

KATARZYNA PAPRZYCKA*

WYJAŚNIENIE EPISTEMICZNEGO EFEKTU KNOBE'A**

Abstract

EXPLANATION OF THE EPISTEMIC SIDE-EFFECT EFFECT

On the omissions account, in Knobe's harm case people attribute an intentional omission, which is linguistically expressed as if it were an intentional action. I extend the omissions account to the epistemic side-effect effect. In the harm case, there is a salient knowledge claim (knowledge of a possible harmful consequence) that is relevant to the intentionality of the omission. In assenting to the claim that the chairman knew that the environment would be harmed, people endorse the claim that the chairman knew that *a possible* harmful consequence of his action was that the environment would be harmed. Such a claim is independent from the predictive claim that the chairman knew that the environment would be harmed; which explains why people assent to the knowledge claim even in the slight chance of harm conditions.

Keywords: epistemic side-effect effect, intentional action, intentional omission, Knobe effect, knowledge

Joshua Knobe (2003a, b) wykazał osobliwą asymetrię w przypisywaniu intencjonalności niezamierzonym, lecz przewidzianym skutkom ubocznym działań w zależności od wartości moralnej tych skutków¹. Na pierwszy rzut oka odkrycie to stanowi problem dla standardowych teorii działania intencjonalnego, ponieważ historyjki, na których Knobe się opiera, są strukturalnie

* Zakład Epistemologii, Instytut Filozofii, Uniwersytet Warszawski, ul. Krakowskie Przedmieście 3, 00-927 Warszawa, kpaprzycka@uw.edu.pl.

** Chciałabym podziękować za uwagi anonimowemu recenzentowi oraz uczestnikom prowadzonego przeze mnie seminarium — szczególnie gorąco pani Marcie Zarębie. Projekt został sfinansowany ze środków Narodowego Centrum Nauki przyznanych na podstawie decyzji numer DEC-2013/11/B/HS1/03939.

¹ Badania Knobe'a (2007) wskazują, że nie chodzi o samą wartość moralną efektów ubocznych, lecz raczej o to, czy stanowią one pogwałcenie jakichś istotnych norm (por. Holton 2010).

identyczne z punktu widzenia tych teorii: jego scenariusze różnią się wartością moralną skutków ubocznych, której standardowe koncepcje działania intencjonalnego w ogóle nie uwzględniają.

Beebe i Buckwalter (2010) odkryli podobną asymetrię, tym razem w przypisywaniu wiedzy w scenariuszach Knobe'a, co z kolei wydaje się podważać standardowe koncepcje epistemologiczne. Historyjki Knobe'a pod względem struktury tradycyjnie rozumianych świadectw epistemicznych są bowiem w pełni symetryczne. Epistemiczny efekt Knobe'a (*epistemic side-effect effect*) mógłby świadczyć o tym, że czynniki moralne odgrywają istotną rolę w przypisywaniu wiedzy. Badanie Beebe'ego i Buckwaltera zostało powielone w kilku różnych wariantach (Beebe, Jensen 2012), także na populacji niemieckojęzycznej (Dalbauer, Hergovich 2013).

W artykule proponuję pewne wyjaśnienie epistemicznego efektu Knobe'a (dalej będę też mówić o „efekcie epistemicznym”, zob. część 2), które odwołuje się do podanego przeze mnie wyjaśnienia klasycznego efektu Knobe'a (dalej będę też mówić o „efekcie Knobe'a”). Pokazałam (Paprzycka 2014, 2015), że efekt Knobe'a można zrozumieć, odwołując się do hipotezy o zaniechaniach (zob. niżej, część 1). Knobe rozważał dwa scenariusze: o uszkodzeniu i o pomaganiu, z których tylko historyjka o uszkodzeniu stanowi problem dla standardowych ujęć działania intencjonalnego (Davidson 1980, Goldman 1970, Mele, Moser 1994). Zgodnie z proponowanym wyjaśnieniem w historyjce o uszkodzeniu znamienne jest to, że można ją rozumieć jako historyjkę nie o działaniu intencjonalnym, lecz o zaniechaniu intencjonalnym, które jest językowo zakamuflowane tak, jak gdyby było intencjonalnym działaniem. Pokażę, że efekt epistemiczny da się wyjaśnić w podobny sposób (część 2). Po pierwsze, można bronić tezy, że dla standardowych ujęć świadectw epistemicznych tylko historyjka o uszkodzeniu jest kłopotliwa (część 3). Po drugie, ponieważ w scenariuszu ze uszkodzeniem istotne jest pewne zaniechanie, to teza o wiedzy odgrywa zasadniczą rolę, skoro zgodnie z normatywną teorią intencjonalnych zaniechań to właśnie warunek wiedzy (a nie warunek intencji) przesądza o intencjonalności zaniechań (por. Paprzycka 2015, 2016a, c). Rozróżniam dwie interpretacje zdania przypisującego wiedzę dyrektorowi w sytuacji uszkodzenia i pokazuję, w jaki sposób odwołanie się do tego rozróżnienia pozwala na wyjaśnienie efektu epistemicznego (część 4), a także pewnych jego wariantów (część 5).

1. HIPOTEZA O ZANIECHANIACH A EFEKT KNOBE'A

Knobe (2003a) przedstawił badanym dwa strukturalnie identyczne scenariusze różniące się jedynie tym, że w jednym mowa jest o szkodzeniu środowisku, a w drugim — o pomaganiu mu (a także tym, że w jednym występuje spójnik „ale”, a w drugim „i”):

Wicedyrektor zwraca się do dyrektora pewnej firmy: „Myślimy o wdrożeniu nowego programu. Pozwoli nam zwiększyć zyski, *ale [i] zaszkodzi [pomocze] środowisku*”.

Dyrektor odpowiada: „Nie obchodzi mnie *szkodzenie [pomaganie] środowisku*. Chcę tylko zwiększyć zyski. Wdrażamy program”.

Program został wdrożony i rzeczywiście *zaszkodził [pomógł] środowisku*.

Czy dyrektor *zaszkodził [pomógł] środowisku* intencjonalnie?

Knobe'owi chodziło o to, by w obu scenariuszach stan umysłu dyrektora był taki sam. W obu ma on intencję zwiększenia zysku i w związku z tym uruchomienia programu, ale nie jest jego intencją szkodzenie czy pomaganie środowisku, choć zdaje sobie sprawę z takich konsekwencji swych działań (został o nich poinformowany). Ze względu na to, że stan umysłu dyrektora jest w obu wypadkach taki sam, wydawałoby się, że na pytanie o to, czy dyrektor intencjonalnie zaszkodził (pomógł) środowisku, badani powinni dawać taką samą odpowiedź. Zgodnie ze standardową koncepcją działania intencjonalnego badani powinni powiedzieć, że nie jest prawdą, iż dyrektor zaszkodził/pomógł środowisku — przecież nie miał takiej intencji. W historyjce o pomaganiu większość badanych (77%) rzeczywiście zaprzeczyła twierdzeniu, że dyrektor intencjonalnie pomógł środowisku. W scenariuszu ze szkodzeniem większość badanych (82%) zgodziła się jednak z twierdzeniem, że:

(A) Dyrektor intencjonalnie zaszkodził środowisku.

	SZKODZENIE	POMAGANIE
Dyrektor działał intencjonalnie	82%	23%
Dyrektor nie działał intencjonalnie	18%	77%

Tabela 1. Efekt Knobe'a (2003a)

Dla standardowego ujęcia działań intencjonalnych trudność stanowi zatem jedynie historyjka o szkodzeniu². Jest ona tym bardziej problematyczna, że badani uzasadniają swój wybór (A) w sposób dla standardowych teorii zaskakujący — powołują się mianowicie na to, że przecież dyrektor dysponował stosowną wiedzą (por. Nichols, Ulatowski 2007).

Według normatywnej teorii zaniechań intencjonalnych (Feinberg 1984, Gorr 1979, Smith 1990, 2005, von Wright 1963, Williams 1995) x dopuszcza się intencjonalnego zaniechania α -owania dokładnie wtedy, gdy:

- (a) x powinien α -ować (np. ma obowiązek α -owania, rozsądne jest normatywne oczekiwanie od x -a, by x α -ował) (warunek normatywny),
- (b) α -owanie leży w mocy x -a (warunek możliwości),
- (c) nieprawda, że x α -uje, tj. x nie wypełnia obowiązku (warunek niespełnienia),
- (d) x zdaje sobie sprawę z (a)-(c) (warunek wiedzy).

W kontekście efektu Knobe'a warto szczególnie zwrócić uwagę na warunek normatywny i warunek wiedzy. Efekt Knobe'a polega bowiem na wskazaniu istotności kontekstu normatywnego dla przypisania intencjonalności. Co więcej, w pewnych sytuacjach przypisuje się intencjonalność, gdy sprawca nie ma stosownej intencji, a jedynie pewne postawy poznawcze. Zatem zarówno warunek normatywny, jak i warunek wiedzy wskazują, że teorię tę można zapewne zastosować do sytuacji szkodzenia.

Problematyczny w tym kontekście okazuje się warunek niespełnienia. Według normatywnej teorii zaniechań mowa jest bowiem o niezrobieniu czegoś, a w sytuacji Knobe'a mowa jest o zrobieniu czegoś — o szkodzeniu środowisku. Warto jednak zwrócić uwagę, że zgodnie z normatywną teorią zaniechań niezrobienie czegoś jest o tyle istotne, o ile niespełniona jest powinność, o której mowa w warunkach normatywnych. Normatywna teoria zaniechań nie musi więc być odczytywana jako metafizyczna teoria zaniechań. W tym rozumieniu dla pojęcia zaniechania istotne jest niespełnienie powinności, a nie brak czynności czy działania sprawcy. Powinność, o której mowa, może być sama w sobie negatywna. Wyobraźmy sobie, że x ma obowiązek nie- β -owania (warunek normatywny). W konsekwencji możemy x -owi przypisać

² Dla pewnych ujęć działań intencjonalnych bardziej problematyczna może być historyjka o pomaganiu (por. Bratman 1987, Ginet 1990, Gorr, Horgan 1982, Wilson 1989). Istnieją też koncepcje działania intencjonalnego, które są zgodne z wynikami Knobe'a (Harman 1976, Holton 2010, Hindriks 2008, 2011). Por. też Paprzycka 2016a.

stosowne zaniechanie, o ile nie- β -owanie leży w mocy x -a (warunek możliwości), lecz x nie dopełnia obowiązku nie- β -owania, np. β -ując (warunek niespełnienia). Innymi słowy, normatywną teorię zaniechań intencjonalnych można stosować również do działań wtedy, gdy działanie (β -owanie) jest niewypełnieniem obowiązku negatywnego (nie- β -owania).

Pokazywałam (Paprzycka 2014, 2015), że jeżeli poszukać intencjonalnego zaniechania w scenariuszach Knobe'a, to można je odnaleźć właśnie w historyjce o szkodzeniu. Dyrektor ma obowiązek nieszkodzenia środowisku (warunek normatywny). Nieszkodzenie środowisku leży w jego mocy — skoro w jego mocy leży niewdrożenie programu (warunek możliwości). Dyrektor szkodzi środowisku, tj. nie wypełnia swego obowiązku nieszkodzenia środowisku (warunek niespełnienia). O ile dyrektor zdaje sobie sprawę z warunków swego zaniechania, wolno mu przypisać intencjonalne zaniechanie nieszkodzenia środowisku:

(B) Dyrektor intencjonalnie zaniechał nieszkodzenia środowisku.

Oczywiście, (B) nie jest zdaniem, którego użylibyśmy potocznie, opisując historyjkę o szkodzeniu. Twierdzę raczej, że jest to zdanie, do którego uznania uprawnia normatywna teoria zaniechań intencjonalnych. Wcale nie jest jednak trudno dokonać dość prostych przekształceń zdania (B) w problematyczne zdanie Knobe'a (A), które rzekomo przypisuje dyrektorowi intencjonalne działanie.

Zdanie (B) zawiera bowiem ukrytą podwójną negację. Jedna negacja jest związana z negatywnym obowiązkiem *nieszkodzenia* środowisku, druga z zaniechaniem — zaniechać zrobienia czegoś to wszakże *nie* zrobić tego czegoś. Te dwie negacje można ujawnić za pomocą rozumowania Γ :

(B) Dyrektor intencjonalnie zaniechał nieszkodzenia środowisku.

(B') Tym, czego dyrektor intencjonalnie zaniechał, było nieszkodzenie środowisku.

(1') Tym, czego dyrektor intencjonalnie nie zrobił, było nieszkodzenie środowisku.

(1) Dyrektor intencjonalnie nie nie-zaszkoził środowisku.

(A) Dyrektor intencjonalnie zaszkodził środowisku.

Zdania (B') i (1') stanowią odmienną konstrukcję gramatyczną odpowiadających im zdań (B) i (1). Należy zwrócić uwagę, że choć przekształcenie zdania (1') w zdanie (1) nie stanowi problemu gramatycznego, gdy mowa o zaniechaniach obowiązków pozytywnych, to przekształcenie (1') w (1) może budzić

wątpliwości, gdy mowa o zaniechaniach obowiązków negatywnych. W zdaniu (1') występuje bowiem podwójna negacja. Nie jest to jednak zdanie niezrozumiałe.

Rozumowanie Γ pozwala pokazać, w jaki sposób można dokonać przejścia od przypisania intencjonalnego zaniechania niezrobienia czegoś do przypisania intencjonalnego działania zrobienia czegoś. Przekonywałam jednak (2014, 2015), że Γ opiera się na wątpliwych przesłankach, a w każdym razie ktoś, kto przyjmuje normatywną teorię zaniechań intencjonalnych i standardową teorię działań intencjonalnych, ma podstawy, by argumentację tę odrzucić. Choć można mieć uzasadnione wątpliwości, czy Γ jest materialnie poprawna, to sama jej prostota może nas — trochę bałamutnie — skłaniać do przyjęcia zdania (A). Tymczasem (A) wydawać się może zdaniem o działaniu intencjonalnym. Jeżeli jednak jedyną podstawą przyjęcia (A) jest (B), to (A) — wbrew pozorom — nie jest zdaniem o intencjonalnym działaniu, lecz zdaniem o intencjonalnym zaniechaniu.

Podsumowując, zgodnie ze standardową teorią działania intencjonalnego scenariusze Knobe'a są symetryczne — w obu mamy powody, by zaprzeczyć tezie o tym, że pomaganie (szkodzenie) środowisku było intencjonalne. W historyjce o szkodzeniu jednak — zgodnie z normatywną teorią zaniechań intencjonalnych — mamy podstawy, by przypisać dyrektorowi intencjonalne zaniechanie nieszkodzenia środowisku, które (zgodnie z rozumowaniem Γ) czasami wyrażamy takimi samymi słowami, jak gdybyśmy przypisywali dyrektorowi działanie intencjonalne.

2. EPISTEMICZNY EFEKT KNOBE'A

Beebe i Buckwalter (2010) powtórzyli badanie Knobe'a, ale zadali badanym pytanie nie o działanie intencjonalne, lecz o wiedzę. Pytali mianowicie, czy dyrektor wiedział, że nowy program zaszkodzi/pomoże środowisku („Did the chairman know that the new program would help/harm the environment?”). Okazuje się, że badani byli bardziej skłonni przypisywać wiedzę dyrektorowi w scenariuszu ze szkodzeniem niż w wariancie z pomaganiem. W pierwszym badaniu (2010) badacze pytali o stopień zgody badanych z twierdzeniem, że dyrektor wiedział, że program zaszkodzi/pomoże środowisku („The CEO knew that the environment would be harmed/helped”)³, pro-

³ Warto zwrócić uwagę, że nie możemy w języku polskim dokonać dokładnego tłumaczenia tego zdania. W badaniach polskojęzycznych testowano różne wersje omawianego twierdzenia, uzyskujące nieco inne rezultaty (badania przeprowadzone przez Martę Zarębę

sząc o zaznaczenie odpowiedzi na skali Likerta od -3 (oznaczonego jako „Dyrektor nie wiedział”) do +3 (oznaczonego jako „Dyrektor wiedział”). W scenariuszu ze szkodzeniem uzyskali średnią +2,25 (SD = 1,50), a w scenariuszu z pomaganiem +0,91 (SD = 2,09).

Dalbauer i Hergovich (2013) powtórzyli ten efekt (a także klasyczny efekt Knobe'a) na populacji niemieckojęzycznej. W innych badaniach używali skali analogowej do pomiaru stopnia zgody badanych — badani przesuwali miarkę na pewnym odcinku, a położenie miarki było następnie kodowane w zakresie od 1 do 100. Badacze uzyskali statystycznie istotną różnicę przypisań wiedzy. W historyjce o szkodzeniu średnia wartość wynosiła 96,07 (SD = 15,71), a w wariancie z pomaganiem średnia wynosiła 87,05 (SD = 26,65).

Beebe i Jensen (2012) powtórzyli też swoje badanie w paradygmacie wymuszonego wyboru. Badani musieli zdecydować, za którą z dwóch odpowiedzi się opowiadają — „Dyrektor wiedział, że program zaszkodzi/pomoże środowisku” czy „Dyrektor nie wiedział, że program zaszkodzi/pomoże środowisku”. W historyjce o pomaganiu 84% osób twierdziło, że dyrektor nie wiedział, że program pomoże środowisku, a w historyjce o szkodzeniu 68% twierdziło, że wiedział, że zaszkodzi.

	SZKODZENIE	POMAGANIE
Dyrektor wiedział	68%	16%
Dyrektor nie wiedział	32%	84%

Tabela 2. Epistemiczny efekt Knobe'a (wymuszony wybór; Beebe i Jensen 2012)

3. EPISTEMICZNY EFEKT KNOBE'A A STANDARDOWE UJĘCIA ŚWIADECTW

Pokażę, że efekt epistemiczny można interpretować w sposób w pełni zgodny z tradycyjnymi podejściami epistemologicznymi, według których w szczególności to, czy coś jest dla x -a świadectwem, czy p , jest niezależne od wartości moralnej działań podmiotu podejmowanych przez niego przy uwzględnieniu przekonania, że p .

Po pierwsze, należy zwrócić uwagę, że — podobnie jak w wypadku klasycznego efektu Knobe'a — tradycyjne teorie epistemiczne dość dobrze radzą

oraz Ignacego Szczeniowskiego i Sławomira Mijasa). W szczególności wydaje się, że efekt epistemiczny nie daje się powielić na gruncie języka polskiego — w każdym razie nie dla czasownika „wiedzieć” (por. artykuł Zaręby w przygotowaniu).

sobie ze zrozumieniem, jak badani odpowiadają w historyjce o pomaganiu. Wszak jedyną podstawą dla przekonania, że program pomoże środowisku, jest informacja uzyskana od wicedyrektora. Nawet jeżeli przyjmiemy, że dyrektor uznaje swego podwładnego za osobę kompetentną i wiarygodną, to przewidywanie tego, co się zdarzy w przyszłości, jest zawsze obarczone ryzykiem. W grę mogą wchodzić rozmaite czynniki, które spowodowałyby, że program — mimo najlepszych prognoz — nie pomoże jednak środowisku. Nie dziwi zatem, że o ile badani są skłonni zgodzić się z twierdzeniem, że dyrektor wiedział, iż program pomoże środowisku, o tyle robią to raczej z rezerwą. Nic dziwnego też, że gdy są przymuszeni do jednoznacznego wyboru, opowiadają się za tym, że dyrektor nie dysponował taką wiedzą. Mamy zatem podstawy, by sądzić, że tradycyjne ujęcia epistemologiczne mogą wyjaśnić zachowanie badanych w historyjce o pomaganiu.

Po drugie, trzeba jasno powiedzieć, że ponieważ tradycyjne ujęcia wiedzy nie uznają zabarwienia moralnego działań jako czynnika istotnego dla wiedzy (którą może kierować się podmiot w tym działaniu), to przewidywałyby dokładnie takie same odpowiedzi badanych w historyjce o szkodzeniu. Innymi słowy, biorąc pod uwagę tradycyjne ujęcia epistemologiczne, przewidywać można, że ludzie będą wybierać tezę:

(~K_s) Dyrektor nie wiedział, że nowy program zaszkodzi środowisku.

Tymczasem badania Beebe'ego i in. pokazują, że ogromna większość badanych wybiera twierdzenie przeciwne:

(K_s) Dyrektor wiedział, że nowy program zaszkodzi środowisku.

Mogłoby to znaczyć, że nasze potoczne pojęcie wiedzy jest — wbrew tradycyjnym koncepcjom epistemologicznym — wrażliwe na czynniki moralne czy normatywne. Mogłoby jednak też znaczyć, że w sytuacji szkodzenia w grę wchodzi np. jakiś inny sąd o wiedzy, który zaburza tę zasadniczą symetrię wynikającą z ujęć tradycyjnych. Spróbuję pokazać, że są podstawy, by przypuszczać, że zachodzi druga z tych możliwości.

4. EPISTEMICZNY EFEKT KNOBE'A W ŚWIETLE HIPOTEZY O ZANIECHANIACH

Jak widzieliśmy, zgodnie z hipotezą o zaniechaniach historyjka o szkodzeniu i historyjka o pomaganiu różnią się od siebie w zasadniczy sposób. Mimo pozorów ścisłego podobieństwa wariant ze szkodzeniem jest historyjką o za-

niechaniu intencjonalnym. Co więcej, pamiętamy, że pytanie o wiedzę ma zasadniczy charakter i musi zostać rozstrzygnięte, aby móc przypisać podmiotowi intencjonalne zaniechanie. Będę starała się wykazać, że jednym z elementów tej wiedzy jest wiedza podmiotu o możliwych negatywnych konsekwencjach działania i że właśnie tak można rozumieć zdanie, które badani przyjmują.

Sprawca dopuszcza się bowiem intencjonalnego zaniechania tylko wtedy, gdy:

- (1) wie, że jego działanie może prowadzić do negatywnych konsekwencji (świadomość możliwości negatywnych konsekwencji),
- (2) wie to o pewnej konkretnej konsekwencji (znajomość konkretnej negatywnej konsekwencji),

a mimo to podejmuje działanie. Należy podkreślić dwie kwestie. Po pierwsze, „może” w zdaniu (1) jest w pełni uzasadnione. Jeśli sprawca ma obowiązek nie- β -owania i wie, że α -owanie *może* doprowadzić do β -owania, to wie również, że α -ując, *może* nie wypełnić swego obowiązku. Sprawca powinien (w pewnych rozsądnych granicach) wypełniać obowiązki, a więc w tym wypadku *ceteris paribus* sprawca powinien nie α -ować. Należy podkreślić, że intencjonalność zaniechania niedoprowadzenia do pewnej szkody nie wymaga, by sprawca wiedział, że jego działanie do tej szkody na pewno doprowadzi. W zupełności wystarczy wiedza, że jego działanie do tej szkody *może* doprowadzić.

Po drugie, możliwa jest sytuacja, w której (1) jest prawdą, a (2) fałszem: sprawca wie, że jego działanie może mieć *pewne* negatywne konsekwencje, ale nie zdaje sobie sprawy, że wśród tych możliwych konsekwencji jest konkretny stan rzeczy *S*, do którego jego działanie *de facto* doprowadzi. Sprawca nie zna wtedy tej negatywnej konsekwencji swego działania i nie można mu przypisać intencjonalnego doprowadzenia do *S* (tj. intencjonalnego zaniechania niedoprowadzenia do *S*). Przeanalizujmy pewien przykład: Antek wie, że jeśli zamknie zawór, to spowoduje uschnięcie roślin, do których podlewania się zobowiązał. Nie wie jednak, że zamknięcie zaworu spowoduje również odłączenie wody w mieszkaniu Beaty. Gdy Antek zamyka zawór, intencjonalnie doprowadza do uschnięcia roślin, ale nieintencjonalnie (nieświadomie, nieumyślnie) doprowadza do odłączenia wody w mieszkaniu Beaty. Aby móc przypisać sprawcy intencjonalne doprowadzenie do pewnego negatywnego stanu rzeczy *S* (tj. intencjonalne zaniechanie niedoprowadzenia do *S*), sprawca musi zdawać sobie sprawę z konsekwencji *S*.

Problematyczne zdanie o wiedzy dyrektora (K_s) można zinterpretować właśnie jako zdanie o znajomości konkretnej konsekwencji (2). Dyrektor za-

szkodził środowisku – zaniechał niedoprowadzenia do negatywnego stanu rzeczy. Co więcej, zrobił to intencjonalnie, ponieważ (1) wiedział, że uruchomienie programu może mieć negatywne konsekwencje, a także (2) wiedział, że dotyczy to tego konkretnego stanu rzeczy, do którego *de facto* doprowadził, tj. zdawał sobie sprawę, że możliwą negatywną konsekwencją uruchomienia programu jest to, że zaszkodzi środowisku. Innymi słowy, zgodnie z hipotezą o zaniechaniach istotnym elementem historyjki o uszkodzeniu, niezbędnym do zrozumienia intencjonalności zaniechania dyrektora, jest to, że:

(ZK_s) Dyrektor wiedział, że [możliwą negatywną konsekwencją jego działania jest to, iż] nowy program zaszkodzi środowisku.

Należy zauważyć, że (ZK_s) nie jest twierdzeniem predyktywnym o tym, co się zdarzy w przyszłości. Twierdzenie predyktywne, które jest naturalnym sposobem odczytania zdania (K_s), można sformułować w następujący sposób:

(PK_s) Dyrektor wiedział, że [jest wysoce prawdopodobne, iż] nowy program zaszkodzi środowisku.

Warto podkreślić, że dyrektor mógł (i) mieć absolutną pewność, że możliwą negatywną konsekwencją jego działania jest to, że nowy program zaszkodzi środowisku, a jednocześnie (ii) nie być w ogóle przekonany, że ten stan rzeczy nastąpi. Teza (ZK_s) jest więc spójna z negacją (PK_s).

Hipoteza o zaniechaniach uprawnia zatem do przypuszczenia, że epistemiczny efekt Knobe'a jest związany z tym, że zdanie o wiedzy jest interpretowane jako zdanie o znajomości konsekwencji w sytuacji uszkodzenia, natomiast jako zdanie predyktywne w sytuacji pomagania. Zgodziliśmy się bowiem, że korzystając ze standardowej teorii wiedzy, możemy przyjąć, że dyrektor miał niewystarczające dane, by uznać, iż jest wysoce prawdopodobne, że program zaszkodzi bądź pomoże środowisku. Jeżeli twierdzenia o wiedzy:

(K_s) Dyrektor wiedział, że nowy program zaszkodzi środowisku,

(K_p) Dyrektor wiedział, że nowy program pomoże środowisku,

miałyby przyjąć interpretację predyktywną:

(PK_s) Dyrektor wiedział, że [jest wysoce prawdopodobne, iż] nowy program zaszkodzi środowisku,

(PK_p) Dyrektor wiedział, że [jest wysoce prawdopodobne, iż] nowy program pomoże środowisku,

to badani (gdyby zostali przymuszeni do wyboru) powinni uznawać oba te twierdzenia z pewnym wahaniem lub nie uznawać ich wcale. Respondenci tak właśnie traktują tezę o wiedzy (K_p) w scenariuszu z pomaganiem.

Hipoteza o zaniechaniach w historyjce o uszkodzeniu pozwala jednak na alternatywną interpretację twierdzenia (K_s) — jako twierdzenia o możliwej negatywnej konsekwencji działania dyrektora, czyli (ZK_s). Co więcej, jak pokazywałam, twierdzenie to odgrywa decydującą rolę przy rozstrzygnięciu, czy zaszkodzenie środowisku przez dyrektora było intencjonalne. W świetle hipotezy o zaniechaniach można przyjąć, że (ZK_s) zaburza naturalne odczytanie tezy o wiedzy (K_s) jako tezy predyktywnej (PK_s), ponieważ (ZK_s) jest istotne dla stwierdzenia intencjonalności zaniechania. Teza (K_s) jest odczytywana jako skrót myślowy zdania (ZK_s).

Można wysunąć zarzut, że przeciw historyjka o pomaganiu jest jednak w pełni symetryczna względem historyjki o uszkodzeniu. W wypadku historyjki o uszkodzeniu zdanie o wiedzy (K_s) można interpretować jako zdanie predyktywne (PK_s) lub jako zdanie o znajomości możliwej negatywnej konsekwencji (ZK_s). W wypadku scenariusza o pomaganiu zdanie o wiedzy (K_p) można jednak również interpretować na dwa sposoby: jako zdanie predyktywne (PK_p) lub jako zdanie o znajomości możliwej — tym razem pozytywnej — konsekwencji działania dyrektora:

(ZK_p) Dyrektor wiedział, że [możliwą pozytywną konsekwencją jego działania jest to, że] nowy program pomoże środowisku.

Zdanie (ZK_p) jest prawdziwe w scenariuszu z pomaganiem. Co więcej, jest ono w taki sam sposób niezależne od zdania predyktywnego (PK_p), w jaki (ZK_s) jest niezależne od (PK_s).

Prawdą jest oczywiście, że również w wypadku pomagania dostępna respondentom (i prawdziwa) jest teza, że dyrektor zdawał sobie sprawę z możliwej pozytywnej konsekwencji swego działania. Jednak tylko w wypadku uszkodzenia teza, że dyrektor zdawał sobie sprawę z możliwej konsekwencji swego działania (ZK_s), odgrywa zasadniczą rolę przy ocenie intencjonalności. W wypadku pomagania decydującą rolę przy ocenie intencjonalności odgrywa zupełnie inna teza, a mianowicie dotycząca tego, czy dyrektor realizował intencję pomagania środowisku. Prawdziwość (ZK_p) nie przesądza o intencjonalności pomagania.

Podsumowując, w historyjce o uszkodzeniu naturalne staje się pytanie, czy owo uszkodzenie było intencjonalne. O jego intencjonalności decyduje m.in. to, czy dyrektor zdawał sobie sprawę, że możliwą konsekwencją jego działania jest to, że zaszkodzi środowisku (ZK_s). Skoro (ZK_s) odgrywa tak istotną rolę w tym scenariuszu, to pytanie o (K_s), czyli o wiedzę dyrektora dotyczącą tego,

czy uruchomienie programu zaszkodzi środowisku, które zwykle byłoby interpretowane predyktywnie jako (PK_s), przybiera w tym wypadku interpretację (ZK_s). (K_s) jest odczytywane jako skrót myślowy narzucającej się w tej sytuacji tezy (ZK_s). W historyjce o pomaganiu nie występuje żadna analogiczna presja, która mogłaby skłaniać do interpretowania (K_p) inaczej niż zgodnie z naturalną interpretacją predyktywną (PK_p).

5. WYJAŚNIENIE EKSPERYMENTÓW Z NISKIM PRAWDOPODOBIENSTWEM SZKODZENIA

W historyjkach Knobe'a prawdopodobieństwo tego, że stan środowiska się pogorszy lub polepszy, jest stosunkowo wysokie – wicedyrektor stwierdza, że program *zaszkodzi* (ew. *pomoże*) środowisku. Jeżeli przedstawiona interpretacja tezy (K_s) jako tezy raczej o znajomości konsekwencji (ZK_s) niż jako tezy o tym, co się stanie (PK_s), jest słuszna, to efekt epistemiczny powinien występować, nawet jeżeli prawdopodobieństwo wystąpienia konsekwencji będzie małe.

W rzeczy samej w jednym z testów przeprowadzonych przez Beebe'ego i Jensena (2012) badacze manipulowali prawdopodobieństwem skutku ubocznego – zmniejszając prawdopodobieństwo wystąpienia skutku negatywnego (szkodzenia), a zwiększając prawdopodobieństwo wystąpienia skutku pozytywnego (pomagania). W warunku niskiego prawdopodobieństwa szkodzenia zastąpiono słowa wicedyrektora wygłaszane w pierwotnej historyjce, tj. „[Program p]ozwoli nam zwiększyć zyski, ale zaszkodzi środowisku”, słowami:

Jesteśmy pewni, że program pozwoli nam zwiększyć zyski, ale istnieje *niewielkie prawdopodobieństwo (slight chance)*, że zaszkodzi środowisku.

W warunku wysokiego prawdopodobieństwa pomagania wicedyrektor mówił:

Jesteśmy pewni, że program pozwoli nam zwiększyć zyski, a istnieje też *bardzo duże prawdopodobieństwo (very strong chance)*, że pomoże środowisku.

Badani odpowiadali na pytanie, czy zgadzają się z tym, że dyrektor wiedział, że program zaszkodzi bądź pomoże środowisku, ponownie zaznaczając swoje odpowiedzi na siedmiostopniowej skali Likerta (gdzie „-3” oznaczone było jako „zdecydowanie się nie zgadzam”, „0” – jako „odpowiedź neutralna”, a „+3” jako „zdecydowanie się zgadzam”). Ponownie wystąpiły statystycznie

istotne różnice między odpowiedziami badanych w scenariuszach o pomaganiu i o szkodzeniu. Wbrew oczekiwaniom efekt epistemiczny nie zniknął. Badani byli skłonni przypisywać wiedzę dyrektorowi nawet w historyjce z niskim prawdopodobieństwem szkodzenia (średnia odpowiedzi wynosiła +1,15) częściej niż w scenariuszu z wysokim prawdopodobieństwem pomagania (średnia odpowiedzi +0,3).

Dalbauer i Hergovich (2013) powtórzyli badanie Beebe'ego i Jensena, używając skali analogowej kodowanej w zakresie od 1 do 100. Ponownie uzyskali istotną statystycznie różnicę między odpowiedziami w historyjce z niskim prawdopodobieństwem szkodzenia (średnia 96,1; SD = 15,71) a odpowiedziami w scenariuszu z wysokim prawdopodobieństwem pomagania (średnia 87,1; SD = 26,65)⁴.

Badania w paradygmacie „niskiego/wysokiego prawdopodobieństwa” były przeprowadzane albo za pomocą skali Likerta, albo skali analogowej. Nie były prowadzone w formie wymuszonego wyboru. Niemniej dane przedstawione przez Beebe'ego i Jensena (2012) można skategoryzować w taki sposób, by zobaczyć, ile procent osób było skłonnych zgodzić się z twierdzeniem o wiedzy (osoby te zaznaczały +1, +2 lub +3 na skali), a jaki procent osób nie zgadzał się z twierdzeniem o wiedzy (wybierając -1, -2 lub -3 na skali). Tabela 3 przedstawia te wyniki. Na jej podstawie można obliczyć „bezwzględny” odsetek osób zgadzających się / niezgadzających się z tezą o wiedzy w stosunku do wszystkich N osób biorących udział w badaniu. Można też obliczyć „względny” odsetek osób zgadzających / niezgadzających się z tezą o wiedzy w stosunku do osób, które nie wybrały odpowiedzi neutralnej, tj. albo zgodziły się, albo nie zgodziły się z tą tezą.

Warto zwrócić uwagę, że odsetek osób, które zgadzają się z twierdzeniem o wiedzy (K_s) w warunkach niskiego prawdopodobieństwa szkodzenia (70%-76%), jest zbliżony do odsetka osób, które zgadzają się z twierdzeniem o wiedzy w oryginalnym eksperymencie Beebe'ego i Jensena przeprowadzonym w formie wymuszonego wyboru (68%, por. Tabela 2). Należy jednak mieć na względzie to, że ani wyliczenie absolutnego, ani względnego odsetka osób

⁴ Dalbauer i Hergovich przeprowadzili to badanie, zmieniając historyjkę na taką, w której badani mieli sami dokonać przypisania odpowiednich prawdopodobieństw (niskiego w historyjce o szkodzeniu i wysokiego w historyjce o pomaganiu). Badacze sądzili bowiem, że jawne stwierdzenia dotyczące prawdopodobieństwa wcale nie muszą być w taki sposób odczytywane przez badanych. Np. komunikat „Jutro będzie padać” może wiązać się z większym prawdopodobieństwem opadów niż komunikat „Jest wysoce prawdopodobne, że jutro będzie padać”. Mimo przeformułowania scenariusza uzyskali jednak wyniki bardzo zbliżone do wyników na podstawie historyjki Beebe'ego i Jensena. W warunkach niskiego prawdopodobieństwa szkodzenia średnia wynosiła 95,4 (SD = 17,86), a w warunkach wysokiego prawdopodobieństwa pomagania 87,2 (SD = 24,07).

zgadzających się z twierdzeniem o wiedzy nie da tych samych wyników co eksperyment przeprowadzony w formie wymuszonego wyboru. Na podstawie danych przedstawionych przez Beebe'ego i Buckwaltera (2010) można podać względny odsetek osób, które zgadzały się z twierdzeniem o wiedzy w wypadku historyjki o szkodzeniu (wynosił on 92%) oraz względny odsetek osób, które zgadzały się z twierdzeniem o wiedzy w wypadku historyjki o pomaganiu (70%)⁵.

	NISKIE PRAWDOPODO- BIENSTWO SZKODZENIA, N=46	WYSOKIE PRAWDOPODO- BIENSTWO POMAGANIA, N=50
Dyrektor wiedział	32 (69,6%) [76,2%]	23 (46%) [53,5%]
Dyrektor nie wiedział	10 (21,7%) [23,8%]	20 (40%) [46,5%]

Tabela 3. Epistemiczny efekt Knobe'a w paradygmacie mieszanego prawdopodobieństwa (Eksperyment 5; Beebe, Jensen 2012). Podane trzy liczby reprezentują liczbę osób, które zgadzają się / nie zgadzają się z tezą o wiedzy. W nawiasie zwykłym podany jest „bezwzględny” odsetek osób zgadzających się / niezgadzających się z tą tezą, w nawiasie kwadratowym „względny” odsetek osób zgadzających się / niezgadzających się z tą tezą.

Wszystkie te wyniki są osobliwe, lecz pozostają w zgodzie z przedstawioną przeze mnie hipotezą. W wypadku historyjki o pomaganiu teza o wiedzy dyrektora (K_p) jest interpretowana predyktownie. Wyniki te są zgodne z ujęciem standardowym i pokazują, że ludzie nie uznają zapewnień o wysokim prawdopodobieństwie wystąpienia pewnego stanu rzeczy za silną rację, by uznać, że ten stan rzeczy wystąpi.

Wyniki uzyskane w warunkach niskiego prawdopodobieństwa szkodzenia należy interpretować przez pryzmat znajomości negatywnych konsekwencji, która jest istotna dla oceny intencjonalności zaniechania niedoprowadzenia do tych negatywnych konsekwencji. Wiedza o negatywnych konsekwencjach jest niezależna od prawdopodobieństwa wystąpienia tych konsekwencji, co wyjaśnia, dlaczego efekt epistemiczny utrzymuje się w warunkach niskiego prawdopodobieństwa szkodzenia.

⁵ Badacze zebrali wyniki osób, które zgadzały się z twierdzeniem o wiedzy (zaznaczały +1, +2, +3), oraz wyniki osób, które z tym twierdzeniem się nie zgadzały (zaznaczały -1, -2, -3). Pierwszej grupie przypisali nową wartość 1, a drugiej grupie wartość 0 i obliczyli nowe średnie i odchylenia standardowe. Stosując tę metodę, otrzymali średnią 0,92 (SD = 0,27) w historyjce o szkodzeniu oraz średnią 0,72 (SD = 0,46) w historyjce o pomaganiu. Na tej podstawie można jedynie oszacować względny odsetek osób zgadzających się / niezgadzających się z tezą o wiedzy.

Wkrada się tu pewna wątpliwość⁶: chociaż w eksperymentach z mieszanym prawdopodobieństwem efekt epistemiczny się utrzymuje, to spada stopień zgody badanych z twierdzeniem przypisującym dyrektorowi wiedzę, że program zaszkodzi środowisku. W wypadku historyjki z niskim prawdopodobieństwem szkodzenia średnia odpowiedzi na pytanie o wiedzę spada z 2,25 do 1,15, a w scenariuszu z wysokim prawdopodobieństwem pomagania — z 0,91 do 0,3. Skoro efekt epistemiczny jest wyjaśniany przez interpretację tezy o wiedzy jako tezy o znajomości konsekwencji (ZK_s), a nie jako tezy o tym, co się stanie (PK_s), i dodatkowo twierdzi się, że (ZK_s) i (PK_s) są od siebie niezależne, to należałoby oczekiwać, że manipulacja prawdopodobieństwem zajścia efektu nie wpłynie na wyniki pytania o wiedzę. Skoro wpływa, to hipoteza o zaniechaniach nie wyjaśnia jednak tego efektu.

Należy zwrócić uwagę, że teza dotycząca niezależności (ZK_s) i (PK_s) jest tezą o logicznej niezależności dwóch zdań w sensie logicznym. Niewykluczone jednak, że zgoda badanych na uznanie bądź odrzucenie takiego zdania jak (K_s) jest wypadkową skłonności badanych do uznania zarówno zdania (ZK_s), jak i zdania (PK_s), a być może jeszcze innych zdań w sensie logicznym, które można by wyróżnić. Jeżeli tak jest, to spadek w średniej odpowiedzi na pytanie o wiedzę w wypadku historyjki o szkodzeniu można wyjaśnić nie tyle zmianą w skłonności badanych do uznania (ZK_s) — ta powinna być niezmienna — ile zmianą w skłonności badanych do uznania (PK_s). Ponieważ w historyjce o szkodzeniu prawdopodobieństwo zaszkodzenia środowisku jest określone jako niewielkie, badani powinni w jeszcze mniejszym stopniu niż w wypadku oryginalnych scenariuszy Knobe'a być skłonni przyjąć twierdzenie (PK_s), zgodnie z którym dyrektor wie, że *jest wysoce prawdopodobne*, iż program zaszkodzi środowisku.

Ponieważ hipoteza o zaniechaniach stosuje się wyłącznie do sytuacji szkodzenia, to nie wyjaśnia wspomnianego spadku w odniesieniu do historyjki o pomaganiu. Natomiast można go tłumaczyć tym, że badani przypisują wyższe prawdopodobieństwo zajścia efektu w oryginalnej historyjce Knobe'a, w której mówi się, że pewien efekt zajdzie, niż w historyjce, w której mówi się, że istnieje wysokie prawdopodobieństwo tego, że zajdzie (por. też Dalbauer, Hergovich 2013).

⁶ Dziękuję anonimowemu recenzentowi za podniesienie tej kwestii.

WNIOSKI

Okazuje się więc, że hipoteza o zaniechaniach pozwala wyjaśnić również epistemiczny efekt Knobe'a. W historyjce o uszkodzeniu badani mają powody, by skupiać się na tym, czy dyrektor zdaje sobie sprawę z możliwych negatywnych konsekwencji swego działania: tylko wtedy, gdy zdawał sobie sprawę z tego, że możliwą konsekwencją jego działania jest zaszkodzenie środowisku, można przypisać jego działaniu intencjonalność (zaniechania nieszkodzenia środowisku). Teza o wiedzy dyrektora (K_s) jest odczytywana właśnie w tym kontekście – mianowicie jako teza o znajomości możliwej negatywnej konsekwencji (ZK_s), a nie jako teza predyktywna (PK_s). Dzięki temu, że teza o znajomości możliwej konsekwencji jest niezależna od tezy o znajomości prawdopodobieństwa wystąpienia tej konsekwencji, można zrozumieć, dlaczego efekt epistemiczny utrzymuje się mimo manipulacji prawdopodobieństwem wystąpienia negatywnego efektu. Nawet jeżeli prawdopodobieństwo to jest niskie, to badani są i tak skłonni przypisywać dyrektorowi wiedzę – przypisują bowiem dyrektorowi znajomość możliwych negatywnych konsekwencji jego działania.

BIBLIOGRAFIA

- Beebe J. R., Buckwalter W. (2010), *The Epistemic Side-Effect Effect*, „Mind and Language” 25(4), 474-498.
- Beebe J. R., Jensen M. (2012), *Surprising Connections between Knowledge and Action. The Robustness of the Epistemic Side-Effect Effect*, „Philosophical Psychology” 25(5), 689-715.
- Bratman M. E. (1987), *Intention, Plans, and Practical Reason*, Cambridge, MA: MIT Press.
- Dalbauer N., Hergovich A. (2013), *Is What Is Worse More Likely? The Probabilistic Explanation of the Epistemic Side-Effect Effect*, „Review of Philosophy and Psychology” 4(4), 639-657.
- Davidson D. (1980), *Essays on Actions and Events*, Oxford: Oxford University Press.
- Feinberg J. (1984), *Harm to Others. The Moral Limits of the Criminal Law*, Oxford: Oxford University Press.
- Ginet C. (1990), *On Action*, Cambridge: Cambridge University Press.
- Goldman A. I. (1970), *A Theory of Human Action*, Princeton, NJ: Princeton University Press.
- Gorr M. (1979), *Omissions* [w:] *Tulane Studies in Philosophy. Studies in Action Theory*, t. 28, R. C. Whittlemore (red.), New Orleans, LA: Tulane University Press, 93-102.
- Gorr M., Horgan T. (1982), *Intentional and Unintentional Actions*, „Philosophical Studies” 41(2), 251-262.
- Harman G. (1976), *Practical Reasoning*, „The Review of Metaphysics” 29(3), 431-463.

- Hindriks F. (2008), *Intentional Action and the Praise–Blame Asymmetry*, „The Philosophical Quarterly” 58(233), 630-641.
- Hindriks F. (2011), *Control, Intentional Action, and Moral Responsibility*, „Philosophical Psychology” 24(6), 787-801.
- Holton R. (2010), *Norms and the Knobe Effect*, „Analysis” 70(3), 417-242.
- 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” 33, 309-324.
- Knobe J. (2007), *Reason Explanation in Folk Psychology*, „Midwest Studies in Philosophy” 31(1), 90-106.
- Mele A. R., Moser, P. K. (1994), *Intentional Action*, „Noûs” 28(1), 39-68.
- Nichols S., Ulatowski J. (2007), *Intuitions and Individual Differences. The Knobe Effect Revisited*, „Mind and Language” 22(4), 346-365.
- Paprzycka K. (2014), *Rozwiązanie problemu Butlera i wyjaśnienie efektu Knobe’a*, „Filozofia Nauki” 22(2) [86], 73-96.
- Paprzycka K. (2015), *The Omissions Account of the Knobe Effect and the Asymmetry Challenge*, „Mind and Language” 30(5), 550-571.
- Paprzycka K. (2016a), *Intention, Knowledge, and Disregard for Norms. The Omissions Account and Holton’s Account of Asymmetrical Intentionality Attributions [w:] Uncovering Facts and Values. Studies in Contemporary Epistemology and Political Philosophy (Poznań Studies in the Philosophy of the Sciences and the Humanities, t. 107)*, A. Kuźniar, J. Odrowąż-Sypniewska (red.), Leiden: Brill–Rodopi, 204-233.
- Paprzycka K. (2016b), *Methodological Reflections on Academic and Experimental Philosophy. The Case of the Omissions Account [w:] Evidence, Experiment and Argument in Linguistics and Philosophy of Language*, M. Hinton (red.), Frankfurt am Main: Peter Lang, 27-42.
- Paprzycka K. (2016c), *O intencjonalności działań i zaniechań, czyli o społecznej naturze sprawstwa*, „Przegląd Filozoficzny — Nowa Seria”, 1 (97), 45-66.
- Smith P. G. (1990), *Contemplating Failure. The Importance of Unconscious Omissions*, „Philosophical Studies” 59(2), 159-176.
- Smith P. G. (2005), *Feinberg and the Failure to Act*, „Legal Theory” 11(3), 237-250.
- Williams B. (1995), *Acts and Omissions, Doing and Not Doing [w:] Making Sense of Humanity and Other Philosophical Essays*, Cambridge: Cambridge University Press, 56-64.
- Wilson G. (1989), *The Intentionality of Human Action*, Stanford, CA: Stanford University Press.
- von Wright G. H. (1963), *Norm and Action*, London: Routledge and Kegan Paul.