

MICHAŁ PIEKARSKI*

EFEKT KNOBE'A, NORMATYWNOŚĆ I RACJE DZIAŁANIA**

Abstract

KNOBE EFFECT, NORMATIVITY, AND REASONS FOR ACTION

The article discusses Joshua Knobe's problem of asymmetry in ascribing intentionality to actions. Knobe claims that, in a way, the asymmetry is related to the human tendency for blaming the agent for negative side effects of her actions rather than praising her for the positive ones. He believes that people are more likely to link intentional actions to side effects when they consider a side effect to be bad than when they regard it as good. The aim of the article is to discuss the problem put forward by Knobe. I try to demonstrate that the asymmetry he identified is founded in a more basic asymmetry between what is normative and what is empirical. In my view, by referring to the concept of recognizing normativity suggested in the article, we can look at the problem raised by Knobe from a totally new perspective. I formulate a research hypothesis whereby knowledge about normative relations is more primary and basic than knowledge about agency and self-knowledge of the subject.

Keywords: Knobe effect, experimental philosophy, intentionality, intention, intentional action, normativity, prediction, rule, theory of predictive coding

W szeroko komentowanym artykule *Intentional Action and Side Effects in Ordinary Language* Joshua Knobe zwrócił uwagę na asymetrię w przypisywaniu intencjonalności działaniom, które są skutkami ubocznymi innych

* Instytut Filozofii, Wydział Filozofii Chrześcijańskiej, Uniwersytet Kardynała Stefana Wyszyńskiego, ul. Wójcickiego 1/3, 01-938 Warszawa, m.piekarski@uksw.edu.pl.

** Pierwsza wersja artykułu została wygłoszona w czasie X Polskiego Zjazdu Filozoficznego w Poznaniu. Pragnę podziękować dr. Andrzejowi Waleszczyńskiemu za wiele pouczających dyskusji dotyczących problemów, które tu poruszam, a także studentom Instytutu Filozofii UKSW, którzy brali udział w moim translatorium „Perception and Action” za wspólną lekturę i analizę wybranych tekstów z filozofii eksperymentalnej. Dziękuję również anonimowemu recenzentowi *Filozofii Nauki* za wszelkie uwagi, które pomogły udoskonalić i poprawić ten tekst, a także dr Katarzynie Kuś za korektę artykułu.

działań¹. Eksperyment przeprowadził na siedemdziesięciu ośmiu osobach, którym przedstawiono jedną z dwóch historii: jedną o uszkodzeniu, a drugą o pomaganiu (Knobe 2003a: 191):

Szkodzenie. Wicedyrektor firmy przyszedł do dyrektora zarządu i powiedział: „Myślimy o wdrożeniu nowego programu. Pozwoli nam zwiększyć zyski, ale zaszkodzi środowisku”. Dyrektor odpowiedział: „Nie interesuje mnie to, że zaszkodzimy środowisku. Najważniejsze są dla nas zyski. Wdrażamy program!”. Program został wdrożony, a środowisko skażone.

Pomaganie. Wicedyrektor firmy przyszedł do dyrektora zarządu i powiedział: „Myślimy o wdrożeniu nowego programu. Pozwoli nam zwiększyć zyski i pomoże środowisku”. Dyrektor odpowiedział: „Nie interesuje mnie to, że pomożemy środowisku. Najważniejsze są dla nas zyski. Wdrażamy program!”. Program został wdrożony i pomógł środowisku.

Przeważająca większość osób (82%), które otrzymały pierwszą historyjkę, uznała, że dyrektor spowodował skutek uboczny (zaszkodzenie środowisku) intencjonalnie. Osoby, które przeczytały drugą historię, w większości (77%) stwierdziły, że dyrektor nie wywołał intencjonalnie skutku ubocznego (pomoc środowisku)². Badani odpowiadali również na pytanie o wielkość winy i zasługi dyrektora. Wyniki wskazują, że w wypadku uszkodzenia środowisku przypisywali mu znacznie więcej winy ($M = 4,8$) niż zasługi w wypadku pomagania środowisku ($M = 1,4$). Przypisywanie winy i zasługi silnie korelowało z przypisywaniem przez nich intencjonalności działaniu dyrektora.

Knobe’a zainteresowało źródło asymetrii w przypisywaniu intencjonalności działaniom podmiotów: dlaczego ludzie są skłonni w pewnych przypadkach przypisywać intencjonalność działaniom stanowiącym skutki uboczne (takim jak uszkodzenie środowisku), a w innych nie? Jego zdaniem istnieje istotny związek między przypisywaniem intencjonalności a przypisywaniem odpowiedzialności za skutki uboczne działań. Podsumowując artykuł, stwierdził:

Wygląda na to, że asymetria opiera się na skłonności ludzi do przypisywania działającemu winy raczej za złe skutki uboczne niż chwaleniu za dobre. Ta asymetria może być źródłem odpowiadającej jej asymetrii co do tego, jak ludzie stosują pojęcie „intencjonalny”: wydają się bardziej skłonni mówić, że skutki uboczne są wywołane intencjonalnie wtedy, gdy uważają je za złe, niż wtedy, gdy oceniają je jako dobre (Knobe 2003a: 193).

¹ Działanie odróżniam tutaj od zachowania. Działanie jest wykonywane z pewną intencją, zachowanie ma zaś charakter automatyczny. Czym innym jest wstrzymanie oddechu w celu zanurkowania, a czym innym sam proces oddychania.

² Knobe (2003: 191) nie był pewien, czy można uznać wyniki pierwszego eksperymentu za wiążące. Mógł na nie wpływać stereotypowy sposób myślenia ludzi o wielkich korporacjach i zanieczyszczaniu środowiska. Zaproponował więc inny eksperyment, w którym przedstawił historię z zakresu działań wojennych. Wyniki drugiego eksperymentu były zbieżne z tymi z pierwszego.

Celem moich rozważań będzie wyjaśnienie wskazanej przez Knobe'a asymetrii dotyczącej przypisywania winy i zasługi oraz jej związków z asymetrią w przypisywaniu intencjonalności. W części pierwszej ogólnie przedstawię projekt filozofii eksperymentalnej oraz jej podstawowe założenia. Następnie omówię istotne dla analiz Knobe'a pojęcie działania intencjonalnego oraz jego pragmatykę. Pokażę, że sama analiza tego pojęcia nie pozwala rozwiązać wskazanego przez Knobe'a problemu. W części czwartej wykazę, opierając się na analizach Hindriksa, że ludzkie przekonania normatywne w sposób ścisły wiążą się ze skłonnością do przypisywania osobom odpowiedzialności za skutki uboczne ich działań. Kolejną część poświęcę pytaniu, czy obie asymetrie wskazane przez Knobe'a można wyjaśnić przez odwołanie się do zależności zachodzących między przekonaniem a pragnieniami działającego. W części szóstej będę dowodzić, że asymetrie wskazane przez autora *Intentional Action and Side Effects* wywodzą się z bardziej podstawowej asymetrii dotyczącej możliwości odróżnienia tego, co normatywne, od tego, co empiryczne. Na koniec wskażę możliwość zastosowania w analizie efektu Knobe'a teorii kodowania predykcyjnego (TKP), której jednym z podstawowych pojęć jest normatywność. Jeżeli skorzysta się z TKP, to można wykazać, że pierwotne intuicje normatywne pojawiają się już na poziomie podejmowania i wykonywania działań oraz że są ściśle zależne od przeprowadzanych przez podmiot predykcji regulujących jego działania i zachowania. Jest to poziom, który nie opiera się na samowiedzy oraz wiedzy o sprawstwie, zasadniczych dla analizowanych przez filozofię eksperymentalną intuicji moralnych.

1. FILOZOFIA EKSPERYMENTALNA I JEJ ZAŁOŻENIA

Filozofię eksperymentalną odróżniają od innych działów filozofii przede wszystkim dwa przekonania: (1) można, a nawet powinno się badać intuicje filozoficzne za pomocą metod eksperymentalnych³, (2) kwestią empiryczną

³ Nie ma zgody co do tego, w jaki sposób należy rozumieć pojęcie intuicji filozoficznych. Niektórzy charakteryzują je ze względu na ich rolę epistemiczną, inni wiążą je z pochodzeniem albo z ich treścią (Ludwig 2010: 431-434). Ludwig (2010: 430) twierdzi, że niezależnie od charakterystyki należy je odróżnić od tzw. intuicji empirycznych obecnych w przyrodoznawstwie. Część zwolenników filozofii eksperymentalnej (m.in. Knobe) wybiera jednak charakterystykę funkcjonalną, która wiąże intuicje z dowolnymi i spontanicznymi reakcjami na eksperymenty myślowe (Ludwig 2010: 435). W tym artykule, mówiąc o intuicjach, będę miał na myśli takie ich rozumienie. W sprawie funkcji, jakie pełnią intuicje w ramach filozofii eksperymentalnej, por. Sosa 2007.

jest to, czy któreś, a jeśli tak, to które z intuicji filozoficznych mają charakter uniwersalny (Nadelhoffer, Nahmias 2007: 125; Sosa 2007)⁴.

Thomas Nadelhoffer i Eddy Nahmias (2007), opierając się na dotychczasowej literaturze, wyróżnili trzy główne projekty w obrębie filozofii eksperymentalnej: eksperymentalną analizę pojęciową (*Experimental Analysis*), eksperymentalny deskrytywizm (*Experimental Descriptivism*) oraz eksperymentalny restrykcyjizm (*Experimental Restrictionism*). Prace Knobe'a należą do drugiego z tych nurtów, ale nie brak w nich pewnych elementów wiążących jego badania z nurtem analitycznym⁵. Eksperymentalny deskrytywizm ma na celu ujawnienie mechanizmów kształtujących potoczne intuicje (*folk intuitions*). Posługuje się w tym celu badaniami empirycznymi, w których uczestnicy konfrontowani są ze scenariuszami i ankietami opartymi na eksperymentach myślowych⁶. Analizy z zakresu filozofii eksperymentalnej odróżnia od klasycznych analiz filozoficznych przekonanie, że postęp badawczy mierzy się nie tyle przez odniesienie do precyzji opisu wzorca charakteryzującego ludzkie intuicje, ile przez stopień wyjaśnienia tych intuicji (Knobe, Nichols 2008a: 5).

Podstawą tego stanowiska jest założenie, że na drodze badań empirycznych można dociekać, jakie procesy psychiczne leżą u podstaw intuicji filozoficznych (Knobe, Nichols 2008a: 3)⁷. Filozofowie eksperymentalni podkreślają odrębność swoich badań od analizy pojęciowej, której celem miałyby być głównie

⁴ Przedstawiciele tzw. restrykcyjnego nurtu filozofii eksperymentalnej (Nichols, Stich, Weinberg 2003; Machery, Mallon, Nichols, Stich 2004) podkreślają, że przynajmniej niektóre intuicje podzielane przez przedstawicieli świata zachodniego nie odpowiadają intuicjom badanych osób pochodzących z kultury wschodnioazjatyckiej. Badania międzykulturowe pozwalają zakwestionować podzielany przez wielu filozofów pogląd, że intuicje filozoficzne mają charakter powszechny (por. także Ludwig 2010: 436).

⁵ Wspólne dla dociekań Knobe'a i nurtu analitycznego jest zainteresowanie analizą intuicji zwykłych ludzi oraz ich odniesienie do debat filozoficznych (Wysocki 2011: 34). Nie ma tutaj miejsca, aby szczegółowo omawiać wszystkie odłamy filozofii eksperymentalnej. W sprawie ich charakterystyki por. Nadelhoffer, Nahmias 2007 oraz Komorowska-Mach 2013 (autorka omawia pozytywny i negatywny program filozofii eksperymentalnej).

⁶ Pouczające w tym kontekście są przeprowadzone przez Joshuę Greena i współpracowników badania neurologiczne (za pomocą fMRI) dotyczące tego, co się dzieje w mózgach osób, które rozważają znane eksperymenty myślowe z zakresu filozofii moralnej. Badania te doprowadziły do postawienia tezy, że różne scenariusze wydarzeń nie tylko mogą wywoływać odmienne intuicje, lecz także mogą aktywować różne procesy poznawcze (Green, Haidt 2002).

⁷ Por. zbiór artykułów wydany pod redakcją Knobe'a i Nicholasa (2008b). W Polsce na temat filozofii eksperymentalnej pisali m.in.: Dębska (2013), Komorowska-Mach (2013), Kuś, Maćkiewicz (2016), Paprzycka (2014, 2016), Wysocki (2011). Ostatnio ukazał się numer *Filozofii Nauki* (4/2016) zawierający zbiór artykułów omawiający różne aspekty problemu sformułowanego przez Knobe'a.

określanie warunków zastosowania odpowiednich pojęć (Knobe, Nichols 2008a: 5). Knobe w jednej z ostatnich prac (2016) podkreśla, że ostatecznie filozofia eksperymentalna ma na celu odkrycie konkretnych mechanizmów poznawczych, które satysfakcjonująco wyjaśnią odpowiednie procesy psychiczne⁸. W tym znaczeniu filozofia eksperymentalna jest częścią nauk kognitywnych.

Podstawowe znaczenie ma tu przekonanie o wartości eksplanacyjnej pojęć z zakresu tzw. psychologii potocznej (PP) (*Folk Psychology*). Jednym z nich jest pojęcie działania intencjonalnego. Filozofia eksperymentalna nie odrzuca PP, traktując ją jako zdroworozsądkowy sposób rozumienia zjawisk psychicznych czy też zdolność do przypisywania innym przekonań, pragnień, intencji czy emocji. Knobe rozumie ją szerzej: nie sprowadza jej do narzędzia pozwalającego przewidywać i wyjaśniać zachowania ludzkie, lecz uznaje, że jest też narzędziem ułatwiającym wydawanie sądów moralnych (Knobe 2004: 270, Knobe, Nichols 2008b: 127).

Filozofia eksperymentalna odwołuje się w swoich analizach do języka oraz metod psychologii eksperymentalnej i społecznej⁹. Wiąże się to z przyjmowaną *implicite* przesłanką, że podzielane przez większość ludzi intuicje moralne mogą być zasadne.

Z tak zarysowanego obrazu filozofii eksperymentalnej wynikają istotne dla naszego dalszego wywodu założenia:

(1) aby zrozumieć, jak ludzie oceniają pewne sytuacje oraz ich skutki uboczne, konieczne jest przeanalizowanie mechanizmów poznawczych stojących za tymi ocenami;

(2) filozoficzne pojęcie działania intencjonalnego powinno ściśle wiązać się ze zdroworozsądkowym pojęciem działania intencjonalnego, czyli takiego, które jest częścią psychologii potocznej;

(3) częścią psychologii potocznej są również kategorie moralne: odpowiedzialność, przypisywanie winy czy przypisywanie zasługi;

(4) odwołując się do psychologii potocznej, możemy wyjaśnić asymetrię ludzkich ocen skutków ubocznych działań (Knobe 2003a, b, 2008: 131-132)¹⁰.

⁸ Nie jest jednak jasne, o jaki rodzaj wyjaśnień ma chodzić (np. mechanistyczne czy dynamiczne).

⁹ Katarzyna Paprzycka (2003) zwraca uwagę na konieczność korzystania ze zdobyczy psychologii społecznej w kontekście dyskusji nad statusem PP.

¹⁰ Knobe (2010) zaproponował takie wyjaśnienie zaobserwowanego efektu za pomocą konkretnego mechanizmu psychologicznego. Nie ma tutaj miejsca, by przeanalizować je szczegółowo.

2. POJĘCIE DZIAŁANIA INTENCJONALNEGO

Ludzie odróżniają zachowania, które są wykonywane z intencją, od takich, którym intencja nie towarzyszy. W języku potocznym rzadko mówi się o działaniu intencjonalnym¹¹. Raczej stosuje się zwroty takie, jak „zrobić coś celowo”, „mieć zamiar” albo „działać świadomie/umyślnie”. Pojęcia zamiaru, celowości i świadomości w codziennym użyciu wiążą się z pojęciem intencji. Nie są one jednak z nim synonimiczne. Obrazuje to przykład snajpera, który chce zabić oficera, lecz ma świadomość, że oddając strzał, zdradzi swoją pozycję. Można powiedzieć, że snajper „nie ma zamiaru zaalarmować wroga” (Harman 1976: 433), lecz świadomie (*intentionally*) to robi. Jeżeli zamiast wyrazów „zamiar” oraz „świadomie” wstawimy słowo „intencja”, cała historia traci sens.

Skutkiem ubocznym oddania przez snajpera strzału jest zdradzenie swojej pozycji, niezależnie od tego, czy oficer zostanie trafiony. Snajper nie miał intencji czy zamiaru powiadomienia wroga o swojej pozycji, ale z całą pewnością jego działanie zabicia wroga było intencjonalne.

Wyobraźmy sobie teraz, że snajper, celując w oficera (z intencją zabicia go), nie ma intencji wywołania alarmu, przypadkiem jednak pociąga za spust i nieintencjonalnie zabija oficera, wywołując jednocześnie alarm (pierwszy skutek uboczny). Jak byśmy teraz ocenili działania snajpera? Skoro wystrzelił przypadkiem, to mimo że zamierzał wystrzelić, nie powiemy, że naciśnięcie spustu było działaniem intencjonalnym, tak jak nie jest działaniem intencjonalnym drżenie rąk w sytuacji stresu. Czy w takim razie możemy również uznać, że snajper jest odpowiedzialny za wywołanie alarmu? Jeżeli czyn snaj-

¹¹ Paprzycka pisze, że terminy „intencjonalny/a/e” występują od niedawna w polskiej (filozoficznej) literaturze fachowej i wiążą się z pracami, w których „autorzy opierają się na literaturze anglosaskiej”. Są to więc słowa słabo „zadomowione” w języku polskim. Częściej używa się w języku polskim takich terminów, jak „świadomie”, „celowo”, „umyślnie” czy „specjalnie” (Paprzycka 2014: 75; por. Kuś, Maćkiewicz 2016). Warto jednak zwrócić uwagę na fakt, że w polskiej filozofii same słowa „intencjonalny” czy „intencjonalność” funkcjonują już od prawie stu lat. Pojawiły się wraz z wczesną recepcją fenomenologii. Najlepszym tego potwierdzeniem jest pochodząca z 1919 roku praca Romana Ingardena pt. *Dążenia fenomenologów* (Ingarden 1963). W niej to autor niejednokrotnie pisał o „intencjonalności świadomości” czy „intencjonalnych aktach”. Zaproponował również ciekawe tłumaczenie niemieckiego *intentionale* jako *intencyjny*. Być może angielskie *intentional action* powinno się właśnie tłumaczyć jako *działanie intencyjne*, w nawiązaniu do znanego z języka potocznego zwrotu „robić coś z intencją”. Odpowiedź na to pytanie znacznie wykracza poza ramy przyjętych tutaj rozważań. Moja uwaga ma charakter przede wszystkim terminologiczny. Należy pamiętać o tym, że pojęcie intencji używane w analitycznej filozofii działania jest różne od fenomenologicznego pojęcia intencjonalności.

pera – zabicie oficera – ocenimy jako zły moralnie, to według Knobe'a będziemy skłonni przypisać mu odpowiedzialność za jego dokonanie oraz za skutki uboczne jego działania. W pewnych przypadkach przypisanie podmiotowi odpowiedzialności za skutki uboczne jego działań będzie prowadzić też do przypisania intencjonalności jego działaniom. Uznajemy, że snajper jest odpowiedzialny za drugi skutek uboczny nieumyślnego oddania strzału – zabójstwo oficera. Snajper, mimo że nieintencjonalnie zabił oficera, ze względu na sam czyn i jego okoliczności będzie traktowany tak, jakby zabił intencjonalnie¹², ponieważ przypisanie odpowiedzialności i intencjonalności snajperowi wiąże się z przekonaniem, że zabijanie jest złe (Harman 1976: 434). Czy w takim razie fakt, że naciśnięcie spustu było działaniem intencjonalnym lub intencji było pozbawione, ma jakieś znaczenie dla przypisania snajperowi odpowiedzialności za wywołanie alarmu i śmierć oficera? Raczej nie. Przykład snajpera pokazuje, że istnieją sytuacje, w których przypisanie odpowiedzialności pewnemu podmiotowi nie wiąże się z przypisywaniem jego konkretnym działaniom intencjonalności¹³.

Pytanie, do którego skłania nas Knobe, brzmi jednak odwrotnie: czy w pewnych przypadkach przypisanie intencjonalności danemu działaniu jest oparte na przypisaniu odpowiedzialności podmiotowi za skutki uboczne tego działania? Jeżeli tak, to dlaczego łatwiej przypisuje się komuś odpowiedzialność za skutki, które uważa się za złe, niż zasługę za skutki uznawane za dobre?

Z rozważań tych wynika, że tzw. potoczne pojęcie działania intencjonalnego (1) nie odpowiada technicznemu pojęciu, którym posługują się filozofowie, oraz (2) nie sprowadza się tylko do pojęć intencji i działania. Działanie intencjonalne wiąże się bowiem również z tym, co uważamy za jego skutki i jak je oceniamy¹⁴.

¹² Według Harmana można również stwierdzić, że snajper „miał szczęście” i dlatego trafił oficera (Harman 1976: 435). Nie wydaje się jednak, aby mogło to być argumentem za jego uniewinnieniem. Wszystkie okoliczności wskazują bowiem, że snajper zamierzał zabić oficera.

¹³ Na temat przypisywania odpowiedzialności za działania nieintencjonalne por. Bratman 1987, Harman 1976, Mele 1992.

¹⁴ Malle i Knobe (1997) za składowe pojęcia działania intencjonalnego uznali pragnienie uzyskania pewnego rezultatu, przekonanie, że dane działanie prowadzi do niego, intencję wykonania danego działania, umiejętność pozwalającą wykonać dane działanie oraz świadomość realizacji danej intencji w czasie wykonywania odpowiedniego działania (Malle, Knobe 1997: 112).

3. PRAGMATYKA POJĘCIA DZIAŁANIA INTENCJONALNEGO

Fred Adams i Annie Steadman podkreślają, że ogólnie rzecz biorąc, istnieją dwie koncepcje pojęcia działania intencjonalnego. Pierwsze (P1) zakłada, że podmiot *S* intencjonalnie wykonuje działanie *A* tylko, gdy *S* zamierza zrobić *A*. Posiadanie intencji jest zatem warunkiem koniecznym działania (intencjonalnego)¹⁵. Zgodnie z drugim (P2) podmiot *S* intencjonalnie wykonuje działanie *A*, nie zamierzając zrobić *A*, o ile działanie *A* jest przewidziane przez podmiot *S* oraz zaakceptowane jako konsekwencja działania *S* (Adams, Steadman 2004a: 173). W wypadku P2 działanie może więc być intencjonalne, mimo że podmiot nie ma intencji wykonania go¹⁶. Zwolennicy P2 utrzymują, że ich koncepcja odnosi się tylko do niektórych ludzkich działań, np. takich jak opisany przez Harmana przypadek snajpera.

Przyjmując P1, nie jesteśmy w stanie wyjaśnić asymetrii w przypisywaniu intencjonalności działaniom. Adams i Steadman (2004a: 176) pokazują jednak, że odrzucenie P1 może być przedwczesne. Można bowiem wskazać dwie interpretacje wyników uzyskanych przez Knobe'a. Pierwsza zakłada, że przeprowadzone przez niego badania umożliwiają badanym pełne wyartykułowanie potocznego pojęcia działania intencjonalnego. Druga interpretacja opiera się na stwierdzeniu, że Knobe odsłania raczej pragmatyczne własności języka intencjonalnego, czyli takie cechy, które należą do społecznego kontekstu użycia danego wyrażenia, a nie do jego treści. Wyniki eksperymentu Knobe'a nie mówią zatem nic na temat semantyki pojęcia działania intencjonalnego, a jedynie wskazują pewne zależności pragmatyczne (Adams, Steadman 2004a, 176-177, 179).

Adams i Steadman twierdzą oprócz tego, że pojęcie winy nie należy do zawartości semantycznej pojęć oznaczających działanie intencjonalne. Nie ma więc potrzeby konstruowania pojęcia działania intencjonalnego tak, jak postulują to niektórzy zwolennicy filozofii eksperymentalnej (w szczególności Knobe). Powinniśmy natomiast zrozumieć pragmatykę języka intencjonalnego. Oddzielenie własności semantycznych od pragmatycznych w języku intencjonalnym pozwala obronić zasadność P1: pojęcia winy i zasługi nie należą do składowych pojęcia działania intencjonalnego¹⁷, lecz do jego użycia, a zatem to ludzie, formułując sądy, wiążą działania z winą lub zasługą.

¹⁵ Jest to pogląd podzielany m.in. przez Adamsa (1986) i McCanna (1986).

¹⁶ Zwolennikami tego poglądu są m.in. Bratman (1987), Harman (1976) i Mele (1992).

¹⁷ Warunki prawdziwości zdania „*S* wykonał działanie *A* intencjonalnie” nie zawierają pojęć zasługi i winy (Adams, Steadman 2004a: 178).

Można teraz zapytać, skąd się bierze wskazana przez Knobe'a asymetria w przypisywaniu intencjonalności pewnym działaniom. Jeżeli, jak sugeruje Knobe, można ją wiązać w jakimś sensie ze „skłonnością ludzi do przypisywania działającemu winy raczej za złe skutki uboczne niż do chwaleń za dobre” (Knobe 2003a: 193), to czy kwestia przypisywania danemu działaniu określeń moralnych jest związana tylko z kontekstem społeczno-pragmatycznym? Jeżeli tak, to w jaki sposób wytłumaczyć fakt, że ludzie w pewnych sytuacjach przypisują komuś winę, a w innych nie chcą uznać czyjeś zasługi? Wypowiedzi, które przypisuje się dyrektorowi w formularzach zysku (np.: „Nie interesuje mnie, że pomoże to środowisku. Najważniejsze są dla nas zyski”), brzmią nienaturalnie. Tak się nie mówi. Wypełniający formularze z pewnością zastanawiali się, dlaczego dyrektor nie zwrócił żadnej uwagi na pozytywne skutki swoich decyzji. Ową obojętność poczytano mu za coś negatywnego (Adams, Steadman 2004a: 179) i dlatego nie zasłużył według badanych na pochwałę. Jest to jednak problem użycia zwrotów intencjonalnych, a nie samych działań¹⁸.

Podjęciu Adamsa i Steadman można zarzucić, że podkreślenie przez nich roli pragmatycznych składników mowy nie wyjaśnia asymetrii Knobe'a. Niezależnie bowiem od tego, czy uznamy, że ocena moralna działań wiąże się z semantyką pojęcia działania intencjonalnego, czy raczej z jego użyciem, to ludzie wydają takie, a nie inne sądy na temat działań. Można zgodzić się z uwagami dotyczącymi pewnej nienaturalności wypowiedzi dyrektora, nie wydaje się jednak, by osoby wypełniające formularz w sytuacji, w której wypowiedzi te byłyby inaczej sformułowane, oceniły je w inny sposób. Niezgoda na propozycje Adamsa i Steadman nie wiąże się z koniecznością odrzucenia P1: najsilniejszym argumentem za P1 jest słabość P2. Istnieją działania, których nie możemy przewidzieć ani zaakceptować jako konsekwencji działania podmiotu, ale które mimo to uznamy za intencjonalne. Przykładem takiego działania może być sytuacja dyrektora, który nie wiedziałby o tym, że jeżeli wprowadzi nowy program, to zaszkodzi środowisku. Czy niewiedza w tym zakresie zwalnia dyrektora z odpowiedzialności za skutki jego działań? Wydaje się, że bez odpowiednich badań empirycznych trudno na to pytanie dać przekonującą odpowiedź¹⁹.

¹⁸ Komentując odpowiedź Knobe'a, Adams i Steadman uznali, że ludzie nie mają żadnego jasnego pojęcia działania intencjonalnego, lecz kierują się intuicją, zgodnie z którą istnieje bliski związek między winą i zasługą a tym, że ktoś robi coś celowo. Związek ten opiera się na pragmatyce zwrotów intencjonalnych (Adams, Steadman 2004a: 275-276; 2004b).

¹⁹ Jego niewiedza może wynikać albo z braku zainteresowania tym problemem, albo z zaniedbania, co w tym przypadku sprowadza się do tego samego. Wydaje się, że zaniedbanie w pewnych sytuacjach jest aktem intencjonalnym polegającym na nieuznaniu (w sensie „niezwrócenia uwagi na coś”) konsekwencji własnych działań i jako takie może być brane pod uwagę w ocenie działania danej osoby (wyłączamy tutaj sytuację zaniedba-

4. PRZYPISYWANIE WINY I ODPOWIEDZIALNOŚCI

Wróćmy do wskazanych przez Adamsa i Steadman dwóch sposobów rozumienia działania intencjonalnego. Czy rzeczywiście jesteśmy zmuszeni wybrać tylko jedną z tych koncepcji? Analiza zaniedbania sugeruje, że ani wiedza, ani intencja nie są warunkami koniecznymi przypisania odpowiedzialności podmiotowi za skutki uboczne jego działań. Z całą pewnością zarówno P1, jak i P2 pozwalają opisać wiele różnych przypadków, nie wszystkie jednak.

W kontekście asymetrii dotyczącej przypisywania winy i zasługi należy odróżnić bycie odpowiedzialnym za coś od bycia winnym (Hołówka 2000: 6). Z faktu, że ktoś za coś odpowiada, nie wynika jeszcze, że ten ktoś ma być uznany za winnego w sytuacji, gdyby pojawił się jakiś negatywnie oceniony przez ludzi skutek uboczny. Czy rodzic, który źle wychował swoje dziecko, ponosi winę za to, że jego syn pobił kolegę lub ukradł komuś pieniądze? Czy przypisalibyśmy odpowiedzialność dyrektorowi za zaszkodzenie środowisku, gdyby był pozbawiony wiedzy o negatywnych skutkach ubocznych wprowadzenia nowego programu? Zapewne tak, ale czy stwierdzimy, że jest winny? Niekoniecznie. Wpierw będziemy chcieli poznać stopień jego kompetencji oraz kompetencji jego podwładnych. Z całą pewnością odpowiedzialność nie zawsze pociąga winę.

Jacek Hołówka (2000: 6) twierdzi, że wina w przeciwieństwie do odpowiedzialności nie może być „konstrukcją społeczną”. Ludzie mogą ustalić, kto jest za co odpowiedzialny, mogą podzielić się zakresem odpowiedzialności. Natomiast wina nie jest czymś, co można przypisać na zasadzie podziału. Przykładowo, jako osoba najbardziej doświadczona i kompetentna w wędrówkach górskich mogę zaproponować, że wezmę na siebie odpowiedzialność za uczestników wyprawy, w której biorę udział. Inne osoby mogą moją propozycję zaakceptować bądź nie. W sytuacji, w której nie uznano by mojej odpowiedzialności za grupę, nie można by mi przypisać winy za ewentualną tragedię na szlaku. Dopiero gdybym był za grupę odpowiedzialny, można by mnie jako samowolnego ochotnika obarczyć winą. Rozważania te uzasadniają konieczność dalszej analizy pojęcia działania intencjonalnego.

nia, które wynika z braku wiedzy czy kompetencji — tych z pewnością nie możemy odmówić dyrektorowi). Między tak rozumianym zaniedbaniem a świadomym brakiem opieki czy zainteresowania nie ma różnicy jakościowej. Matkę, która zaniedbała dziecko, uznamy za odpowiedzialną za ten czyn, ale nie przypiszemy specjalnej zasługi matce, która wykonuje swoje obowiązki sumienne. Pytanie o to, dlaczego tak jest, odsyła do problemu Knobe'a. Kwestia zaniedbania ściśle wiąże się z czymś, co Katarzyna Paprzycka (2016) nazywa intencjonalnością zaniechań.

Frank Hindriks (2008) zgadza się z Knobem, że asymetria w przypisywaniu intencjonalności działaniom ściśle wiąże się z asymetrią między zasługą a winą (*Praise–Blame Asymmetry*). Wskazuje przy tym, że pojęcie działania intencjonalnego ma składnik normatywny. Danej osobie przypisuje się zasługę wtedy, gdy wykonując pewne działanie, jest odpowiednio zmotywowana do tego, żeby jej działanie nie prowadziło do złego skutku lub by doprowadziło do dobrych konsekwencji²⁰. Dzięki temu spostrzeżeniu można wyjaśnić asymetrię między przypisywaniem winy i zasługi: zasługę wolno przypisać tylko komuś, kto jest odpowiednio zmotywowany, podczas gdy stwierdzenie odpowiedniej motywacji nie jest potrzebne do obarczenia kogoś winą (Hindriks 2008: 632). Ludzkimi działaniami muszą kierować pewne motywacje i normatywne racje. Racje motywacyjne — podkreśla Hindriks — odgrywają ważną rolę przy wyjaśnianiu działań, racje normatywne służą zaś ich uzasadnieniu. Te ostatnie są brane pod uwagę wtedy, kiedy decydujemy się na wykonanie danego działania lub z niego rezygnujemy. Wiążą się ściśle z analizą skutków działań.

Hindriks jest skłonny uznać, że to raczej nasze sądy dotyczące moralnej odpowiedzialności kształtują sądy na temat intencjonalnych działań, a nie odwrotnie, jak jego zdaniem twierdzi Knobe²¹. Z rozważań tych można wyciągnąć wniosek, że ludzkie przekonania normatywne w sposób ścisły wiążą się ze skłonnością do przypisywania osobom odpowiedzialności za skutki uboczne ich działań. Nie jest jednak jasne, czy przypisywanie odpowiedzialności danym podmiotom jest tym samym co przypisywanie intencjonalności ich działaniom.

5. PRZEKONANIA I PRAGNIENIA

Fiery Cushman i Alfred Mele przekonują, że problem Knobe'a można przeformułować, zadając pytanie, dlaczego pewni ludzie uważają zarówno szkodenie, jak i pomaganie za intencjonalne, podczas gdy inni uważają je za nieintencjonalne (Cushman, Mele 2008: 172). Tak postawione pytanie pozwala im szukać rozwiązania efektu Knobe'a w analizach psychologicznych

²⁰ Nie jest jasne, dlaczego Hindriks nadaje skutkom i konsekwencjom charakter moralny, tzn. przypisuje im własność bycia dobrymi lub złymi.

²¹ Istnieją koncepcje opierające się na podobnej interpretacji do tej przedstawianej przez Hindriksa — por. Malle, Nelson 2003 oraz Nadelhoffer 2004. W sprawie zależności rozważań nad przypisywaniem intencjonalności od rozważań moralnych por. Nadelhoffer 2006.

zaproponowanych przez Nicholasa i Ulatowskiego (2007). Cushman i Mele twierdzą, że należy rozróżnić dwa potoczne pojęcia działania intencjonalnego: (1) związane z pojęciem przekonania (*belief*) oraz (2) związane z pojęciem pragnienia (*desire*). Przez „przekonania” wspomniani autorzy rozumieją „uzasadnione przekonania działających, że *wykonają dane działania*, gdy są *pewni, że tak właśnie jest*” (Cushman, Mele 2008: 173). Przez „pragnienie” zaś rozumie się to, że działający chce wykonać dane działanie. Aby właściwie opisać intencjonalne działanie, należy rozważyć cztery możliwe przypadki:

- (1) w działaniu intencjonalnym obecne są zarówno przekonanie, jak i pragnienie;
- (2) w działaniu intencjonalnym obecne jest tylko przekonanie;
- (3) w działaniu intencjonalnym obecne jest tylko pragnienie;
- (4) w działaniu intencjonalnym nie jest obecne ani przekonanie, ani pragnienie (Cushman, Mele 2008: 173).

Cushman i Mele, analizując te przypadki, stwierdzili, że wyniki ich badań są zgodne z następującymi regułami działań:

- (RP) Reguła pragnienia: działanie jest intencjonalne, jeśli jest wykonane z pewnym pragnieniem, które wyznacza konieczne warunki wstępne (*background conditions*) działania (posiadanie pewnego pragnienia jest warunkiem wystarczającym działania intencjonalnego, nie wyklucza przy tym posiadanie przekonania);
- (RPZ) Reguła przekonania: działanie jest intencjonalne, jeśli jest wykonane z pewnym przekonaniem (przekonanie jest warunkiem wystarczającym działania intencjonalnego);
- (RMZ) Przekonanie jest wystarczającym warunkiem do uznania działania za intencjonalne tylko i wyłącznie w przypadku moralnie złych działań;
- (RMD) Pragnienie jest warunkiem koniecznym do uznania działania za intencjonalne w przypadkach działań, których rezultaty są opisywane jako moralnie dobre (Cushman, Mele 2008: 175).

Zauważalna jest różnica między regułami RP oraz RPZ a RMZ i RMD. Dwie pierwsze reguły odnoszą się do struktury działania, dwie ostatnie wiążą zaś intencjonalność działania (która zgodnie z RP i RPZ jest warunkowana strukturalnie) z moralną predykacją. Zdaniem Cushmana i Melego ludzie przyjmują różne rozumienia działania intencjonalnego i ten fakt tłumaczy asymetrię ocen w przypisywaniu odpowiedzialności za skutki uboczne działań.

Na podstawie analizy statystycznej Cushman wraz z Mele (2008: 183) sformułowali trzy hipotezy:

- (H1) Każdy używa RP (choć ludzie mogą się nie zgadzać w ocenie jakości pragnienia);
- (H2) Tylko 20% osób używa RPZ;
- (H3) Wielu ludzi używa RMZ, być może w pewnych jakościowo różnych sytuacjach niż wtedy, gdy korzysta z RP i RPZ.

Problematyczne jednak jest powiązanie pragnień i przekonań z kwalifikacją moralną danych działań. Wydaje się, że fakt istnienia działań opartych na słabej woli, często określanych mianem działań niepowściągliwych, może podważyć sens wszystkich reguł sformułowanych przez Melego i Cushmana. Działania takie polegają na tym, że sprawca, wykonując daną czynność *A*, postępuje niepowściągliwie, gdy (1) robi *A* w sposób zamierzony, (2) jest przekonany, że może wybrać inny sposób postępowania *B*, oraz (3) wydaje sąd, że wzięwszy wszystkie okoliczności pod uwagę, lepiej byłoby zrobić *B* aniżeli *A* (Davidson 1997: 82)²². W takim ujęciu okazuje się, że mogą mieć pewne moralnie słuszne przekonanie co do potrzeby wykonania danego moralnie dobrego działania, a jednak chcieć wykonać zupełnie inne działanie. Mogą też chcieć wykonać to właśnie działanie, a mimo to wykonać zupełnie inne. Istnienie różnego rodzaju konfliktów moralnych ściśle powiązanych ze zjawiskiem słabej woli zobowiązuje nas do ponownego przemyślenia warunków, w których przypisuje się intencjonalność działaniom²³.

6. TRZECIA ASYMETRIA

Zdaniem Knobe'a asymetria w przypisywaniu intencjonalności działaniom podmiotów wiąże się z tym, czy badani oceniają dany skutek uboczny jako dobry, czy jako zły (Knobe 2003a: 193). W późniejszych rozważaniach Knobe doprecyzował pojęcie skutku ubocznego. Pewien rezultat jest skutkiem ubocznym, jeśli (1) sprawca nie próbował specjalnie go osiągnąć, ale (2) podjął pewne działanie, przewidując, że przyniesie również ten rezultat (Knobe 2008: 132). Skutek uboczny wykonanego działania jest więc niezamierzony

²² Pouczająca jest w tym miejscu dyskusja dotycząca statusu działań niepowściągliwych. Odsyłam tutaj do bloku tekstów zatytułowanego *Słabość woli* w pracy *Filozofia moralności* (Hołówka 1997).

²³ Problem ten jednak znacznie wykracza poza ramy badawcze tego artykułu.

i nie był przedmiotem pragnienia, mimo że był w jakimś sensie przewidziany. Czy w takim ujęciu może on być oceniony jako dobry? Mówienie o dobrych skutkach ubocznych jest podobne do mówienia o dobrych katastrofach. Uważamy, że użycie pojęcia „skutek uboczny”, które ujmuje się w różnych kontekstach pragmatycznych wskazuje jednoznacznie na — jeśli można się tak wyrazić — negatywną semantykę tego wyrażenia. Mówiąc o skutkach ubocznych, mówi się nie tylko o tym, co nie było przedmiotem pragnienia, jak twierdzi Knobe, ale także o czymś przykrym, niechcianym, szkodliwym i niepotrzebnym. Knobe, uznając w swoich badaniach nieszkodzenie środowisku za skutek uboczny, dokonuje zniekształcenia znaczenia tego potocznego pojęcia. Ludzie są bowiem nauczani, aby osoby odpowiadały za skutki swoich słów, decyzji i działań, zwłaszcza te negatywne. Na tym opiera się możliwość uczenia kogoś odpowiedzialności. Patrząc jednak od strony filozofii eksperymentalnej, można znów zapytać o to, czy nie jest tak, że owa ludzka skłonność do tego, aby odpowiadać za skutki własnych działań i decyzji, nie jest także ufundowana na pewnych intuicjach, które również można badać empirycznie.

Przedstawiona dyskusja nad efektem Knobe’a prowadzi do interesujących wniosków i pytań. Najciekawszym z nich jest wielokrotnie stawiane pytanie o charakter związku między skłonnością do przypisywania intencjonalności działaniom a przypisywaniem odpowiedzialności za ich skutki. Z całą pewnością taki związek istnieje, jednak jego natura nie jest jasna, skoro można łatwo wskazać sytuacje, w których przypisanie komuś odpowiedzialności za skutki pewnych działań nie będzie pociągać za sobą przypisania intencjonalności.

Analiza obu asymetrii wskazanych przez Knobe ujawnia — jak sądzę — istnienie innej asymetrii, która leży u ich podstaw. Inspirujące są tutaj uwagi Hindriksa dotyczące roli tzw. racji normatywnych w przypisywaniu win i zasług. Pokazują one, że rozważania moralne oparte są na bardziej podstawowym poziomie normatywnym, wiążącym się z przyjmowaniem takiej racji działania, która pozwoli uniknąć niepożądanych przez podmiot konsekwencji lub skutków ubocznych.

Choć Hindriks ma słuszość, to w sposób nieuzasadniony zawęża normatywność tylko do relacji motywacji. „Pole normatywne” ma zasadniczo dużo większy zasięg. Obejmuje nie tylko motywacje działań oraz sferę moralności (z którą najczęściej jest związane), lecz także różnego rodzaju związki normujące, czyli takie, które zobowiązują nas do czegoś lub mają charakter regularno-wzorcowy. Zanim coś zostanie określone jako dobre lub złe moralnie czy jako pożyteczne lub szkodliwe, musi być wpierw odróżnione od tego, co jest tylko faktyczne. Aby móc powiedzieć, że np. szkodzenie środowisku jest godne potępienia, należy wpierw zdać sobie sprawę z tego, że takie działanie ma zupełnie inny charakter niż np. chodzenie na spacer. Szkodzenie komuś lub

czemuś ma charakter *przede wszystkim* moralny²⁴. Działanie takie jak spacer ma charakter moralny jedynie wtórnie. Powinniśmy szukać ujęcia umożliwiającego nam wykazanie różnicy między zdarzeniami, które są *przede wszystkim* problematyczne moralnie, a takimi, które problematycznie moralne są wtórnie. Możliwość owego rozróżnienia chciałbym wiązać z asymetrią zachodzącą między tym, co normatywne, a tym, co empiryczne, lub mówiąc inaczej: między tym, co normatywne *przede wszystkim*, a tym, co normatywne *wtórnie*²⁵. Według mnie wskazane przez Knobe'a asymetrie wywodzą się z bardziej podstawowej asymetrii dotyczącej możliwości odróżnienia tego, co normatywne, od tego, co faktyczne czy empiryczne.

7. PREDYKCJE I DZIAŁANIA

W pracy *Experimental Philosophy as Cognitive Science* Knobe stwierdza, że celem filozofii eksperymentalnej jest odkrywanie mechanizmów poznawczych, dzięki którym można będzie wyjaśnić konkretne procesy psychiczne. Przyjmując ten postulat, uważam jednocześnie, że pełne wytłumaczenie wykrytych przez Knobe'a asymetrii będzie możliwe dopiero wtedy, gdy odniesiemy się do sposobu, w jaki podmiot poznaje rzeczywistość. Mówiąc inaczej: istotne jest tu zrozumienie tego, w jaki sposób umysł wchodzi w interakcje ze środowiskiem. Współczesne badania nad poznaniem z zakresu teorii umysłu wcielonego i rozszerzonego jednomyślnie podkreślają, że nie da się ściśle rozdzielić kodowania percepcyjnego od kodowania dotyczącego kontroli działania. Percepcja — zgodnie z twierdzeniami Andy'ego Clarka (2013) — jest zorientowanym na działanie procesem predykcyjnym.

Według teorii kodowania predykcyjnego (TKP) (Clark 2013, 2016, Friston 2003, Hohwy 2013) mózg należy rozumieć w kategoriach zaawansowanej maszyny testującej różne hipotezy na temat poznawanego świata. Poznanie jest procesem inferencyjnym, którego celem jest minimalizacja tzw. błędu predykcyjnego. Błąd ten pojawia się pomiędzy hipotezami tworzonymi na podstawie modelu świata a informacjami pochodzącymi ze świata i dostarczonymi przez zmysły. Przychodzące sygnały zmysłowe traktowane są jako

²⁴ Przekonanie to oczywiście może być oparte na potocznych intuicjach pewnej grupy osób żyjących w konkretnym czasie i miejscu.

²⁵ Przykładowo: teoria affordancji Anthony'ego Chemera i Erandy Jayawickreme (2008) zapoznaje tę asymetrię i nie pozwala podać kryteriów odróżnienia tzw. affordancji moralnych od takich, które moralne nie są. Problemem affordancji zajmuję się w artykule *Does Theory of Perception Assume Anti-representationalism? The Role of Affordances in the Explanation of Perception* (w trakcie recenzji).

pewnego rodzaju hipotezy. Na podstawie wiedzy na temat przyczynowych zależności zachodzących w świecie oszacowane zostaje, która z nich jest najbardziej prawdopodobna (Hohwy 2013: 16). Ucieleśniony system poznawczy (por. Allen, Friston 2016, Clark 2015) korzysta przy tym z wielopoziomowego modelu generatywnego, który odgórnie tworzy hipotezy interpretujące oddolnie przychodzące informacje, które pochodzą z sygnału zmysłowego. Każdy poziom tego modelu minimalizuje błędy predykcyjne poziomu niższego — od procesów neuronalnych aż po świadomie podejmowane decyzje i działania (Clark 2013, 2016, Friston 2003).

Dla teorii kodowania predykcyjnego zasadniczą kwestią jest normatywność²⁶. Minimalizacja błędu predykcyjnego zakłada wprost „niskopoziomową” normatywność biologiczną, która odnosi się do utrzymywaniu systemu żywego w stanie dalekim od równowagi termodynamicznej. Friston określa ją mianem *free-energy principle*, ponieważ zakłada ona minimalizację wolnej energii w celu osiągnięcia homeostazy (Friston i in. 2010, Friston, Stephan 2007). Na poziomach wyższych normatywność można (1) wiązać z wzorcami pobudzeń neuronów opartymi o minimalne predykcje oraz (2) z funkcją, którą predykcje pełnią w procesach decyzyjnych i kontroli działań. W tym miejscu — jak sądzę — teoria kodowania predykcyjnego może istotnie wzbogacić analizy Knobe’a.

Rozważmy prosty przykład: spoglądam przez okno i widzę zachmurzone niebo. Na podstawie moich wcześniejszych doświadczeń mogę wnioskować, że będzie padać. Przewiduję zatem, że może spaść deszcz. Owo predyktywne przekonanie sprawia, że zaczynam się zastanawiać, czy mam wziąć parasol, czy może zaryzykować i wyjść bez niego. Niezależnie jednak od tego, jaką decyzję podejmę, to dokonana przeze mnie predykcja staje się racją mojego przyszłego działania. Znaczący to, że niezależnie od tego, co zrobię, moje działanie jest uwarunkowane dokonaniem przeze mnie wcześniej przewidywaniem, które dotyczy przyszłego stanu rzeczy lub zbioru stanów rzeczy. Skoro zatem działania podejmowane przez dany system poznawczy są oparte na predykcjach, to można je uznać za swoiste reguły działania, czyli za ich obiektywne racje, a nie tylko przyczyny czy motywy. W proponowanej interpretacji predykcje są czynione ze względu na rozpoznawane wzorce. Wzorzec zostaje rozpoznany jako to, co pełni określoną funkcję w dokonywanych przez system poznawczy przewidywaniach na temat prawdopodobieństwa zajścia pewnych zdarzeń, w odniesieniu do swoich możliwych reakcji. Dopiero wtórnie wzór jest łączony z jakąś własnością przedmiotów lub ich zespołem czy też tzw. prototypem.

²⁶ Problem ten podkreśla Paweł Gładziejewski (2017).

Można więc powiedzieć, że pierwotne intuicje normatywne pojawiają się już na poziomie podejmowania i wykonywania działań. Wiązą się one wprost z rozpoznaniem normatywnego wymiaru predykcji regulującej i warunkującej działania. Jeżeli dokonywane predykcje, rozumiane tutaj jako racje działania, są powiązane z percepcją odpowiednich własności świata, to z punktu widzenia predykcji pewne zdarzenia mają już charakter normatywny, inne zaś nie: z perspektywy dokonywanej przez mnie predykcji istotne jest zachmurzenie, a nie na przykład pora dnia. Jest ono istotne, ponieważ na podstawie jego rozpoznania dokonam takiej, a nie innej predykcji. Dopiero po tym normatywnym rozpoznaniu będę skłonny powiedzieć coś w rodzaju: „Dobrze, że dzisiaj pada” itd.²⁷

Możliwość rozpoznania asymetrii normatywne–empiryczne pozwala na nowo przeanalizować asymetrie odkryte przez Knobe'a oraz zachodzące między nimi zależności. Możliwość pojawienia się asymetrii w przypisywaniu intencjonalności i przypisywaniu odpowiedzialności opiera się na wiedzy dotyczącej zależności normatywnych, które są już obecne — w myśl teorii kodowania predykcyjnego — w każdym jednorazowym akcie percepcji, która jest zawsze zorientowana na działanie. Możemy komuś przypisać intencjonalność czy odpowiedzialność, ponieważ opieramy się na pewnych intuicjach normatywnych, które, jak wykazałem, są nam dostępne już na poziomie percepcyjnym. Owe intuicje, w przeciwieństwie do analizowanych przez Knobe'a intuicji moralnych, nie zakładają jeszcze żadnej formy wiedzy o podmiocie i sprawstwie²⁸.

ZAKOŃCZENIE

Odwołanie się do zaproponowanej tutaj koncepcji rozpoznawania normatywności pozwala spojrzeć w zupełnie nowy sposób na problem postawiony przez Knobe'a. Wskazana przez niego asymetria wywodzi się z bardziej podstawowej asymetrii dotyczącej możliwości odróżnienia tego, co normatywne, od tego, co empiryczne, lub inaczej mówiąc: tego, co normatywnie pierwotne, od tego, co normatywnie wtórne. Możemy postawić tutaj niebanalną hipotezę badawczą, zgodnie z którą wiedza o stosunkach normatywnych jest bardziej podstawowa niż wiedza o sprawstwie czy samowiedza. Pro-

²⁷ Na ten temat por. także Piekarski 2016.

²⁸ Problem ten jest tutaj naszkicowany. Wymaga on dalszych badań dotyczących pojęcia intuicji oraz normatywności w ujęciu TKP. Konieczne jest również uwzględnienie i przeanalizowanie dalszej literatury przedmiotowej dotyczącej efektu Knobe'a i filozofii eksperymentalnej.

ponowana hipoteza ma jednak charakter czysto filozoficzny. Z perspektywy filozofii eksperymentalnej powinno się ją przebadac empirycznie, czyli sprawdzić, czy istnieją takie intuicje normatywne, których nie da się sprowadzić do intuicji moralnych. Zadanie to znacznie wykracza jednak poza ramy przyjęte dla tych rozważań i wyznacza dalsze perspektywy badań.

BIBLIOGRAFIA

- Adams F. (1986), *Intention and Intentional Action. The Simple View*, „Mind and Language” 1(4), 281-301.
- Adams F., Steadman A. (2004a), *Intentional Action in Ordinary Language. Core Concept or Pragmatic Understanding?*, „Analysis” 64(282), 173-181.
- Adams F., Steadman A. (2004b), *Intentional Action and Moral Considerations. Still Pragmatic*, „Analysis” 64(283), 268-276.
- Allen M., Friston K. J. (2016), *From Cognitivism to Autopoiesis. Towards a Computational Framework for the Embodied Mind*, „Synthese” 1-24, DOI: 10.1007/s11229-016-1288-5.
- Bratman M. (1987), *Intention, Plans and Practical Reason*, Cambridge, MA: Harvard University Press.
- Chemero A., Jayawickreme E. (2008), *Ecological Moral Realism. An Alternative Theoretical Framework for Studying Moral Psychology*, „Review of General Psychology” 12(2), 118-126.
- Clark A. (2013), *Whatever Next? Predictive Brains, Situated Agents, and the Future of Cognitive Science*, „Behavioral and Brain Sciences” 36(3), 181-204.
- Clark A. (2015), *Embodied Prediction [w:] Open MIND: 7(T)*, T. Metzinger, J. M. Windt (red.), Frankfurt am Main: MIND Group, 267-287.
- Clark A. (2016), *Surfing Uncertainty. Prediction, Action and the Embodied Mind*, Oxford: Oxford University Press.
- Cushman F., Mele A. (2008), *Intentional Action. Two-and-a-Half Folk Concepts? [w:] Experimental Philosophy*, J. Knobe, S. Nichols (red.), Oxford: Oxford University Press, 171-188.
- Davidson D. (1997), *Jak jest możliwa słabość woli? [w:] Filozofia moralności. Postanowienie i odpowiedzialność moralna*, J. Hołówka (red.), Warszawa: Spacja, 81-106.
- Dębska A. (2013), *Wnioskowanie na temat intencjonalności działań w ujęciu filozofii eksperymentalnej*, „Filozofia Nauki” 21(3) [83], 143-155.
- Friston K. J. (2003), *Learning and Inference in the Brain*, „Neural Networks” 16(9), 1325-1352.
- Friston K. J., Stephan K. E. (2007), *Free-Energy and the Brain*, „Synthese” 159, 417-458.
- Friston K. J. i in. (2010), *Action and Behavior. A Free-Energy Formulation*, „Biological Cybernetics” 102(3), 227-260.
- Gładziejewski P. (2017), *The Evidence of the Senses. A Predictive Processing-Based Take on the Sellarsian Dilemma [w:] Philosophy and Predictive Processing*, T. Metzinger, W. Wiese (red.), Frankfurt am Main: MIND Group, DOI: 10.15502/9783958573161.
- Green J., Haidt J. (2002), *How (and Where) Does Moral Judgment Work?*, „Trends in Cognitive Sciences” 6(12), 517-523.

- Harman G. (1976), *Practical Reasoning*, „Review of Metaphysics” 29(3), 431-463.
- Hindriks F. (2008), *Intentional Action and the Praise–Blame Asymmetry*, „Philosophical Quarterly” 58(233), 630-641.
- Hołówka J. (red.) (1997), *Filozofia moralności. Postanowienie i odpowiedzialność moralna*, Warszawa: Spacja.
- Hołówka J. (2000), *Przedmowa* [w:] *Filozofia moralności. Wina, kara, wybaczenie*, J. Hołówka (red.), Warszawa: Aletheia.
- Hohwy J. (2013), *The Predictive Mind*, Oxford: Oxford University Press.
- Ingarden R. (1963), *Dążenia fenomenologów* [w:] *Z badań nad filozofią współczesną*, Warszawa: Państwowe Wydawnictwo Naukowe.
- Knobe J. (2003a), *Intentional Action and Side Effects in Ordinary Language*, „Analysis” 63(3), 190-194.
- Knobe J. (2003b), *Intentional Action in Folk Psychology. An Experimental Investigation*, „Philosophical Psychology” 16, 309-324.
- Knobe J. (2004), *Folk Psychology and Folk Morality. Response to Critics* „Journal of Theoretical and Philosophical Psychology” 24(2), 270-279.
- Knobe J. (2008), *The Concept of Intentional Action. A Case Study in the Uses of Folk Psychology* [w:] *Experimental Philosophy*, J. Knobe, S. Nichols (red.), Oxford: Oxford University Press, 129-148.
- Knobe J. (2010), *Person as Scientist, Person as Moralist*, „Behavioral and Brain Sciences” 33(4), 315-329.
- Knobe J. (2016), *Experimental Philosophy Is Cognitive Science*, [w:] *A Companion to Experimental Philosophy*, J. Sytsma, W. Buckwalter (red.), Malden, MA: Willey-Blackwell, 37-52.
- Knobe J., Nichols S. (2008a), *An Experimental Philosophy Manifesto* [w:] *Experimental Philosophy*, J. Knobe, S. Nichols (red.), New York, NY: Oxford University Press, 3-16.
- Knobe J., Nichols S. (red.) (2008b), *Experimental Philosophy*, New York, NY: Oxford University Press.
- Komorowska-Mach J. (2013), *Negatywny program filozofii eksperymentalnej a odwołanie do intuicji w argumentacji filozoficznej*, „Filozofia Nauki” 21(3) [83], 156-165.
- Kuś K., Maćkiewicz B. (2016), *Z rozmysłem, ale nie specjalnie. O językowej wrażliwości filozofii eksperymentalnej*, „Filozofia Nauki” 24(3) [95], 73-101.
- Ludwig K. (2010), *Intuitions and Relativity*, „Philosophical Psychology” 23(4), 427-445.
- Machery E., Mallon R., Nichols S., Stich S. P. (2004), *Semantics, Cross-cultural Style*, „Cognition” 92(3), 1-12.
- Malle B., Knobe J. (1997), *The Folk Concept of Intentional Action*, „Journal of Experimental Social Psychology” 33(2), 101-121.
- Malle B., Nelson S. E. (2003), *Judging Mens Rea. The Tension between Folk Concepts and Legal Concepts of Intentionality*, „Behavioral Sciences and the Law” 21(5), 563-580.
- McCann H. (1986), *Rationality and the Range of Intention*, „Midwest Studies in Philosophy” 10(1), 191-211.
- Mele A. (1992), *Springs of Action*, New York, NY: Oxford University Press.
- Mele A. (2001), *Acting Intentionally. Probing Folk Notions* [w:] *Intentions and Intentionality. Foundation of Social Cognition*, B. Malle, L. Moses, D. Baldwin (red.), Cambridge, MA: MIT Press, 27-43.
- Nadelhoffer T. (2004), *Blame, Badness, and Intentional Action. A Reply to Knobe and Mendlow*, „Journal of Theoretical and Philosophical Psychology” 24(2), 259-269.

- Nadelhoffer T. (2006), *Bad Acts, Blameworthy Agents, and Intentional Actions. Some Problems for Jury Impartiality*, „Philosophical Explorations” 9(2), 203-220.
- Nadelhoffer T., Nahmias E. (2007), *The Past and Future of Experimental Philosophy*, „Philosophical Explorations” 10(2), 123-149.
- Nichols S., Stich S., Weinberg J. (2003), *Metaskepticism. Meditations in Ethno-Epistemology* [w:] *The Sceptics*, S. Luper (red.), Burlington, VT: Ashgate, 227-247.
- Nichols S., Ulatowski J. (2007), *Intuitions and Individual Differences. The Knobe Effect Revisited*, „Mind and Language” 22(4), 346-365.
- Paprzycka K. (2003), *Falszywa świadomość psychologii intencjonalnej, czyli o pewnym dogmacie współczesnej analitycznej filozofii umysłu*, „Filozofia Nauki” 11(1) [41], 89-116.
- Paprzycka K. (2014), *Rozwiązanie problemu Butlera i wyjaśnienie efektu Knobe’a*, „Filozofia Nauki” 22(2) [86], 73-96.
- Paprzycka K. (2016), *O intencjonalności działań i zaniechań, czyli o społecznej naturze sprawstwa*, „Przegląd Filozoficzny – Nowa Seria” 25(1) [97], 45-65.
- Piekarski M. (2016), *Can Action Be Normative? Some Preliminary Remarks*, „Studia Philosophiae Christianae” 52(3), 97-108.
- Sosa E. (2007), *Experimental Philosophy and Philosophical Intuition*, „Philosophical Studies” 132(1), 99-107.
- Young L. i in. (2006), *Does Emotion Mediate the Effect of an Action’s Moral Status on Its Intentional Status? Neuropsychological Evidence*, „Journal of Cognition and Culture” 6(1-2), 265-278.
- Wysocki T. (2011), *Filozofia eksperymentalna jako metodologia filozoficzna*, <http://www.xphi-europe.org/>.