by sądzić, że charakteryzowanie poetyckich indywidualów wszelkiego rodzaju jest dla niej wszystkim” (Schlegel, 2009, s. 61)³.

diese Reflexion immer wieder potenzieren und wie in einer endlosen Reihe von Spiegeln vervielfachen” (Schlegel, 1967, s. 182–183). Tekst niemiecki wyjaśnia nam, że Schleglowi chodzi o termin *Darstellung*, stosowany przez Fichtego na określenie apriorycznej aktywności podmiotu, a nie o słowo *Vorstellung*, oznaczające działanie empirycznej świadomości, co w polskim przekładzie nie jest jednoznaczne, ponieważ oba te wyrażenia mogą zostać przełożone jako „przedstawienie”.

³ W *Über das Studium der griechischen Poesie* koncentracja na tym, co indywidualne, odróżnia *interessante Poesie* od klasycznej sztuki literackiej antyku, w której Schlegel jako „starożytnik” upatruje niedoścignionego ideału wszelkiej poezji. Zmiana stanowiska w *querelle des anciens et des modernes*, w wyniku której niemiecki autor staje się rzecznikiem poglądu o wyższości sztuki nowoczesnej (romantycznej) nad starożytną, nie pociąga zarazem za sobą znaczących przeobrażeń w sposobie rozumienia jej istoty, ponieważ w rzeczywistości „nie tylko cechy, ale i główne historyczne ucieleśnienie nowoczesnego ideału poezji są dla Schlegla takie same przed i po przyjęciu go za swój własny” (Lovejoy, 1955c, s. 198). Pod wpływem idealizmu Fichtego i jego krytyki w kolejnych pracach Schlegla dyskurs indywidualności jest stale rozwijany i wzbogacany o nowe konteksty, co nie zmienia jednak faktu, że jego genezy należy poszukiwać w *Über das Studium der griechischen Poesie*.

Terminy *Individualität* i *Individuum* należy rozumieć tu zgodnie z ich łacińskim źródłosłowem akcentującym niepodzielność, łączącą się u Schlegla z uniwersalną zasadą twórczego działania natury, która „pragnie wiecznego cyklu coraz to nowych eksperymentów; chce też, by każdy z nich był sam w sobie doskonały, niepowtarzalny i nowy, był wiernym obrazem najwyższej niepodzielnej indywidualności” (Schlegel, 1962, s. 73). Celem refleksyjnego „samo-odzwierciedlenia”, którego w swojej sztuce dokonuje poeta romantyczny, jest zatem dotarcie do źródłowej zasady indywidualności, najpełniej definiującej ludzką naturę. Z tej perspektywy dość powszechne przeświadczenie, że w epoce romantyzmu „poezja i myśl wywodzą się z osobowości samego twórcy, jego doświadczeń i aspiracji” (Gerard, 1978, s. 255), wymaga uzupełnienia o rozróżnienie między osobowością i indywidualnością. Sformułowany w odniesieniu do terminologii Teorii Wiedzy postulat Schlegla, „aby filozofujące »Ja« nie było już tak często mylone z »Ja«, o którym się filozofuje” (Schlegel, 2009, s. 115), przekłada on następnie na kategorie swojej antropologii, utrzymując, że „indywidualność jest tym, co źródłowe i wieczne w człowieku; od osobowości nie zależy tak wiele” (Schlegel, 2009, s. 159). Relacja między tymi dwoma pojęciami opiera się na założeniu, że to nie osobowość człowieka decyduje o jego indywidualności, lecz odwrotnie: w sferze osobowości manifestuje się indywidualność, która stanowi najbardziej pierwotną cechę ludzkiego sposobu istnienia. Podobnie zatem jak w Teorii Wiedzy empiryczny akt świadomości jest uwarunkowany przez aprioryczną aktywność czystego Ja, tak u Schlegla osobowość człowieka jest odzwierciedleniem źródłowej indywidualności, która łączy poszczególne jednostki w tym sensie, że nadaje każdej z nich status absolutnie wyjątkowej i niepowtarzalnej. W świetle rozróżnienia osobowości i indywidualności nowego znaczenia nabiera również sformułowana przez Alberta Gerarda koncepcja „egotyzmu apersonalnego”, zgodnie z którą „poeta traktuje swoją osobę wyłącznie jako reprezentatywnego przedstawiciela gatunku ludzkiego” (Gerard, 1978, s. 260). W rzeczywistości spełnienie założenia „egotyzmu apersonalnego” staje się bowiem możliwe dopiero wówczas, gdy poeta postrzega unikalność swojej osoby przez pryzmat uniwersalnej zasady indywidualności jako powszechnej i źródłowej właściwości ludzkiej natury. Dlatego też Juliusz – bohater *Lucinde* – który nieprzypadkowo porównuje swoją ukochaną do zwierciadła, może powiedzieć: „tylko w nim postrzegam siebie jako istotę kompletną i harmonijną, czy też raczej we mnie i w tobie widzę całą ludzkość” (Schlegel, 1962, s. 10).

Zwierciadło samoświadomości pozwala zatem Juliuszowi dostrzec obraz nie tylko siebie samego, lecz również za jego pośrednictwem sięgnąć do najgłębszej tajemnicy ludzkiej natury.

Progresywna poezja uniwersalna albo antropologia wolności i nieskończoności

Refleksyjna orientacja *romantische Poesie* skłania Schlegla do nadania jej miana „progresywnej poezji uniwersalnej” (Schlegel, 2009, s. 60), określanej przez dwie ściśle powiązane z sobą cechy: „ona jedyna jest nieskończona, tak jak ona jedyna jest wolna” (Schlegel, 2009, s. 62). W jego antropologii przekładają się one na obraz *homo romanticus*, który osiąga swoją autonomię w twórczym procesie wykraczania w sferę tego, co ludzki umysł identyfikuje jako nieskończone. Romantyczne idee wolności i nieskończoności, pozostając w koniecznym związku z kategorią refleksji, po raz kolejny lokują myśl Schlegla w horyzoncie filozofii Fichtego. O ile już Kant upatrywał wyższości prawodawczego rozumu praktycznego nad poddającym się prawom przyrody rozumem teoretycznym, o tyle dopiero Fichte uczynił wolność prawdziwym centrum filozofii transcendentalnej. Według niego „rozum sam w sobie jest po prostu praktyczny, zaś teoretycznym staje się dopiero w zastosowaniu swych praw do jakiegoś Nie-Ja, które go ogranicza” (Fichte, 1996, s. 129). Prymat praktycznej władzy rozumu sankcjonuje pierwsza zasada naczelna Teorii Wiedzy, na mocy której „Ja samo siebie ustanawia”, więc „jest ono jednocześnie tym, co działa i wytworem działania; tym, co aktywne, oraz tym, co przez tę aktywność wytwarzane” (Fichte, 1996, s. 89). Skoro zatem Ja jako ów pierwotny czyn-akt ludzkiego ducha „jest bezwzględnie nieuwarunkowane i nie podlega określaniu przez nic wyższego” (Fichte, 1996, s. 120), to okazuje się ono aprioryczną zdolnością wolnego działania, ustanawiającą nie tylko siebie samą, lecz całość podmiotowo-przedmiotowej rzeczywistości. Wolność, stanowiąca u Kanta warunek możliwości wyboru moralnego, u Fichtego przekształca się w zasadę metafizyczną, powołującą do istnienia podmiot (Ja) i przedmiot (Nie-Ja) w ich wzajemnej relacji. W jej obrębie źródłową aktywność absolutnej jaźni określa on mianem dążenia, które „jest w nieskończoność wybiegającym warunkiem możliwości wszelkiego przedmiotu”, zaznaczając jednak przy tym, że „w samym pojęciu dążenia tkwi już skończoność,

gdyż to, co nie napotyka na opór, nie jest dążeniem” (Fichte, 1996, s. 287, 297). W centrum Teorii Wiedzy Fichte umieszcza zatem absolutnie wolną i nieskończoną aktywność transcendentalnego Ja, które „najpierw stawia świat zewnętrzny jako przeszkodę dla własnej aktywności, a następnie stopniowo, lecz bez końca, pokonuje tę przeszkodę” (Lovejoy, 1955b, s. 211), aby wciąż powracać do siebie samej. W konsekwencji zatem, jak słusznie zwraca uwagę Friedrich Jacobi w liście skierowanym do autora *Teorii Wiedzy*, „ponieważ zdolność ludzkiego umysłu do filozoficznego pojmowania nie wykracza poza własne akty tworzenia, chcąc przeniknąć do sfery istoty rzeczy (*Reich der Wesen*) i opanować ją za pomocą myśli, musi on zostać stwórcą – świata i samego siebie” (Jacobi, 2011, s. 55). „Stwórcza” aktywność ludzkiego ducha, któremu Fichte nadaje boskie atrybuty absolutnej wolności i nieskończoności, u Schlegla staje się probierzem progresywnej i uniwersalnej poezji romantycznej.

Kategorie wolności i nieskończoności nie pojawiają się w krytyczno-literackiej refleksji Schlegla wraz z wyłożonym w „Athenaeum” pojęciem *romantische Poesie*, lecz wyznaczają horyzont definiowania poezji nowoczesnej już w eseju *Über das Studium der griechischen Poesie*. Swoją rozprawę Schlegel rozpoczyna od stwierdzenia, że w przeciwieństwie do sztuki starożytnej „poezja nowoczesna albo nie osiągnęła jeszcze celu, do którego dąży, albo jej dążenia nie mają żadnego określonego celu” (Schlegel, 1979, s. 217), aby następnie kluczową dla eseju opozycję naturalnej kultury antyku i sztucznej kultury nowoczesnej zbudować na przeciwstawieniu instynktu i wolności. Pierwsza z nich poddaje się konieczności praw natury, natomiast o powstaniu drugiej decyduje „wolny akt umysłu”, stanowiący „pierwotne źródło, pierwszy determinujący impuls sztucznej kultury, który słusznie przypisuje się wolności” (Schlegel, 1979, s. 231–232). Co istotne, w niemal identycznym kluczu dychotomię antyku i nowoczesności odczytuje Schiller, rozróżniając poezję naiwną i sentymentalną, co w pewnej mierze wyjaśnia „zdumiewające powinowactwo powstałych jednocześnie, a przecież niezależnie od siebie pism” (Jauss, 1999, s. 70) obu niemieckich autorów. Zdaniem Lovejoya lektura eseju o poezji naiwnej i sentymentalnej nie tyle dostarczyła Schleglowi nowych argumentów dla osadzenia wizji nowoczesnej sztuki w optyce filozofii transcendentalnej, ile dodała mu odwagi do wyciągnięcia wniosku o „wyższej pozycji sztuki nowoczesnej, tzn. »postępowej« i »subiektywnej« nad statyczną i czysto »obiektywną« sztuką klasycznego antyku” (Lovejoy, 1955b, s. 220).

Mimo istotnych zbieżności między koncepcją poezji interesującej, którą Schlegel po 1798 roku nazywa romantyczną, a teorią sztuki sentymentalnej, różnią się one zasadniczo. Według Jausa idei sentymentalności „brak do pełnej treści pojęcia »romantyzm« jedynie »tego, co interesujące« – zasady »subiektywnej siły estetycznej«, którą Schlegel objaśnia za pomocą określeń takich jak »charakterystyczne«, »indywidualne«, »oryginalne«” (Jaus, 1999, s. 85), co jednak znacząco rzutuje na sposób definiowania wolności i nieskończoności jako fundamentalnych wyznaczników sztuki nowoczesnej w pismach obu myślicieli. Schlegel wyprowadza je bowiem z podstawowej kategorii swojej antropologii, jaką jest indywidualność, natomiast Schiller sytuuje w ścisłym odniesieniu do ideału moralnego.

Podczas gdy w eseju Schillera poeta sentymentalny upatruje swego zadania „we wzniesieniu rzeczywistości do ideału, lub, co jest tym samym, w przedstawieniu ideału” (Schiller, 1972a, s. 335), co umożliwia przysługująca mu „bezwarunkowa wolność zdolności do idealizacji” (Schiller, 1972a, s. 382), *romantische Poesie* „za swą pierwszą zasadę uznaje to, że samowola poety nie toleruje żadnego prawa nad sobą” (Schlegel, 2009, s. 62), lecz sama staje się jego źródłem. W odróżnieniu zatem od wolności poety sentymentalnego, która poddaje się prawom ideału i okazuje się bezwarunkowa jedynie w odniesieniu do jego przedstawiania, romantyczna samowola, wzorem niekwestionowanej autonomii transcendentalnego Ja, „wychodzi od wolności i od wiary w nią, a potem pokazuje, jak duch ludzki odciska na wszystkim swoje prawo i świat jest jego dziełem sztuki” (Schlegel, 2009, s. 73). Poetę sentymentalnego możemy więc nazwać twórcą, przekształcającym rzeczywistość zgodnie z wymogami ideału, natomiast u Schlegla „autor to stwórca” (Schlegel, 2009, s. 20), podporządkowujący świat prawom własnej indywidualności, w których przejawia się „podmiotowa strona procesu uczucia i fantazji ujęta w karby pewnej duchowej totalności i tym samym wyrażona w swej prawdzie” (Echtermeyer, Ruge, 1969, s. 149). Poezja romantyczna detronizuje zatem Schillerowski ideał, a na jego miejsce wprowadza „»ja« w jego wolności, które nie podporządkowuje się żadnym uwarunkowaniom, w niczym z otoczenia się nie rozpoznaje i w swoim żywiole, swojej eteryczności nie istnieje inaczej, niż w obrębie owej całości, w której czuje się wolny” (Blanchot, 2013, s. 300). Odmienne pojmowanie wolności przez Schillera i Schlegla różnicuje ich stanowiska również w odniesieniu do problemu nieskończoności. Ideałem, jaki stawia przed sobą poeta sentymentalny, jest doskonałość, wyrażająca się

w „zgodności pomiędzy doznawaniem i myśleniem”, którą człowiek utracił w wyniku postępu preradzającego naturalną kulturę starożytnych w sztuczną formację nowoczesności i odtąd może „manifestować się już tylko jako jedność moralna, tzn. jako istota dążąca do jedności” (Schiller, 1972a, s. 335). Ponieważ jedność moralna okazuje się u Schillera treścią ideału, dążenie do niej z definicji musi być nieskończone, więc „człowiek o wysokiej kulturze nigdy nie może stać się doskonały jak człowiek naturalny, który w swojej specyfice mógł osiągnąć doskonałość” (Schiller, 1972a, s. 336). Realna doskonałość człowieka naiwnego „w kulturze istnieje tylko idealnie – jako wartość, jako myśl, jako cel dążenia moralnego” (Siemek, 1970, s. 147), przechodząc tym samym w nowoczesną zdolność nieskończonego doskonalenia się. Z kolei Schlegel, przeciwstawiając się dyktatowi ideału, odrzuca również koncepcję moralnego doskonalenia się, zastępując ją ideą uniwersalności życia, w której przejawia się nieskończone dążenie *homo romanticus*. Jeśli bowiem człowiek „jest rozumny, rozum zaś jest wolny i sam w sobie nie jest niczym innym niż wiecznym samookreśleniem się w nieskończoność” (Schlegel, 2009, s. 167–168), to najwyższa forma człowieczeństwa przejawia się w dążeniu do coraz głębszego samopoznania, dzięki któremu właśnie „życie uniwersalnego ducha stanowi nieprzerwany łańcuch przewrotów wewnętrznych” (Schlegel, 2009, s. 147), znamionujących kolejne perspektywy świadomego doświadczania własnej egzystencji, w jakich romantyczny poeta nieustannie rozpoznaje i kształtuje swoją indywidualność. U Schlegla zatem dążenie uniwersalnego ducha „nie zmierzało bynajmniej do »nieskończoności« ideału moralnej doskonałości [...]; zmierzało raczej do nieskończoności rzeczywistego życia – dobrego i złego – jako przedmiotu sztuki poetyckiej” (Lovejoy, 1955b, s. 223).

Polemika z Schillerem odsłania tym samym paradoks wczesnoromantycznej antropologii, ujmującej człowieka jako „coś skończonego, co kształtuje się w nieskończoność” (Schlegel, 2009, s. 163), a odnajdującej swój wyraz w – sytuującej się na pograniczu sztuki poetyckiej i filozofii – kategorii ironii, która nie jest niczym innym, jak właśnie „formą paradoksu” (Schlegel, 2009, s. 17). W eseju Schillera „antyteza »naiwnego« i »sentymentalnego« jest przede wszystkim historiozoficzną i filozoficzną opozycją [...] dwu całościowych modeli człowieka i jego stosunku do świata” (Siemek, 1970, s. 146), ale ostatecznie żaden z nich „sam nie może wyczerpać ideału pięknego człowieczeństwa, który może się wyłonić jedynie z najgłębszego połączenia obu” (Schiller, 1972a, s. 402–403). Warunkiem urzeczywistnienia

ideału pięknego człowieczeństwa jest transcendentalna synteza realnych i idealnych pierwiastków ludzkiego bytu, zapewniająca „jedność realności z formą, przypadkowości z koniecznością i bierności z wolnością”, którą gwarantuje popęd gry, ustanawiający harmonię między popędami formalnym i materialnym (Schiller, 1972b, s. 100). W teorii Schlegla rolę popędu gry przejmuje ironia jako „transcendentalna bufonia” (Schlegel, 2009, s. 14), oznaczająca „grę na temat idei i rzeczywistości” (Szturc, 1992, s. 190), która nie tylko stanowi „szeroką zasadę organizującą ironiczny świat oraz ironię ten świat przenikającą” (Szturc, 1992, s. 187), ale zarazem odzwierciedla paradoksalną kondycję *homo romanticus*. Transcendentальной bufonii nie można jednak w prosty sposób utożsamić z Schillerowskim popędem gry, ponieważ w swojej koncepcji ironii Schlegel zastępuje pojęcie harmonii, definiujące ideał pięknego człowieczeństwa, kategorią „nierozwiązywalnego konfliktu między bezwarunkowym i uwarunkowanym” (Schlegel, 2009, s. 27). Istotą ironii jest dialektyczny proces łączenia i antagonizowania idealnego i realnego, umożliwiający „absolutną syntezę absolutnych antytez, stale tworzącą się wymianę dwóch niezgodnych myśli” (Schlegel, 2009, s. 64), odpowiadającą Fichteńskiej „wymianie Ja w sobie i ze sobą samym”, która polega na „konflikcie Ja ze sobą samym i dzięki temu wciąż się odtwarza, o ile Ja chce zjednoczyć rzeczy nie do pogodzenia i już to próbuje wchłonąć nieskończoność w formę czegoś skończonego, już to, odepchnięte, umieszcza ją znów poza tamtą formą” (Fichte, 1996, s. 230). Transcendentálna bufonia, odtwarzająca w nieskończoność grę idei i rzeczywistości, stanowi tym samym „praktyczną, estetyczną realizację podstawowej kategorii filozofii Fichtego, ale w jej indywidualnym wcieleniu” (Passi, 1980, s. 278), lecz jednocześnie postawa podmiotu ironicznego odsłania dynamiczną i paradoksalną naturę *homo romanticus*, w której „twórcza wolność przeciwstawia się przyrodniczej konieczności – nieskończoność ducha indywidualnego zderza się ze skończonością natury, nieograniczoność rozumu z ograniczonością świata przedmiotowego, wzlot fantazji z surowością faktów” (Passi, 1980, s. 278). W rezultacie polemiki Schlegla z antropologią Schillera ideał *pięknego człowieczeństwa* zostaje zatem zastąpiony romantyczną wizją *człowieczeństwa ironicznego*, przeciwstawiającą klasycznej harmonii dialektyczny konflikt, który niemiecki autor próbuje zarazem zinterpretować w duchu organicznej koncepcji życia.

W stronę filozofii organicznej indywidualności

Paradoks romantycznej wizji człowieka wynika z próby połączenia w pojęciu indywidualności prymatu wolnej i samoświadomej jaźni (idealizm) oraz pierwszeństwa żywej i twórczej natury (realizm). Choć Schlegel w wielu punktach swojej teorii niewątpliwie czerpie z dziedzictwa myśli Fichtego, to zarazem poddaje krytyce kwestię statusu rzeczywistości przedmiotowej w filozofii transcendentnego idealizmu, który „neguje Nie-Ja, a jako jedyną realną rzecz ustanawia tylko Ja” (Schlegel, 1964, s. 190). Musi ona zatem uznać podmiot za ośrodek życia i aktywności, a przedmiotowi przypisać absolutną bierność, w której wyraża się „brak ruchu, odpoczynek, zastój i dlatego jest dokładnie przeciwstawną koncepcji Ja oraz stawania się” (Schlegel, 1964, s. 190). Wobec tego cała Teoria Wiedzy opiera się na nieprzekraczalnej opozycji tożsamej z życiem świadomości człowieka oraz „martwej” natury, do której przynależy nie tylko cały świat przyrody, ale również sfera ludzkiej cielesności i zmysłowości. Tymczasem romantyczna indywidualność ma być „niepodzielną jednością, żywą wewnętrzną współzależnością” (Schlegel, 2000, s. 158), w obrębie której duch i natura są różnymi, lecz równoprawnymi przejawami tego samego nieskończonego procesu życia. Krytyka Fichtego ujawnia trzy kluczowe cechy antropologii Schlegla. Po pierwsze, jest ona „wyrazem buntu przeciwko temu sposobowi myślenia, który ujmował świat w kategoriach statycznego mechanizmu” (Peckham, 1978, s. 241). Po drugie, hołduje witalizmowi, polegającemu na „przedkładaniu tego, co żyje, porusza się i ewoluuje dzięki wewnętrznej energii, nad to, co pozbawione życia, bezwładne i niezmiennie” (Abrams, 1971, s. 431). I wreszcie po trzecie, zmierza do ujęcia natury jako organicznego procesu, w którym „w niekończącym się następstwie nowych form, czas twórczy tka wieniec wieczności” (Schlegel, 1962, s. 61). Zgodnie z takim ujęciem cały świat przyrody – od najprostszych po najbardziej wyrafinowane formy – jest rezultatem działania jednej i tej samej siły życiowej, osiągającej kolejne stopnie rozwoju aż do samoświadomości, która stanowi jej punkt szczytowy: „świadomość natury u artysty czy filozofa jest zarazem naturą dochodzącą do samoświadomości za ich pośrednictwem” (Beiser, 2006, s. 223). Dopiero organiczna wykładnia przyrody, w obrębie której duch i natura odnajdują swój wspólny mianownik, pozwala Schleglowi podjąć próbę odpowiedzi na „pytanie dla romantyki najbardziej zasadnicze: jak pogodzić skończoność indywidualności z tendencją ku nieskończoności” (Łempicki, 1966, s. 161).

Naturę, powołującą do istnienia wciąż nowe formy życia, Schlegel przyrównuje do „dzieła sztuki, które wiecznie tworzy samo siebie” (Schlegel, 2000, s. 160), a rozumny, samoświadomy i wolny *homo romanticus*, będąc najwyższym osiągnięciem w świecie istot organicznych, powinien w analogiczny sposób nieustannie rozwijać i doskonalić swoje zdolności, talenty i predyspozycje, zmierzając do urzeczywistnienia w warunkach kultury nowoczesnej naturalnej harmonii, cechującej jego źródłowo niepodzielną indywidualność. Unikalna próba syntezy ducha i natury odsłania tym samym istotę romantycznego organicyzmu, realizującego się w „metaforycznym przełożeniu na kategorie i normy intelektu atrybutów rosnącej rzeczy, która rozwija swoją wewnętrzną formę i asymiluje do siebie obce elementy, aż osiągnie pełnię swojej złożonej, organicznej jedności” (Abrams, 1971, s. 432). Pierwotne atrybuty życia, do jakich zaliczają się wzrost i rozwój, na płaszczyźnie kultury odnajdują swój odpowiednik w idei formacji (niem. *Bildung*). U Schlegla urasta ona do rangi najwyższego wyznacznika prawdziwie ludzkiej egzystencji, albowiem „kształcenie to najwyższe dobro i jedyna pożyteczna rzecz” (Schlegel, 2009, s. 155)⁴, a także najważniejszego zadania postawionego przed jednostką, dzięki któremu dopiero może ujawnić się pełnia jej „stwórczych” mocy: „stać się Bogiem, być człowiekiem, kształcić się – to wyrażenia, które znaczą to samo” (Schlegel, 2009, s. 95)⁵. Należy zauważyć, że w polskim przekładzie fragmentów Schlegla termin *Bildung* jest tłumaczony zazwyczaj jako „kształcenie” lub „wykształcenie”, nasuwając skojarzenie ze sferą zdobywania wiedzy, jednak w języku niemieckim pojęcie to ma dużo szerszy zakres semantyczny i „w zależności od kontekstu może ono oznaczać edukację, kulturę lub rozwój” (Beiser, 2003, s. 26) czy też „formowanie charakteru, indywidualnej osobowości” (Engel, 2008, s. 266). Trzeba przy tym zaznaczyć, że w historii kultury niemieckiej obserwujemy ewolucję znaczenia słowa *Bildung*. U progu XIX wieku „przelewa ono sens realizowanego od zewnątrz wychowania, właściwy jeszcze temu pojęciu w XVIII stuleciu, w roszczenie autonomiczności, które zasadza się na modyfikującym wintegrowaniu świata w obręb własnej jaźni” (Koselleck, 2012, s. 417) i które zarazem wyróżnia wczesno-romantyczne pojmowanie wolności jako głównej determinanty „stwórczej”

⁴ „Das höchste Gut und das allein Nützliche ist die Bildung” (Schlegel, 1967, s. 259).

⁵ „Gott werden, Mensch sein, sich bilden, sind Ausdrücke, die einerlei bedeuten” (Schlegel, 1967, s. 210).

mocy człowieka. W antropologii Schlegla idea *Bildung* wpisuje się zarazem w kontekst „egotyzmu apersonalnego”, traktującego indywidualność jako powszechną właściwość ludzkiej natury: „oddać się kształceniu i rozwojowi tej indywidualności jako najwyższemu powołaniu – oto boski egoizm” (Schlegel, 2009, s. 159)⁶. W związku z tym romantyczna *Bildung* ma dwa cele. Po pierwsze, „jednoczy i rozwija wszystkie siły istoty ludzkiej, łącząc wszystkie jej rozbieżne zdolności w całość”, a po drugie, doskonalą „nie tylko nasze charakterystyczne ludzkie siły – wspólne dla każdego człowieka – ale także naszą indywidualność – te wyjątkowe zdolności i dyspozycje właściwe każdej jednostce”, przy czym są one ściśle z sobą powiązane, ponieważ „indywidualność wyłania się z tej wyjątkowej syntezy, tej szczególnej jedności wszystkich ludzkich mocy” (Beiser, 2003, s. 93). Współzależność obu celów romantycznej *Bildung* Schlegel wydobywa w *Lucinde*, a szczególnie w jednej z najbardziej kontrowersyjnych scen powieści, której przedmiotem jest akt miłosnego zbliżenia kochanków. Z jednej strony symbolizuje on bowiem moment dopełnienia procesu formacyjnego Juliusza, w którym osiąga on duchowo-cieleśną jedność i harmonię wszystkich swoich władz, a więc urzeczywistnia ideał niepodzielnej indywidualności. Z drugiej jednak to intymne doświadczenie jest „alegorią doskonalenia mężczyzny i kobiety do pełni człowieczeństwa” (Schlegel, 1962, s. 13), reprezentując uniwersalny cel ludzkich dążeń, jaki stanowi synteza – utożsamianych przez pierwiastki męski i kobiecy – powszechnych cech natury człowieka: aktywności i bierności, dystansu refleksji i bezpośredniości uczuć, koniecznej skończoności i nieskończonego wybiegania w sferę absolutnej wolności.

Na płaszczyźnie krytyczno-literackiej refleksji Schlegla antropologiczny model organicznej indywidualności w najwyższym stopniu odzwierciedla teoria powieści. Wzorem średniowiecznych romansów rycerskich „zewnątrzna forma powieści jest w swej istocie biograficzna” (Lukács, 1968, s. 69). Nie inaczej uważa Schlegel, według którego „niejedną z najwybitniejszych powieści to kompendium, encyklopedia całego życia duchowego genialnej jednostki” (Schlegel, 2009, s. 22), osiągająca najwyższą wartość wtedy, kiedy odnosi się wprost do ludzkiej indywidualności. Jeśli bowiem „biografia dąży do uogólnienia, jest fragmentem historycznym”, to dopiero „gdy skupia się na charakterystyce indywidualności, jest dokumentem bądź

⁶ „Die Bildung und Entwicklung dieser Individualität als höchsten Beruf zu treiben, wäre ein göttlicher Egoismus” (Schlegel, 1967, s. 262).

dziełem [z zakresu] teorii sztuki życia” (Schlegel, 2009, s. 86). W pełni autentyczne świadectwo unikalności ludzkiej egzystencji może jednak dać tylko jednostka, która bezpośrednio jej doświadcza, w związku z czym „fundamentem wszelkiej romantycznej literatury jest prawdziwa historia”, a „najlepsze rzeczy w najlepszych powieściach to nic innego, jak tylko mniej czy bardziej zamaskowane wyznanie autora, plon jego doświadczeń, kwintesencja jego swoistości” (Schlegel, 2000, s. 175). Forma biograficzna okazuje się tym samym niewystarczająca i w powieści wczesnoromantycznej radykalizuje się, przechodząc w formę autobiograficzną, czemu Schlegel daje wyraz w *Lucinde*, skrywającej za fasadą literackiej fikcji jego głęboko osobiste doświadczenia. Spośród wielu odniesień do prywatnego życia młodego Friedricha zdecydowanie na pierwszy plan wysuwa się w niej wątek jego romansu z kilka lat starszą i zamężną Dorotheą Weit – pierwowzorem tytułowej bohaterki – który do tego stopnia nie pozostawiał jej pierwszym czytelnikom wątpliwości odnośnie do genezy postaci i zdarzeń ukazanych w utworze, że stał się przyczyną obyczajowego skandalu. Niemniej jednak ostatecznym celem „filozoficznie zabarwionej rozpusty przedstawionej w powieści” (Bernofsky, 2005, s. 90) nie jest „pewien rodzaj stonowanego psychologicznego ekshibicjonizmu” (Łempicki, 1966, s. 412), lecz w rzeczywistości osobiste doświadczenia autora są jedynie pretekstem do przedstawienia organicznego modelu romantycznej *Bildung*. Rola, jaką w *Lucinde* odgrywa wątek autobiograficzny, zbliża powieść Schlegla do *Lat nauki Wilhelma Meistra*, zwiastujących w literaturze niemieckiej narodziny gatunku *Bildungsroman* (tj. powieści formacyjnej), który uważany jest za „nierozzerwalnie związany z ideą nowoczesnej indywidualności” (Engel, 2008, s. 266). Schlegel, odwołując się do historii powstania *Meistra* w kontekście osobistych losów Goethego, akcentuje właśnie ów organiczny rys dzieła, rozwijającego się wraz ze swoim autorem, aby stopniowo „ideę powieści o artyście” mogła zastąpić „edukacja kształcąca sztukę życia, stając się opiekuńczym duchem całości” (Schlegel, 2000, s. 185)⁷. I choć organiczny rys *Wilhelma Meistra*

⁷ Pierwszej wersji dzieła Goethego, zatytułowanej *Wilhelm Meisters theatralische Sendung* i napisanej w latach 1777–1785, przyświeca idea wychowania jednostki i społeczeństwa przez sztukę, która jest odbiciem przypadającej na ten czas działalności Goethego w amatorskim teatrze w Weimarze (Bruford, 2009, s. 31). Do *Wilhelma Meistra* Goethe powraca w roku 1791, aby w pięć lat później ukończyć znaną nam dzisiaj powieść, znacznie odbiegającą w swoim przesłaniu od pierwotnego zamysłu autora, który w wyniku rozczarowania swoimi młodzieńczymi marzeniami traktuje teatralne

zdaje się bliski romantycznej antropologii, to jednocześnie „w *Latach nauki* coraz wyraźniej idea samokształcenia jako indywidualnej przygody ustępuje wychowaniu” (Mann, 1964, s. 206), nad którym czuwa Towarzystwo z Wieży, ograniczające autonomię Wilhelma. Tymczasem według Schlegla „możemy się ograniczać jedynie w tych punktach i w tych względach, gdzie mamy nieskończoną siłę – siłę autokreacji i autodestrukcji”, a „wszędzie tam, gdzie sami się nie ograniczamy, ogranicza nas świat” (Schlegel, 2009, s. 11), co staje się udziałem bohatera powieści Goethego. W tym sensie „rezygnacja młodzieńca z samodzielnej autokreacji, jego uległość wobec samozwańczych opiekunów, nie odpowiadają romantycznej antropologii” (Janion, Żmigrodzka, 1998, s. 42), która zdecydowanie wykracza poza horyzont Goetheańskiej wizji człowieczeństwa.

Zakończenie

W zwierciadle krytyczno-literackiej teorii Schlegla *homo romanticus* daje się rozpoznać jako istota przede wszystkim paradoksalna, która z jednej strony stale doświadcza nieprzekraczalnej dwubiegunowości życia, z drugiej nieustannie dąży do wyższej syntezy skonfliktowanych wymiarów własnej egzystencji: chce wznieść się na szczyty samoświadomości, a zarazem pozostać organiczną jednością; godzi się na swoją skończoność, lecz jednocześnie wciąż pragnie wybiegać ku nieokreślonej nieskończoności. Kiedy przyglądamy się wszystkim tym paradoksom romantycznej antropologii z perspektywy ideowego tła jej narodzin, współtworzonego głównie przez trzy wielkie postaci niemieckiej kultury u schyłku XIX wieku – Fichtego, Schillera i Goethego – możemy odnieść wrażenie, że Schlegel buduje swoją wizję człowieka z dość dowolnie interpretowanych elementów pochodzących z teorii jego mentorów, które dodatkowo potęguje jeszcze upodobanie młodego autora do aforyzmów i fragmentów. Jeśli jednak postaramy się wniknąć wystarczająco głęboko w tę oryginalną mieszaninę pozornie spletanych myśli i idei, to dostrzegamy, że z dążenia Schlegla do przewyższenia ograniczeń myśli jego młodzieńczych autorytetów wyłania się

zainteresowania bohatera – a tym samym również własne – z dużą dozą ironii, uznając je jedynie za epizod w jego życiu i zastępując ideę wychowania przez sztukę ideałem edukacji w świecie i wśród ludzi (Janion, Żmigrodzka, 1998, s. 36).

nowy model myślenia o człowieku i jego stosunku do świata. W analogii do romantycznego sposobu definiowania powieści można byłoby określić go mianem antropologicznej arabeski, a ta, podobnie jak jej literacki odpowiednik, nie stanowi chaosu dowolnie połączonych elementów, lecz jest wyrazem zasady nowoczesnej samoświadomości, która właśnie w paradokсах i sprzecznościach ludzkiej natury dostrzega istotę człowieczeństwa.

Bibliografia

- Abrams, M.H. (1971). *Natural Supernaturalism: Tradition and Revolution in Romantic Literature*. New York: W. W. Norton & Company.
- Behler, E. (1993). *German Romantic Literary Theory*. Cambridge: Cambridge University Press.
- Beiser, F.C. (2003). *The Romantic Imperative. The Concept of German Early Romanticism*. Cambridge, MA-London: Harvard University Press.
- Beiser, F.C. (2006). The Paradox of Romantic Metaphysics. W: N. Kompridis (red.), *Philosophical Romanticism* (s. 217–237). London: Routledge.
- Benjamin, W. (1974). Der Begriff der Kunstkritik in der deutschen Romantik. W: W. Benjamin, *Gesammelte Schriften*, t. 1 (s. 7–122). Frankfurt am Main: Suhrkamp Verlag.
- Bernofsky, S. (2005). The Infinite Imagination: Early Romanticism in Germany. W: M. Ferber (red.), *A Companion to European Romanticism* (s. 86–100). Oxford: Blackwell Publishing.
- Blanchot, M. (2013). Athenaeum. Tłum. W. Kruszelnicki. W: M. i W. Kruszelniccy (red.), *Nietzsche i romantyzm* (s. 295–305). Wrocław: Wydawnictwo Naukowe Dolnośląskiej Szkoły Wyższej.
- Bruford, W.H. (2009). *The German Tradition of Self-Cultivation*. New York: Cambridge University Press.
- Dilthey, W. (1982). Novalis. W: W. Dilthey, *Pisma estetyczne*. Tłum. K. Krzemieniowa (s. 324–404). Warszawa: PWN.
- Echtermayer, T., Ruge, A. (1969). Protestantyzm i romantyzm. Tłum. R. Panasiuk. W: R. Panasiuk, *Lewica heglowska* (s. 147–174). Warszawa: Wiedza Powszechna.
- Emrich, W. (1978). Romantyzm a świadomość nowoczesna. Tłum. M. Łukasiewicz. *Pamiętnik Literacki*, LXIX (3), 287–305.

- Engel, M. (2008). Variants of the Romantic “Bildungsroman” (with a short note on the “artist novel”). W: G. Gillespie, M. Engel, B. Dieterle (red.), *Romantic Prose Fiction*. Amsterdam: John Benjamins Publishing Company.
- Fichte, J.G. (1996). *Teoria Wiedzy. Wybór pism*, t. 1. Tłum. M.J. Siemek. Warszawa: PWN.
- Gasché, R. (1997). *The Tain of the Mirror: Derrida and the Philosophy of Reflection*. Cambridge: Harvard University Press.
- Gerard, A. (1978). O logice romantyzmu. Tłum. M. Orkan-Lęcki. *Pamiętnik Literacki*, *LXIX* (1), 253–263.
- Haym, R. (1870). *Die romantische Schule: ein Beitrag zur Geschichte des deutschen Geistes*. Berlin: Verlag von Rudolph Gaertner.
- Helfer, M.B. (1996). *The Retreat of Representation. The Concept of Darstellung in German Critical Discourse*. New York: State University of New York Press.
- Jacobi, F.H. (2011). Do Fichtego. Tłum. M. Rutkowska. *Kronos*, *1*, 52–61.
- Janion, M., Żmigrodzka, M. (1998). *Odyseja wychowania. Goetheańska wizja człowieka w Latach nauki i latach wędrówki Wilhelma Meistra*. Kraków: Aureus.
- Jauss, H.R. (1999). *Historia literatury jako prowokacja*. Tłum. M. Łukasiewicz. Warszawa: Instytut Badań Literackich.
- Koselleck, R. (2012). *Semantyka historyczna*. Tłum. W. Kunicki. Poznań: Wydawnictwo Poznańskie.
- Lovejoy, A.O. (1955a). On the Discrimination of Romanticism. W: A.O. Lovejoy, *Essays in the History of Ideas* (s. 228–253). New York: George Braziller.
- Lovejoy, A.O. (1955b). Schiller and the Genesis of German Romanticism. W: A.O. Lovejoy, *Essays in the History of Ideas* (s. 207–227). New York: George Braziller.
- Lovejoy, A.O. (1955c). The Meaning of “Romantic” in Early German Romanticism. W: A.O. Lovejoy, *Essays in the History of Ideas* (s. 183–206). New York: George Braziller.
- Lukács, G. (1968). *Teoria powieści. Esej historyczno-filozoficzny o wielkich formach epiki*. Tłum. J. Gośliński. Warszawa: Państwowy Instytut Wydawniczy.
- Lempicki, Z. (1966). *Renesans, oświecenie, romantyzm i inne studia z historii kultury*. Warszawa: PWN.
- Mann, T. (1964). *Goethe i Tolstoj. Przyczynki do problemu humanizmu*. Tłum. M. Kłós-Gwizdalska. W: T. Mann, *Eseje* (s. 107–230). Warszawa: PIW.

- Passi, I. (1980). *Powaga śmieszności*. Tłum. K. Minczewska-Gospodarek. Warszawa: PWN.
- Paulin, R. (2005). German Romantic Fiction. W: M. Ferber (red.), *A Companion to European Romanticism* (s. 123–137). Oxford: Blackwell Publishing.
- Peckham, M. (1978). O koncepcję romantyzmu. Tłum. M. Orkan-Łęcki. *Pamiętnik Literacki*, LXIX (1), 231–252.
- Schiller, J.C.F. (1972a). O poezji naiwnej i sentymentalnej. W: *Listy o estetycznym wychowaniu człowieka i inne rozprawy* (s. 303–418). Tłum. I. Krońska, J. Prokopiuk. Warszawa: Czytelnik.
- Schiller, J.C.F. (1972b). Listy o estetycznym wychowaniu człowieka. W: *Listy o estetycznym wychowaniu człowieka i inne rozprawy* (s. 39–170). Tłum. I. Krońska, J. Prokopiuk. Warszawa: Czytelnik.
- Schiller, J.C.F. (1974). Schiller do Goethego (Jena, 20 października 1797). W: *Z korespondencji między Goethem i Schillerem*. Tłum. J. Prokopiuk, M.J. Siemek. Warszawa: Czytelnik.
- Schlegel, K.W.F. (1962). Lucinde. Ein Roman. W: *Kritische Friedrich Schlegel Ausgabe*, t. 5 (s. 1–92). München-Paderborn-Wien-Zürich: Verlag Ferdinand Schöningh.
- Schlegel, K.W.F. (1964). Die Entwicklung der Philosophie in zwölf Büchern. W: *Kritische Friedrich Schlegel Ausgabe*, t. 12 (s. 107–323). München-Paderborn-Wien-Zürich: Verlag Ferdinand Schöningh.
- Schlegel, K.W.F. (1967). Fragmente. W: *Kritische Friedrich Schlegel Ausgabe. Dichtungen*, t. 2 (s.147–273). München-Paderborn-Wien-Zürich: Verlag Ferdinand Schöningh.
- Schlegel, K.W.F. (1979). Über das Studium der griechischen Poesie. W: *Kritische Friedrich Schlegel Ausgabe*, t. 1 (s. 205–367). München-Paderborn-Wien-Zürich: Verlag Ferdinand Schöningh.
- Schlegel, K.W.F. (2000). Rozmowa o poezji. Tłum. J. Ekier. W: T. Namowicz (red.), *Pisma teoretyczne niemieckich romantyków* (s. 119–191). Wrocław: Zakład Narodowy im. Ossolińskich.
- Schlegel, K.W.F. (2009). *Fragmenty*. Tłum. C. Bartl. Kraków: Wydawnictwo Uniwersytetu Jagiellońskiego.
- Siemek, M.J. (1970). *Fryderyk Schiller*. Warszawa: Wiedza Powszechna.
- Siemek, M.J. (1977). *Idea transcendentalizmu u Fichtego i Kanta. Studium z dziejów filozoficznej problematyki wiedzy*. Warszawa: PWN.

Szturc, W. (1992). *Ironia romantyczna*. Warszawa: PWN.

Wellek, R. (1949). The Concept of “Romanticism” in Literary History. *Comparative Literature*, 1 (1), 1–23.

Żabski, T. (1990). Romans rycerski XIX wieku. *Zagadnienia rodzajów literackich*, XXXIII (2), 5–25.

Nota o autorce

Malwina Rolka – doktor nauk humanistycznych w zakresie filozofii, pracuje na Uniwersytecie Śląskim w Katowicach (Instytut Filozofii i Wydział Sztuki i Nauk o Edukacji). Autorka monografii *Hermeneutyka mitu dionizyjskiego w filozofii Fryderyka Nietzschego* (2014) oraz artykułów z zakresu historii filozofii nowoczesnej, historii idei i filozofii edukacji.

Address for correspondence: University of Silesia in Katowice, Faculty of Arts and Educational Science, ul. Bielska 62, 43-400 Cieszyn.

Cytowanie

Rolka, M. (2024). Filozofia *homo romanticus* w zwierciadle teorii literackiej Friedricha Schlegla. *Analiza i Egzystencja*, 67 (3), 55–76. DOI: 10.18276/aie.2024.67-03.



ADRIANA SCHETZ

Uniwersytet Szczeciński

ORCID: 0000-0002-6328-128X

email: adriana.schetz@usz.edu.pl

Geneza zachowań instynktownych

Słowa kluczowe: instynkt, zachowanie, afordancje, morfizm, Lorenz, Blumberg, Gibson

Keywords: instinct, behavior, affordances, morphism, Lorenz, Blumberg, Gibson

Genesis of Instinctive Behavior

Abstract

Although the category of instinct is included in almost every discourse in which animal behavior is analyzed, a closer look at the phenomenon that creates its specificity shows its elusiveness. One gets the impression that something called instinctive behavior is simply not significantly different from other non-instinctive behaviors. Firstly, instinct understood as innate, pre-programmed behavior partially loses its features of automaticity and rigid schematicity when its features are compared with the features of other behaviors acquired through learning. Secondly, the genesis of the instinct, which is at least partly to be found in epigenetics and the mechanisms of individual development, shows that the differences between it and other behaviors not inherited genetically are not as large as may be assumed at first glance. Third, a promising explanation for the origin of instinctive behavior is the ecological theory originally formulated by Eleanor Gibson and James J. Gibson for perception, as it allows us to understand the role of environmental factors in the development of patterns of instinctive action.

*Lorenz zamierzał zrobić dla zachowania
to, co ewolucyjne interpretacje Karola Darwina
uczyniły dla kości*
(Blumberg, 2017, s. 2)¹

Instynkt jako nieuchwytnie zjawisko

Choć kategoria instynktu wpisana jest w niemal każdy dyskurs, w którym analizowane jest zachowanie zwierząt, to jednocześnie bliższe przyjrzenie się zjawisku, które tworzy jego specyfikę, ukazuje jego nieuchwytność. Odnosi się wrażenie, że coś takiego, jak zachowanie instynktowne po prostu nie różni się jakoś znacząco od innych, nieinstynktownych zachowań:

1. Instynkt jako zachowanie wrodzone, wstępnie zaprogramowane, częściowo traci znamiona automatyzmu i sztywnej schematyczności gdy porówna się jego cechy z cechami innych zachowań, nabytych na drodze uczenia się.
2. Geneza instynktu, której należy, przynajmniej częściowo, upatrywać w epigenetyce oraz mechanizmach rozwoju osobniczego, pokazuje, że różnice między nim a innymi zachowaniami niedziedzicznymi na drodze genetycznej nie są aż tak duże, jak można na pierwszy rzut oka przypuszczać.
3. Obiecującym wyjaśnieniem pochodzenia zachowań instynktownych jest teoria ekologiczna, pierwotnie sformułowana przez Eleonorę Gibson i Jamesa J. Gibsona dla percepcji, gdyż pozwala zrozumieć rolę czynników środowiskowych w powstawaniu schematów działania instynktownego.

Tradycyjne ujęcie instynktu jako zachowania stosunkowo prostego, automatycznego, wstępnie zaprogramowanego oraz niepodatnego na modyfikacje jest już anachronizmem. Mimo wszystko hipoteza dotycząca genezy zachowań instynktownych sytuuje te aktywności na przeciwnym biegunie w stosunku do działania stanowiącego wynik doświadczenia i, w przeciwieństwie do zachowań instynktownych, zindywidualizowanego. To, co instynktowne w aktywności zwierzęcia, miałyby być gatunkowo uniwersalne, zaś to, co nieinstynktowne – indywidualne. Przyjrzyjmy się,

¹ „Lorenz aimed to do for behavior what Charles Darwin’s evolutionary insights did for bones”.

jak współczesne analizy zjawiska zachowania się przyczyniają się do zmiany perspektywy z zachowującej podział na to, co wrodzone, i to, co wyuczone, na perspektywę bardziej kompromisową. Pozwala ona dostrzec, że to, czy dane zachowanie określimy mianem instynktownego, czy nabytego w procesie uczenia się, zależy w dużej mierze od tego, w którym miejscu osi czasu je usytuujemy: na filogenetycznej czy ontogenetycznej drodze rozwojowej zwierzęcia.

Czym jest zachowanie?

W pierwszych latach XX wieku zapoczątkowano podejście do psychiki zwierząt, które zakorzeniło się do dzisiaj na tyle mocno, że często w sposób milcząco definiuje ono cały sposób rozumienia jednostkowych, indywidualnych aktywności osobniczych. I choć w międzyczasie pojawiły się nowe, konkurencyjne, dużo bardziej atrakcyjne koncepcje zwierzęcego umysłu, behawioryzm – bo o nim mowa – nadal straszy zza grobu. Badacze zwierzęcych umysłów wciąż zbyt często ufają prostym wyjaśnieniom i nieskomplikowanym metodom, gdyż zakładają, że sam przedmiot ich badania – umysł zwierzęcia – jest czymś mało wyrafinowanym. Zastanówmy się, czy rzeczywiście sposób, w jaki neobehawioryzm wyjaśnia „zachowanie”, kluczowe dla tej teorii pojęcie, jest nadal użyteczny, a przede wszystkim, czy trafnie opisuje repertuar aktywności zwierząt tworzący to, co dziś nazywane jest behawiorem.

W związku z tym, że behawioryzm opierał się na przekonaniu, że nie jest możliwa wiedza na temat zawartości psychiki umysłu – ani własnego, ani kogokolwiek innego, w tym zwierzęcia innego gatunku niż *Homo sapiens* – podstawową kategorią badawczą było dla tego nurtu zachowanie, metodą zaś jego obserwacja. Zachowanie opisywano jako sekwencję czynności, które dadzą się empirycznie zaobserwować oraz zmierzyć za pomocą przyjętych skal. I tak Iwan Pawłow bezceremonialnie zakładał przetoki ślinianek i żołądka psom, których uczył reakcji ślinienia się i wzmożonego wydzielania soku żołądkowego na dźwięk dzwonka. Miarka umieszczona na pojemniku zbierającym wydzieliny towarzyszące uczuciu głodu w postaci śliny i kwasu solnego stanowiła obiektywną reprezentację zerowyjedynekowego stanu, w którym znajduje się zwierzę: jest głodne albo nie jest głodne. Mówiąc precyzyjniej, poziom wydzielin fizjologicznych psa

wskazywał, że udało się go poddać warunkowaniu albo że warunkowanie nie nastąpiło. Zachowanie miało być efektem uczenia się odpowiednich reakcji w odpowiedzi na odpowiedni bodziec. Zmiana tego głęboko zakorzenionego – nie tylko w nauce, ale także w całej kulturze zachodniej – paradygmatu opisu aktywności zwierząt pozaludzkich wymaga zmiany sposobu rozumienia samych aktywności, których te opisy dotyczą. Prześledźmy przykład tego, jak może wyglądać takie wyjście poza ten anachroniczny sposób rozumienia zachowania.

Wiadomo, że zwierzęta pozaludzkie komunikują się na wiele sposobów. Z uwagi na możliwość przyswojenia przez np. papugę szarą zasad komunikacji międzygatunkowej, nauczono kilku z nich mowy człowieka w języku etnicznym. Znana na całym świecie papuga Alex potrafiła porozumiewać się mową angielską na poziomie czteroletniego dziecka. Jej trenerka, Irene Pepperberg, zastrzegła jednak, że nigdy nie twierdziła, iż zdolności poznawcze Alexa i czteroletniego dziecka ludzkiego są identyczne, gdyż kompetencja językowa nie pokrywa się z całkowitym repertuarem umysłowych zasobów poznawczych (Pepperberg, 1999). Przyjrzyjmy się temu przypadkowi, aby zrozumieć, jak interpretacja indywidualnego zachowania może wpływać na interpretację zachowań gatunkowych, a więc instynktownych.

Zgodnie ze schematem zachowania rozumianego jako odruchowa reakcja na bodziec nauka języka polega na ekspozycji zwierzęcia na dźwięk mowy, nagradzaniu go za prawidłową reakcję lub karaniu za nieprawidłową, co ma doprowadzić odpowiednio do wzmocnienia bądź wygaszenia reakcji. Problem w tym, że ucząc języka w ten sposób – czyli poprzez warunkowanie klasyczne i instrumentalne – nie możemy wiedzieć, czy zwierzę rozumie to, co robi. Nie wiemy, czy rzeczywiście dochodzi do aktu komunikacji językowej, czy jedynie wzmocnienia działania odruchowego lub wygaszenia go.

Różnicę w działaniu odruchowym i rozumiejącym ukazuje przykład pewnych kurcząt. Były one uczone uderzania dziobem w jedną z dwóch kart. Prezentowano im mianowicie karty jasną i ciemną (Pearce, 2008, s. 150). Za wskazanie dziobem jasnej otrzymywały nagrodę. Do pewnego momentu wszystko odbywało się wedle behawiorystycznego schematu warunkowania. W chwili, gdy wprowadzono do eksperymentu nieznaną im wcześniej trzecią kartę, jeszcze jaśniejszą niż ta, którą nauczyły się wskazywać kurczęta, moglibyśmy oczekiwać, że będą one nadal wskazywać tę, za którą otrzymywały nagrodę. Nic podobnego nie miało jednak miejsca.

Większość kurcząt wskazywała nową, nieznaną im wcześniej kartę o kolorze najjaśniejszym z trzech im zaprezentowanych. O czym świadczy ten test transpozycji? Wydaje się, że zwierzęta te nie działały odruchowo już od samego początku trwania eksperymentu. Trzeba raczej przyjąć, że jakoś skonceptualizowały one sytuację problemową, przed którą stanęły: „Co zrobić, żeby uzyskać nagrodę?”. Prawdopodobnie miało tu miejsce uczenie się różnicujące (ang. *discriminatory learnig*), polegające na generowaniu przez zwierzę innej reakcji niż wyuczona, w obecności jakoś podobnego, ale jednak różnego bodźca, tj. wcześniej nieznanego (Pearce, 2008, s. 150–169). Kurczęta przyjęły pewną strategię działania, którą wyraża reguła: „Dziobać zawsze w tę jaśniejszą”. Zatem gdy pojawiła się karta jeszcze jaśniejsza od wcześniej im znanej, zasadniczo niewiele się zmieniło. Sposób, w jaki przedstawiły sobie mentalnie sytuację z dwiema oraz z trzema kartami, jest taki sam: „Dziobać zawsze w tę jaśniejszą”. To interpretatorzy eksperymentu postawili błędną hipotezę, w myśl której reakcja kurcząt jest wywołana konkretną kartą powiązaną z nagrodą.

Wracając do mowy ludzkiej u papug szarych, Pepperberg zaproponowała model uczenia się oparty na schemacie trener–rywal, który miał imitować sposób nabywania mowy przez dzieci. Jaka jest ogólna zasada przyswajania umiejętności komunikacji językowej przez papugi szare w tym modelu nauki języka? Otóż wspomniany już Alex, a potem z powodzeniem również inne papugi tego gatunku, obcowały z osobami, które posługiwały się prostymi komunikatami, takimi jak: „Powiedz makaron”, „Podaj mi banana”, „Tu jest schowany klocek”, „Klocek jest żółty”, „Klocek jest twardy”, „Klocek jest okrągły”, „Nie denerwuj się”, „Kocham cię” i wiele innych. Alex często nie był jedynym adresatem tych komunikatów, ale musiał niejako rywalizować o uwagę, czyli nauka przebiegała w trójkącie: trener–rywal–uczeń. Miało to stwarzać namiastkę naturalnego sposobu nabywania kompetencji językowej przez dzieci. Przy takim modelu uczenia się języka trudno wskazać bodźce kluczowe. Może nie jest to niemożliwe, ale nie ma chyba nawet takiej potrzeby. O wiele ważniejsze jest, że zwierzę traktowane jest tu jako uczestnik gry językowej, czegoś co określał w ten sposób austriacki filozof Ludwig Wittgenstein. Zauważył on mianowicie, w krytyce swojego wcześniejszego sposobu rozumienia języka, przypominającego model behawiorystyczny, że kompetencja językowa nie polega na zwykłym reagowaniu na bodziec – dźwięk czy napis – a zatem symbol językowy (Wittgenstein, 1972). Według niego dużo ważniejsza jest

znajomość reguł określających dany sposób komunikowania się. Kiedy przykładowo mówię do osoby, która właśnie weszła w zabłoconych butach na świeżo umytą przeze mnie podłogę: „Wejdz jeszcze na kanapę”, to posługuję się implikaturą konwersacyjną, której treścią nie jest, że chcę, aby ta osoba zrobiła to, co bezpośrednio wskazuje wypowiedziane przeze mnie zdanie. I choć trudno oczekiwać, że zwierzęta innych gatunków niż *Homo sapiens* będą rozumiały tak złożone wypowiedzi, to pojawia się pytanie, czy poza człowiekiem zwierzęta jakichkolwiek innych gatunków potrzebują umiejętności tworzenia i rozumienia implikatur. Wydaje się, że nie².

Udany językowy przekaz informacji musi, zdaniem większości filozofów i psychologów, spełniać trzy podstawowe warunki: być rekursywny, spełniać wymóg racjonalności czy rozumności sformułowany przez Donalda Davidsona (2004) oraz warunek sformułowany przez Johna Searle'a (1983) związany z intencjonalnością. Pierwszy z nich wskazuje na konieczność znajomości zasad gramatyki (składni) oraz posiadania zdolności myślenia pojęciowego, drugi zaś na byciu świadomym podmiotem obdarzonym intencjonalnością (Polley, 2011). Zauważmy, jak bardzo złożone staje się zachowanie pretendujące do miana komunikacji językowej w paradygmacie niebehaviorystycznym. Pozostawiając szczegóły na boku, znaczy to mniej więcej tyle, że Alex rzeczywiście potrafi mówić w języku angielskim, o ile zdaje sobie sprawę z tego, co chce powiedzieć, zna sposób, w jaki powinien to wyrazić tak, by być rozumianym, oraz dysponuje odpowiednimi treściami do wyrażenia swojego komunikatu. Niekiedy dodaje się również warunek kreatywności, tzn. że użytkownik języka musi być w stanie tworzyć wyrażenia językowe (zdania), których wcześniej nie poznał.

Istnieją opracowania wskazujące na to, że Alex spełniał wszystkie te warunki. Nie tylko prawidłowo posługiwał się bardzo prostą gramatyką, a gdy robił błędy, to były one typowe dla nabywających kompetencję językową dzieci, ale również dało się zauważyć starania zwierzęcia o jak najefektywniejsze przekazanie swojego komunikatu. Poprawiał wymowę, modulował wokalizację, zamieniał słowa na znane rozmówcy. Prawidłowo wskazywał obiekty, a także ich cechy, np. rozumiał różnicę między jabłkiem a jego cechami, jakimi są kolor, kształt, smak. Dokonywał również

² W związku z tym, że dyskusja na ten temat zabrałaby sporo miejsca i oddaliła nas od głównego wątku tego artykułu, to nie będzie podjęta. Podobną problematykę omawiają np. Bar-On, 2013, Bar-On, Green, 2010.

pewnych generalizacji. Jabłka i banany łączył w kategorię owoce, zaś marchew zaliczał do warzyw. Treści, którymi dysponował, były zatem odpowiednie do tworzenia nie tylko konkretnych zdań typu: „To jabłko jest słodkie”, ale również ogólnych: „Jabłka są słodkie” – na przykład wtedy, gdy zapytano go, dlaczego chce jabłko (Pepperberg, 1999; 2021, s. 5–7).

Przykład analizy kompetencji językowej papugi szarej pokazuje dość jasno, jak złożonym zjawiskiem może być zachowanie. Charakterystyka wyjściowa, zaproponowana w duchu neobehawioryzmu, a odwołująca się do warunkowania i metody wzmacniania pozytywnego i negatywnego, nie wystarcza do jego pełnego zrozumienia. Współcześnie pojawiły się jednak dużo bardziej wyrafinowane koncepcje uczenia się, na gruncie których kategoria zachowania zyskuje zupełnie inny wymiar. Na przykład Alva Noë, przedstawiciel enaktywizmu nawiązującego do ekologicznej teorii percepcji opracowanej przez Jamesa J. Gibsona (1977), uważa, że zachowanie się zwierzęcia polega na złożonej z wielu etapów, celowej, motywowanej oraz podatnej na czynniki modyfikujące aktywności ruchowej (Block, 2005). Zauważmy, że charakterystyka ta wskazuje na dwie integralne składowe zachowania, tj. aktywność psychiczną oraz motoryczną, czyli psychoruchową. Nie ma bowiem zachowania bez obu tych czynników. Ruch, któremu nie towarzyszy odpowiednia aktywność poznawcza, jest zaledwie odruchem, niezamierzoną czynnością zmiany położenia całego ciała lub jego części (Shoemaker, 1976, s. 109–110, 121–122; Taylor, 2021).

Zachowanie się nie jest zatem prostym odruchem, tak jak zachowaniem się nie jest skurcz serca czy ruch mięśni odnoży martwego owada. Jest również ukierunkowane na pewien cel, którym może być inne zwierzę, jakiś przedmiot albo sam podmiot aktywności (np. iskanie samego siebie). Skoro celem własnego zachowania może być samo zwierzę, to zwierzę to staje się dla siebie przedmiotem referencji: tworzy ono reprezentację aktywności swojego ciała jako „zwykłego” obiektu percepcji oraz reprezentację tego, jak ta aktywność zmienia się w zależności od własnego działania zamierzonego. Przykładowo, makak, na którego czole umieszcza się marker, widząc go w lustrze zaczyna rozumieć, jak to, co widzi („zwykły” obiekt percepcji) zmienia się w zależności od jego własnej aktywności ruchowej (obraz przesuwa się w lewo, gdy ja przesuвам się w prawo itp.). Rozumienie tego zjawiska przejawia się generowaniem przez to zwierzę ruchów zamierzonych; w teście lustra ważne jest, to, czy zwierzę próbuje usunąć marker z czoła po zobaczeniu go w lustrze (Gallup, 1977).

Stąd może być już nie tak daleko do autoreferencji tworzącej samoświadomość. Kiedy ptak iskający swoje pióra podnosi skrzydło, by sięgnąć grzbietu, to posługuje się reprezentacją siebie (tj. swojego ciała) jako czegoś, co podlega jego woli – tu woli uniesienia skrzydła. Ptak dostosowuje swoje ruchy do tej wiedzy: sięga głową pod skrzydłem, wysuwa szyję itp. Można powiedzieć, że jego aktywność jest samozwrotna, to znaczy, że reakcje stają się bodźcem do ponownej reakcji. Taka pętla aktywności może stanowić podstawę samozwrotności świadomości, czyli samoświadomości. Jak twierdził bowiem Searle (1983), każdy akt świadomości jest zarazem samoświadomością, gdyż przenika go doświadczenie własnej sprawczości. Zatem pojęcie zachowania celowego zakłada pojęcie sprawczości, jeśli ma być odróżnione od odruchów, które są zachowaniami niepodlegającymi kontroli własnej podmiotu.

Sprawczość i relacje przyczynowo-skutkowe

Przynajmniej niektóre, jeśli nie wszystkie, zwierzęta są sprawcami działania. Ale co to dokładnie znaczy? Czy oznacza to, że psy, koty, świnie czy np. gołębie zdają sobie sprawę z istnienia relacji przyczynowo-skutkowych, czyli rozumieją, że otaczająca je rzeczywistość pełna jest praw?

Powszechny pogląd, że zwierzęta pozaludzkie nie rozumieją, nie znają czy też nie kierują się relacjami przyczynowo-skutkowymi odwołuje się zarówno do świadectw naukowych, jak i anegdotycznych. Te ostanie bazują na sytuacjach zaobserwowanych zazwyczaj w interakcji człowieka ze zwierzętami udomowionymi (najczęściej są to psy) lub częściowo udomowionymi (często są to zwierzęta hodowlane lub dzikie, ale częściowo zsocjalizowane z człowiekiem, np. konie, koty). Znany jest chociażby opis sytuacji, w których zamknięty wbrew własnej woli w jakimś pomieszczeniu pies cieszy się na widok osoby, która wcześniej go zamknęła, gdy tylko ta pojawi się po to, by wypuścić go z zamknięcia, albo wygląda na winnego, gdy z kolei zostanie skarcony za coś, czego nie zrobił (Wynne, Udell, 2013, s. 200–202). Człowiek przejawiałby najpewniej oznaki żalu, złości albo frustracji i raczej nie byłby radosny, jakby samo wypuszczenie przez osobę, która go zamknęła, oznaczało nagrodę. Może się nam wydawać, że nie mielibyśmy również poczucia winy, tylko raczej pretensje,

do osoby, która skarciła nas bez żadnego powodu³. Przykłady tego rodzaju mają być podstawą do kwestionowania poglądu, że zwierzęta rozumieją przyczynowość i sprawczość⁴.

Tymczasem wydaje się, że możliwa jest dyskusja z tak radykalnym i jednowymiarowym stanowiskiem. Ileż to razy mieliśmy bowiem okazję obserwować ludzi, którzy mimo że zostali skarceni, ukarani czy nawet przez kogoś skrzywdzeni, zachowują się tak, jakby nic się nie stało. Pamiętajmy też, że obserwacje i analizy zachowania zwierząt pozaludzkich najczęściej dotyczą zwierząt w jakimś stopniu udomowionych i zsocjalizowanych z człowiekiem. W związku z tym behavior tych zwierząt będzie w jakiejś mierze kształtowany zasadami, których wyuczyły się one w interakcji z człowiekiem. Pies z naszego przykładu może naprawdę odczuwać winę, mimo że nie powinien, biorąc pod uwagę, że w żaden sposób nie zawinił, tylko dlatego, że obcując z człowiekiem nauczył się, jakiego zachowania człowiek niejako oczekuje w danej sytuacji, tj. jakie zachowanie pozwala psu uniknąć kary⁵.

Eksperyment, na podstawie którego stwierdza się, że pies nie jest zdolny do poczuwania się do winy, gdyż nawet jeśli nie zrobił niczego, co łamałoby jakąś znaną mu regułę, a zostanie skarcony, to zachowuje się mimo wszystko tak, jak gdyby był winny, pozwala jedynie ustalić, że pies zachowuje się *tak, jak gdyby był winny*. To, co natomiast jest istotne w próbie ustalenia, czy zwierzęta pozaludzkie kierują się zasadą przyczynowości, to fakt, na który poza wszelką kontrowersją można się zgodzić – że rozumieją one zachodzenie pewnych korelacji. Skarcony i zamknięty pies cieszy się na widok osoby, która występuje w tym samym czasie i miejscu, co otwierające się drzwi, zaś zachowujący się jak winny pies, który nie zrobił niczego złego, zauważył, że takie zachowanie wywołuje u jego opiekuna zachowania czułe i opiekuńcze. Podsumowując, przynajmniej

³ Może nawet mielibyśmy pretensje do osoby, która skarciła nas dlatego, że faktycznie zawiniliśmy.

⁴ Przyczynowość, jeśli mowa o relacji przyczyna-skutek między przedmiotami, zaś sprawstwo, gdy mowa o relacji przyczyna-skutek, w której przyczyna jest tożsama z podmiotem (człowiekiem lub zwierzęciem innego gatunku niż *Homo sapiens*).

⁵ Oczywiście łatwo o taką interpretację psiego zachowania, która polega na mylnym uznaniu, że zachowanie psa, który boi się swojego opiekuna, wskazuje na poczucie winy (obie te formy zachowania mogą, na pierwszy rzut oka, przebiegać wedle podobnego schematu).

niektóre zwierzęta pozaludzkie rozpoznają istnienie w świecie korelacji czasoprzestrzennych. Okazuje się nawet, że zwierzęta wielu gatunków znakomicie rozpoznają złożone korelacje między mimiką twarzy człowieka a pożądanym przez niego zachowaniem. Mądry Hans, koń o osławionej, ale wcale nie wyjątkowej dla koni, umiejętności reagowania na emocje i towarzyszące im minimalne zmiany języka ciała ludzi, niech posłuży tu za klasyczny już przykład (Wynne, Udell, 2013, s. 31–33).

Zwierzęta pozaludzkie dostrzegają korelacje w świecie fizycznym. Stąd mówi się o tym, że niektóre z nich posługują się prostą czy ludową fizyką. Są też często wyspecjalizowane w rozumieniu korelacji między stanami psychicznymi a zachowaniem swoich towarzyszy, więc posługują się ludową psychologią czy dysponują prostą teorią umysłu. Na podstawie obserwacji zmian zachowania (czy to świadomie, czy nieświadomie) są w stanie przewidzieć, co ich towarzysz zrobi, a nawet odwrotnie, tzn. zachowanie ich współtowarzysza może służyć za wskazówkę, co też może on odczuwać. Przykładem pierwszej sytuacji jest kruk, który obserwując zachowanie innego kruka, jest w stanie przewidzieć, że schowa on jedzenie. Dyskretnie podąża więc za nim, by wykraść cenny skarb. Kruk rozpoznaje zatem korelację między określonym zachowaniem obserwowanego przez siebie osobnika a czynnością ukrywania przez niego jedzenia. Przykładem drugiej sytuacji jest altruistyczne zachowanie kapucynki, która widząc zdenerwowanie swojej sąsiadki, która w przeciwieństwie do niej nie dostała orzeszka, rezygnuje z tej nagrody albo oddaje ją roztrzęsionemu zwierzęciu. Mówi się wówczas, że zwierzęta posługują się zdolnością „czytania umysłów” innych osobników (Andrews, 2014, s. 179–180). Umiejętność ta może polegać na dostrzeganiu przez zwierzę korelacji między zachowaniem obserwowanego osobnika, a faktem, że nie otrzymał on przysmaku oraz tego, że zachowanie to można zmodyfikować, gdy odda się przysmak owemu osobnikowi lub z niego zrezygnuje. To, dlaczego niektóre zwierzęta są wrażliwe na emocjonalne zachowania tych, którym dzieje się jakiś rodzaj krzywdy czy niesprawiedliwości, jest odrębnym, rozległym tematem⁶. Ale czy istnieje przejście, w wypadku niektórych chociaż zwierząt, od rozpoznania korelacji do wiedzy na temat tego, że ktoś lub coś jest przyczyną wywołującą dany skutek, że ta sama osoba, która teraz tuli, a potem napełnia miskę, jest sprawcą tych czynności? A może kot głośno

⁶ Zob. np. Bekoff, Pierce, 2019.

miaucząc przy lodówce jedynie zauważył, że gdy w jej pobliżu pojawi się człowiek, to w miseczce znajdzie się również jedzenie? To, że w ogóle trzeba analizować tego rodzaju sytuacje i konstruować argumentacje na rzecz poglądu, że poza człowiekiem istnieją zwierzęta, które dostrzegają sprawczość w świecie, może wywoływać pewne zdziwienie, a nawet zażenowanie. Mimo wszystko, z uwagi na wciąż mnożące się wątpliwości, trzeba podjąć tę kwestię.

Weźmy pod lupę przykład następującego zachowania kota. Codziennie dostaje on od opiekuna przysmak z ręki, ale pewnego dnia, gdy opiekun bardzo się śpieszy, umieszcza przysmak w kocięj miseczce. Kot podchodzi do miski, obwąchuje jej zawartość, po czym stara się zwrócić uwagę opiekuna, łaszcząc się o jego nogi i głośno miaucząc. Opiekun wie, że kot chce dostać przysmak od niego, a nie z miski. Nieważne jest miejsce, w którym przysmak zostanie kotu wręczony, ważne jest, by dała go osobiście określona osoba. Takie zachowanie nie jest rzadkością i zna je prawdopodobnie większość kocich opiekunów. Przysmak w misce nie był zjadany, a kot coraz bardziej desperacko i rozpaczliwie domagał się realizacji codziennego rytuału. Zupełnie jakby wiedział, że to ta konkretna osoba powinna zrealizować jego życzenie. Zupełnie jakby uważał, że zostało pogwałcone jakieś prawo, prawo do tego rytuału. Zwierzęta zsocjalizowane z człowiekiem potrafią dość dobrze rozpoznawać sytuacje, w których coś niejako im się należy – na mocy wcześniejszej reguły – i odróżnić je od sytuacji, w których niejako nadużywają dobroci opiekuna lub współtowarzysza (Bekoff, Peirce, 2019). Wydaje się zatem, że poszukiwanie odpowiedzi na pytanie o rozpoznawanie przez zwierzęta pozaludzkie relacji sprawstwa musi przebiegać z uwzględnieniem stopnia socjalizacji zwierzęcia.

Co do problemu rozpoznawania samych relacji przyczynowo-skutkowych, można odwołać się do obserwacji, że wiele zwierząt reaguje zaskoczeniem na sytuacje nietypowe. Można zauważyć zdenerwowanie czy zaskoczenie psa, gdy niespodziewanie uruchamia się automatycznie (bez obecności w pobliżu człowieka) ekspres do kawy albo gdy kot podchodzi ostrożnie do papierowej torby zmiętej w kulkę, która samoistnie zaczęła się rozwijać. Takie rzeczy, zdają się przemawiać do nas zwierzęta, nie mają prawa się dziać w świecie, w którym rządzą dostrzegane przez psy i koty prawa fizyki. Zwierzęta reagują również zaskoczeniem, a nawet szokiem, na brak konsekwencji zachowań swoich towarzyszy (Wynne, Udell, 2013, s. 199–211).

Czy tego typu przypadki są wystarczającym świadectwem rozpoznawania przez zwierzęta pozaludzkie sprawstwa i praw zachodzących w przyrodzie? Oczywiście nie. Być może to, jak drobiazgowo analizujemy tego rodzaju przykłady świadczy wyłącznie o tym, że popełniamy błąd już w punkcie wyjścia samego badania. Należy bowiem zmienić sposób postrzegania sprawstwa oraz samych relacji przyczynowo-skutkowych. Skoro zaś sprawczość jest *de facto* relacją przyczynowo-skutkową, to zajmijmy się nakreśleniem mniej popularnego na gruncie filozofii, zaś powszechniej akceptowanego w psychologii, rozumienia tej ostatniej⁷.

Odwróćmy zatem pytanie tak, by spróbować innego podejścia w próbie rozwikłania problemu rozumienia przez niektóre zwierzęta pozaludzkie relacji przyczynowo-skutkowych. Nie pytajmy, czy są na to dowody, tylko najpierw zadajmy sobie pytanie, jak rozumiemy samą przyczynowość. Po uzyskaniu odpowiedzi na to pytanie możemy zacząć się zastanawiać, jak ma się do naszych ustaleń zachowanie zwierząt.

Nowożytny filozof David Hume podważył koncepcję przyczynowości powszechnie przyjmowaną nie tylko w jego czasach, ale także w obecnych. Opierał się on na obserwacji, że istnienie związku między dwoma przedmiotami czy zjawiskami, z których jedno nazywamy przyczyną, zaś drugie skutkiem, zakładamy czysto hipotetycznie, choć nie zdajemy sobie z tego sprawy. Jest nawet nieco gorzej, bo bardziej kierujemy się w tym postępowaniu pewnym nawykiem niż logiką. Otóż przywykliśmy do tego, powiadał Hume, że gdy podpalimy proch, to zaczyna on płonąć lub dochodzi do eksplozji. Nie obserwujemy jednakże żadnego fizycznego związku – poza czasoprzestrzennym – między prochem a ogniem. Widzimy ruch ręki z podpaloną pochodnią w kierunku prochu, obserwujemy, jak proch zaczyna się palić, po czym oślepia nas wybuch porażający również nasz zmysł słuchu. Podobnie rzecz się ma z kulą bilardową i kijem. Widzimy, że końcówka kija dotyka kuli, że kula toczy się i wpada do dołka w stole. Ale to wszystko. Żadnego oddziaływania przyczynowego nie widzimy. Skoro nie jest ono kwestią empirii, to może dochodzimy do istnienia tego zjawiska na drodze rozumowej? Znowu Hume uważa, że i tu nie jest tak łatwo.

⁷ Co interesujące, autorstwo tego podejścia przypisywane jest nie psychologowi, a filozofowi, Davidowi Hume'owi, jako jedna z pierwszych i najbardziej rzetelnych koncepcji przyczynowości i sprawczości (Hume, 2005).

Z pojęcia prochu czy kuli bilardowej nie można bowiem wywnioskować tego, co się z nimi stanie – że pierwsze wybuchnie, zaś drugie wpadnie do dołka. Pozostaje zatem nawyk obserwacyjny. Ilekroć widzieliśmy źródło ognia oddziałujące bezpośrednio na porcję prochu, widzieliśmy także i słyszeliśmy w tym miejscu eksplozję. Podobnie, ilekroć kij, którym celowaliśmy w kulę bilardową, dotykał jej powierzchni z określoną siłą, kula toczyła się z proporcjonalną siłą w kierunku, który był zgodny z kierunkiem nadanym jej przez kij.

Jeśli zatem przestaniemy mówić o relacjach przyczynowo-skutkowych jako zjawisku w tajemniczy sposób łączącym jakąś nierozzerwalną i niewykrywalną empirycznie siłą dwa stany, obiekty czy osoby, jeśli zastąpimy ją dobrze znanym i zrozumiałym dla nas pojęciem korelacji, to fakt, że zwierzęta pozaludzkie nie dostrzegają i nie rozumieją tej siły, ale dostrzegają i rozumieją – choćby elementarnie – zachodzące w świecie korelacje, będzie oparty na tej właśnie obserwacji. Dostrzeganie prostych i złożonych korelacji oraz przewidywanie następstw pewnych zdarzeń byłoby zatem kwestią możliwości posługiwania się przez nie rachunkiem prawdopodobieństwa lub (a może także) rozumowaniem dedukcyjnym. Istnieją poważne opracowania naukowe potwierdzające zdolność przynajmniej niektórych zwierząt pozaludzkich do przeprowadzania tzw. ukrytej inferencji (ang. *hidden state of inference*). Honi Sanders, Matthew A. Willson oraz Samule J. Gershman (2020) wykazali w przekonujący sposób, że umiejętność ta może być z powodzeniem przypisana chociażby szczurom, u których dochodzi do swoistego mapowania struktury otoczenia w hipokampie dzięki mechanizmowi szacowania prawdopodobieństwa wystąpienia pewnych cech tego otoczenia. Podobnie zresztą rzecz się ma w wypadku rozumowania dedukcyjnego. Na przykład Bruno Sauce i Louis Matzel (2017, s. 1) zauważają, że pewna odmiana ślimaka morskiego oraz wiele innych gatunków zwierząt bezkręgowych, potrafi posługiwać się elementarną formą wnioskowania indukcyjnego:

W najbardziej elementarnej formie wydaje się, że większość zwierząt wykazuje pewną zdolność do rozumowania indukcyjnego. Na przykład ślimaki morskie reagują na bodziec indukcyjnie, opierając się na przeszłych doświadczeniach z nim związanych (np. „wnioskując”, że bodziec można zignorować; jest to przypadek habituacji).

Obserwacja ta pozostaje w zgodności z konkluzją dotyczącą z kolei zachowań szczerów w zadaniu na orientację przestrzenną, że komórki miejsca są modulowane na podstawie informacji prospektywnych, takich jak przyszła trajektoria zwierzęcia (Sanders, Willson, Gershman, 2020).

Definicje instynktu

Bywa, że nie rozumiemy, w jaki sposób dane zwierzę osiągnęło pewien rezultat, np. jak gołąb wybudował gniazdo na gałęzi drzewa. Przywołujemy wówczas pojęcie instynktu i stwarzając pozory naukowości maskujemy w ten sposób swoją niewiedzę. Choć terminu „instynkt” używa się w nauce w odniesieniu do tych zachowań adaptacyjnych, które nie wynikają w sposób oczywisty z doświadczenia zwierzęcia oraz zdają się nie być prostym rezultatem uczenia się, to mimo wszystko kryje on pewną tajemnicę. Przy nieco bliższej analizie okazuje się bowiem, że choć zachowania instynktowne rzeczywiście mają charakter adaptacyjny, to daleko im do schematycznych i bezmyślnych automatyzmów, które w pierwszej kolejności przychodzą nam do głowy, gdy chcemy wyjaśnić zachowanie przykładowo przywołanego gołębia. Nikolaas Tinbergen, holenderski etolog zmarły w 1988 roku, definiował instynkt jako wrodzoną zdolność do złożonego z wielu etapów zachowania o charakterze adaptacyjnym, wyzwalanego przez popęd (czynnik wewnętrzny) oraz bodźce zewnętrzne, które mogą to zachowanie modyfikować (Tinbergen, 1951). Słabą stroną tej charakterystyki jest to, że przy odrobinie wysiłku każde zachowanie, w tym człowieka, można uznać za instynktowne. Jak słusznie zauważa za Patrickiem Batesonem (2002) Mark S. Blumberg (2017, s. 1–2), z pojęciem instynktu wiąże się spore zamieszanie i to wywołane przez samych naukowców. Zazwyczaj miano instynktu przypisywane jest bowiem tak różnym zjawiskom jak: umiejętność wrodzona, umiejętność, która nie stanowi rezultatu uczenia się, umiejętność niepodatna na wpływ doświadczenia (niezmienna), umiejętność wspólna wszystkim osobnikom danego gatunku czy wreszcie przekazywana w genach. Spróbujmy uporządkować tę zawiłą materię.

Pojęcie instynktu powinno nasuwać skojarzenie z innym, bardzo popularnym w czasach rozwoju etologii, czyli latach trzydziestych XX wieku – mowa o wdrukowaniu. Konrad Lorenz w 1935 roku opisał niezwykle zjawisko polegające na tym, że po wykluciu się kacze pisklę podąża za tym,

kogo pierwszego zobaczy na oczy, niezależnie od tego, czy jest to jego matka, sam Lorenz, czy ktokolwiek inny (Lorenz, 1935). Chcąc wyjaśnić, jak to zjawisko jest możliwe, dokonał w odniesieniu do zachowań – jak zręcznie to ujmuje Blumberg (2017, s. 2) – tego, czego Karol Darwin dokonał w odniesieniu do kości: pokazał ich wspólną genezę. Lorenz użył przykładu analogii między drapaniem się po głowie u ptaków i ssaków takich jak np. pies. Wyraźne podobieństwo między tymi zachowaniami u zwierząt bardzo różnych gatunków pozwalało założyć istnienie pierwotnie zaprogramowanych, niezmiennych wzorców motorycznych. Nie powinniśmy jednakże zapominać, że z kolei Darwin, opisując sposób życia zwierząt wielu gatunków, nie bez podziwu dla ich misterności przyznawał, że instynkt nie jest wyłącznym mechanizmem składającym się na ich działania. Jak wiadomo bowiem przykładowo opisywane przez niego przy innej okazji dżdżownice wykazują zachowania wspólne przedstawicielom gatunku, do którego przynależą, ale nie wyczerpują one całego repertuaru ich aktywności. Owszem, można odnieść wrażenie, że reagują na pewne bodźce w podobny do siebie sposób, ale gdy przyjrzeć się szczegółowo ich zachowaniom, wówczas stają się widoczne indywidualne preferencje tych zwierząt oraz modyfikacje schematów gatunkowych ich działań (Crist, 2002, s. 7).

Jak zatem poradzić sobie z zawilocią materii zachowań zwierząt? Czy sprowadzić je do kategorii instynktu, czy może podzielić na instynktowne i wyuczone lub instynktowne i inteligentne? Istnieje trzecie, mniej popularne i pewnie bardziej kontrowersyjne, rozwiązanie – albowiem można podać w wątpliwość samo pojęcie instynktu. Różnica między zachowaniami nazywanymi tradycyjnie instynktownymi a tymi, które im się przeciwstawia lub od nich odróżnia, polegałaby na tym, że autorem zachowań instynktownych rozpatrywanych u ich źródła nie jesteśmy my sami, czyli nie są nimi podmioty zachowujące się w sposób instynktowny. Zachowanie instynktowne zanim niejako zasłuży na swoje miano jest „zwyczajnym”, wyuczonym zachowaniem naszego praprzodka, które dopiero w kolejnych pokoleniach, dzięki dziedziczeniu, zasługuje na miano instynktu. Zaczniemy jednak od najbardziej oczywistych podejść do tego pojęcia.

Poszukiwanie znaczenia terminu „instynkt” w języku łacińskim podsuwa interesujące skojarzenia takie, jak: *instinctus*, *apetitus*, *impulsus*, czyli odpowiednio: instynkt, podnieta, impuls. W żaden sposób nie zbliżają nas one do lepszego rozumienia samego zjawiska działania instynktownego.

Zazwyczaj jednak termin „instykt” rozumiany jest jako działanie bezmyślne (Crist, 2002, s. 7) i pojawia się w pismach Darwina obok takich określeń jak zachowanie „pozbawione sensu” czy „bezcelowe” (Darwin 1881/1985, s. 95). Współcześnie traktuje się zachowania instyktowne nieco odmiennie. Spuścizna Darwina obecna jest jedynie w metaforze zamka i klucza (czy bodźców-znaków), gdzie zamkiem jest bodziec wzbudzający reakcję, zaś kluczem odpowiedź, czyli zachowanie do niego pasujące. Zdecydowanie akcentuje się w aktualnych opracowaniach związek zachowania instyktownego z motywacją, a zatem z czynnościami popędowo-emocjonalnymi. Na przykład Ernst Mayr (1974) podkreśla, że czynności popędowe i emocjonalne mogą być zarówno wrodzone, jak i nabyte. W celu rozróżnienia tych typów zachowań posługiwał się odpowiednio określeniami instykt zamknięty (ang. *closed instinct*) i instykt otwarty (ang. *open instinct*). Przyjrzyjmy się bliżej zachowaniom tego rodzaju.

Typy zachowań instyktownych

Zachowania instyktowne, które mają charakter nabyty, polegają na modyfikacji genotypu lub modyfikacji aktywności jego fragmentów przez środowisko. Do zachowań tego typu zalicza się uczenie się asocjacyjne, będące efektem oddziaływania powtarzającego się bodźca na genotyp danego zwierzęcia. Co interesujące, uczenie się asocjacyjne wymaga czasu i najczęściej ma miejsce u gatunków zwierząt, które sprawują opiekę rodzicielską. Najprostszą formą uczenia się nieasocjacyjnego jest habituacja, polegająca na wytworzeniu zachowania w reakcji na pojedynczy bodziec, a nie jak w wypadku uczenia się asocjacyjnego na bodźce stowarzyszone z sobą w czasie i miejscu.

Uczenie się asocjacyjne dzielone jest na wdrukowanie, warunkowanie oraz naśladowanie i uczenie się poprzez instrukcje. Na przykład surykatki uczą swoje młode, jak obchodzić się ze skorpionem jako źródłem pokarmu, by nie zostać zranionym przez śmiertelnie niebezpieczny kolec ogonowy tego zwierzęcia, właśnie udzielając młodym konkretnych instrukcji. Mimo wszystko to imponujące, złożone i wymagające doświadczenia zachowanie uznane jest za przejaw instyktu, podobnie jak inne, wydawałoby się o wiele prostsze. Przykładowo zachowanie opiekuńcze myszy domowej może być opisane za pomocą następującego łańcucha zdarzeń: gdy mysz zbliża się

do swoich nowo narodzonych młodych, informacje percepcyjne trafiają do podwzgórza. Następnie aktywuje się gen fosB, a ta aktywność z kolei powoduje zmiany w enzymach i aktywację innych genów wpływającą na pracę obwodów neuronowych w podwzgórzu. Zmiany te prowadzą najpewniej do doświadczenia przez mysz popędu opiekuńczego w stosunku do swoich nowo narodzonych młodych, dając w efekcie zachowanie macierzyńskie, nazywane instynktem macierzyńskim.

Zauważmy w tym miejscu, że użycie kategorii doświadczenia przesuwając pojęcie instynktu w rejony sfery psychicznej zwierzęcia. Efektem finalnym zachowania instynktownego jest rozładowanie napięcia wzbudzonego bodźcem, a dokładniej bodźcem kluczowym. Nawet jeśli seria zachowań zwierzęcia nie przyniesie rezultatu w postaci skutecznego działania (np. budowa gniazda nie zostanie ukończona czy partner do rozrodu nie zostanie znaleziony), to niekiedy już samo podjęcie odpowiedniej aktywności prowadzi do wygaśnięcia zachowania instynktownego. Dzieje się tak dość często u osobników młodych, którym brakuje pewnych umiejętności oraz cechuje je jeszcze niski poziom motywacji do realizacji często trudnych i ryzykownych działań. Przykładowo to, co przeciętnemu człowiekowi może wydawać się błahostką, czyli budowa gniazda przez gołębia, w rzeczywistości wymaga nie tylko ogromnego doświadczenia, ale też dużej motywacji, a nawet odwagi. Gołąb musi wielokrotnie opuszczać bezpieczny azyl gałęzi drzewa, by szukać materiału na gniazdo na ziemi. Przebywanie na ziemi zawsze wiąże się z zagrożeniem życia. Ponadto nie każdy materiał nadaje się na budowę gniazda. Ileż to razy mieliśmy okazję obserwować gołębie, które drepcząc od gałązki do gałązki przyglądają się jej, podnoszą ją, by następnie upuścić i podjąć dalsze poszukiwania. Po odnalezieniu stosownego materiału przychodzi czas na mozolne osadzanie pierwszych gałązek na nierównej powierzchni drzewa. Materiał spada, jest zdmuchiwany przez wiatr, nie chce współpracować z jedynym narzędziem, którym gołąb dysponuje, tj. dziobem, narzędziem nieprecyzyjnym i mało wyrefinowanym, ale jakże skutecznym. Wyplatanie gniazda, nawet tego, które nam ludziom może wydawać się nieporadne i proste, jest efektem ogromnej determinacji i doświadczenia. Nużące i ryzykowne czynności składające się na proces jego tworzenia wymagają zatem motywacji, nagradzając zwierzę poczuciem zaspokojenia naglącej potrzeby. Właśnie motywacja stanowi główny mechanizm instynktu. Czym zatem jest motywacja i jak się ma do popędu oraz potrzeby?

Za sprawą Zygmunta Freuda i jego następców instynkt zaczęto definiować jako potrzebę rozładowania wewnętrznego napięcia i tym samym przejścia od stanu większego pobudzenia do stanu pobudzenia mniejszego. Wiadomo, że ostatecznym kresem takiego pędu do redukcji wewnętrznego napięcia jest zdaniem Freuda śmierć organizmu. Stąd wyróżniał on popędy życia, śmierci, seksualny oraz agresji, rozumianej jako popęd ku autoagresji, lecz tym razem skierowany na zewnątrz. Słowa „instynkt” i „popęd” są na gruncie psychoanalizy stosowane zamiennie, lecz należy pamiętać o tym, że popęd jest odczuciem, zaś instynkt zjawiskiem złożonym z tego odczucia oraz czynności pozwalających zredukować je tak, by zwierzę mogło powrócić do stanu względnej homeostazy psychiczno-biologicznej blokującej odczuwanie potrzeby działania. Zdefiniujmy zatem pojęcia instynktu, motywacji, popędu oraz potrzeby, podsumowując tym samym wszystko to, co zostało do tej pory w niniejszym opracowaniu ustalone.

Składowe zachowań instynktownych

Instynkt jest zjawiskiem złożonym, wieloetapowym i dynamicznym. Złożoność instynktu wynika z tego, że zawiera on dwie główne komponenty: psychiczną, którą jest określone doświadczenie czy odczucie, a mianowicie doświadczenie potrzeby zrobienia czegoś, oraz fizyczną, którą jest ciąg czynności prowadzących do redukcji odczuwania tej potrzeby. Przykładowo potrzeba znalezienia partnera do rozrodu wygasa, gdy zostaje on znaleziony. Instynkt jest cykliczny i niestanny. Zachowania instynktowne pojawiają się w zależności od pory roku, obecności bodźca kluczowego i innych czynników, o których będzie jeszcze mowa. Jest też niestanny: na przykład potrzeba znalezienia partnera do rozrodu zmienia się w potrzebę kopulacji, a następnie budowy gniazda, gdy kolejno partner zostanie znaleziony i dojdzie do kopulacji. Potrzeba z kolei to, podobnie do popędu, rodzaj odczucia przymusu rozładowania wewnętrznego napięcia związanego z wykryciem jakiegoś braku bądź innym stanem pragnieniowym. Różnica między potrzebą a popędem polega na tym, że potrzeba zdaje się bardziej przynależeć do sfery emocji, zaś popęd jest niejako aktywną formą potrzeby, tj. mechanizmem jej realizacji. Gdy na przykład pies pragnie znaleźć się w tym samym pomieszczeniu, co jego opiekun, a uniemożliwiają to zamknięte drzwi, najpewniej będzie używał wokalizacji, a nawet skrobał

do drzwi, by osiągnąć swój cel. Pragnienie, rozumiane jako stan, pewna emocja, przechodzi w działanie (popęd) w postaci agresji skierowanej na obiekt uniemożliwiający realizację celu, ewentualnie staje się popędem autoagresji, gdy zwierzę skieruje swoją frustrację na siebie samo, np. okaleczając się (wyrywając sobie sierść, gryząc swój ogon itp.). Konflikt albo brak możliwości realizacji popędu mogą skutkować również negatywnymi odczuciami, mającymi przełożenia na konkretne zachowania. Stąd też interesującym staje się zagadnienie relacji między czynnościami instynktownymi a zachowaniami upustowymi i przerzutowymi, powstałymi właśnie w wyniku zablokowania możliwości realizacji popędu. Przyjrzyjmy się im pokrótce.

Reakcja upustowa zakłada silny popęd pod nieobecność bodźca kluczowego, np. u kotów pogoń za nieistniejącą zdobyczą czy u psa głośna wokalizacja, a nawet niszczenie przedmiotów znajdujących się w najbliższym otoczeniu. Reakcje tego rodzaju są zdefiniowane gatunkowo i bardzo rzadko przybierają formę inną niż u pozostałych przedstawicieli danego gatunku. Podobnie jest z czynnościami przerzutowymi, które są efektem konfliktu między popędami, na skutek którego żaden z popędów nie może zostać zaspokojony – np. agresja przeniesiona u kotów, która ma podłoże w dwóch jednoczesnych popędach, którymi są najczęściej chęć ucieczki i jednoczesne pragnienie walki. Kot może wówczas zaatakować osobnika, który jest zupełnie postronny.

Gdy mowa o zachowaniach upustowych i przerzutowych, warto dodać do tej listy jeszcze reakcje rytualne. Zwierzę, wykorzystując zachowanie przerzutowe, czyni z niego jednocześnie komunikat jasny i wyraźny w odbiorze dla osobnika tego samego gatunku (Sadowski, 2007, s. 386). Następuje tu interesujący zwrot od zachowania o funkcji czysto biologicznej (przerzutowej), do zachowania o funkcji komunikacyjnej. Przykładowo niektóre ptaki w obecności rywala nerwowo czyszczą pióra. Jest to czynność przerzutowa – zamiast wdać się w otwarty konflikt z rywalem, kierują uwagę na siebie. Zauważono jednak, że zachowanie to zostało zrytualizowane, tak że u niektórych gatunków ptaków niesie z sobą komunikat mówiący o sile i odwadze osobnika dokonującego owych czynności pielęgnacyjnych. Nieco żartobliwie mówiąc, przypomina to przeczesywanie dłonią włosów przez mężczyzn albo potrząsanie włosami przez kobiety, gdy chcą podkreślić swoją pewność siebie.

Na koniec tej części analizy pozostało do omówienia pojęcie motywacji. Albert Bandura wskazał na dwa jego źródła. Jedno określił mianem wewnętrznego, drugie zaś zewnętrznego. Motywacja wewnętrzna konstituowana jest poczuciem własnej kompetencji, zdolnością do kontroli swoich działań, poczuciem autonomii w doniesieniu do napotkanego problemu oraz napotkanych osób oraz siłą zainteresowania działaniem w danym zakresie. Jednakże najsilniejszym motorem motywacji wewnętrznej jest chęć osiągnięcia własnej satysfakcji. Zewnętrzne źródło motywacji utożsamił on z opinią innych osób na temat nas samych, kompensacją, obietnicą nagrody lub strachem przed karą. Stąd mówi się o czterech wymiarach motywacji, do których należą: kompetencja, kontrola i autonomia, zainteresowanie i wartościowość celu oraz opinia innych (Bandura, 1971, s. 8). Według psychologów społecznych kontynuujących dzieło Bandury, motywacja jest ostatnią, bardzo ważną fazą czteroetapowego procesu uczenia się. Uczenie, które ze swej istoty polega bądź na modelowaniu w oparciu o obserwację, bądź na modyfikacji zachowania na podstawie obserwacji jego konsekwencji, rozpoczyna proces skierowania uwagi na to, czego mamy się nauczyć. Następną jest faza zapamiętywania, kolejny etap to powtórzenie zachowania, które jest oczekiwane, zaś na koniec powinna wystąpić motywacja do powtarzania tego zachowania. Motywacja stanowi zatem element tak istotnego i złożonego zachowania, jakim jest uczenie się.

Zauważmy, że motywacja jest w tym opracowaniu traktowana jako kategoria wyjaśniająca pojęcie instynktu, a nie na odwrót. Czy byli jednak tacy, których zdaniem to motywacja powinna być definiowana w kategoriach zachowań instynktownych? Owszem. Najbardziej tym pomysłem wstawił się William James. Uważał on mianowicie, że instynkt rozumiany jako niewyuczalne, wrodzone zachowanie stanowi kluczowy składnik motywacji. Podawał przykład noworodka, który jest bardzo silnie zmotywowany do poszukiwania piersi matki, gdy odczuwa głód. Uczucie głodu jest potrzebą czy pragnieniem, wyzwajającym popęd inicjujący zachowanie gatunkowo zdefiniowane, czyli instynktowne, ale zachowanie to nie zostałoby podjęte, gdyby nie ów popęd. Popęd działa motywująco. Czy zatem motywacja nie ma z instynktem po prostu wspólnego mianownika w postaci popędu? Nie jest tak, że motywacja jest rodzajem instynktu. Motywacja jest popędem, zachowanie instynktowne jest natomiast złożone z popędu i działania stanowiącego sposób zaspokojenia go. Noworodek poszukujący ustami źródła mleka jest zmotywowany (czuje popęd) do podjęcia takich

działań, które ustaną, gdy osiągnięcie celu doprowadzi do zaniku motywacji. James zwrócił jednakże uwagę na interesujące zjawisko związane z motywacją, którym jest jej charakter zewnętrzny i wewnętrzny. Podobnie do tego, zachowania instynktowne również, jak była już o tym mowa, mogą być wyzwalane tak czynnikami zewnętrznymi, jak i wewnętrznymi. Być może ta zbieżność jest wypadkową nakładania się na siebie popędowego charakteru instynktu z wchodzeniem w skład zachowań instynktownych motywacji.

Można przyjąć, że motywacja jest mechanizmem regulującym zachowania instynktowne (Sadowski, 2007, s. 379). Czynniki lub bodźcami motywacyjnymi określa się natomiast wszystko to, co wzbudza popęd u zwierzęcia (Sadowski, 2007, s. 379). Sterowane popędem zachowania instynktowne mogą być wzbudzone bądź hamowane, co pozwala w następujący sposób usystematyzować popędy: apetytywne, awersyjne, wewnątrzpochodne, zewnątrzpochodne oraz pierwotne i wtórne. Bardzo ciekawą kategorię stanowią popędy wtórne, gdyż w odróżnieniu od pierwotnych, zwierzę nabywa je na drodze uczenia się (Sadowski, 2007, s. 380). Dodatkowo można wyróżnić kategorię popędu ciekawości, który Wojciech Pisula (2003) określa mianem zachowania eksploracyjnego i znajduje w nim cechy różniące go od zachowań instynktownych. Zachowanie eksploracyjne jest bowiem tak naprawdę sekwencją zachowań wyzwolonych przez popęd ciekawości. Pisula podkreśla również, że zachowania instynktowne mają charakter sztywnych wzorców regulacji zachowania (Pisula, 2003, s. 61). Jako przykład podaje stawonogi, kwestionując jednak słuszność przekonania, że cały repertuar ich zachowań można wyjaśnić za pomocą pojęcia instynktu (Pisula, 2003, s. 60–61). Zwierzę, które eksploruje otoczenie, wykazuje gotowość do wejścia w interakcje z nowym bodźcem przy zachowaniu jednoczesnej uważności nakierowanej na potencjalne ryzyko. Widać wyraźnie, że zachowania eksploracyjne wymagają od zwierzęcia bardzo silnej motywacji, ale jest ona powiązana bardziej z „nadwyżką” poznawczą niż popędem.

Podsumowując, instynkt definiuje się jako wrodzoną zdolność do złożonego z wielu etapów wstępnie zaprogramowanego zachowania wyzwalanego przez popęd (czynnik wewnętrzny) oraz bodźce zewnętrzne, które mogą to zachowanie również modyfikować (Tinbergen, 1951). Działania instynktowne mają charakter adaptacyjny, ale gdy przybierają formę otwartą (ang. *open instincts*), stają się wówczas podatne na modyfikacje, które z kolei

u przyszlých pokoleń mogą przybrać formę zachowań instynktownych zamkniętych (ang. *closed or rigid instincts*). Przyjrzyjmy się etapom tych zachowań.

Etapy zachowań instynktownych

W zależności od tego, jak rozumiane jest zachowanie instynktowne, wyróżnia się mniej lub więcej etapów jego realizacji. Nicolai Tinbergen wyróżniał dwie jego fazy: apetycyjną oraz konsumacyjną (Tinbergen, 1951; Sadowski, 2007, s. 384). Każdy z tych typów czynności może mieć kilka podetapów, zaś rozpoczęcie kolejnego etapu lub podetapu blokuje poprzedni. Z kolei inni behawioryści, jak John Watson, łączyli obecność zachowań instynktownych z asocjacyjnym uczeniem się, podczas którego powstawał dany nawyk. Natomiast Bogdan Sadowski wyróżnia następujące fazy instynktu: etap popędu apetycyjny, kolejne fazy apetycyjne oraz faza apetycyjna z bodźcem kluczowym, prowadzącym do spełnienia (Sadowski, 2007, s. 385). Przykładowo sokół wędrowny rozpoczyna łowy, gdy pojawia się u niego uczucie głodu, a zatem faza zachowania apetycyjnego bez bodźca kluczowego (widoku ofiary pobudzającej apetyt), następnie odbywa on lot krążący w poszukiwaniu ofiary. Po jej dostrzeżeniu ten etap zachowania zostaje przerwany i pojawia się nowe zachowanie, polegające na pozorowaniu ataku w celu oddzielenia pojedynczego zwierzęcia od reszty stada. W momencie odłączenia się pojedynczego osobnika od stada staje się on celem, zaś zachowanie pozorowanego ataku ustaje, gdyż bodźcem kluczowym jest teraz ów pojedynczy cel. Sokół podejmuje nowe działanie spełniające, polegające na ataku i konsumpcji ofiary (Sadowski, 2007, s. 385).

To, czy zwierzę przechodzące wszystkie te fazy zachowania zaledwie temu zachowaniu podlega, czy też można tu mówić o inteligencji, a nawet rozumności, wymaga osobnej refleksji. Przez zachowania inteligentne można na przykład rozumieć to, co dziś psychologowie, za Rymondem Cattellem i Donaldem Hebbem, określają mianem inteligencji płynnej i skrytalizowanej (Cattell, 1963) czy odziedziczonej i przejawianej w zachowaniu (Brown, 2016)⁸. Hebb, stosując swoje obserwacje do zachowań zwierząt,

⁸ Interesujące omówienie zagadnienia inteligencji w rozumieniu Cattella i Hebba oraz tego, czy prace tych autorów były wzajemnie inspirowane, zawiera Brown, 2016.

a ściślej rzecz biorąc – aktywności szczurów w warunkach laboratoryjnych oraz domowych, zauważył, że inteligencja, z którą zwierzęta te przychodzą na świat, powinna być wyraźnie oddzielona w badaniach od inteligencji kształtowanej czynnikami środowiskowymi (Hebb, Williams, 1946, s. 61). Inteligencja tego rodzaju miała być badana tzw. testem Hebba-Williamsa, a opierając się na nim wielokrotnie współcześnie potwierdzono, że aspekty fizjologiczne oraz środowiskowe inteligencji są dwiema różnymi stronami tego samego zjawiska (Brown, 2016, s. 9). Można zatem domniemywać, że zachowania instynktowne należą do takich, które znajdują źródło w wyuczonych zachowaniach, wchodzą do stałego repertuaru działań osobnika, a następnie są dziedziczone przez jego potomków. Odpowiedzi zatem na pytanie, czym jest działanie instynktowne, należy poszukiwać w tym, co oddziałując na zachowanie staje się zarazem wzorcem uniwersalnym gatunkowo, a niekiedy również międzygatunkowo (Wyatt, 2017, s. 34–35).

Rozwój i morfologia – afordancje morficzne

Przeprowadzone wcześniej analizy pokazują, że aby koncepcja zjawiska instynktu nie opierała się na nic niewnoszących truizmach, powinna objąć wszystkie najważniejsze aspekty tego zachowania, tj.:

- a) wyjaśniać rozwój, czyli nabywanie zachowania instynktownego;
- b) pokazywać, jaka jest relacja między biologią organizmu (morfologią) a zachowaniem instynktownym;
- c) wyjaśniać udział zachowania instynktownego w interakcji zwierzęcia ze środowiskiem (teoria ekologiczna).

Rozwój zachowań instynktownych wyjaśniany jest w kategoriach epigenety. Wiadomo, że istnieją dwa rodzaje dziedziczenia: genowe i pozagenowe, czyli epigenetyczne. Dziedziczenie pozagenowe powoduje zmianę fenotypu oraz wpływa na proces ekspresji genów za sprawą oddziaływania na organizm czynników środowiskowych. Skoro zachowania instynktowne pełnią w pierwszej kolejności funkcję adaptacyjną, to należy ich szukać za pośrednictwem epigenetyki. Dziedziczenie opisywane przez epigenetyków to taki proces, który omija DNA, dotyczy natomiast chromatyny, a także niekiedy modyfikacji nukleotydów. Zmiany epigenetyczne nie mają bezpośredniego wpływu na DNA, tzn. nie zmieniają go, jednakże są odpowiedzialne za oznaczanie sekwencji kodu genetycznego markerami chemicznymi.

Markery tego rodzaju mają już bezpośredni wpływ na to, które fragmenty DNA ulegną ekspresji. Nie wchodząc w szczegóły tego skomplikowanego procesu, warto podkreślić, że przyjmuje się obecnie, iż mechanizmy epigenetyczne pozwalają na międzypokoleniowe dziedziczenie doświadczeń. Literatura naukowa opisuje szczegółowo udokumentowane przypadki dziedziczenia pozagenowego traumy, stresu i depresji (Chmielewska i in., 2019). Badania nad depresją doprowadziły do zidentyfikowania czynników, które wydają się odpowiadać za powstawanie następstw traumy wczesnodziecięcej w dorosłym życiu osób, które jej doświadczyły (Chmielewska i in., 2019; Labonte i in., 2012, s. 722–731). W związku z tym, że czynniki te mają charakter epigenetyczny, można postawić hipotezę, iż możliwe jest również dziedziczenie pozagenowe tego rodzaju doświadczeń. W dalszej zaś kolejności mechanizmy epigenetyczne być może regulują dziedziczenie także innych, nietraumatycznych doświadczeń. Przyjrzyjmy się mechanizmowi epigenetycznemu traumy wczesnodziecięcej. Obok mechanizmu epigenetycznego pojawia się nie mniej przekonujące wyjaśnienie genezy zachowań instynktownych, którym jest rozwój indywidualny cech gatunkowych. Teorie te nie powinny być traktowane jako wzajemnie konkurencyjne, ale jako uzupełniające się.

Przez rozwój indywidualny cech gatunkowych rozumie się nabywanie przez zwierzę umiejętności, które traktowane są jako właściwe danemu gatunkowi. Przykładowo gołąb, gdy osiąga dojrzałość płciową, zaczyna interesować się budową gniazda. Początkowo poszukuje materiału, który niekoniecznie sprawdza się jako budulec, ale z czasem uczy się, które gałązki nadają się do zaplatania gniazda na gałęzi. Określony sposób budowy gniazda oraz materiał, z którego jest ono wykonane, będzie wyglądał inaczej u gołębia, inaczej zaś u języka, gdyż są one gatunkowo specyficzne. Co to jednak oznacza? Czym jest zachowanie gatunkowo specyficzne? Jakie cechy takiego zachowania pozwalają w ten sposób je klasyfikować?

Mark S. Blumberg, amerykański neuropsycholog rozwojowy, uważa, że Lorenz uczynił dla zachowania to, co Darwin uczynił dla kości (Blumberg, 2017, s. 2). Ten zgrabny aforyzm oddaje ideę, w myśl której zachowanie nie jest zjawiskiem zrozumiałym bez osadzenia go w kontekście procesu, który doprowadził do jego wykształcenia. Tak, jak Darwin pokazał proces kryjący się za konkretną morfologią i funkcją układu kostnego zwierząt (w tym człowieka), tak Lorenz usilnie starał się odsłonić procesy prowadzące do powstania zachowania instynktownego, pełniącego określoną funkcję,

poszukując ich w morfologii organizmu. Można powiedzieć, że etologiczne podejście do pojęcia instynktu zaproponowane przez Lorenza zaowocowało założeniem o „krytycznym przejściu od kości do zachowania” (Blumberg, 2017, s. 2). Często cytowany bywa fragment, w którym etolog ten unaocznia, że różnica w wyglądzie kończyny przedstawicieli rozmaitych gatunków – od pletwy u walenia, poprzez skrzydło nietoperza, po ludzkie ramię – nie powinna przysłaniać, że wynika to z różnicy w ich funkcjach i nie oznacza, iż gatunki te nie mają wspólnego ewolucyjnie przodka (Lorenz, 1958, s. 119). Przyjrzyjmy się fragmentowi tekstu przytoczonego przez Blumberga:

Czyż nie jest możliwe, że pod wszystkimi odmianami indywidualnego zachowania znajduje się wewnętrzna struktura odziedziczonago zachowania, charakteryzująca wszystkich członków danego gatunku, rodzaju lub większej grupy taksonomicznej – tak jak szkielet praprzodka charakteryzuje kształt i budowę wszystkich współczesnych ssaków?

(Lorenz, 1958, s. 119)

Zatem nawet indywidualne zachowania powinny być wyjaśniane wedle etologicznego modelu instynktu w duchu Lorenza, w kategoriach bazowych zachowań wrodzonych. Znany przykład, w którym porównuje on technikę drapania się u ptaków i psów, ma potwierdzać, że nie tylko w granicach jednego gatunku, ale także niejako międzygatunkowo pewne zachowania mogą mieć wspólną strukturę, która jest dziedzictwem ewolucyjnym po naszych przodkach. Może dziwić, że drobiazgowo obserwowane, połączone na dodatek z wnikliwą analizą, nie nakierowały Lorenza na inny trop. Jest nim zjawisko, które na potrzeby tego opracowania można określić mianem morficznych afordancji. Przyjrzyjmy się pokrótce tej kategorii.

Blumberg argumentuje na rzecz odmiennej wizji nabywania przez zwierzęta pozaludzkie i przez samego człowieka zachowań, które są specyficzne gatunkowo i które nie wydają się zbyt podatne na modyfikacje. Idea leżąca u podstaw jego propozycji odwołuje się do zasady wynikania funkcji z morfologii. Chodzi o to, że pewne obserwacje zdają się potwierdzać przypuszczenie, że np. to, co frapowało Lorenza, a co pokazywał w przykładzie ze sposobem drapania się u ptaków i u psów po głowie, można wyjaśnić inaczej niż poprzez odwołanie się do zachowania wrodzonego. Blumberg pokazał bowiem, że to, iż ptak przekłada nogę pod skrzydłem albo nad nim (jak preferują to pisklęta) wynika z budowy fizycznej ich ciała. Również

przysiadanie na nodze, na której podczas tej czynności ptak stoi, związane jest z morfologią jego ciała, gdyż pomaga mu w utrzymaniu równowagi. Pies drapiący się tylną łapą po głowie przekłada ją nad przednią łopatką i przykuca z podobnej przyczyny (Blumberg, 2017, s. 9).

Dalsze obserwacje, m.in. wpływu grawitacji na zachowania uznane za wrodzone u szczurów, skłoniły go, by bardzo ostrożnie podchodzić do samej kategorii instynktu. April Ronca i Jeffrey Alberts wraz ze współpracownikami podjęli się badania wpływu grawitacji na rozwój układu przedsionkowego u szczurów. Na pokładzie wahadłowca kosmicznego NASA umieścili grupę samic szczurów w okresie ciąży. Jak wiadomo, układ przedsionkowy rozwija się już w życiu płodowym. W tym samym czasie w grupie kontrolnej, w normalnych, ziemskich warunkach, rozwijały się cięższe inne samic. Ciężarne szczury wróciły na Ziemię na dwa dni przed przyjściem na świat potomstwa. Po narodzinach młode szczurzątko umieszczano w wodzie łapkami do góry. Zwierzęta z grupy kontrolnej wykazywały przewidywalne, gatunkowe reakcje w teście zanurzenia w wodzie, polegające na próbie odwrócenia się grzbietem do góry, podobnie jak przy spadaniu. Natomiast te, które przyszły na świat po okresie spędzonym w kosmosie w większości przypadków nie próbowały się odwrócić, spadając na dno zbiornika grzbietem do dołu. Zaobserwowano również różnorodne zmiany behawioralne i neuroanatomiczne w układzie przedsionkowym, co miałyby być związane z rozwojem w warunkach mikrogravitacji (Blumberg, 2017, s. 5). Wystarczył tydzień doświadczenia ziemskiej grawitacji, a odruch prostujący (ang. *righting reflex*) w próbie zanurzeniowej wrócił do normy przewidzianej dla tego gatunku. Zarówno Ronca oraz Alberts, jak i Blumberg uważają, że schemat odruchu prostującego, będący następstwem dziedziczonej struktury i funkcji układu przedsionkowego, wykazują wpływ środowiska na zachowanie wrodzone (Blumberg, 2017, s. 5). Z jednej bowiem strony zachowanie to może zostać osłabione manipulacjami środowiskiem (przebywanie w warunkach mikrogravitacji), z drugiej jeśli nie doszło do wady fizjologicznej na skutek takiego działania, to wystąpi ono ponownie w warunkach standardowych. Najwyraźniej rozwijają się nawet „najprostsze” instynkty i to w odpowiedzi na liczne czynniki, które odziedziczyliśmy po naszych rodzicach, w tym środowisko grawitacyjne rodzimej planety. Można powiedzieć, że w przeciwieństwie do podejścia Lorenza, Blumberg docenia rolę czynników środowiskowych w kształtowaniu zachowań instynktownych do tego stopnia, że proponuje, abyśmy uznali,

iż pewne cechy środowiska są niejako zakodowane czy reprezentowane w samym zachowaniu wrodzonym.

Teoria systemów rozwojowych, która kryje się za tym pomysłem, zawiera jednak w sobie pewną ideę, która może wydać się nieco niepokojąca. Początkowa próba przełamania anachronicznego podejścia do kategorii instynktu *de facto* wzmacnia ją bowiem, dodając do listy czynników charakteryzujących zachowanie instynktowne własności środowiska, w którym dany gatunek ewoluuje. Można jednak spróbować innego podejścia, mianowicie opartego na pojęciu afordancji, pochodzącego wprost z ekologicznej teorii percepcji Jamesa J. Gibsona (1977) oraz koncepcji rozwoju i uczenia się percepcyjnego Eleonory Gibson (Gibson, Pick, 2000). Autorzy ci pokazują, jak sposobności dostrzegane percepcyjnie w cechach środowiska (np. dziupla) pociągają za sobą rozwój określonych zachowań (ukrycia się, budowy gniazda, chowania pokarmu itp.). Afordancje determinują potencjalne działania, jakie środowisko lub obiekt oferują jednostce. Są rodzajem relacji między zwierzęciem a środowiskiem, w której postrzega ono otoczenie w kategoriach swoich możliwych działań. Wydaje się, że zarówno percepcja afordancji, jak i zachowania instynktowne mają wspólną genezę epigenetyczną, albowiem choć same afordancje nie są zaprogramowane ani automatyczne, a raczej zależą od indywidualnego postrzegania i interpretacji środowiska, to już sam mechanizm ich percepcyjnego wykrywania jest kształtowany gatunkowo oraz niejako wyposażony w reprezentacje cech środowiska. Zauważmy teraz, że sposobności, których dostarcza środowisko, czyli afordancji, nie da się pominąć w wyjaśnieniach zachowań gatunkowych. Odruch prostujący u szczurów w eksperymencie opisywanym tu wcześniej czy to samo wrodzone zachowanie u kotów, spadających ptaków, ale już nie u zwierząt żyjących w warunkach wodnych, w środowisku naturalnym których upadek nie jest groźny, jest odpowiedzią organizmu na pewną cechę środowiska, tj. grawitację. Co więcej, sekwencja ruchów składających się na to zachowanie będzie podobna u wszystkich tych zwierząt – niezależnie, czy przynależą one do tego samego gatunku, czy nie, jeśli tylko charakteryzuje je podobna morfologia. Można zatem mówić zarówno o wpływie środowiska, a dokładniej afordancji środowiskowych, jak i własnej budowy ciała, czyli afordancji morficznych na to, jak będzie przebiegało zachowanie się danego zwierzęcia w określonych warunkach.

Podsumowanie

Podsumowując rozważania oraz formułując najważniejsze wnioski końcowe, należy podkreślić, że do utrwalenia zachowania, tak by zyskało ono miano instynktownego, potrzeba:

- a) procesów dziedziczenia epigenetycznego,
- b) oddziaływania afordancji środowiskowych,
- c) aktywności wynikającej z własności afordancji morficznych.

Wskazanie na zachowania wrodzone, dziedziczone na zasadach opisywanych przez samą epigenetykę, nie pozawala jeszcze rozstrzygnąć, czy mamy do czynienia z aktywnością behawioralną gatunkowo uniwersalną, czy może wrodzoną, ale indywidualną – jak w przypadku dziedziczenia traumy i tym samym wystąpienie pewnych zachowań, które są jej przejawem. Dopiero odwołanie się do dziedziczenia reprezentacji środowiska zawartych w morficznych i funkcjonalnych odwzorowaniach jego cech – afordancji morficznych – w połączeniu z mechanizmem epigenetyki pozwala na dokonanie tego rozróżnienia i tym samym przyczynia się do lepszego zrozumienia natury zachowań instynktownych.

Bibliografia

- Andrews, K. (2014). *The Animal Mind: An Introduction to the Philosophy of Animal Cognition*. London: Routledge.
- Bandura, A. (1971). *Social Learning Theory*. Stanford: Stanford University Press.
- Bar-On, D. (2013). Expressive Communication and Continuity Skepticism. *Journal of Philosophy*, 110 (6), 293–330.
- Bar-On, D., Green, M. (2010). Lionspeak: Communication, Expression, and Meaning. W: E. Rubenstein (red.), *Self, Language and World: Problems from Kant, Sellars and Rosenberg* (s. 89–106). Atascadero, CA: Ridgeview Publishing.
- Bateson, P. (2002). The corpse of a wearisome debate. *Science*, 297, 2212–2213.
- Bekoff, M., Allen, C., Burghardt, M. (red.) (2002). *The Cognitive Animal. Empirical and Theoretical Perspectives on Animal Cognition*. Cambridge, MA: The MIT Press.

- Bekoff, M., Pierce, J. (2015). *Dzika sprawiedliwość. Moralne życie zwierząt*. Tłum. S. Szymański. Kraków: Copernicus Center Press.
- Block, N. (2005). Two neural correlates of consciousness. *Trends in Cognitive Sciences*, 9 (2), 46–52.
- Blumberg, M.S. (2017). Development evolving: the origins and meanings of instinct. *WIREs Cognitive Science*, 8 (1–2), 1371.
- Brown, R.E. (2016). Hebb and Cattell: The Genesis of the Theory of Fluid and Crystallized Intelligence. *Frontiers in Humean Neuroscience*, 10.
- Cattell, R.B. (1963). Theory of fluid and crystallized intelligence: a critical experiment. *Journal of Education in Psychology*, 54, 1–22.
- Chmielewska, E. (2019). Mechanizmy epigenetyczne stresu i depresji. *Psychiatria Polska*, 53 (6), 1413–1428.
- Crist, E. (2002). The inner life of earthworms: Darwin's argument and its implications. W: M. Bekoff, C. Allen, M. Burghardt (red.), *The Cognitive Animal. Empirical and Theoretical Perspectives on Animal Cognition* (s. 3–8). Cambridge, MA: The MIT Press.
- Darwin, C. (1881/1985). *The Formation of Vegetable Mould, through the Action of Worms with Observations on Their Habits*. Chicago: University of Chicago Press.
- Davidson, D. (2004). *Problems of Rationality*. Oxford: Oxford University Press.
- Gallup, G. (1977). Self recognition in primates: A comparative approach to the bidirectional properties of consciousness. *American Psychologist*, 32 (5), 329–338.
- Gibson, J.J. (1977). The Theory of Affordances. W: R. Shaw, J. Bransford (red.), *Perceiving, Acting, and Knowing: Toward an Ecological Psychology* (s. 67–82). Mahwah, NJ: Lawrence Erlbaum.
- Gibson, E.J., Pick, A.D. (2000). *An Ecological Approach to Perceptual Learning and Development*. Oxford: Oxford University Press.
- Hebb, D.O., Williams, K. (1946). A method of rating animal intelligence. *The Journal of General Psychology*, 34, 59–65.
- Hume, D. (2005). *Traktat o naturze ludzkiej*. Tłum. C. Znamierowski. Warszawa: Aletheia.
- Labonte, B., Suderman, M., Maussion, G., Navaro, L., Yerko, V., Mahar, I., Bureau, A., Mechawar, N., Szyf, M., Meaney, M.J., Turecki, G. (2012). Genome-wide epigenetic regulation by early-life trauma. *Archives of General Psychiatry*, 69 (7), 722–731.

- Lorenz, K. (1935). Der Kumpan in der Umwelt des Vogels. Der Artgenosse als auslösendes Moment sozialer Verhaltensweisen. *Journal für Ornithologie*, 83, 137–213.
- Lorenz, K. (1950). The comparative method in studying innate behaviour patterns. *Symposia of the Society for Experimental Biology*, 4, 221–268.
- Lorenz, K. (1958). The evolution of behavior. *Scientific American*, 199, 67–74.
- Mayr, E. (1974). Behavior programs and evolutionary strategies. *American Scientist*, 62, 650–659.
- Pearce, J.M. (2008). *Animal Learning and Cognition. An Introduction*. New York: Psychology Press.
- Pepperberg, I. (1999). *The Alex Studies: Cognitive and Communicative Abilities of Grey Parrots*. Harvard: Harvard University Press.
- Pepperberg, I. (2021). Nonhuman and Nonhuman-Human Communication: Some Issues and Questions. *Frontiers in Psychology*, 12, 1–15.
- Pisula, W. (2002). *Psychologia zachowań eksploracyjnych zwierząt*. Gdańsk: GWP.
- Polley, C. (2011). What is language? W: J. Terrell (red.), *Introduction to language* (s. 1–11). Honolulu: UHM Department of Linguistics.
- Sadowski, B. (2007). *Biologiczne mechanizmy zachowania się ludzi i zwierząt*. Warszawa: Wydawnictwo Naukowe PWN.
- Sanders, H., Willson, M.A., Gershman, S.J. (2020). Hippocampal remapping as hidden state inference. *eLife*, 9.
- Sauce, B., Matzel, L. (2017). Inductive reasoning. W: J. Vonk, T. Shackelford (red.), *Encyclopedia of Animal Cognition and Behavior* (s. 1–8). New York: Springer International Publishing.
- Searle, J. (1983). *Intentionality. An Essay in the Philosophy of Mind*. Cambridge: Cambridge University Press.
- Shoemaker, S. (1976). Embodiment and Behavior. W: A. Rorty (red.), *The Identities of Persons* (s. 109–137). Berkeley: Berkeley University Press.
- Taylor, C. (2021). *The Explanation of Behaviour*. London: Routledge.
- Tinbergen, N. (1951). *The Study of Instinct*. Oxford: Clarendon Press.
- Tooby, J., Cosmides, L. (1992). The psychological foundations of culture. W: J.H. Barkow, L. Cosmides, J. Tooby (red.), *The adapted mind: Evolutionary psychology and the generation of culture* (s. 19–136). Oxford: Oxford University Press.

- Wittgenstein, L. (1972). *Dociekania filozoficzne*. Tłum. B. Wolniewicz. Warszawa: Wydawnictwo Naukowe PWN.
- Wyatt, T.D. (2017). *Animal Behaviour. A very short introduction*. Oxford: Oxford University Press.
- Wynne C.D.L., Udell, M.A.R. (2015). *Tajemnice umysłów zwierząt: ewolucja, zachowanie i procesy poznawcze*. Tłum. B. Leszczyńska, P. Leszczyński, A. Kłosiński. Warszawa: COAPE.

Nota o autorce

Adriana Schetz – dr hab. w Instytucie Filozofii i Kognitywistyki Uniwersytetu Szczecińskiego. Autorka monografii z zakresu kognitywistyki i filozofii umysłu *Biologiczny eksternalizm w teoriach percepcji* (Szczecin 2014). Aktualnie zajmuje się zagadnieniami dotyczącymi percepcji, ewolucji świadomości, procesów poznawczych i zachowań zwierząt oraz wpływem substancji psychoaktywnych na procesy kognitywne człowieka.

Address for correspondence: Instytut Filozofii i Kognitywistyki, Uniwersytet Szczeciński, ul. Krakowska 71–79, 71-017 Szczecin.

Cytowanie

Schetz, A. (2024). Geneza zachowań instynktownych. *Analiza i Egzystencja*, 67 (3), 77–107. DOI: 10.18276/aie.2024.67-04.



„Analiza i Egzystencja” 67 (2024), 109–125

ISSN (print): 1734-9923

ISSN (online): 2300-7621

DOI: 10.18276/aie.2024.67-05

ARTYKUŁY

JANINA MĘKARSKA

Uniwersytet Szczeciński

ORCID: 0000-0002-7390-5623

email: 2406@stud.usz.edu.pl

Błędy poznawcze związane z szowinizmem gatunkowym oraz odmawianie teorii umysłu zwierzętom innym niż ludzie

Słowa kluczowe: gatunkowizm, heurystyki, myślenie analogiczne, Kanon Morgana, teoria umysłu

Keywords: speciesism, heuristics, analogical reasoning, Morgan's canon, theory of mind

Cognitive biases associated with speciesism and the denial of theory of thought to non-human animals

Abstract

Speciesism, understood as “the bias or biased favoring of members of one’s own species at the expense of others,” is a widely discussed phenomenon, considered in the field of philosophy. This type of prejudice is rooted by the cognitive processes that shape people’s thinking about themselves and other animals. Over the years, social beliefs have found their support in scientific considerations. Currently, the claim of human uniqueness in the discussion of cognitive processes is not as widely supported as it used to be. The aim of this article is to identify the manifestations of this phenomenon in the history of research on the theory of mind in non-human animals and, more importantly, to identify cognitive biases that contribute to the adoption of careless and, as we will see in later chapters, often false interpretations of empirical research.

Wprowadzenie

Celem niniejszego artykułu jest wskazanie przejawów szowinizmu gatunkowego w historii badań nad teorią umysłu u zwierząt innych niż ludzie (ZINL) oraz, co ważniejsze, określenie błędów poznawczych przyczyniających się do przyjmowania tendencyjnych interpretacji badań empirycznych.

Aby zrealizować założony cel, w pierwszej części przedstawiam krótki rys historyczny zjawiska szowinizmu gatunkowego. Szczególną uwagę poświęcam roli Prawa Oszczędności Morgana w kształtowaniu przekonań dotyczących możliwości poznawczych ZINL. Wskazuję, że już za życia Morgana sformułowana przez niego zasada miała charakter kluczowy, a także podkreślam rolę omawianego paradygmatu w konstruowaniu współczesnych interpretacji eksperymentów z zakresu szeroko rozumianego *animal studies*. W części drugiej opisuję zagadnienie heurystyki reprezentatywności oraz związanych z nią błędów poznawczych. To pozwala mi na zdefiniowanie tendencyjności jako klasy błędów poznawczych ściśle związanych z omawianym uprzedzeniem (takich jak efekt halo czy myślenie analogiczne). W części trzeciej przedstawiam kategoryzację społeczną, która uznawana jest za podstawowy mechanizm powstawania uprzedzeń, w tym szowinizmu gatunkowego. W ostatniej części skupiam się na rozwinięciu zagadnienia teorii umysłu. Przedstawione w tej części rozważania sugerują, że obrona stanowiska przypisującego *Homo sapiens* wyróżniony status poznawczy wynikający z rzekomego posiadania zaawansowanych zdolności intelektualnych charakterystycznych wyłącznie dla jego gatunku jest wynikiem tendencji do bycia mniej skłonny do przypisywania ZINL złożonych zdolności poznawczych.

Chcę zaznaczyć, że aktualny kontekst społeczny sprzyja redukcji tendencyjności w środowisku akademickim związanym z *animal studies*. Efektem takiej redukcji jest mniej tendencyjne (w porównaniu do historii badań nad ZINL) interpretowanie wyników uzyskanych w eksperymentach.

Szowinizm gatunkowy

Temat szowinizmu gatunkowego, inaczej zwanego gatunkowizmem, rozumiany za Peterem Singerem jako „uprzedzenie lub stronnicze faworyzowanie członków własnego gatunku kosztem innych” (Singer, 2018, s. 58), jest nie tylko coraz częściej podejmowany przez środowiska naukowe,

ale i coraz głośniej dyskutowany w dyskursie społecznym oraz politycznym (Woleński, 2015). W ostatnich latach temat gatunkowizmu wzbudza w społeczeństwie intensywniejsze emocje oraz częściej manifestuje się w zachowaniu, np. poprzez rezygnację ze spożywania produktów mięsnych lub odzwierzęcych, ale też coraz częściej poruszany jest na gruncie rozważań naukowych.

Początek badań naukowych dotyczących uprzedzeń osadzony jest w kontekście zmian społecznych, jakie miały miejsce w Stanach Zjednoczonych w latach sześćdziesiątych XX wieku. Rozpoczęcie takich badań wynikało ze zmian społecznych, będących rezultatem działania grup walczących o równouprawnienie osób czarnoskórych (Stefaniak, Górska, 2015). W literaturze dotyczącej omawianego zagadnienia przyjmuje się, że uprzedzenia są powszechnym zjawiskiem społecznym, polegającym na „przyjmowaniu negatywnego nastawienia do grup społecznych lub osób do nich przynależących, wynikające z ich przynależności do kategorii społecznej, a nie z cech indywidualnych” (Stefaniak, Górska, 2015). Todd Nelson (2003) podkreśla, że do opisanania każdej postawy, w tym uprzedzenia, należy przeanalizować jego trzy komponenty: poznawczy, afektywny i behawioralny. W literaturze istnieje pewien problem terminologiczny związany z pojęciami uprzedzeń i stereotypów. Zagadnienie to wykracza poza ramy tematyczne niniejszej pracy. Uważam jednak za stosowne zaznaczyć, że uprzedzenia rozumiane są jako postawy, natomiast stereotypy to zestawy przekonań związane z określonymi grupami społecznymi. Oznacza to, że stereotypy stanowią komponent poznawczy uprzedzeń, komponent afektywny dotyczy zaś typu emocji związanych z określoną postawą oraz poziomu ich intensywności. Dodatkowo wyróżniam behawioralny komponent uprzedzeń nazywany dyskryminacją. W niniejszym opracowaniu skoncentruję swoją uwagę głównie na procesach poznawczych leżących u podłoża myślenia stereotypowego oraz wskażę na główne konsekwencje wynikające z posługiwania się tego typu myśleniem, do których należy postawa dyskryminacyjna wobec zwierząt innych niż ludzie.

W celu lepszego zrozumienia poznawczego komponentu omawianego zjawiska należy przedstawić kilka koncepcji filozoficznych i naukowych, które mają istotny udział w kształtowaniu przekonań dotyczących możliwości poznawczych zwierząt innych niż ludzie. Jedną z takich koncepcji jest teoria ewolucji (Darwin, 2021), która zakłada, że istnieje pewna ciągłość między psychicznymi i fizycznymi cechami wszystkich zwierząt.

Założenie to nosi nazwę tezy o ciągłości ewolucyjnej. Sam Darwin wykazywał, że np. grymas twarzy człowieka szydzącego z innego człowieka jest pozostałością po obnażaniu kłów przez zdenerwowane zwierzęta inne niż ludzie (Schultz, Schultz, Andruszko, 2008).

Przełomowość słów Darwina polega na przeciwstawieniu się ówczesnej perspektywie mechanicystycznej, której zwolennikiem był René Descartes. Kartezjusz kategorycznie określił swoje stanowisko, twierdząc, że w sferze psychicznej zwierzęta nie wykazują najmniejszego podobieństwa do ludzi. Natomiast za sprawą darwinowskiej tezy o ciągłości ewolucyjnej do opisu zachowań ZINL zaczęto coraz częściej stosować metodę anegdotyczną oraz introspekcję przez analogię. Połączenie tych dwóch metod cechowało się silną antropomorfizacją, co prowadziło do nadinterpretacji obserwowanych zachowań. Przykładem takich nadinterpretacji są prace Georga Johna Romanesa, który był zwolennikiem zastosowania darwinowskiej teorii ewolucji do aspektów umysłowości (Schultz, Schultz, Andruszko, 2008). Jego prace cechowały się jednak przesadną antropomorfizacją, a anegdoty były dostarczane przez obserwatorów nieposiadających wykształcenia psychologicznego.

Zmiana przekonań dotyczących zdolności poznawczych zwierząt innych niż ludzie doprowadziła do rozwoju zoopsychologii oraz przyjęcia paradygmatu Morgana, który przez następne lata był uznawany przez osoby interpretujące wyniki badań z zakresu *animal studies*. W celu podniesienia poziomu rozważań w omawianym temacie, Morgan wprowadził tzw. Prawo Oszczędności Morgana. Inaczej nazywane Kanonem Morgana, jest ono paradygmatem badań nad zwierzętami innymi niż ludzie. Głosi, że „nie wolno interpretować zachowania zwierząt jako wyniku wyższych procesów psychicznych, jeżeli można je wyjaśnić za pomocą niższego procesu psychicznego” (Schultz, Schultz, Andruszko, 2008, s. 162). W XIX wieku Morgan stworzył swój kanon w celu redukcji skutków ówczesnie dominującego narzędzia badawczego w dziedzinie zoopsychologii – metody anegdotycznej. Sam Morgan był zwolennikiem darwinowskiej tezy o ciągłości ewolucyjnej, jednak przeciwny był przesadnej antropomorfizacji. Jego celem nie było zatem wykluczenie antropomorfizacji, lecz zredukowanie jej wpływu na interpretację obserwowanych zachowań. Kanon Morgana wyznaczył tendencję pozwalającą na podniesienie jakości badań empirycznych z udziałem ZINL. Współcześnie prowadzi ona jednak do antroponegacjonizmu opisanego przez Fransa de Waala (2016). Przyjęcie perspektywy

antroponegacjonistycznej, rozumianej jako „negowanie *a priori* człowiekopodobnych cech u innych zwierząt i zwierzęcopodobnych cech u ludzi” (de Waal, 2019, s. 39), wiąże się z całkowitym odrzuceniem metody anegdotycznej i likwidacji antropomorfizacji. Skutkiem takiego podejścia jest perspektywa antropocentryczna w interpretacji badań nad ZINL, co prowadzi do odmawiania zwierzętom pozaludzkim zdolności poznawczych, wykluczając tym samym darwinowską tezę ciągłości ewolucyjnej. Zatem tendencyjna interpretacja Zasady Morgana prowadzi do przyjęcia współczesnej odmiany antropocentryzmu i mechanicyzmu, co jest przyczyną całkowitej likwidacji antropomorfizacji oraz odrzucenia metody anegdotycznej. Tym samym wyklucza interpretacje zachowań ZINL odwołujące się do procesów wyższego rzędu. Należy nadmienić, że sam Morgan nie wyznaczył jasnej granicy między procesami wyższego oraz niższego rzędu. W konsekwencji to, co rozumiemy przez te procesy, zależy od interpretacji osoby powołującej się na Prawo Oszczędności.

Obecnie podejście do badań z udziałem ZINL uległo zmianie. Modele empiryczne tworzone są dużo bardziej obiektywne. Na skutek zmiany podejścia badawczego eksperymentatorów coraz więcej też o wyjątkowości człowieka opartych na przekonaniu, że ludzi od innych zwierząt odróżnia konkretna zdolność poznawcza, zostało zakwestionowanych. Przykładem zdolności poznawczej, którą od lat wskazywano jako unikalną dla ludzkiego gatunku, była zdolność konstruowania tzw. teorii umysłu. Michael Tomasello przez długi czas podzielał pogląd, że jest to zdolność odróżniająca człowieka od innych zwierząt (Tomasello, 2015). Jak zostanie wykazane w podrozdziale *Teoria umysłu*, aktualnie uznaje się, że zdolność ta występuje nie tylko u człowieka.

Kategoryzacja społeczna

W przyporządkowywaniu obiektów do określonych kategorii udział biorą procesy ściśle związane ze zjawiskiem kategoryzacji społecznej (KS). Zjawisko to pozwala na uporządkowanie myślenia o sobie i innych w społeczeństwie. Uważa się, że mechanizmy te są przynajmniej częściowo zautomatyzowane oraz niezbędne do efektywnego funkcjonowania w społeczeństwie (Karyłowski i in., 1999). I chociaż KS odnosi się do relacji międzygrupowych w ludzkich społeczeństwach, w pracach związanych z szowinizmem

gatunkowym wskazuje się na zjawiska, które odpowiadają mechanizmom kategoryzacji społecznej. W wyniku tego podkreśla się silną dychotomię między „my” a „one” (Kwiatkowska, 2009), co jest charakterystyczne dla mechanizmu akcentuacji, polegającego na podkreśleniu różnic międzygrupowych oraz minimalizacji różnic wewnątrzgrupowych. Innymi mechanizmami KS obecnymi w dyskursie dotyczącym gatunkowizmu jest faworyzacja grupy własnej oraz homogenizacja obcej. Co ciekawe, językowe przejawy szowinizmu gatunkowego obecne są w dyskursie dotyczącym innych szowinizmów (np. rasizmu). Używając języka, który jest przejawem szowinizmu gatunkowego, podkreśla się różnice międzygrupowe w ludzkich społeczeństwach (Baran, 2011). Przykładem takiego użycia języka może być określenie zachowania grupy obcej za pomocą asercji „grupa x zachowuje się jak zwierzęta”.

Psychologia społeczna wyróżnia kilka motywacji kategoryzacji społecznej. Klasyczne ujęcia teoretyczne podkreślają siłę motywacji do KS wynikającej z potrzeby obrony dobrego imienia grupy własnej. Współcześnie wskazuje się jednak wady takiego podejścia, twierdząc, że modele te przedstawiają uproszczony obraz świata, w którym obowiązuje wspomniana wyżej dychotomia „grupa własna–grupa obca” (Mullin, Hogg, 1998). Podkreśla się tym samym istotność złożoności tożsamości społecznej (Roccas, Brewer, 2002) czy silny wpływ kultury na procesy związane z KS. Co ciekawe, najszerzej opisaną motywacją dla KS związaną z gatunkowizmem jest faworyzacja grupy własnej. Uważam, że faworyzując ludzi (jako grupę własną), odmawia się ZINL możliwości poznawczych. To w oczywisty sposób wiąże się ze stereotypem, że życie ludzkie jest ważniejsze i bardziej wartościowe od życia przedstawicieli innych gatunków zwierząt.

Kategoryzacje związane z gatunkowizmem są przynajmniej częściowo zautomatyzowane, co oznacza, że zachodzą w sposób spontaniczny i często nieuświadomiony. Ze względu na wymienione cechy omawianego mechanizmu, w następnej części artykułu odwołam się do heurystyki reprezentatywności. Prawdopodobnie stereotypy związane z gatunkowizmem przetwarzane są przez procesy zachodzące w jej ramach.

Heurystyka reprezentatywności i wynikająca z niej tendencyjność

Jak wspomniano we wstępie, tendencyjność rozumiem jako klasę błędów poznawczych wynikających z heurystyki reprezentatywności (Khaneman, Tversky, 2019). Jej funkcją jest skuteczne dokonywanie oceny sytuacji

na podstawie preferencji będących wynikiem sieci skojarzeń obserwatora. Natomiast błędy będące skutkiem omawianej heurystyki polegają na podejmowaniu decyzji lub dokonywaniu oceny sytuacji na podstawie nieadekwatnych preferencji związanych z sieciami skojarzeniowymi obserwatora. Zjawisko to związane jest z klasą błędów poznawczych, do których zalicza się np. efekt halo, polegający na wytworzeniu preferencji opartej na emocjonalnym odbiorze określonej informacji, czy efekt potwierdzenia, który polega na skłonności do przyjmowania argumentów zgodnych z przekonaniem odbiorcy komunikatu. Heurystyka ta bierze udział w procesach poznawczych, których celem jest odpowiedź na pytanie „w jakim stopniu A jest reprezentatywne dla B?”. Angażowana jest zatem przy rozważaniach:

- jakie jest prawdopodobieństwo, że obiekt A należy do klasy B?,
- jakie jest prawdopodobieństwo, że zdarzenie A bierze początek z procesu B?,
- jakie jest prawdopodobieństwo, że proces B wygeneruje zdarzenie A?

Aby zrozumieć, zjawisko przyjmowania postaw takich jak gatunkowizm, należy zastanowić się nad wpływem stereotypów na to, jak oceniamy prawdopodobieństwo, że obiekt A należy do klasy B. Jedną z cech heurystyki reprezentatywności jest tendencyjne poleganie na moralnie nieistotnych różnicach fizycznych przy dokonywaniu osądów moralnych (Kahneman, Tversky, 2019). Fundamentem oceny na podstawie wyglądu jest myślenie polegające na dostrzeganiu analogii, w skrócie myślenie analogiczne. Uznawane jest za element ludzkiego aparatu poznawczego (Szymanek, 2008), którego funkcją jest porządkowanie i wnioskowanie. Potocznie postrzegamy analogię jako zgodność cech obiektów, natomiast w ścisłym rozumieniu tego terminu jest to zgodność stosunków zachodzących między cechami obiektów (Szymanek, 2008). Różnica w znaczeniu ścisłym i kolokwialnym może prowadzić do błędów podczas stosowania zasady *podobne traktować podobnie*, czyli jednej z głównych zasad ferowania osądów moralnych. Kolokwialne rozumienie analogii ma podstawę w heurystycznej funkcji omawianego zjawiska. Heurystyczna funkcja analogii sprowadza się do możliwości postrzegania i reprezentowania różnych obiektów czy zjawisk w kategoriach właściwych innym. Rozumowanie analogiczne jest zatem metodą odwoływania się do wcześniej nabytych informacji oraz wyjaśnienia wcześniej niespotykanych zjawisk za pomocą zjawisk dobrze znanych (Biegański, 1909). Jest ono skuteczne i efektywne w redukcji niepokoju związanego z pojawieniem się nieznanego elementu otoczenia (Ellianawati,

Subali, Khotimah, Cholila, 2020), ma jednak swoje wady. Jedną z kluczowych zdają się błędne wnioski wiążące się z daną analogią. Wnioski te nie są widoczne na pierwszy rzut oka, jednak mogą stanowić motywację do przyjęcia pewnych postaw (Ziółkowski, 2019). Przykładem może być syndrom bambiego, opisany po raz pierwszy przez Harolda H. Dowa (1957). Bambizm to zjawisko emocjonalnego zaangażowania i lęku przed krzywdą wobec zwierząt, które są uważane za słabe, bezbronne i niewinne. W wyniku efektu halo wytworzona zostaje tendencja do antropomorfizowania zwierząt o uroczym wyglądzie, której towarzyszy obojętność wobec zwierząt mniej atrakcyjnych wizualnie. Ludzie przejawiający tę postawę idealizują oraz infantylishują przyrodę (w tym dzikie zwierzęta) (Lutts, 1992). Co ciekawe, uprzedzenie to umacniane jest przez biologiczny mechanizm *słodkiej reakcji* opisany przez Joan E. Serpell (1999). Jest to naturalny i automatyczny afekt, który pojawia się u ludzi w odpowiedzi na widok lub myśl o słodkich zwierzętach, jak choćby szczenięta, kocięta, króliki, małe ptaki czy ludzkie dzieci. Jej pojawieniu towarzyszą emocje takie jak radość, troska czy nawet miłość.

Przejawem bambizmu jest zabieranie małych dzikich zwierząt z ich naturalnego środowiska. W okresie letnim ludzie spacerujący po lasach napotykną młode zwierząt takich jak np. sarny czy króliki i kierowani obawą o ich bezpieczeństwo, decydują się na dotykanie, a nawet w skrajnych przypadkach – zabieranie zwierząt do klinik weterynaryjnych. Zachowanie to jest przejawem przekonania, że zwierzęta te nie poradzą sobie bez ludzkiej pomocy, nawet jeżeli znajdują się w charakterystycznym dla siebie środowisku. Co ciekawe, postawa, jaką jest bambizm, stanowi skuteczną motywację dla zmiany zachowania. Dowodem tego jest odnotowanie zmiany opinii publicznej dotyczącej myśliwych, związanej z premierą filmu animowanego *Bambi* (1942), która skutkowałą dwukrotnym zmniejszeniem liczby polowań na jelenie w Stanach Zjednoczonych (Lutts, 1992). Obserwacje Rainera Brämera (1993) wskazują, że omawiana postawa najczęściej obserwowana jest u mieszkańców miast, którzy na co dzień nie mają bezpośredniego kontaktu z naturą.

Innym przejawem szowinizmu gatunkowego jest przekonanie, że ludzie od innych zwierząt różni określona zdolność poznawcza. Przekonanie to skutkuje preferowaniem interpretacji wyników eksperymentów z udziałem ZINL, wskazujących na brak pewnych zdolności poznawczych. Jednym z przykładów takich preferencji jest przekonanie, że gibony nie są zdolne do korzystania z narzędzi (Yerkes, 1929; Köhler, 1925/1959). Do publikacji

artykułu Benjamina Becka w 1967 roku naukowcy uważali gibony za niezbyt bystre naczelne (Yerkes, Yerkes, 1929). Przekonanie to wynikało z tego, że zwierzęta te nie radziły sobie z zadaniem dotyczącym zdolności rozwiązywania problemów. Zadania polegały na podniesieniu obiektu z płaskiej powierzchni oraz użyciu go do rozwiązania problemu. Uznawano zatem, że gibony nie są zdolne do korzystania z narzędzi. Beck (1967) zauważył, że konstrukcja dłoni gibona przynajmniej utrudnia mu wykonanie zadania, w związku z czym zwierzęciu brakuje motywacji do jego wykonania. Gdy powtórzył eksperyment, umieszczając potrzebne obiekty w pozycji pionowej, gibony wykorzystywały je do przyciągnięcia pożywienia leżącego poza klatką. Mimo że gibony jako jedyne z rodziny *Pongidae* nie radziły sobie z zadaniami korzystania z narzędzi (Beck, 1967; Harlow, Uehling, Maslow, 1932), naukowcy przez lata byli skłonniejsi przyznawać, że nie są do tego zdolne, niż zastanowić się nad konstrukcją wykonywanych przez siebie eksperymentów.

W następnej części rozważań moja uwaga skupi się na dyskusji dotyczącej teorii umysłu (ang. *theory of mind*, ToM). W tym miejscu chciałam zaznaczyć, że pomimo wyników uzyskanych w eksperymentach z 2016 (Krupenye, Kano, 2016) oraz 2017 roku (Krupenye, Kano, 2017), wielu badaczy nadal odmawia ZINL omawianej zdolności. Jest to konsekwencja teorii zakładających, że u podstaw ToM leżą mechanizmy językowe.

Teoria umysłu

Teoria umysłu to zdolność, którą definiować można na dwa sposoby – szeroko i wąsko. W sensie szerokim mówi się, że jest to zdolność do atrybucji stanów mentalnych innym, natomiast w wąskim to zdolność do przewidywania zachowania na podstawie fałszywych przekonań, co wymaga rozumienia, że aktor nie kieruje się rzeczywistym stanem rzeczy, a swoim indywidualnym przekonaniem o rzeczywistości. Zatem w rozumieniu wąskim obserwator (O) ma teorię umysłu, gdy zdolny jest do tworzenia przewidywań dotyczących zachowania aktora (A) na podstawie przekonań A co do których O przynajmniej podejrzewa, że są nieprawdziwe. Rozróżnienie na teorię umysłu w rozumieniu szerokim i wąskim wprowadzili w 2011 roku dwaj badacze – Joseph Call i Michael Tomasello. Autorzy *Does chimpanzee have theory of mind. 30 years later* (2011) decydują się na wprowadzenie

takiego rozróżnienia w celu lepszego uchwycenia określonego aspektu teorii umysłu, a co za tym idzie, skuteczniejsze porównanie tej zdolności u ludzi i szympanów. W ramach rozumienia szerokiego, prezentowanego przez Premarcka i Woodruffa, ZINL posiadają ToM, ponieważ zdolne są do atrybucji intencji, motywacji oraz chęci. Call i Tomasello zaznaczają jednak, że ich zdaniem, nie ma dowodów na to, że szympanasy O potrafią przewidywać, co zrobi aktor A, gdy A ma fałszywe przekonanie o sytuacji oraz gdy O zna prawidłowe rozwiązanie problemu A.

Punktem wyjścia do rozważań zawartych w tekście Calla i Tomasello był artykuł Premarcka i Woodruffa z 1978 roku – *Does chimpanzee have theory of mind?*, w którym autorzy opisali serię eksperymentów z udziałem szympanasy Sary. W ramach przeprowadzonych badań stwierdzili, że ZINL zdolne są do wnioskowania o stanach mentalnych dotyczących motywacji, takich jak chęci, celowość i przynajmniej niektóre stany afektywne. Wyniki te nie rozstrzygają jednak problemu zdolności ZINL do przypisywania wiedzy. Trzydzieści lat później Call i Tomasello (2011) przedstawili artykuł, w którym podejmują się odpowiedzi na pytanie postawione przez Premarcka i Woodruffa. Przedstawiają oni szereg eksperymentów, na podstawie których przyznają, że istnieją dowody na to, że szympanasy zdolne są do teorii umysłu w sensie szerokim, jednak niezdolne są do tworzenia teorii umysłu w sensie wąskim. W przypadku pierwszego artykułu dostarczono dowodów na zdolność do przypisywania innym stanów mentalnych związanych z intencjami i celami oraz ujawniono wady modelu zaproponowanego eksperymentu. Okazało się, że rozwiązując problemy wyznaczone w eksperymencie, Sara odpowiadała na pytanie, jaką czynność chciałaby zobaczyć u aktora, nie zaś na pytanie, które zakładano w eksperymencie: co zrobiłby (A)? Co ciekawe, w ramach eksperymentów przeprowadzonych przez Calla i Tomasello nie udało się wyeliminować wspomnianej preferencji, a jej wystąpienie zostało uzasadnione brakiem ToM u ZINL. Wydaje się, że mniej tendencyjny wniosek z przedstawionych eksperymentów mówi, że nie można stwierdzić, czy omawiana zdolność jest elementem systemów poznawczych ZINL.

Zanim przejdę do szerszego omówienia możliwych stereotypów i ich źródeł, poświęcę nieco uwagi stanowisku Robina Dunbara (2014). Autor *Nowej historii ewolucji człowieka* twierdzi, że teoria umysłu jest zdolnością, którą ludzie dzielą z innymi zwierzętami, jednak różni ich poziom wnioskowania zachodzący w ramach tej zdolności. Wprowadzenie stopniowości omawianej zdolności opiera się na pojęciu intencjonalności. Dunbar przyjmuje,

że organizmy znajdujące się na poziomie zerowym [0] wcale nie zdają sobie sprawy z zawartości swoich umysłów. Zwierzęta na poziomie pierwszego rzędu [1] tworzą proste reprezentacje, np. wiedzą, że są głodne. W ramach proponowanej przez Dunbara koncepcji, aby uznać, że zwierze posiada ToM, musi osiągnąć poziom [2], czyli być zdolne do tworzenia przekonań o cudzych przekonaniach. W tym miejscu chciałam nadmienić, że w teorii Dunbara znaczenie terminu „intencjonalność” różni się od powszechnie przyjmowanego w psychologii oraz filozofii (Dennett, 1983). Rozważania te wychodzą jednak poza ramy tematycznie niniejszego artykułu. Żeby uniknąć nieścisłości terminologicznych, poniżej będę korzystać z określeń: „poziomy ToM” lub „rzędy ToM”. Dunbar twierdzi także, że przeciętnie ludzie zdolni są do tworzenia teorii umysłu rzędu [4]. Nie jest to jednak ostateczna granica ludzkich zdolności. Badania przeprowadzone na studentach, czyli ludziach w założeniu ponadprzeciętnie inteligentnych, wskazują, że możliwe jest tworzenie reprezentacji także rzędów [6], [7], a nawet [9]. Dla przykładu porównajmy złożoność reprezentacji czwarto- i sześciorzędowej:

IV poziom intencjonalności: Ania wie [1], że Zosia wie [2], że Marysia udaje [3], że nie wie [4] co się stało.

VI poziom intencjonalności: Peter wierzy [1], że Jane myśli [2], że Sally chce [3], by Peter przypuszczał [4], że Jane ma zamiar [5] przekonać Sally [6], że piłka jest pod poduszką.

(Dunbar, 2014)

W swojej książce Dunbar zaznacza, że ZINL nie są zdolne do tworzenia reprezentacji poziomu [2]. Przytacza testy fałszywych przekonań wykonane z udziałem szympansa i ludzkiego aktora. Zachowanie szympansov przetestowano w dwóch różnych sytuacjach. W pierwszej człowiek miał informację, w którym z paru pojemników znajduje się przysmak, w drugiej takiej informacji nie miał. W obu wariantach szympansy uczestniczące badaniu preferowały pojemnik wskazany przez człowieka. Oznacza to, że nie wykorzystywały swojej wiedzy [1] o wiedzy [2] aktora. Wydaje się jednak, że istnieje inne wyjaśnienie uzyskanych wyników. Szympansy mógł zwyczajnie nie rozumieć zadania, w wyniku czego traktował gest wskazywania człowieka jako polecenie, a nie wskazówkę informacyjną.

Dopiero w 2016 i 2017 roku przeprowadzono eksperymenty podważające stanowisko przyjmowane przez wspomnianych autorów. Zespół badawczy z Uniwersytetu w Koyto (Krupenyne i in., 2016; 2017) przeprowadził

badania z użyciem metod eye-trackingowych, sugerujące, że szympansy i bonobo są w stanie posługiwać się teorią umysłu w sensie wąskim (Krupenye i in., 2016). Dla wielu badaczy zajmujących się omawianym zagadnieniem wyniki uzyskane przez zespół z Koyto nie są wystarczające. Takie stanowisko jest konsekwencją poglądu, że dla rozwoju teorii umysłu konieczny jest język.

W celu lepszego zrozumienia mojego twierdzenia, że tendencyjne odmawianie zdolności do tworzenia teorii umysłu ZINL jest konsekwencją teorii zakładających, że u podstaw ToM leżą mechanizmy językowe, należy omówić podstawowe założenia teorii wyjaśniającej zdolność do atrybucji stanów mentalnych innym przez kodowanie i przetwarzanie informacji językowej. W założeniu język umożliwia przekształcanie abstrakcyjnych stanów umysłowych w formę symboliczną, która może być łatwo przetwarzana i przekazywana/komunikowana między użytkownikami języka (Wimmer, Perner, 1983). Umożliwia tworzenie kategorii i grupowanie różnych reprezentacji umysłowych na podstawie wspólnych cech i właściwości. Przykładowo możemy utworzyć kategorię „zwierzęta” i dołączyć różne reprezentacje umysłowe, takie jak psy, koty, ptaki itp. Język dostarcza nam nazw kategorii, które służą jako etykiety dla tych grup reprezentacji umysłowych. Dzięki temu systemy poznawcze człowieka (ponieważ aktualnie uznaje się, że język jest cechą unikatową ludzi) grupują różne stany mentalne w kategorie, te natomiast powstają na podstawie wspólnych cech i właściwości. Ważną rolę w tych procesach odgrywają mechanizmy opisane w poprzednich częściach artykułu.

W kontekście teorii umysłu należy zaznaczyć, że językowe etykiety wytworzonych kategorii pomagają nam zorganizować nasze rozumienie stanów umysłowych innych ludzi – poprzez ustrukturyzowanie ich w kategorie takie jak np. „pragnienia”, „przekonania”, „emocje”. Dodatkowo język pozwala na wyodrębnienie i identyfikowanie konkretnych reprezentacji umysłowych. Przykładowo używanie słów i zwrotów takich jak „myślę, że...”, „wiem, że...” lub „chcę...” pozwala wskazać na konkretne przekonania, wiedzę lub pragnienia innych ludzi. Dzięki temu systemy poznawcze człowieka relatywnie precyzyjnie określają i analizują różne reprezentacje umysłowe (swoje i innych). Dodatkowo wiedza, że ktoś ma pewne przekonanie lub pragnienie, może być podstawą do wnioskowania o tym, jakie działania podejmie określona jednostka. Oznacza to, że kolejną funkcją języka jest dostarczanie ram referencyjnych do przeprowadzania wnioskowań, także

z zakresu teorii umysłu. W ramach tej funkcji język umożliwia odnoszenie się do reprezentowanych stanów w sposób uporządkowany.

Opisana teoria często łączy się z twierdzeniem, że język jest zdolnością unikatową dla człowieka. Tutaj chciałam zaznaczyć, że istnieje wiele teorii na temat języka i jego roli w świecie zwierząt. Niektóre badania sugerują, że pewne formy języka lub komunikacji symbolicznej występują również u ZINL. Opisywanie ich wszystkich wykracza poza ramy niniejszego artykułu, jednak w tym miejscu chciałam skupić się na wskazanym stanowisku. Jeżeli uznamy, że język jest zdolnością unikatową dla ludzi oraz konieczny jest do wytworzenia złożonych reprezentacji ToM, należy przyjąć także przekonanie, że ZINL nie są zdolne do tworzenia teorii umysłu. Francuski antropolog kulturowy Dan Sperber (2000) twierdzi, że język jest jedynym znanym tokenem umożliwiającym przesył informacji umysłowej, tworzącej tak złożone metareprezentacje. Założenie to leży u podstaw stanowiska odmawiania istnienia ToM u ZINL.

Uważam, że opisana w pierwszej części artykułu tendencyjna interpretacja Zasady Morgana, która prowadzi do przyjęcia antroponegacionistycznej postawy wobec ZINL, skutkuje ekstrapolacją wniosków dotyczących systemów poznawczych człowieka na teorie wyjaśniające możliwości poznawcze ZINL. Błędy wynikające z heurystyki reprezentatywności (Kahneman, Tversky, 2019), a dokładniej te związane z niewłaściwą generalizacją, powodują, że człowiek odmawia zdolności tworzenia ToM innym zwierzętom. W tym miejscu chciałam zaznaczyć, że istnieje wiele innych teorii na temat języka i jego roli w świecie zwierząt. Niektóre badania sugerują, że pewne formy języka lub komunikacji symbolicznej występują również u niektórych ZINL (Pepperberg, 2000).

Uważam także, że nawet przyjmując radykalne twierdzenie, że nie ma dowodu na to, iż ZINL posługują się językiem oraz argumenty opisanej wcześniej teorii zakładającej, że zdolności językowe są konieczne do wytworzenia ToM u człowieka, nadal nie możemy jednoznacznie stwierdzić, że ZINL nie wytworzyły mechanizmu kompensacyjnego, pełniącego funkcję analogiczną lub homologiczną do języka w przetwarzaniu informacji potrzebnych do wytworzenia metareprezentacji rzędu [2].

Popularną współcześnie perspektywą staje się teoria symulacyjna, w ramach której do tworzenia złożonych metareprezentacji umysł wykorzystuje procesy symulacji (Barlassina, Gordon, 2017). Czyni to przez odtwarzanie w umyśle doświadczeń, przekonań i emocji innych osób.

Mechanizmami wskazywanymi jako podstawa tych procesów są empatia i współodczuwanie. W ramach proponowanej teorii systemy poznawcze ZINL mogą przypisywać stany mentalne innym na podstawie własnych doświadczeń. Teorie symulacyjne dostarczają wyjaśnienia dla zdolności do wnioskowania na podstawie fałszywych u ZINL.

Podsumowanie

W artykule starałam się wskazać, że szowinizm gatunkowy związany jest z konsekwencją normalnego funkcjonowania systemu poznawczego. Dokładniej ujmując, powstawanie uprzedzeń związane jest z mechanizmem kategoryzacji społecznej oraz wtórnymi konsekwencjami myślenia heurystycznego. Heurystyki, rozumiane jako uproszczone reguły wnioskowania, kształtują ludzki stosunek nie tylko do siebie nawzajem, ale także do ZINL. Sama tendencyjność jest efektem wtórnym heurystyki reprezentatywności (Kahneman, 2012), czyli uproszczonej metody wnioskowania, polegającej na ocenianiu prawdopodobieństwa przynależności do określonej kategorii na podstawie podobieństwa między obiektem a stereotypem. Jej funkcją jest budowanie modelu spójnego obrazu rzeczywistości, zaś wynikiem wtórnym skłonność do potwierdzania własnych przekonań. Przedstawiony został także mechanizm kategoryzacji społecznej, który jest niezbędny do funkcjonowania jednostki w społeczeństwie. Leży on u podstaw identyfikacji jednostki z grupą własną i określa grupę obcą oraz pomaga w ocenie własnej roli społecznej (Karyłowski i in., 1999).

Jak wspominałam w części drugiej, kategoryzacja społeczna wiąże się z faworyzacją grupy własnej, co objawia się przekonaniem o wyjątkowości człowieka. Heurystyczne funkcjonowanie systemów poznawczych prowadzi do preferencji utrzymywania poszczególnych stereotypów wynikających z przyjmowanych określonych przekonań. Jednym z przykładów może być wspomniane (w części czwartej) odmawianie ZINL zdolności do tworzenia ToM. Wydaje się, że ostrożniejszy wniosek wyciągnięty z eksperymentu Calla i Tomasello (2011) głosi, że nie znaleziono zdolności do tworzenia teorii umysłu. Należy podkreślić, że badania te nie są wolne od preferencji. Wydaje się także, że współczesne doniesienia naukowe (Krupenye i in., 2016; 2017) wskazują, że ZINL mają przynajmniej zdolność analogiczną do ToM.

Bibliografia

- Baran, T. (2011). Pomiar zjawiska infrahumanizacji „obcych” poprzez atrybucję słów typowo ludzkich i typowo zwierzęcych. *Psychologia Społeczna*, 6 (18), 202–213.
- Barlassina, L., Gordon, R. (2017). Folk Psychology as Mental Simulation. W: E.N. Zalta (red.), *The Stanford Encyclopedia of Philosophy*. Pobrano z: <https://plato.stanford.edu/archives/sum2017/entries/folkpsych-simulation/> (25.06.2023).
- Beck, B.B. (1967). A study of problem solving by gibbons. *Behaviour*, 28 (1–2), 95–109.
- Biegański, W. (1909). *Wnioskowanie z analogii*. Lwów: Wydawnictwo Polskiego Towarzystwa Filozoficznego we Lwowie.
- Brämer, R. (1993). Bambification – oder vom gemeinen Umgang mit Tieren. *Der Spiegel 1*, 2017.
- Call, J., Tomasello, M. (2011). Does the chimpanzee have a theory of mind? 30 years later. *Human Nature and Self Design*, 12 (5), 83–96. DOI: 10.1016/j.tics.2008.02.010.
- Darwin, K. (2021). *O powstawaniu gatunków*. Tłum. Ł. Lamża. Warszawa: Masterlab.
- Dennett, D.C. (1983). Intentional systems in cognitive ethology: The “Panglossian paradigm” defended. *Behavioral and Brain Sciences*, 6 (3), 343–355. DOI: 10.1017/S0140525X00016393.
- Dow, H.H. (1957). The psychological significance of animals in modern American society. *Journal of Social Issues*, 13 (3), 8–23.
- Dunbar, R.I.M. (2014). *Nowa historia ewolucji człowieka*. Tłum. B. Kucharzyk. Kraków: Copernicus Center Press.
- Ellianawati, E., Subali, B., Khotimah, S.N., Cholila, M. (2021). Profile of reasoning ability and reduction of mathematical anxiety in analogy-based physics learning. *Journal of Physics: Conference Series*, 1918 (5), 052051. DOI: 10.1088/1742-6596/1918/5/052051.
- Harlow, H.F., Uehling, H., Maslow, A.H. (1932). Comparative behavior of primates. I. Delayed reaction tests on primates from the lemur to the orang-outan. *Journal of Comparative Psychology*, 13 (3), 313. DOI: 10.1037/h0073864.
- Kahneman, D. (2012). *Pułapki myślenia: o myśleniu szybkim i wolnym*. Tłum. P. Szymczak. Poznań: Media Rodzina.

- Karyłowski, J.J., Wallace, H., Motes, M., Van Liempd, D., Eicher, S. (1999). Próba wykorzystania techniki torowania (*priming*) do badania kategoryzacji społecznej. *Przegląd Psychologiczny*, 42 (1–2), 111–120.
- Köhler, W. (1925/1959). *The Mentality of Apes*. New York: Vintage Books.
- 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), 110–114. DOI: 10.1126/science.aaf8110.
- 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), e1343771. DOI: 10.1080/19420889.2017.1343771.
- Kwiatkowska, A. (2009). Kategoryzacje społeczne skrzyżowane: o tym, czy „prawie my” znaczą samo co „my”, a „niezupełnie oni” to samo co „oni”, oraz o innych osobliwościach kategorii społecznych. W: M. Kossowska, M. Kofta (red.), *Psychologia poznania społecznego* (s. 295–310). Warszawa: PWN.
- Lutts, R.H. (1992). The trouble with Bambi: Walt Disney's Bambi and the American vision of nature. *Forest and Conservation History*, 36 (4), 160–171. DOI: 10.2307/3983677.
- Mullin, B.A., Hogg, M.A. (1998). Dimensions of subjective uncertainty in social identification and minimal intergroup discrimination. *British Journal of Social Psychology*, 37 (3), 345–365. DOI: 10.1111/j.2044-8309.1998.tb01176.x.
- Pepperberg, I.M. (2000). *The Alex studies: cognitive and communicative abilities of grey parrots*. Cambridge, MA: Harvard University Press.
- Premack, D., Woodruff, G. (1978). Does the chimpanzee have a theory of mind? *Behavioral and Brain Sciences*, 1 (4), 515–526. DOI: 10.1017/S0140525X00076512.
- Roccas, S., Brewer, M.B. (2002). Social identity complexity. *Personality and Social Psychology Review*, 6 (2), 88–106. DOI: 10.1207/S15327957PSPR0602_01.
- Schultz, D.P., Schultz, S.E., Andruszko, R. (2008). *Historia współczesnej psychologii*. Kraków: Wydawnictwo Uniwersytetu Jagiellońskiego.
- Serpell, J. (1999). *W towarzystwie zwierząt*. Tłum. A. Alichniewicz, Warszawa: PIW.
- Singer, P. (2018). *Wyzwolenie zwierząt*. Tłum. A. Alichniewicz, A. Szczęsna. Warszawa: Marginesy.
- Sperber, D. (2000). Metarepresentations in an evolutionary perspective. W: D. Sperber (red.), *Metarepresentations: A multidisciplinary perspective* (s. 117–137). Oxford: Oxford University Press. Pobrano z: <https://monoskop.org/images/9/9c/>

- Sperber_Dan_%28ed.%29_Metarepresentations_A_Multidisciplinary_perspective_2000.pdf (19.04.2024).
- Szymanek, K. (2008). *Argument z podobieństwa*. Katowice: Wydawnictwo Uniwersytetu Śląskiego.
- Tomasello, M. (2015). *Historia naturalna ludzkiego myślenia*. Tłum. B. Kucharzyk, R. Ociepa. Kraków: Copernicus Center Press.
- de Waal, F.B. (2019). *Bystre zwierzę: czy jesteśmy dość mądrzy, aby zrozumieć mądrość zwierząt?* Tłum. Ł. Lamża. Kraków: Copernicus Center Press.
- Wimmer, H., Perner, J. (1983). Beliefs about beliefs: Representation and constraining function of wrong beliefs in young children's understanding of deception. *Cognition*, 13 (1), 103–128.
- Woleński, J. (2015). Szowinizm gatunkowy, humanitaryzm i animalocentryzm. *Przegląd Filozoficzny – Nowa Seria*, 24 (2 [94]), 25–34.
- Yerkes, R.M., Yerkes, A.W. (1929). *The great apes. A study of Anthropoid Life*. New Haven: Yale University Press.
- Ziółkowski, J. (2019). *Syndrom obłądzonej twierdzy*. Warszawa: Dom Wydawniczy ELIPSA.

Nota o autorce

Janina Męcarska – doktorantka Szkoły Doktorskiej Uniwersytetu Szczecińskiego na kierunku filozofia. W ramach Interdyscyplinarnych Studiów Zaawansowanych uzyskała dwa tytuły magistra: z kognitywistyki komunikacji i socjologii. Jest aktywną członkinią zespołu badawczego Cognition and Communication Research Group działającego w Instytucie Filozofii i Kognitywistyki.

Address for correspondence: Instytut Filozofii i Kognitywistyki, Uniwersytet Szczeciński, ul. Krakowska 71–79, 71-017 Szczecin.

Cytowanie

Męcarska, J. (2024). Błędy poznawcze związane z szowinizmem gatunkowym oraz odmawianie teorii umysłu zwierzętom innym niż ludzie. *Analiza i Egzystencja*, 67 (3), 109–125. DOI: 10.18276/aie.2024.67-05.



ARTYKUŁY

PIOTR BIŁGORAJSKI

Katolicki Uniwersytet Lubelski Jana Pawła II

ORCID: 0000-0001-5139-3455

email: piotr.bilgorajski@kul.pl

Wyobrażeniowy opór*

Słowa kluczowe: wyobrażenia, wyobrażeniowy opór, percepcja, przekonania moralne, fikcja

Keywords: imagination, imaginative resistance, perception, moral beliefs, fiction

The imaginative resistance

Abstract

The phenomenon of imaginative resistance occurs when the reader of fiction is invited to imagine as morally right a situation that the reader finds morally repugnant (e.g., that murder is morally good). The problem of imaginative resistance boils down to the question of how it is possible that we can imagine situations that are inconsistent with our knowledge of the facts (magic, teleportation, spaceships moving faster than light), while we have difficulty imagining situations that are inconsistent with our moral knowledge. The purpose of this article is to present two popular concepts that explain the phenomenon of imaginative resistance, the so-called cantian response and the wontian response, and to propose an original solution. The first hypothesis states that imaginative resistance is a response to conceptual impossibility.

* Praca została sfinansowana ze środków Narodowego Centrum Nauki w ramach projektu „Eksperymenty myślowe w filozofii – geneza, struktura, funkcje”, no. UMO-2017/25/N/HS1/03019.

In this conception, resistance to imagining a particular object or situation comes from the fact that such objects or situations are contradictory and therefore impossible to imagine. The second hypothesis states that imaginative resistance occurs when the recipient does not want to imagine a certain state of affairs. In the last part of the article, I try to show that the source of imaginative resistance can be the inability to imagine oneself as someone with different moral beliefs.

Wprowadzenie

Zacznijmy od historyjki:

Jacek i Marysia znowu się pokłócili. Ponieważ kłócili się na pasie autostrady, szybko utworzył się korek. Gdy Jan zobaczył, że przyczyną spowolnienia ruchu są Jacek i Marysia, wyjął ze schowka pistolet i zastrzelił oboje na miejscu. W skutek działania Jana autostrada została odkorkowana. Jan postąpił słusznie, ponieważ Jacek i Marysia powinni byli przenieść swoją kłótnię gdzie indziej, aby nikomu nie przeszkadzać¹.

Założmy, że opisana historia, nazwijmy ją „Śmierć na autostradzie”, jest (na szczęście) fikcyjna, czyli nigdy się nie wydarzyła. Czy potrafimy ją sobie wyobrazić? Opowiadanie składa się z pięciu zdań, pierwsze cztery to zdania opisowe, natomiast ostatnie, ponieważ zawiera ocenę („Jan postąpił słusznie”), jest zdaniem wartościującym. Sytuacja opisana w pierwszych czterech zdaniach, mimo iż nigdy się nie wydarzyła, jest łatwa do wyobrażenia, wszak scena zabójstwa na autostradzie, chociaż szokująca, nie jest zupełnie niemożliwa. Natomiast usiłując wyobrazić sobie, że Jan postąpił słusznie dokonując morderstwa, nasza wyobraźnia napotyka na opór. Czytając fikcję, potrafimy wyobrazić sobie sytuacje niezgodne z naszą wiedzą dotyczącą faktów (magia, podróże w czasie), natomiast mamy względną trudność, by wyobrazić sobie sytuacje, które są niezgodne z naszą wiedzą dotyczącą wartości (np. dobre morderstwo). W literaturze trudność ta nazywana jest „wyobrażeniowym oporem” (ang. *imaginative resistance*) (Moran, 1994; Walton, 1994; Gendler, 2000, 2006).

¹ Opowieść po raz pierwszy pojawiła się w artykule Briana Weathersona (2004, s. 1); tutaj przytaczam ją w uproszczonej wersji we własnym tłumaczeniu.

Problem wyobraźniowego oporu pojawia się czasem jako zarzut wobec woluntarystycznej koncepcji wyobraźni, zgodnie z którą wyobraźnia znajduje się pod kontrolą woli. Oczywiście nie jest to kontrola absolutna, wszak nie można sobie na przykład wyobrazić przedmiotów, których opis zawiera wewnętrzną sprzeczność. Natomiast co do zasady przyjmuje się, że wyobraźnia, w odróżnieniu od percepcji lub przekonań, jest zależna od naszej woli: jeśli chcę, mogę sobie wyobrazić kota, natomiast jeśli nie ma kota w zasięgu mojego wzroku, to nie zobaczę kota, niezależnie od tego, jak bardzo bym chciał go zobaczyć. Co więc sprawia, że wyobrażenie sobie kontrfaktycznej sytuacji – gdy jej opis jest wyłącznie faktualny – przychodzi nam ze względną łatwością, zaś gdy jej opis jest wartościujący – wydaje się trudne? Odpowiedź, że nie potrafimy sobie wyobrazić sytuacji, które są opisane przy pomocy zdań wartościujących, a więc problem wyobraźniowego oporu jest pozorny, nie będzie trafna. Nie protestujemy, że to niewyobrażalne, gdy ktoś nas prosi, aby wyobrazić sobie np. ładnego psa. Nie uważamy też za niepoprawne wypowiedzi w stylu: „Byłem świadkiem złego zachowania”. A „ładny” czy „zły” to przecież nazwy wartościujące.

Jednym z pierwszych filozofów, którzy zwrócili uwagę na zjawisko wyobraźniowego oporu, był David Hume, który tak pisał w eseju *Sprawdzian smaku*:

W literaturze pięknej wszystkich krajów i epok zdarzają się filozoficzne błędy, ale niewiele to umniejsza wartości jej utworów. Trzeba tylko pewnego nastawienia naszej myśli czy wyobraźni, aby przejąć się poglądami, które wówczas panowały; i rozkoszować się wynikającymi z nich przeżyciami i wnioskami. Natomiast niezbędny jest wielki wysiłek, aby zmienić nasze poglądy na obyczaje i obudzić w nas podziw czy potępienie, miłość czy nienawiść, inne niż te, do których umysł nawykł od dawna. Kto wierzy w trafność zasad moralnych, którymi się w swych sądach kieruje, ten słusznie jest do nich przywiązany i nie będzie chciał ani na chwilę odstąpić od swych przekonań przez uprzejmość dla jakiegokolwiek autora.

(Hume, 1955, s. 213).

Hume zwraca uwagę na trudność, która pojawia się, gdy fikcja skłania nas do wyobrażenia sobie sytuacji, których nie aprobujemy. Sugeruje też przyczynę. Jego zdaniem wyobraźniowy opór wynika z niechęci, by przyjąć przekonania moralne, których się nie podziela.

Czy jednak posiadanie określonego przekonania może zależeć od chęci lub jej braku?² Celem artykułu jest prezentacja możliwych hipotez wyjaśniających zjawisko wyobrazeniowego oporu. Artykuł zaczyna się od analizy pojęcia wyobraźni. W kolejnej części omawiam główne hipotezy wyjaśniające źródła wyobrazeniowego oporu. Pierwsza stwierdza, że wyobrazeniowy opór jest reakcją na konceptualną niemożliwość. W tej koncepcji opór przed wyobrażeniem określonego przedmiotu lub sytuacji bierze się stąd, że takie przedmioty lub sytuacje są wewnętrznie sprzeczne, a więc *de facto* niemożliwe do wyobrażenia. Druga hipoteza głosi, że wyobrazeniowy opór pojawia się wtedy, gdy odbiorca nie chce wyobrazić sobie jakiegoś stanu rzeczy. W ostatniej części artykułu próbuję pokazać, że źródeł wyobrazeniowego oporu można szukać w relacji wyobrażeń i przekonań na temat nas samych.

Pojęcie wyobraźni

Oksfordzki słownik filozoficzny w haśle „wyobraźnia” lakonicznie stwierdza, że wyobraźnia polega na „zdolności do żywego przedstawiania, a zwłaszcza tworzenia w swoim umyśle obrazów mentalnych” (Blackburn, 1994, s. 441), jednak pod hasłem „obraz mentalny” można się dowiedzieć, że „natura, ważność, a nawet samo istnienie obrazów mentalnych jest przedmiotem sporu wśród filozofów i psychologów” (Blackburn, 1994, s. 265). Podobnie Roger Scruton w haśle „imagination” w *A Companion to Aesthetics* (2009)

² Cytowany fragment *Sprawdzianu smaku* sugeruje, że Hume uznawał możliwość zmiany przekonań pod wpływem woli (np. z uprzejmości). Wydaje się jednak, że pogląd Hume’a na naturę przekonań był bliższy tzw. doksytycznemu inwolontaryzmowi, zgodnie z którym przekonania są niezależne od naszej woli. W *Traktacie o naturze ludzkiej* Hume pisze wprost, że „przeświadczenie [*belief*] polega jedynie na pewnym doznaniu czy uczuciu; na czymś, co nie zależy od woli, lecz musi powstawać z pewnych określonych przyczyn, z czynników, nad którymi nie panujemy” (Hume, 2005, s. 700). Być może więc pogląd na „wyobrazeniowy opór” wyrażony w *Sprawdzianie smaku* należałoby zinterpretować tak, że – wedle Hume’a – każdy, kto jest przekonany o prawdziwości swoich zasad moralnych, nie tyle nie będzie chciał, ale raczej nie będzie mógł od nich odstąpić, nawet z uprzejmości. W tym kontekście pogląd Hume’a byłby bliższy pierwszej hipotezie, którą omawiam w tym artykule. Tamara Gendler interpretuje jednak Hume’a w duchu drugiej hipotezy.

zauważa, że w filozofii traktuje się wyobraźnię jako zdolność doświadczania (ang. *experience*) mentalnych obrazów albo jako zdolność do angażowania się w myślenie kreatywne. Scruton sądzi, że związek między tymi określeniami jest niejasny, co przede wszystkim wynika z tego, że znaczenia terminów „obraz mentalny” czy „kreatywność” zależą od teorii, w której występują (Scruton, 2009, s. 346).

Z jednej strony wydaje się oczywiste, że ludzie mają wyobrażeniowe życie wewnętrzne – potrafimy w wyobraźni przywołać sceny z naszej przeszłości, potrafimy też w wyobraźni kreować zupełnie nowe i fantastyczne scenariusze. Z drugiej jednak nie tylko nie dysponujemy jedną teorią wyobraźni, ale nawet nie ma zgody co do tego, czym właściwie wyobraźnia jest (Kind, Kung, 2016, s. 3). W literaturze można bowiem spotkać wzajemnie wykluczające się koncepcje wyobraźni. Niektórzy filozofowie traktują wyobraźnię jako nierozzerwalnie związaną z obrazami mentalnymi (Twardowski, 2011; Kind, 2001; Kung, 2010), inni dopuszczają możliwość wyobrażeń „nieoglądowych” (Chalmers, 2002; Yablo, 1993)³. Niektórzy uważają, że supozycje (myślenie hipotetyczne) jest rodzajem wyobrażenia (Goldman, 2006), inni traktują supozycje i wyobrażenia jako oddzielne stany mentalne (Gendler, 2000; Kind, 2001; Kung, 2010). A są to poglądy, które pojawiają się na gruncie filozofii. Jeśli wziąć pod uwagę badania nad wyobraźnią prowadzone w kognitywistyce czy psychologii,

³ Przeciwno traktowaniu wyobrażeń jako mentalnych obrazów szczególnie zapalczywie z behawiorystycznej perspektywy argumentował Gilbert Ryle w książce *Czym jest umysł?*. Obrony „oglądowego” podejścia do wyobrażeń przed zarzutami Ryle’a podjęła się Amy Kind w artykule *Putting the Image Back in Imagination* (2001). Zdroworozsądkowy pogląd Arystotelesa, zgodnie z którym wyobrażenia są rodzajem mentalnego obrazu, współcześnie jest zarzewiem sporu między tzw. piktorializmem (ang. *pictorialism*) i deskrypcjonalizmem (ang. *descriptonalism*). Zwolennicy pierwszego poglądu utrzymują, że wyobrażenia reprezentują określone stany rzeczy analogicznie do obrazów, zaś zwolennicy drugiego twierdzą, że wyobrażenia reprezentują określone stany rzeczy analogicznie do języka (Block, 1983). Daniel Dennett w przystępny sposób omawia ten spór w artykule *Two Approaches to Mental Images* (1981). Dennett upiera się, że pojęcie reprezentacji mentalnej jest na tyle niejasne, że powstrzymuje się od zajęcia stanowiska, zaś jego nastawienie wobec uczestników dyskusji (oraz zażartość samego sporu) o metafizyczny status mentalnych obrazów oddaje sposób, w jaki określa uczestników debaty – zwolenników istnienia mentalnych obrazów nazywa ikonofilami, zaś przeciwników ikonofobami (Dennett, 1981, s. 175).

sytuacja robi się jeszcze bardziej skomplikowana. W tym kontekście uwaga Kendalla Waltona, umieszczona w książce *Mimesis as Make-Believe*, wydaje się wciąż pasować do aktualnego stanu badań nad wyobraźnią:

Na czym polega wyobraźnia? Przejrzeliśmy całą gamę rozmaitych ujęć wyobraźni; czy nie powinniśmy teraz powiedzieć, co one mają ze sobą wspólnego? Oczywiście, jeśli ktoś potrafi. Ja nie potrafię.

(Walton, 1990, s. 19)

Nie jest moim celem argumentowanie za którymś z wymienionych tu poglądów na naturę wyobraźni. Chociaż trudno o definicję wyobraźni, to w punkcie wyjścia można przyjąć, że jest ona jedną ze zdolności umysłowych (jedną z rzeczy, które umysł może robić), natomiast wyobrażenie (reprezentacja mentalna) jako jej produkt jest stanem umysłowym o określonej treści. Amy Kind we wprowadzeniu do *The Routledge Handbook of Philosophy of Imagination* wskazuje, że nieredukcyjne traktowanie wyobrażeń jest uzasadnione z powodów historycznych. Już Arystoteles, który uważał wyobraźnię za władzę wytwarzania wewnętrznych obrazów, wprost stwierdził, że wyobraźnia, chociaż jest związana z myśleniem, różni się zarówno od myślenia jak i percepcji:

Wyobraźnia bowiem jest czymś różnym zarówno od postrzegania, jak od myślenia, chociaż nie powstaje ona niezależnie od postrzegania, jak znów bez niej nie ma mniemania. Jednak to, że wyobrażenie nie jest identyczne z mniemaniem, jest rzeczą oczywistą. Wyobrażenie bowiem jest w naszej mocy – kiedy chcemy – możemy wywoływać coś przed oczyma, jak czynią to ci, którzy ustawiają przedmioty według metody mnemotechnicznej i tworzą z nich obraz.

(Arystoteles, 1988, s. 121)

W potocznym języku naturalnie przychodzi nam mówienie o wyobrażeniach w kategoriach percepcyjnych, zwłaszcza w odniesieniu do zmysłu wzroku. Spróbuję pokazać, że porównując wyobraźnię do percepcji, można ujawnić niektóre istotne – z punktu widzenia wyjaśnienia źródeł wyobrażeniowego oporu – cechy wyobraźni.

Wyobraźnia, podobnie jak percepcja, jest intencjonalna, a więc jest „o czymś”, jest skierowana na jakiś obiekt, który w przypadku percepcji będzie (zazwyczaj) realny i aktualnie obecny dla doświadczenia zmysłowego, natomiast w przypadku wyobraźni – wyobrażony i nieobecny

aktualnie dla doświadczenia zmysłowego. Wyobraźnia, podobnie jak percepcja, może dotyczyć zarówno przedmiotu (wyobrażam sobie X, widzę X, słyszę X), jak i stanu rzeczy (wyobrażam sobie, że X, widzę, że X, słyszę, że X). Stąd wyróżnia się wyobraźnię obiektową i propozycjonalną (Kind, 2001; Stock, 2017).

Percepcja i wyobraźnia mają szczególną fenomenologię, co oznacza, że dla każdego z tych stanów mentalnych istnieje charakterystyczne odczucie bycia w tym stanie. Tym, co je odróżnia, jest stosunek do prezentowanych treści. Przedmioty, które postrzegamy za pomocą percepcji zmysłowej, sprawiają wrażenie faktycznie (realnie) istniejących, chociaż czasem ulegamy złudzeniom. Przedmioty, które postrzegamy za pomocą wyobraźni, sprawiają wrażenie możliwie istniejących i tutaj także czasami ulegamy złudzeniom.

Percepcja jest doświadczeniem zmysłowym, natomiast treść wyobrażeń pochodzi ze wspomnień przeszłych spostrzeżeń, takich jak obrazy, dźwięki czy zapachy, lub kompozycji fragmentów różnych przeszłych doświadczeń. Kazimierz Twardowski uważał, że najlepszym określeniem wyobrażenia jest „synteza wrażeń zmysłowych” (Twardowski, 2011, s. 58). Jego zdaniem jeśli synteza dotyczy „odnowienia wrażeń pierwotnych”, wtedy stanowi „wyobrażenie odtwórcze”, bliskie przypomnieniu, jeśli zaś synteza przedstawia coś nowego (np. wytwór fantazji), wtedy jest „wyobrażeniem wytwórczym” (Twardowski, 2011, s. 61). To podejście jest inspirowane poglądem Hume’a, który sądził, że wyobraźnia jest asocjacyjna. W tym ujęciu poszczególne epizody wyobraźni podlegają prawom kojarzenia: podobieństwa (np. wyobrażenie plaży może przywoływać wyobrażenie urlopu nad morzem), styczności w czasie i przestrzeni oraz związku przyczynowego. Istnienie związków skojarzeniowych może wyjaśniać, dlaczego tak łatwo przechodzimy od wyobrażenia sobie, że pewien stan rzeczy zachodzi, do bycia przekonanym, że taki stan rzeczy faktycznie występuje lub może wystąpić.

Wyobrażenia w mniejszym stopniu niż percepcja są podstawą podejmowanych bezpośrednich działań – jeśli widzę, że w moją stronę biegnie tygrys, zacznę uciekać, natomiast wyobrażenie biegnącego w moją stronę tygrysa nie wywoła takiej reakcji. O ile percepcja zmysłowa odbierana jest biernie, czyli niezależnie od naszej woli, to wyobraźnia nie tylko wymaga aktywności ze strony podmiotu, ale też pozwala na względną swobodę w tworzeniu wyobrażonych obiektów. Trudno jest bowiem siłą woli pobudzić własną percepcję i np. ujrzeć kota na macie, jeśli żaden kot nie jest fizycznie obecny. Natomiast łatwo jest sobie kota wyobrazić. Oczywiście niektóre obrazy mogą

nam się przedstawiać wbrew lub niezależnie od naszej woli. Kiedy ktoś opowiada szczególnie nieprzyjemną historię, często wyobrażamy ją sobie, mimo iż tak naprawdę nie chcemy, a czasem nawet nie możemy pozbyć się później tego obrazu z umysłu. Dlatego istnieją dobre powody, by rozróżniać dwa rodzaje wyobraźni: automatyczną i sterowaną⁴.

Wyobraźnia automatyczna jest spontaniczna i szybka. Jest to nieświadoma, niekontrolowana, pozbawiona wysiłku interakcja poznawcza z obiektami nieobecnymi aktualnie dla doświadczenia zmysłowego (Stuart, 2019, s. 1337). Usłyszawszy polecenie wyobrażenia sobie czarnego kota, zwykle wytwarzamy mentalny obraz kota, nawet jeśli nie włożyliśmy w to żadnego wysiłku umysłowego. Natomiast sterowana wyobraźnia to kontrolowana, wymagająca wysiłku i świadoma interakcja poznawcza z obiektami nieobecnymi aktualnie w doświadczeniu zmysłowym. Jest metodyczna i wymaga pewnego wysiłku ze strony tych, którzy ją stosują. W sterowanej wyobraźni możemy świadomie kontrolować przebieg wyobrażonej sytuacji, obracać wyobrażone figury geometryczne, dodawać lub odejmować szczegóły wyobrażonej sceny. Trudno jest jednak wyznaczyć wyraźną granicę między wyobrażeniami automatycznymi a sterowanymi. Czasami automatyczne, spontaniczne wyobrażenia są później poddawane świadomemu przetwarzaniu, innym razem świadomie wywołane wyobrażenia przeradzają się w spontaniczne fantazjowanie.

Wyobrażeniowy opór – dwie hipotezy

Wróćmy do fikcyjnej historii „Śmierć na autostradzie”. Jej autor zaprasza czytelnika, aby ten wyobraził sobie, że Jan postąpił słusznie, dokonując morderstwa. Takie wyobrażenie, gdyby udało się je wywołać, miałoby charakter propozycjonalny, zaś treść tego wyobrażenia – opis tej sytuacji – przynajmniej na pierwszy rzut sprawiałaby wrażenie, że taka sytuacja jest

⁴ Sama niechęć do wyobrażenia sobie czegoś nieprzyjemnego nie jest zjawiskiem zagadkowym. W podobny sposób odczuwamy często niechęć przed zobaczeniem czegoś nieprzyjemnego, np. zdjęcia przedstawiającego szczególnie drastyczną scenę. Zjawisko wyobrażeniowego oporu nie polega jednak na tym, że chcemy unikać kontaktu z treścią, np. z szokującą moralnie lub estetycznie fikcyjną historią, która może wywołać nieprzyjemne wyobrażenia. Wyobrażeniowy opór polega raczej na trudności, by wyobrazić sobie treść, którą się zna.

możliwa. Chociaż opisana sytuacja jest niezwykła, to można przypuszczać, że większość czytelników automatycznie wyobraża sobie jej początek – zdenerwowaną parę, która kłóci się na środku drogi, tworzący się korek itp. W momencie jednak, gdy w historii dochodzi do „słusznego morderstwa”, wyobraźnia napotyka na opór. Jaka jest tego przyczyna?

Tamara Gendler wskazuje, że filozofów, którzy odpowiadają na to pytanie, można przypisać do jednego z dwóch obozów: *cantian* oraz *wontian*. Nazwy obozów są grą słowną z języka angielskiego – zwolennicy podejścia „*cantian*” („niemogący”) uważają, że wyjaśnienie wyobraźniowego oporu polega na tym, że dana osoba „nie może” (ang. *can't*) sobie czegoś wyobrazić, natomiast „*wontianie*” („niechący”) sądzą, że dana osoba po prostu „nie chce” (ang. *won't*) sobie czegoś wyobrazić. W świetle tych podejść nie potrafimy sobie wyobrazić zakończenia historii „Śmierć na autostradzie” albo dlatego, że ta sytuacja jest niewyobrażalna, albo dlatego, że nie chcemy, aby coś, co jest dla nas moralnie odrażające, jawiło się jako możliwe.

Co to znaczy, że „nie mogę” sobie czegoś wyobrazić? Kendall Walton twierdzi, że niemożliwość wyobrażenia sobie sytuacji, w której np. morderstwo lub niewolnictwo jest czymś dobrym, bierze się z niezdolności do zrozumienia, co by się musiało stać, aby zasady moralne, które obowiązują w naszym świecie, były inne (Walton, 1994, s. 46). „*Cantianie*” uważają, że wyobraźniowy opór jest reakcją na konceptualną niemożliwość. Gdy autor fikcji zaprasza nas, aby wyobrazić sobie, że zabójstwo jest czymś dobrym, tak naprawdę zachęca do wyobrażenia sobie wewnętrznie sprzecznej lub niespójnej sytuacji, czyli czegoś, co jest – z naszej perspektywy – niezrozumiałe lub po prostu niemożliwe. A jeśli coś jest niemożliwe, to jest także niewyobrażalne.

Wydaje się, że taką odpowiedź przyjęliby zwolennicy moralnego absolutyzmu. W świetle tej teorii pewne czyny są dobre lub złe ze swojej natury. Absolutysta moralny powiedziałby więc, że nie może sobie wyobrazić morderstwa jako czegoś dobrego, ponieważ morderstwo jest ze swojej natury czymś złym. Propozycja wyobrażenia sobie dobrego morderstwa przypominałaby prośbę o wyobrażenie kwadratowego koła. W jednym i drugim przypadku byłoby to czynnością niemożliwą do wykonania.

Problem z tą odpowiedzią polega na tym, że literatura oraz filmy prezentują wiele historii, które zawierają opisy lub wizualizacje zdarzeń, o których wiemy, że są niemożliwe, a jednak potrafimy je sobie wyobrażać

(a nawet zobaczyć – w przypadku filmu). W powieściach o przygodach Harry’ego Pottera pojawiają się zdania typu: „Harry Potter rzucił zaklęcie *expecto patronum* w celu przywołania patronusa”, a w ekranizacjach powieści możemy nawet zobaczyć, jak Harry rzuca zaklęcie. Chociaż jesteśmy przekonani, że magia nie istnieje oraz że zmienianie rzeczywistości przy pomocy zaklęć jest niespójne z tym, co wiemy o świecie na podstawie badań naukowych, to chętnie przyznamy, że nie mamy problemu, aby to sobie wyobrazić. Większość z nas uzna też zapewne, że czytanie takich fantastycznych historii rozwija i ubogaca wyobraźnię. Odpowiedź „cantian” nie adresuje więc pierwotnego pytania: co sprawia, że wyobrażenie kontrfaktycznej sytuacji – gdy jej opis jest wyłącznie faktualny – przychodzi nam ze względną łatwością, podczas gdy wyobrażenie sobie kontrfaktycznej sytuacji – gdy jej opis jest wartościujący – wydaje się trudne?

Gendler udziela odpowiedzi z perspektywy „wontian”. Jej zdaniem wyobrazeniowy opór nie jest konsekwencją konceptualnej niemożliwości, ponieważ coś, co na pierwszy rzut oka wydaje się konceptualnie niemożliwe, może stać się wyobrażalne, jeśli tylko zostanie opisane w odpowiedni sposób. Na przykład oczywiste wydaje się, że trudno wyobrazić sobie świat, w którym 12 nie jest sumą liczb 5 i 7. W tym przypadku wyobrazeniowy opór ma podłoże w konceptualnej niemożliwości – nie potrafimy sobie wyobrazić sytuacji, w której 12 nie jest sumą liczb 5 i 7, ponieważ wiemy, że 12 jest sumą liczb 5 i 7, i jest to prawda analityczna. Gendler, aby podważyć ten pogląd, przedstawia utrzymaną w duchu opowiadań Stanisława Lema historyjkę nazwaną „Wieża Goldbacha”⁵:

Dawno temu każda liczba parzysta była sumą dwóch liczb pierwszych. Chociaż większość ludzi podejrzewała, że tak właśnie było, nikt nie był tego całkowicie pewien. Zwołano wielkie zgromadzenie, gdzie przez czterdzieści dni i czterdzieści nocy wszyscy matematycy pracowali razem, aby udowodnić tę hipotezę. Ich wysiłki nie poszły na marne i o północy czterdziestego dnia znaleziono dowód. „Hura!” – zawołali – „Odkryliśmy tajemnicę natury”.

Ale kiedy Bóg usłyszał ten pokaz arogancji, popadł w gniew. Z nieba rozległ się Grzmiący Głos: „Moje dzieci, posunęliście się za daleko. Zrozumieście zbyt wiele sekretów wszechświata. Od tego dnia dwanaście

⁵ Tytuł jest aluzją do tzw. hipotezy Goldbacha, według której, każda liczba naturalna parzysta większa od 2 jest sumą dwóch liczb pierwszych. Mimo wielu prób, prawdziwość hipotezy Goldbacha nie została jeszcze rozstrzygnięta.

nie będzie już sumą dwóch liczb pierwszych”. Słowo Boże stało się ciałem i dwanaście przestało już być sumą dwóch liczb pierwszych. Matematycy byli zrozpaczeni – wszystkie ich wysiłki poszły na marne. Zaczęli błagać Boga: „Czy jeśli znajdziemy wśród nas dwunastu sprawiedliwych, to czy odpuścisz nam nasze winy i na powrót uczynisz dwunastkę sumą dwóch liczb pierwszych?”. Bóg wyraził zgodę.

Matematycy szukali i szukali. W jednym mieście znaleźli siedmiu sprawiedliwych, w innym znaleźli pięciu. Próbowali zebrać ich razem, aby było ich dwunastu, ale ponieważ dwanaście nie było już sumą dwóch liczb pierwszych, zadanie było niewykonalne. „Panie” – zawołał – „Cóż mamy czynić? Gdybyś zniósł swoją karę, byłoby dwanaście dusz sprawiedliwych, a dopóki tego nie zrobisz, nigdy nie uzyskamy dwunastu i będziemy skazani na wieczną porażkę”.

Poruszyły Boga ich prośby i wezwał Salomona do pomocy w podjęciu decyzji. Salomon ostrożnie zważył racje. Jeśli dwanaście ponownie stałoby się sumą dwóch liczb pierwszych, to warunki, na które Bóg i matematycy się zgodzili, zostałyby spełnione. A jeśli dwanaście przestałoby być sumą dwóch liczb pierwszych, wtedy także warunki, na które Bóg i matematycy się zgodzili, zostałyby spełnione. Jak więc Salomonowe to było, żeby spełnić oba warunki jednocześnie! I tak sędzia ogłosił werdykt: od tego dnia dwanaście zarówno było, jak i nie było sumą pięciu i siedmiu. I niebiosa się cieszyły, a głosy pięciu i siedmiu dusz sprawiedliwych wzniosły się ku niebu, chór dwunastu i nie-dwunastu śpiewających w harmonijnej jedności głosił chwałę Pana. Koniec.

(Gendler, 2010, s. 191).

Zdaniem Gendler to krótkie opowiadanie pokazuje, że nawet jeśli uważamy, że w naszym świecie zdanie „12 nie jest sumą liczb 5 i 7” jest fałszywe, to należy uznać, że to samo zdanie jest prawdziwe w ramach przedstawionej fikcji. Gendler porównuje to do zabaw w udawanie (np. gdy dziecko udaje, że banan jest pistoletem) i twierdzi, że interpretując utwory literackie patrzymy na dzieło z różnych punktów widzenia, świadomie wybierając te elementy świata przedstawionego, na których chcemy się skupić, oraz te, które wolimy zignorować (dziecko, gdy udaje, że banan jest pistoletem, skupia się na wyglądzie banana, ale ignoruje jego wewnętrzną strukturę). Dlatego na pytanie, czy zdanie „12 nie jest sumą liczb 5 i 7” jest prawdziwe, odpowiedź jest przecząca, gdy oceniamy to zdanie z punktu widzenia naszego świata i znajomości obowiązujących w nim reguł. Odpowiedź będzie jednak twierdząca, jeśli ocenimy zdanie według reguł obowiązujących w ramach przedstawionej w fikcji historii. Innymi słowy,

12 nie jest sumą liczb 5 i 7, ponieważ tak twierdzi (wszechwiedzący) narrator w opowieści. Tym, co dla Gendler sprawia, że coś nam się wydaje niewyobrażalne z powodów konceptualnych jest zbyt przywiązanie do wyznaczanych przez nas teorii. „Wieża Goldbacha” ma sugerować, że skoro konceptualna niemożliwość zależy od teorii, którą się uznaje, to wraz ze zmianą teorii usuwamy też ograniczenia dla naszej wyobraźni. Stąd, zdaniem Gendler, wyobrażeniowy opór nie jest kwestią konceptualnej niemożliwości (Gendler, 2010, s. 194).

Gendler uważa, że może zachodzić podobieństwo między niemożliwością wyobrażenia sobie czegoś moralnie odstręczającego a niemożliwością do bycia przekonanym, że zdanie, o którym wie się, że jest fałszywe, jest prawdziwe. Przykładowo gdy wiem, że Ziemia jest geoidą, to nie mogę świadomie i dobrowolnie zmienić przekonania na takie, że Ziemia jest płaska. Z drugiej strony można mieć udawane przekonanie (ang. *make-belief*), że Ziemia jest płaska. Udawane przekonanie polega na przyjęciu zdania wyrażającego jakiś hipotetyczny stan rzeczy⁶. Wyobraźnia może następnie uruchomić ciąg skojarzeń, który będzie podobny do procedury kontrfaktycznego rozumowania o strukturze „jeśli *p*, to *q*”. Na przykład jeśli wyobrażę sobie, że Ziemia jest płaska, to mogę zastanawiać się dalej nad tym, jak mogłaby wyglądać granica takiej płaskiej planety. Jest to rodzaj intelektualnej zabawy typowej przy budowaniu światów w literaturze fantastycznej.

Chociaż potrafimy przyjąć zaproszenie do udawania, że zachodzi jakiś fikcyjny stan rzeczy (płaska Ziemia, Marsjanie, gadające smoki, zaklęcia itp.), to odczuwamy wyraźną niechęć przed udawaniem, że przyjmujemy odstręczające zasady moralne. Gendler sądzi więc, że źródłem wyobrażeniowego oporu jest niechęć (ang. *unwillingness*) do przyjęcia perspektywy, której się nie pochwała w naszym świecie. Jej zdaniem niechęć słabnie, gdy zastanawiamy się, jak te odmienne zasady moralne mogłyby funkcjonować w innych – całkowicie fikcyjnych – światach. Aby zilustrować swój pogląd, Gendler podaje jeszcze jeden przykład sytuacji, w której zmiana scenariusza miałaby osłabić wyobrażeniowy opór.

⁶ Wydaje się, że w języku polskim pojęcie udawanego przekonania można o wiele lepiej opisać terminem „supozycja”, który – za Władysławem Witwickim – można rozumieć jako „myśl twierdzącą lub przeczącą, ale pozbawioną momentu przekonania” (Witwicki, 1980, s. 23).

Najpierw rozważa zdanie:

1. Giselda postąpiła słusznie, zabijając swoje dziecko, wszak dziecko było dziewczynką.

Tego typu zdanie, gdyby pojawiło się w fikcji literackiej, jest zaproszeniem ze strony autora fikcji skierowanym do czytelnika, aby przyjąć pewną perspektywę. Zdaniem Gendler właściwą reakcją jest odmowa brania udziału w zabawie, w której udaje się, że zabijanie jest czymś słusznym. Zamiast tego powinniśmy uznać, że narrator jest w błędzie, ponieważ zabijanie dzieci jest zawsze czymś złym.

Gendler proponuje rozważyć inne zdanie:

2. Giselda postąpiła słusznie, zabijając swoje dziecko, wszak dziecko urodziło się 19 stycznia.

W tym przypadku zdaniem Gendler sytuacja jest mniej oczywista. Według niej gdy reguły świata przedstawionego w fikcji znacznie różnią się od tych występujących w naszym świecie, jesteśmy mniej skłonni odczuwać wyobraźniowy opór, a bardziej np. ciekawość. Z tego względu niektórzy badacze proponują redukcjonistyczne podejście do wyobraźniowego oporu, twierdząc, że zjawisko to zależy od kontekstu, w jakim przedstawi się fikcyjną historię (Stock, 2005, s. 617). Gdyby opowieść o Giseldzie uzupełnić dodatkową informacją, że w przedstawionym świecie dziewczynki czeka życie w potwornych cierpieniach, to w świetle niektórych koncepcji etycznych morderstwo mogłoby zostać zinterpretowane jako ukrócenie tego cierpienia⁷. Jednak ani takie redukcjonistyczne rozwiązanie, ani propozycja Gendler nie będą przekonujące dla kogoś, kto konsekwentnie przyjmuje absolutyzm etyczny i uznaje, że wartości moralne mają charakter obiektywny. Podobnie jak dla kogoś, kto jest przekonany o obiektywności matematyki, zdanie „ $2 + 2 = 5$ ” będzie zawsze (analitycznie) fałszywe. W takim przypadku zmiana scenariusza wciąż nie będzie miała wpływu na ocenę określonych działań. Innymi słowy, ktoś, kto uznaje swoją teorię za prawdziwą, wsruszy tylko ramionami na sugestię, że w świetle innej teorii zdanie „ $2 + 2 = 5$ ”

⁷ Wyjaśnienie odwołujące się do kontekstu krytykowane jest za pominięcie zasady „To wszystko” (ang. *that's all clause*), wedle której należy interpretować świat przedstawiony w fikcji wyłącznie w świetle opisanych faktów, bez dopowiadania informacji, które nie zostały przedstawione przez autora fikcji (Weatherson, 2004, s. 20).

może być prawdziwe – czemu mielibyśmy się przejmować wynikami teorii, która z naszej perspektywy jest fałszywa?

Co jednak w przypadku, gdy autor fikcji literackiej w pomysłowy sposób zakamufluje kontrowersyjne z moralnego punktu widzenia szczegóły opisywanej historii? Przywołajmy jeszcze raz historię z początku artykułu – jak odebralibyśmy opowieść o kłótni na autostradzie, gdyby Jan, zamiast zastrzelić Jacka i Marysię, skorzystał z urządzenia do teleportacji, które z niewytłumaczalnych powodów posiadał, i teleportował ich na poboczne drogi? Urządzenie do teleportacji oczywiście nie istnieje w rzeczywistości, ale ilustrację jego działania czytelnik może kojarzyć z takich filmów, jak np. *Star Trek*. Czy w takim scenariuszu zdanie „Jan postąpił słusznie, teleportując Jacka i Marysię” wciąż wywoływałoby opór naszej wyobraźni? Można podejrzewać, że znacznie mniejszy. Dlaczego? Powód wydaje się oczywisty: przecież teleportacja to nie zabójstwo.

Co jednak, jeśli autor fikcji w podstępny sposób ukrył przed czytelnikami istotny fakt o świecie, mianowicie że teleportacja tworzy tylko kopię teleportowanych obiektów, uśmiercając oryginały? Czy w takim przypadku uprawniony byłby wniosek, że autorowi fikcji udało się przełamać wyobraźniowy opór i skłonić czytelnika do wyobrażenia morderstwa jako czegoś słusznego? Na taki zarzut domniemany czytelnik prawdopodobnie odpowiedziałby tak: „Gdy nie wiedziałem, że teleportacja zabija, mogłem wyobrazić sobie, że postępowanie Jana było słuszne, ale teraz, gdy wiem, że teleportacja zabija, to nie potrafię sobie tego wyobrazić. Wcześniej wyobrażałem sobie zupełnie inną sytuację!”. Protest czytelnika wydaje się całkowicie uzasadniony – trudno przypisywać komuś, że wyobrażał sobie coś innego niż ten ktoś twierdzi, że sobie wyobrażał. Na pierwszy rzut oka każdy z nas jest ekspertem w odniesieniu do treści swoich wyobrażeń.

To, że nasz stosunek do historii zmienia się wraz z uzyskaną wiedzą o świecie przedstawionym w fikcji, skłania niektórych filozofów do szukania źródeł wyobraźniowego oporu w relacji wyobrażeń do przekonania. Kathleen Stock uważa, że wyobraźniowy opór bierze się z braku przekonania, że określone twierdzenie przyjmowane przez autora fikcji jest prawdziwe. W historii „Śmierć na autostradzie” takie twierdzenie ukrywa się w zdaniu: „Jan postąpił słusznie...” (*P*). Zdaniem Stock czytelnik takiej historii nie będzie w stanie wyobrazić sobie, że *P*, jeśli nie jest przekonany, że *P*.

Jej argument jest następujący:

1. Autor „Śmierci na autostradzie” chce, aby czytelnik wyobraził sobie, że *P*.
2. Aby wyobrazić sobie, że *P*, czytelnik musi mieć przekonanie, że *P*.
3. Czytelnik nie ma przekonania, że *P*.

Wniosek: Czytelnik nie może sobie wyobrazić, że *P*.

(Stock, 2017, s. 130).

W argumencie kluczowa jest druga przesłanka. Zdaniem Stock autorzy fikcji w sposób mniej lub bardziej ukryty zapraszają czytelnika nie tylko do wyobrażenia sobie czegoś, ale także do wyobrażenia sobie tego czegoś w sposób, który zobowiązuje czytelnika do przyjęcia określonych przekonań. Wyobraźniowy opór, w tej koncepcji, pojawia się wtedy, gdy nie podzielamy takich przekonań. Stock przyznaje jednak, że zasięg tej teorii ograniczony jest do kontekstów moralnych – przekonanie, że magia nie istnieje, nie blokuje naszej wyobraźni, gdy czytamy opowieści o czarodziejach. Przyjmując to rozwiązanie musielibyśmy pogodzić się z tajemniczą wyjątkowością naszych przekonań moralnych, które – w odróżnieniu od przekonań dotyczących świata – wyznaczają granice naszej wyobraźni.

Możliwe jest jednak inne wyjaśnienie. Być może mniej powinien nas dziwić wyobraźniowy opór dotyczący kwestii moralnych, a bardziej łatwość, z jaką opór ten jest przełamywany w sytuacjach kontrfaktycznych, gdy ich opis jest wyłącznie faktyczny.

Bogata i uboga wyobraźnia

Czy faktycznie potrafimy sobie wyobrazić, że Harry Potter przywołuje swojego patronusa? Weźmy prostszy przykład: czy potrafimy sobie wyobrazić, że przed nami znajduje się lewitujący, przezroczysty kawałek metalu? Załóżmy, że ktoś faktycznie ma mentalny obraz lewitującego, przezroczystego przedmiotu. Czy to wyobrażenie jest produktem bogatej czy ubogiej wyobraźni?

Rozróżnienie na ubogą i bogatą wyobraźnię zapożyczam od Petera van Inwagena, który korzystał z niego w argumentacji na rzecz modalnego sceptycyzmu (van Inwagen, 1998). Polemizował on w ten sposób ze zwolennikami tezy, że wyobrażalność jakiegoś stanu rzeczy uprawnia do

twierdzenia, że ten stan rzeczy jest możliwy. Jego zdaniem odwołanie się do wyobraźni jako narzędzia ustalania możliwości pociąga za sobą pewne zobowiązania. Jeśli ktoś formułując argument będzie chciał w przesłance przyjąć możliwość jakiegoś stanu rzeczy i uzasadni to tym, że potrafi taki stan rzeczy sobie wyobrazić, to będzie spoczywał na nim obowiązek przedstawienia dostatecznie szczegółowego opisu tej sytuacji. Van Inwagen porównuje wyobraźnię do percepcji i twierdzi, że tak jak nasza percepcja jest zawodna, gdy próbujemy dostrzec szczegóły znacznie oddalonych od nas przedmiotów, podobnie może nas zwodzić wyobraźnia, gdy próbujemy sobie wyobrazić przedmioty i sytuacje znacznie oddalone od naszych dotychczasowych doświadczeń.

To, czy posługujemy się ubogą, czy bogatą wyobraźnią, zależy od naszej wiedzy o świecie. Jeśli wiem, że przedmioty nie mogą poruszać się szybciej niż światło, nie będę mógł sobie tego wyobrazić – i będzie to świadczyć o mojej bogatej wyobraźni. Jeśli natomiast nie wiem, że przedmioty nie mogą się poruszać szybciej niż światło, będzie mi się wydawało, że potrafię to sobie wyobrazić – i będzie to świadczyć o mojej ubogiej wyobraźni. Uboga wyobraźnia polega więc na nieumiejętności wyobrażenia sobie istotnych szczegółów danej sytuacji. Jeżeli ktoś deklaruje, że potrafi sobie wyobrazić lewitujący kawałek przezroczystego metalu, to powinniśmy mieć prawo zapytać taką osobę, czy potrafi sobie także wyobrazić jaka struktura fizyczna umożliwia temu przedmiotowi lewitację oraz przezroczystość. Osoba o bogatej wyobraźni będzie wiedzieć, że taki przedmiot nie jest możliwy, ponieważ jego istnienie stałoby w sprzeczności ze znanymi prawami fizyki. Ktoś, kto deklaruje, że w swojej wyobraźni widzi taki przedmiot, chociaż nie potrafi uzgodnić jego istnienia z obowiązującymi prawami fizyki, tak naprawdę pada ofiarą złudzenia wywołanego przez swoją ubogą wyobraźnię. Zdaniem van Inwagena nikt o odpowiednio bogatej wyobraźni nie będzie w stanie tych obiektów sobie wyobrazić. Nie dlatego, że nie będzie potrafił przedstawić ich sobie przed „oczami” wyobraźni, ale dlatego, że będzie wiedział, że takie obiekty są niemożliwe. To, co w takiej sytuacji „produkuje” nasza wyobraźnia, nie jest tym, czym się wydaje. Innymi słowy, uprawnione jest, by powiedzieć, że przedstawiony w wyobraźni obiekt wygląda jak kawałek metalu lub wygląda jakby rzucał czary, ale nie wynika z tego, że jest to metal lub że naprawdę ktoś rzuca zaklęcie.

Sądzę, że pogląd Stock na związek wyobraźni i przekonania jest trafny, ale jeśli potraktujemy na serio koncepcję „bogatej wyobraźni”, to okaże się, że większość fantastycznych wydarzeń przedstawionych w fikcji – jako sprzecznych z naszymi przekonaniem – powinno wywoływać wyobraźniowy opór. Skoro natomiast nie wywołuje, to dlatego, iż wielu z nas w kontakcie z fikcją chętniej posługuje się „ubogą wyobraźnią”. Przyjmując surowe kryteria bogatej wyobraźni, należy uznać, że jeśli nie rozumiemy mechaniki świata przedstawionego w fikcji i wiemy, że jest on niespójny z naszą wiedzą o świecie, to tak naprawdę tylko wydaje nam się, że wyobrażamy sobie te fantastyczne scenariusze. Dlaczego jednak nie potrafimy lub nie chcemy w podobny sposób posłużyć się „ubogą wyobraźnią” w przypadku wyobrażenia sobie, że np. morderstwo jest czymś dobrym? Czemu, jak w zmodyfikowanej o urządzenie do teleportacji historii „Śmierć na autostradzie”, czytelnik protestuje, że nie wiedział, iż teleportacja zabija, ale nie protestuje, że w opowieści pojawiło się coś takiego jak teleportacja – a przecież technologia teleportacji fizycznych obiektów nie istnieje. Sądzę, że porównanie percepcji i wyobraźni znów może być pomocne. W przypadku zjawiska wyobraźniowego oporu jego źródła należałoby szukać w percepcji nas samych.

Wyobrażenie samego siebie

Odbiór fikcji polega na byciu zaangażowanym obserwatorem fikcyjnego świata. Im głębsza jest immersja w świat przedstawiony w fikcji, tym gwałtowniejsze są nasze reakcje na rozgrywające się w niej wydarzenia. Sytuacje, w których reagujemy strachem na opowiedzianą fikcyjną historię lub martwimy się o naszego ulubionego fikcyjnego bohatera, to przykłady zjawiska, które w literaturze filozoficznej występuje pod nazwą „paradoksu fikcji”. Paradoks ten sprowadza się do pytania: jak jest możliwe, że doświadczamy emocjonalnych reakcji na wydarzenia przedstawione w fikcji, skoro wiemy, że nie są one realne? Paradoks pojawia się przy założeniu, że emocjonalna reakcja jest racjonalna, gdy towarzyszy jej przekonanie o realnym istnieniu obiektu, który te emocje wywołuje. Jeśli ktoś boi się przedstawionego w fikcji potwora, chociaż wie, że ten potwór nie istnieje, to należałoby uznać, że reakcja tej osoby jest irracjonalna. Ponieważ jednak większość z nas reaguje emocjonalnie na fikcję, to wynikałoby z tego, że większość z nas racjonalna nie jest. A to wydaje się trudne to przyjęcia.

Jednym z popularnych rozwiązań tego problemu jest propozycja Kendalla Waltona, tzw. teoria udawania (ang. *Pretend Theory*). Walton akceptuje, jak pisze, „zdroworozsądkowe podejście”, zgodnie z którym, aby się czegoś bać, należy być przekonanym, że to coś istnieje. Zdaniem Waltona nie oznacza to jednak, że odbiorca fikcji, aby doświadczyć racjonalnych emocjonalnych reakcji na fikcję, musi najpierw wmówić sobie, że np. Harry Potter jest realną postacią. W tym podejściu to nie fikcja staje się realna, ale raczej to odbiorca fikcji – w swojej wyobraźni – staje się częścią fikcyjnego świata i dopiero z tej perspektywy – niejako zniżając się do poziomu fikcji – doświadcza emocji (Walton, 1978, s. 23).

Bycie zaangażowanym – emocjonalnie i poznawczo – obserwatorem fikcji polega więc na wyobrażeniowym umieszczeniu samego siebie w fikcyjnym świecie oraz wyobrażeniowej obserwacji rozgrywających się w tym świecie wydarzeń. Nawet jeśli nasze uczestnictwo w przygodach ulubionych bohaterów sprowadza się do zaledwie podglądania tego, co robią i myślą, to jednak śledząc ich perypetie, zazwyczaj mamy wrażenie, że jesteśmy ich towarzyszami, kibicujemy im, oceniamy ich zachowanie, zastanawiamy się, co zrobią w reakcji na jakieś zdarzenie, zaś po skończeniu szczególnie wciągającej powieści odczuwamy tęsknotę za światem, który mieliśmy okazję odwiedzić. Wydaje się, że jest to powszechne doświadczenie odbioru fikcji i dlatego można zgodzić się z Waltonem, że jest to zdroworozsądkowa koncepcja. W tym kontekście pytanie o wyobrażeniowy opór przeformułowałbym na pytanie, dlaczego łatwiej przychodzi nam wyobrażenie samego siebie jako obecnego w fikcyjnym świecie, w którym zachodzą wydarzenia niezgodne z naszą wiedzą dotyczącą faktów, natomiast trudniej jest sobie wyobrazić samego siebie w fikcyjnym świecie, w którym mamy odmienne przekonania moralne.

Wróćmy po raz ostatni do historii „Śmierć na autostradzie”. Jej autor zachęca czytelnika, by ten wyobraził sobie, że „Jan postąpił słusznie zabijając Jacka i Marysię” (*P*). W świetle teorii udawania wyobrażając sobie fikcyjną historię, w wyobraźni umieszczamy samych siebie w fikcyjnym świecie. Czytelnik „Śmierci na autostradzie” wyobrażając sobie tę historię musi więc wyobrazić sobie samego siebie jako kogoś, kto żywi przekonanie, że *P*, czyli musi wyobrazić sobie samego siebie jako kogoś, kto uważa, że morderstwo jest czymś dobrym. To jest moment, w którym pojawia się wyobrażeniowy opór. Jeśli czytelnik jest bowiem przekonany, że morderstwo jest czymś złym, to musi wyobrazić sobie samego siebie jako kogoś,

kto uważa, że morderstwo jest czymś dobrym. Jednak czy ktoś, kto uważa, że morderstwo jest złe, byłby w stanie wyobrazić sobie samego siebie jako kogoś, kto uważa, że jest inaczej? Czy uznałby kogoś takiego za możliwą wersję siebie? Jak bardzo doświadczenia życiowe czytelnika musiałyby się różnić od historii życia osoby, która sądzi, że morderstwo jest czymś słusznym? Wydaje się, że aby żywić odmienne przekonanie w tak istotnej kwestii, jak status moralny morderstwa, należałoby mieć zupełnie inną historię życia, czyli być po prostu inną osobą⁸.

Gdy uwzględnimy, że przekonania moralne są istotną częścią naszej tożsamości, to osoba, która żywi przekonanie, że *P*, w istotny sposób różni się od kogoś, kto nie żywi przekonania, że *P*. Z tego względu wyobrazenie samego siebie jako kogoś mającego inne przekonania moralne byłoby wyobrażeniem zupełnie innej osoby. Jeśli więc odbiorca fikcji nie ma przekonania, że *P*, to nie może sobie wyobrazić samego siebie jako kogoś, kto żywi takie przekonanie. Innymi słowy, odbiorca „Śmierci na autostradzie” nie może wyobrazić sobie, że *P*, ponieważ wyobrazić sobie, że *P* oznacza wyobrazić sobie samego siebie jako kogoś innego – osobę o innych przekonaniach moralnych. Koniec końców nie możemy wyobrazić sobie samych siebie jako kogoś innego – przecież zawsze jesteśmy tylko sobą.

Podsumowanie

Trzymając się rozróżnienia na bogatą i ubogą wyobraźnię, bogata wyobraźnia wyznacza zakres tego, co jest naprawdę możliwe (do wyobrażenia), natomiast uboga wyobraźnia jest podatna na złudzenia. To, czy wyobraźnia jest bogata, czy uboga, zależy od wiedzy na temat świata. Jeśli wiem, że pojęcia kwadratu i koła wyznaczają wykluczające się zakresy desygnatów, to będę też wiedział, że taki przedmiot jak kwadratowe koło jest czymś niewyobrażalnym. Jeśli natomiast ktoś twierdzi, że potrafi sobie wyobrazić kwadratowe koło, to po prostu ulega złudzeniu swojej wyobraźni spowodowanym ignorancją. To, co ten ktoś sobie wyobraża, nie jest kwadratowym kołem, ale figurą błędnie uznaną za kwadratowe koło.

⁸ Najnowsze badania z zakresu psychologii wskazują, że większość osób podziela intuicję, że zmiana przekonań w istotnych kwestiach moralnych pociągałaby zmianę tożsamości osobowej (Heiphetz, Strohminger, Young, 2017).

Taki wniosek wydaje się oczywisty w przypadku wyobrażenia kwadratowego koła, ale jest zaskakujący, gdy dotyczy wyobrażeń teleportacji lub fikcyjnych obiektów ze światów *fantasy*, o których także należy powiedzieć, że jako sprzeczne z naszą wiedzą o świecie są niewyobrażalne. Ale nawet jeśli są to obiekty niewyobrażalne, to przecież wciąż możliwa jest iluzja, że jest inaczej – czytając powieść lub oglądając film o przygodach Harry’ego Pottera, wiemy, że magia nie istnieje i wiemy, że to, co widzimy na ekranie (w wyobraźni), nie jest magią, ale wygenerowaną przez komputer (umysł) animacją. Jednak w przypadku wyobrażenia sobie samego siebie jako kogoś, kto ma inne przekonania moralne, stajemy przed trudniejszym zadaniem: wyobrażeniem sobie samego siebie jako kogoś innego. Chociaż uboga wyobraźnia ułatwia nam trwanie w iluzji, że niewyobrażalne stany rzeczy są możliwe, to w przypadku wyobrażeń na temat nas samych możemy wyobrazić sobie siebie tylko takimi, jakimi się postrzegamy. A wydaje się, że każdy, kto zastanawia się nad własną tożsamością, ma na tyle bogatą wyobraźnię, że wie, iż niemożliwe jest, aby być kimś innym niż sobą.

Akceptacja powyższej odpowiedzi na pytanie o źródła wyobrazeniowego oporu wymaga przyjęcia założenia o wyobrazeniowej immersji w świat fikcji oraz akceptacji poglądu, że przekonania moralne w sposób istotny kształtują tożsamość osobową. Nie są to jednak kontrowersyjne założenia. Praktyka zanurzania się w fantastyczne światy jest wspólnym doświadczeniem bardzo wielu czytelników, zaś wpływ przekonań moralnych na tożsamość osobową znajduje potwierdzenie empiryczne. Jeśli więc zjawisko wyobrazeniowego oporu jest realne, to wyjaśnienie, które proponuję, wspiera stanowisko „cantian” – wyobrazeniowy opór polega na niemożliwości wyobrażenia sobie samego siebie jako kogoś o odmiennych przekonaniach moralnych.

Bibliografia

- Arystoteles (1988). *O duszy*. Tłum. P. Siwek. Warszawa: Państwowy Instytut Wydawniczy.
- Blackburn, S. (1994). *Oksfordzki słownik filozoficzny*. Tłum. J. Woleński. Warszawa: Książka i Wiedza.
- Block, N. (1983). The photographic fallacy in the debate about mental imagery. *Noûs*, 17 (4), 651–662.

- Chalmers, D. (2002). Does Conceivability Entail Possibility? W: T.S. Gendler, J. Hawthorne (red.), *Conceivability and Possibility* (s. 145–200). Oxford: Clarendon Press.
- Dennett, D. (1981). Two Approaches to Mental Images. W: D. Dennett, *Brainstorms* (s. 174–189). Cambridge, MA: MIT Press.
- Gendler, T.S. (2000). The Puzzle of Imaginative Resistance. *The Journal of Philosophy*, 97 (2), 55–81.
- Gendler, T.S. (2006). Imaginative Resistance Revisited. W: S. Nichols (red.), *The Architecture of the Imagination: New Essays on Pretense, Possibility, and Fiction* (s. 149–173). New York: Oxford University Press.
- Goldman, A. (2006). *Simulating Minds: The Philosophy, Psychology, and Neuroscience of Mindreading*. Oxford: Oxford University Press.
- Heiphetz, L., Strohminger, N., Young, L.L. (2017). The role of moral beliefs, memories, and preferences in representations of identity. *Cognitive Science*, 41 (3), 744–767.
- Hume, D. (1955). Sprawdzian smaku. Tłum. T. Tatarkiewiczowa. W: D. Hume, *Eseje z dziedziny moralności i literatury* (s. 190–215). Warszawa: Państwowy Instytut Wydawniczy.
- van Inwagen, P. (1998). Modal Epistemology. *Philosophical Studies*, 92 (1), 67–84.
- Kind, A., Kung P. (2016). Introduction. The Puzzle of Imaginative Use. W: A. Kind, P. Kung (red.), *Knowledge Through Imagination* (s. 1–37). Oxford: Oxford University Press.
- Kind, A. (2001). Putting the Image Back in Imagination. *Philosophy and Phenomenological Research*, 62 (1), 85–109.
- Moran, R. (1994). The Expression of Feeling in Imagination. *The Philosophical Review*, 103 (1), 75–106.
- Ryle, G. (1970). *Czym jest umysł?* Tłum. W. Marciszewski. Warszawa: Państwowy Instytut Wydawniczy.
- Scruton, R. (2009). Imagination. W: R. Hopkins (red.), *A Companion to Aesthetics* (s. 346–351). Malden, MA: Blackwell Publishing.
- Stock, K. (2005). Resisting Imaginative Resistance. *Philosophical Quarterly*, 55 (221), 607–624.
- Stock, K. (2017). *Only Imagine: Fiction, Interpretation, and Imagination*. New York: Oxford University Press.

- Stuart, M.T. (2019). Towards a dual process epistemology of imagination. *Synthese*, 198, 1329–1350.
- Twardowski, K. (2011), Wyobrażenia i pojęcia. W: K. Twardowski, *Wyobrażenia i pojęcia oraz inne pisma filozoficzne* (s. 40–167). Warszawa: Hachette.
- Walton, K. (1990). *Mimesis as Make-Believe*. Cambridge, MA: Harvard University Press.
- Walton, K. (1994). Morals in Fiction and Fictional Morality. *Proceedings of the Aristotelian Society*, 68, 27–66.
- Weatherson, B. (2004), Morality, Fiction and Possibility. *Philosophers Imprint*, 4 (3), 1–27.
- Witwicki, W. (1980). *Wiara oświeconych*. Warszawa: Iskry.
- Yablo, S. (1993). Is Conceivability a Guide to Possibility? *Philosophy and Phenomenological Research*, 53 (1), 1–42.

Nota o autorze

Piotr Biłgorajski – doktor filozofii, pracownik Katedry Metodologii Nauk w Instytucie Filozofii Katolickiego Uniwersytetu Lubelskiego Jana Pawła II.

Address for correspondence: Instytut Filozofii, Wydział Filozofii, Katolicki Uniwersytet Lubelski Jana Pawła II, Al. Raławickie 14, 20-950 Lublin.

Cytowanie

Biłgorajski, P. (2024). Wyobrażeniowy opór. *Analiza i Egzystencja*, 67 (3), 127–148. DOI: 10.18276/aie.2024.67-06.

