

The Problem of Intention and the Evaluative Properties of Effects in the Knobe Effect


Andrzej Waleszczyński* – Michał Obidziński** – Julia Rejewska***

Received: 30 July 2020 / Revised: 20 July 2021/ Accepted: 1 December 2021

Abstract: In the article, we present analyses and findings which add precision to the role of intentions and the relation between effects in attributing the intentionality of causing a side effect. Our research supplements and modifies numerous findings regarding the appearance of the so-called Knobe effect. The experiments and analyses show that the very originality of the story used by Knobe and the relationship between the evaluative properties of the main effect and the side effect results in an asymmetry of responses and contributes

* Cardinal Stefan Wyszyński University


 <https://orcid.org/0000-0003-0426-3919>

 Corresponding author. Institute of Philosophy, Faculty of Christian Philosophy, Cardinal Stefan Wyszyński University, Wóycickiego 1/3, 01-938 Warsaw, Poland

 a.waleszczyński@uksw.edu.pl

** Cardinal Stefan Wyszyński University


 <https://orcid.org/0000-0002-7854-3123>

 Institute of Philosophy, Faculty of Christian Philosophy, Cardinal Stefan Wyszyński University, Wóycickiego 1/3, 01-938 Warsaw, Poland

 m.m.obidziński@gmail.com

*** Cardinal Stefan Wyszyński University

 <https://orcid.org/0000-0001-5740-6608>

 Institute of Philosophy, Faculty of Christian Philosophy, Cardinal Stefan Wyszyński University, Wóycickiego 1/3, 01-938 Warsaw, Poland

 julia.rejewska@gmail.com



to the occurrence of the side-effect effect. Because of this, we reject the thesis that the mode of attitude of the agent to the caused side effect or that the social expectation of this attitude determine the attribution of the intentionality of the caused effect. On the contrary, we defend the thesis that it is the relationship between the evaluative properties of the main effect and those of the side effect, as well as the impact of a side effect on the main effect, that significantly influence the attribution of intentionality in causing a side effect.

Keywords: Side-effect effect; Knobe effect; intention; intentionality; evaluative properties.

1. Introduction

Studies on intentionality are of key importance for the philosophy of action (Mele 1992, 199). Standard accounts of intentionality show that the attribution of intentionality is based on clearly definable descriptive properties of the situation or of the agent (Malle and Knobe 1997; Mele and Sverdlik 1996). From this perspective, intentional causation of an effect is possible if the agent had the intention of causing it (Adams 1986; McCann 1987).

In recent years, a considerable contribution to studies on intentionality has been made by experimental philosophy, in particular studies on the attribution of intentionality in causing side effects. Philosophers and psychologists are very much interested in those experiments which analyze the so-called side-effect effect, also called the Knobe effect (Knobe 2003, 2006; Nadelhoffer 2005, 2006; Nichols and Ulatowski 2007; Nado 2008; Guglielmo and Malle 2010; Uttich and Lombrozo 2010). The effect reveals an asymmetry in the attribution of intentionality. It turns out that people are more apt to attribute intentionality in causing a side effect when the effect is negative than when it is positive. This has led to an alternative view which suggests that the attribution of intentionality in causing a side effect may also depend on its morally negative properties.

A detailed analysis of the impact of the agent's attitude on the side effects caused and the relationship between the moral weight of the main effect and the side effect was presented by Joshua Shepherd (Shepherd

2012). Further studies describing and analyzing the significance of the relationship between the main effect and the side effect (Waleszczyński, Obiźniński, and Rejewska 2019) have provided new interesting data on the emergence and disappearance of the Knobe effect. They have shown that the attribution of intentionality in causing a side effect is also possible in the case of positive effects and when the agent does not care about causing it. These findings are problematic for the explanations of the Knobe effect provided so far (Nadelhoffer 2006b; Nichols and Ulatowski 2007; Wright and Bengson 2009; Holton 2010; Sripada 2010; Sripada and Konrath 2011; Cova et al. 2012, Paprzycka 2015; Hindriks 2011, 2014; Hindriks, Douven, and Singmann 2016).

In this article, we will focus first on examining the relevance of the intention of causing a side effect for the attribution of intentionality to the agent. Our analyses will be concerned with the attribution of intentionality in causing a side effect, and not the attribution of intentionality in the action as such. Consequently, our conclusions will not apply to the intentionality of causing the main effect, but to that of causing side-effects. We will present analyses and findings which add precision to the role of intentions in attributing the intentionality of causing a side effect. Second, we will focus on further empirical investigations of the interaction between the importance of the main effect and the valence of the side effect (Cova et al. 2012, 402). We will defend the thesis that it is the relationship between the moral evaluative properties of the main effect and those of the side effect, and the impact of a side effect on the main effect that significantly influence the attribution of intentionality in causing a side effect.

2. Concepts of intentional action

Analyses in the field of experimental philosophy, despite having analytical elements, are classified as experimental descriptivism (Nadelhoffer and Nahmias 2007). Therefore, it is necessary to emphasize the distinctiveness of research in the field of experimental philosophy from conceptual analysis, the aim of which would be mainly to determine the conditions for the application of appropriate concepts (Knobe and Nichols 2008, 5). In his research, Knobe searches for cognitive mechanisms shaping popular intuitions

that would satisfactorily explain the relevant mental processes (Knobe 2016). For this purpose, he uses empirical research in which participants are confronted with scenarios based on thought experiments. From this perspective, experimental philosophy is part of the cognitive sciences (Piekarski 2017, 112).

Research on the concept of intentional action conducted within the framework of experimental philosophy focuses on the study of the relationship between the folk concept of intentional action, i.e. one that is part of folk psychology, and the philosophical concept of intentional action. They show that the folk concept of intentional action does not correspond to the concept used by philosophers and is not limited to the concepts of intention and prediction. In the folk concept of intentional action, its effects and how we evaluate them are also taken into account (Piekarski 2017, 115).

According to Fred Adams and Annie Steadman, there are basically two concepts of intentional action. The first assumes that person S intentionally performs action A only when S intends to do A. In this case, having an intention is a necessary condition of an (intentional) action. According to the second concept, person S intentionally performs action A, not intending to do A, as long as action A is predicted by person S and accepted as a consequence of action S. In this case, the action may be intentional, even though the person has no intention of doing it (Adams 1986; Adams and Steadman 2004a, 2004b). However, the multiplicity and complexity of the results of research conducted as part of the experimental philosophy on the concept of intentional action generate the formulation of new explanations that go beyond the above two concepts of intentional action. Taking into account publications from recent years, one should also mention the explanations regarding the intentional actions provided by the representatives of responsibility (Paprzycka 2012, 476–77). Herbert Hart, in his famous article, ‘The Ascription of Responsibility and Rights’ (1949), that action statements are not action statements, but rather ascriptive statements. In other words, they assign responsibility for certain events in the world before assigning them intentionality. Although this position was rightly criticized, it is worth emphasizing the main intuition presented by Hart, which may be significant for the analysis of the folk concept of intentional action. It con-

cerns the fact that the concept of action is a secondary concept to the concept of responsibility (Sneddon 2006). This leads to a situation where the intentionality of the action may be attributed to person S due to the prior attribution of responsibility to person S for action A, even in a situation in which he had no intention of causing action A, and did not even anticipate it. As a consequence, intentionality may be assigned due to a breach of obligations or applicable standards (Paprzycka 2015). This means that apart from intention and prediction, there may appear normative and evaluative factors that affect the attribution of intentionality not only to actions, but also to the side effects they cause.

3. Knobe Effect

Joshua Knobe presented a story in which the “HARM” condition was as follows:

The vice-president of a company went to the chairman of the board and said, ‘We are thinking of starting a new program. It will help us increase profits, but it will also harm the environment.’ The chairman of the board answered ‘I don’t care at all about harming the environment. I just want to make as much profit as I can. Let’s start the new program.’ They started the new program. Sure enough, the environment was harmed. (Knobe 2003, 191)

In the “HELP” condition, the structure of the story was the same, with the only difference being that the side effect is positive—it will “help the environment.”

The vice-president of a company went to the chairman of the board and said, ‘We are thinking of starting a new program. It will help us increase profits, but it will also help the environment.’ The chairman of the board answered ‘I don’t care at all about helping the environment. I just want to make as much profit as

I can. Let's start the new program.' They started the new program. Sure enough, the environment was helped. (Knobe 2003, 191)

After being presented with the story, depending on the scenario each respondent was asked a question: "Did the chairman intentionally harm the environment?" or "Did the chairman intentionally help the environment?". It turned out that in the HARM condition, 82% of the respondents attributed intentionality of causing the side effect. In the HELP condition, on the other hand, only 23% of respondents believed that the environment was helped intentionally.

The results of this experiment have been repeatedly confirmed (Knobe 2004b, 2004a; Nadelhoffer 2004b, 2004a; Mele and Cushman 2007) also in studies on children (Leslie, Knobe, and Cohen 2006) and with adults with high functioning autism or Asperger's syndrome (Zalla and Leboyer 2011) as well as in various languages, including Hindi (Knobe and Burra 2006), German (Dalbauer and Hergovich 2013), Polish (Kuś and Maćkiewicz 2016; Waleszczyński, Obidziński, and Rejewska 2018).

4. Structure of the story and the problem of intention

Knobe's story is based on a predefined structure. The CEO is only concerned with A (the main effect) and does not care about B (the side effect). In the HARM condition the side effect (B1) is negative, and in the HELP condition the side effect (B2) is positive. It does not result from the structure of the story itself that there is any difference between the likelihood of effects A or B, or that of effects B1 and B2. And yet, studies have shown (Nakamura 2018) that the kind of side effect (positive, negative) alone determines the attribution of likelihood. Negative effects are considered more likely than positive ones. From this perspective, the difference between the two stories is not only related to the content of the story itself, but also to expectations concerning the likelihood of a side effect. This suggests that certain notions and descriptions which refer to (morally) positive and negative effects are related not only to the purely descriptive function of the language, but also to its evaluative and normative one. Authors studying

the Knobe effect point out that it is not so much the difference on the descriptive plane of the language, but that on the evaluative and normative one that causes the Knobe effect.

The problem of intention plays an important role in the problem of intentionality. Therefore, it is necessary to analyze the attitude of the agent to a particular side effect. This attitude may be threefold: (1) I want it; (2) I do not want it; (3) I do not care about it. In the last case, the attitude of the agent may be treated as indirect intention or consent for a particular side effect (Głowala 2013). There is also a fourth special case that should be considered, where the agent does not know about some possible consequences of his or her actions. More specifically, this describes a situation when the agent does not foresee a particular effect. In such a situation, the agent cannot be regarded as having intentionally caused that effect, since he or she did not expect to cause such an effect when taking the action, and thus could not have had the intention of causing it. Consider the following example. A and B are playing ball, throwing it to each other. While they are playing, a factor distracting the attention of B appears as a firefighting siren. Consequently, B is hit by the ball thrown by A, falls to the ground, and hurts himself on the head. Did A foresee such a complex coincidence? He did not. We may say that he had the intention to play ball with B, but not the intention to hurt his head.

The above distinction refers to the issue of intention. The situation is somewhat different, however, when we take into account the existence of obligations, that is, the influence of the normative factor on judging the intentionality of the agent, which needs to be determined in order to decide whether the agent is to be blamed or not. In such a situation, the attribution of obligation burdening the agent creates a fourth attitude to the side effect (4) (non-intentionally) causing an effect, for example by negligence or failure to act. This happens, for example, when a steersman leaves the wheel, goes below deck to play cards, and the ship is wrecked. In such circumstances, even if he did not expect any danger to occur, he will be charged with intentional default on a duty. In other words, the agent's attitude is referred to the (intentional) negligence to fulfill an obligation which resulted in a particular side effect (the main effect is playing cards, the side effect is the shipwreck). In situation (4), however, the intention is referred not so

much to causing an effect (since the agent did not foresee effect [3]), but rather to violating a norm resulting from a duty. In some explanations of the Knobe effect (Holton 2010; Paprzycka 2015), it is argued that the attribution of intentionality results precisely from the attribution of the intention to violate a norm, or the intention to fulfill a (social) norm. Such a simplification seems to be unsubstantiated, however, as reference to a norm helps evaluate the (moral or legal) responsibility for causing an effect, rather than the intentionality of causing it, even if we consider the side effect to have been foreseeable (3). The story used in Knobe's experiment cannot be referred directly to situation (4), since it results from the structure of his story that the agent foresaw the occurrence of a particular effect. Consequently, it is not necessary to refer to a norm in order to attribute intentionality or to judge a side effect as having been caused intentionally. Unless we assume, in line with what is argued by Holton and Paprzycka, that the power of a (social) norm affects the attribution of intentionality to such an extent that it overrides (suspends) the intention explicitly stated in the analyzed stories. It should also be assumed that the attribution of intentionality is secondary to the attribution of responsibility; this, however, would stand in opposition to other studies (Nadelhoffer 2004c) which have shown that causation of a negative side effect and attribution of responsibility does not need to be coupled with attribution of intentionality.

Studies by Waleszczyński, Obidziński and Rejewska (2019) have shown, however, that the thesis about the key significance of the norm for the attribution of intentionality in causing a side effect in the Knobe effect does not apply in one particular case – namely, when there is a significant difference¹ in weight between the evaluative properties of the main effect and the evaluative properties of the side effect. This does not mean that (social) norms do not influence the attribution of intentionality in causing a side effect. Their impact is indirect, however, and relates to the determination of the evaluative-normative meaning of a particular situation and the

¹ Assessment of this difference is problematic. Authors of this article make it based on prevailing convictions about the value of particular effects as seen against other effects. In the stories used in the experiments, the emphasis was mainly on the relationship between the moral weight of the main effect and the moral weight of the side effect.

caused effects. Thence the importance of studies which aimed at a more precise identification of the role of intentions in the attribution of intentionality in causing a side effect.

5. The influence of moral factors—research findings to date and solutions

The observed side-effect effect has been explained by Knobe (Knobe 2004a, 2006) as the influence of the moral valuation of side effects on the judgment of intentionality in causing them. In his opinion, people have a tendency to attribute intentionality when the side effect is negative. The attribution of intentionality in such circumstances may be influenced by moral factors. Studies by Hindriks (Hindriks et al. 2016) show that the influence of moral valuation is only observable in the HARM condition and does not explain judgments made in the HELP condition. Hindriks's comment is significant in that it shows that we should not refer to the influence of moral valuation in general, but only to the influence of (morally) negative effects on the attribution of intentionality. In other words, only a part of moral reality may influence judgments concerning intentionality. Negative moral valuation may be the very factor which influences the occurrence of asymmetric judgments in attributing the intentionality of causing a side effect.

Some of the hypotheses proposed so far have linked the occurrence of asymmetry in the attribution of intentionality in causing a side effect with a praise-blame asymmetry (praise-blame asymmetry) (Malle and Nelson 2003; Nadelhoffer 2004a, 2006b, 2006a; Nado 2008; Hindriks 2008, 2011), and with the attribution of responsibility for the caused negative side effects (Wright and Bengson 2009). These explanations only explain the HARM condition in the Knobe effect, and make the attribution of intentionality dependent on the attribution of blame or responsibility. Other hypotheses have tried to explain the side-effect effect by referring to social norms, which in the story proposed by Knobe would concern the prohibition of harming the environment. It is the intentional violation of such a norm (Holton 2010), or its intentional neglect (Paprzycka 2015, 2016), that is supposed

to determine the attribution of intentionality in causing a side effect. An important question that may be asked concerns the character of the norm to which the respondents are supposed to be referring to. Is it a legal, social, or moral norm, and does the character of this norm matter? Finding an answer to these questions would be relevant to the hypothesis regarding the influence of evaluative properties on the attribution of intentionality. If the violated norm did not need to be a moral, but could for example be a legal one, one could hardly claim that moral norms affect the attribution of intentionality.

There are also other explanatory hypotheses which point to the relationship between the good or the evil of side effects and the values to which the agent refers when making a judgment (Sripada 2010, 2012; Sripada and Konrath 2011). The values and attitudes endorsed by the agent form the structure of his or her deep Self. Consequently, the respondents who attribute anti-environmental values to the CEO also attribute intentionality in the HARM condition. This hypothesis also naturally allows for a gradation in the attribution of intentionality. As pointed out by Hindriks (2016, 215). However, this hypothesis is symmetrical, since when it is interpreted in categorical terms, it only explains the HELP condition, and when it is interpreted in graded terms, it only explains the HARM condition. It may be observed, however, that in both interpretations, the role of intentions and of the reference to the side effect itself are secondary. What moves to the forefront is the attribution of a particular structure of values to the agent.

A different explanation of the Knobe effect is proposed by Hindriks. He points out that the CEO is attributed a degree of indifference to the side effect he causes. On this basis, Hindriks proposes his Normative Reason Hypothesis. He suggests the existence of a normative reason which is the obligation to care about the side effect caused. The attribution of indifference to the agent is gradable and may extend from the attitude of neutrality to that of full care. Hindriks's hypothesis does not so much refer to intentions as it does to the attitude (care) of the agent to (about) the side-effect he causes. It is the degree of this attitude that is supposed to explain the attribution of intentionality in causing a side effect. The problem of this explanation lies on the theoretical side, however. It assumes that no normative reason exists for positive side effects. In other words, it presents a

reality in which people are supposed to care about negative side effects, but are not required to care about positive ones.

The findings made so far are complicated by the results of studies which show that it is possible for intentionality of causing a side effect to be attributed in the HELP condition although the agent does not care about the side effect. Waleszczyński, Obidziński and Rejewska (2019) have pointed out that the influence on the attribution of intentionality in causing a side effect may depend not so much on the moral weight of the side effect, but on the relationship between the moral weight of the main effect and that of the side effect. Their studies show that if in Knobe's story we swap the main effect for the side effect, then despite a statistical difference in answers provided in the two conditions, they become symmetrical and reveal a tendency in the attribution of intentionality.² Such experimental findings undermine most of the hypotheses explaining the Knobe effect proposed so far. The attribution of intentionality in causing a side effect in the HELP

² The results below present data combined from two experiments downloaded from: <https://osf.io/ky3re/>. They have been combined for the purposes of further comparative analyses. Waleszczyński, Obidziński and Rejewska presented a study of the three types of the stories.

The first story was taken from the Knobe study (2003: 191). Results: $M_{\text{Harm}} = 1.645$, $M_{\text{Help}} = -1.113$, $t(122) = 7.552$, $p < 0,001$, $d = 0.864$.

The second one, was the original story created after the pattern of the Knobe's story: "The Deputy of Experimental Oncology Hospital asks the director: "We can produce a drug that will heal patients with pancreatic cancer, but it would cause pneumonia/and cure pneumonia". The director responds: "I want to primarily cure patients with pancreatic cancer. We start production and give medicine to patients." The drug has been given and has caused/cured pneumonia." Results: $M_{\text{Harm}} = 0.603$, $M_{\text{Help}} = -0.587$, $t(124) = 3.195$, $p = 0,002$, $d = 0.317$.

The third story was the modification of the Knobe's story: "The vice-president of a company went to the chairman of the board and said, 'We are thinking of starting a new program. It will help the environment, but it will also increase/decrease profits.' The chairman of the board answered 'I don't care at all about profits. I just want to help the environment. Let's start the new program.' They started the new program. Sure enough, the profits were increase/decrees." Results: $M_{\text{Harm}} = 0.952$, $M_{\text{Help}} = 0.274$, $t(122) = 1.997$, $p = 0,048$, $d = 0.359$.

condition can no longer be explained by the attribution of blame, as suggested, e.g., by Malle and Nelson, Nadelhoffer, Nado, or Hindriks. In addition, as pointed out by Holton, it would be difficult to demonstrate the existence of a norm which could be violated in the HELP condition. Also, focusing on the agent's attitude to the side effect and duty to care about the side effect will not help much in explaining the observed relations. Similarly, the standard understanding of intentionality does not help explain the observed result, as the agent is indifferent to the expected side effect.

Joshua Shepherd (2012) presented studies that broadly analyzed the impact of the relationship between the main and side effects and the agent's attitude to said side effects. However, the appearance of new interesting data in studies on the Knobe effect makes it necessary not only to reconsider the significance of the relationship between the main effect and the side effect, but also to reconsider the significance of the role of intention in the attribution of intentionality, or more generally, the attitude of the agent to the side effect caused. Therefore, in our study we have focused on the significance of intention, or more precisely, lack of intention, in the attribution of intentionality in causing a side effect, and the importance of moral evaluation of the relationship between the main and side effects in attributing intentionality to causing a side effect.

6. The experiment

The experiment employed modified versions of three scenarios used in experiments performed by Waleszczyński, Obidziński and Rejewska, diversified in terms of the evaluative properties of the effects. In order to see what role for the attribution of intentionality in causing a side effect is played by the agent's attitude to that side effect, the stories had been modified so that it could not be inferred what the attitude of the agent was to the side effect he caused. The investigated hypothesis assumed that if it is not possible to determine the attitude of the agent to causing a side effect, and thus to know his intention, then the model response would be "Hard to say." The experiments were aimed at verifying this hypothesis. Removing

the agent's attitude to causing a side effect from the experiment will complement Shepherd's (2012) research and will also verify the attributing intentionality account proposed by Cova, Dupoux and Jacob (2012).

Verifying the hypothesis is crucial for the direction of further research. In philosophical conceptions, the intention of an action or the prediction of that action determines an intentional action. However, taking the example of responsibilism, we have indicated that the concept of intentional action (causing an effect), which uses the concept of intention, can be said in a way that is secondary to, for example, previously ascribed responsibility. Therefore, in our study, we pose only one question concerning the intentionality of causing a side effect in a situation where the respondents do not know either the intention or the attitude of the acting subject towards the side effect caused. Ascribing or not ascribing intentionality to inducing a side effect would mean that additional factors that influence the ascribing of intentionality to induce side effects, beyond those of intention and prediction, should be sought.

6.1. Methodology

The experiment was performed at two locations: the Cardinal Stefan Wyszyński University in Warsaw among students of various faculties, and at Warsaw Metro stations among randomly selected passengers. For the purposes of this experiment, results from both locations have been combined, and all analyses are performed on the resulting samples. Every quizzed person responded only once, after reading one story in one of the study conditions. Thus, the sum total of all respondents was 372.

In the experiment described above, the following stories and questions were used:

S1_Help

The vice-president of a company went to the chairman of the board and said, 'We are thinking of starting a new program. It will help us increase profits, but it will also help the environment.' The chairman of the board answered 'I just want to make as much profit as I can. Let's start the new program.' They started the new program. Sure enough, the environment was helped.

Question: Did the chairman intentionally help the environment?

S1_Harm

The vice-president of a company went to the chairman of the board and said, ‘We are thinking of starting a new program. It will help us increase profits, but it will also harm the environment.’ The chairman of the board answered ‘I just want to make as much profit as I can. Let’s start the new program.’ They started the new program. Sure enough, the environment was harmed.

Question: Did the chairman intentionally harm the environment?

S2_Help

The vice-president of an experimental oncological hospital went to the chairman of the board and said, ‘We are thinking of starting the production of a new medicine. It will help us cure patients of pancreatic cancer but it will also cure them of pneumonia.’ The chairman of the board answered, ‘I just want to cure patients of pancreatic cancer. Let’s start the production of a new medicine.’ They started the production of the new medicine. Sure enough, the patients were cured of pneumonia.

Question: Did the chairman intentionally cure pneumonia?

S2_Harm

The vice-president of an experimental oncological hospital went to the chairman of the board and said, ‘We are thinking of starting the production of a new medicine. It will help us cure patients of pancreatic cancer but it will also cause pneumonia.’ The chairman of the board answered, ‘I just want to cure the patients of pancreatic cancer. Let’s start the production of a new medicine.’ They started the production of the new medicine. Sure enough, the patients came down with pneumonia.

Question: Did the chairman intentionally cause pneumonia?

S3_Help

The vice-president of a company went to the chairman of the board and said, ‘We are thinking of starting a new program. It will help us help the environment, but it will also increase profits.’ The chairman of the board answered, ‘I just want to help the environment as much as I can.

Let's start the new program.' They started the new program. Sure enough, profits were increased.

Question: Did the chairman intentionally increase profits?

S3_Harm

The vice-president of a company went to the chairman of the board and said, 'We are thinking of starting a new program. It will help us help the environment, but it will also cause losses.' The chairman of the board answered, 'I just want to help the environment as much as I can. Let's start the new program.' They started the new program. Sure enough, losses were caused.

Question: Did the chairman intentionally cause losses?

Responses were given on a seven-point scale ranging from -3 to 3, where -3 meant "Absolutely not," 0 meant "Hard to say," and 3 meant "Absolutely yes." Participants responded using the classic paper-and-pencil method.

6.2. Results³

Tests for distribution normality from the results of our study show that all distributions are significantly different from the normal distribution. Thus, we have started our statistical analysis using nonparametric tests of differences. However, after the first analysis, we made another one using a parametric test and compared their results. We have found that obtained results are similar in both significance and effect size. Therefore, in the analysis report, we present the results of parametric t-tests.

To test our hypothesis, we have used a one-sample t-Student test that allows us to compare collected data with assumed values (e.g. from previous data or data based on a theoretical approach). We have compared observed means with scale grade 0 ("Hard to say") testing if there was a significant difference between the results of our studies and the value that is the model response of the tested hypothesis. There was a statistically significant difference in the distribution of results compared to standard normal distribution. Results of the one sample Wilcoxon test corresponded to the results

³ The results can be downloaded from the repository at: (Obidziński and Waleszczyński 2021) <https://osf.io/8c5qk>.

of t-Student test results though. For this reason, we decided to present the results using a parametric test. Descriptive statistics and test results are presented in Table 1.

	<i>M</i>	<i>SD</i>	<i>t</i>	<i>df</i>	<i>p</i>	<i>Cohen's d</i>
<i>S1_Harm</i>	2.2	1.277	13.887	64	< 0.001	1.723
<i>S1_Help</i>	-0.754	1.723	-3.527	64	< 0.001	0.298
<i>S2_Harm</i>	-0.613	2.059	-2.343	61	0.022	0.283
<i>S2_Help</i>	-0.661	2.172	-2.397	61	0.02	0.438
<i>S3_Harm</i>	0.565	1.997	2.266	61	0.03	0.304
<i>S3_Help</i>	0.629	1.84	-2.692	61	0.009	0.342

Table 1. The difference between mean results and scale grade 0

The difference between mean results in the HARM and HELP conditions were then compared for each story. The findings are presented in Table 2.

	<i>M_{Harm}</i>	<i>SD_{Harm}</i>	<i>M_{Help}</i>	<i>SD_{Help}</i>	<i>t</i>	<i>Df</i>	<i>p</i>	<i>Cohen's d</i>
<i>S1</i>	2.2	1.277	-0.754	1.723	11.102	118.01	< 0.001	1.153
<i>S2</i>	-0.613	2.059	-0.661	2.172	0,127	121.65	0.899	-
<i>S3</i>	0.565	1.997	0.629	1.84	-0,187	121.19	0.852	-

Table 2. Comparison of the HARM and HELP conditions

The analyses show that results for all three stories both in the HARM and the HELP condition differ significantly from 0, meaning “Hard to say.” A statistical significance in differences between conditions in each study was only observed for story N1, corresponding to the standard story used in the Knobe experiment. Effect sizes for the observed differences were: high for the difference between conditions in story S1, and between 0 and the mean *S1_Harm* result; and low for the other statistically significant differences.

Finally, in order to see whether the manipulation performed on the study material (modified scenarios) significantly affected the results obtained in the experiment, differences have been analyzed between results in the corresponding stories of the above experiment and studies performed by

Waleszczyński, Obidziński, and Rejewska (2019)⁴ designated in the Table as “Old” – separately for the HARM and HELP conditions. Table 3 presents the results of analyses performed for the t-Student test on two independent samples.

	M_{Old}	SD_{Old}	M_{New}	SD_{New}	t	Df	p	<i>Cohen's d</i>
<i>S1_Harm</i>	1.645	1.9	2.2	1.277	1.94	125	0,055 ^T	0.102
<i>S1_Help</i>	-1.113	2.159	-0.754	1.723	1.038	125	0.301	-
<i>S2_Harm</i>	0.603	2.174	-0.613	2.059	-3.21	123	0.002	0.314
<i>S2_Help</i>	-0.587	2.005	-0.661	2.172	-0.198	123	0.843	-
<i>S3_Harm</i>	0.952	1.741	0.565	1.997	-1.151	122	0.252	-
<i>S3_Help</i>	0.274	2.026	0.629	1.84	3.447	122	0.001	0.619

T – statistical tendention

Table 3. Comparison of results before and after story modification

The analyses have revealed two significant differences between the results of studies using stories before and after modification. One difference concerns story S2 in the HARM condition – the effect size for this difference is low. The other difference concerns story S3 in the HELP condition – it is characterized by a medium effect size. One statistical tendency has also been observed concerning story S1 in the HARM condition – however, the very low effect size suggests that this potential difference between the studied groups is insignificant.

6.3. Discussion

In his experiment, Knobe (2003) used a typical structure of a situation analyzed using the Doctrine of Double Effect to evaluate responsibility for causing a side effect. Person X takes an action to achieve a particular goal, which in Knobe’s experiment is increasing the company’s profits. In order to achieve the intended (main) goal A, a side-effect B is caused. Knobe introduced what turned out to be a significant modification by creating an

⁴ The results can be downloaded from the repository at: (Obidziński and Waleszczyński 2019) <https://osf.io/ky3re>.

alternative situation where the side effect is positive and by asking an appropriate question. The question concerned not so much responsibility for causing the side effect, but the intentionality of causing it. It is also significant that it is clear from the story that X does not care about the side effects. Results of the experiment proved to be very interesting and became problematic for the understanding of the notion of intentional action. The greatest difficulty concerns results in the HARM condition, as they suggest that respondents attribute intentionality in causing the side effect even though the agent says he does not care about side effects. In other words, the problem consists in that intentionality is attributed even though the agent does not have the intention of causing a particular effect.

Our study presents results which provide a new perspective on Knobe's experiment. We used Knobe's story, designated as S1, where the part "I do not care that it will help [harm] the environment" has been removed. The goal was to see if once the explicit statement of the agent's intentions with regard to the expected side effect is removed, the attribution of intentionality in causing this effect will be significantly affected. The proposed structure of story S1 does not provide information on the intention to cause a side effect, and consequently the expected result was the answer "Hard to say." The results are presented in Tables 1 and 2. First, it turned out that responses for the HARM and HELP conditions significantly differ statistically from the expected response "Hard to say," designated as "0". Second, responses for the S1_Harm and the S1_Help condition proved to be asymmetrical, and the difference between them is also statistically significant. The study shows that concealing the agent's intention as regards causing the side effect does not significantly influence the asymmetry in responses previously observed by Knobe. This is also confirmed by a comparison with the earlier (*Old*) studies presented in Table 3 for the S1_Harm and S1_Help conditions.

In the context of the above data concerning story S1, new data is provided by responses concerning stories S2 and S3. They are modified stories taken from studies by Waleszczyński, Obidziński and Rejewska (2019). First, in both stories, responses for both conditions are significantly statistically different from the "Hard to say" response. This coincides with the results for story S1. This would confirm the conclusion that revealing or

concealing the intention is not the only, or the most important, element affecting the respondents' attribution of intentionality in causing a side effect. Second, unlike those in story S1, the differences between HARM and HELP conditions in stories S2 and S3 are not statistically significant (Table 2). On the contrary, the responses are largely consistent and symmetrical. It is interesting that the symmetry in story S2 points to the non-intentionality of causing a side effect, and in story S3 – to its intentionality. This means that for three different stories based on the same scheme, three different results have been received. They seem to undermine the so-called Norm Violation Hypothesis (Holton 2010; Paprzycka 2015) which says that the attribution of intentionality is caused by attributing the intention to violate or neglect to conform to a particular norm. It would be difficult to find a universally valid norm which prohibits the increase of profits in companies which help the environment (S3_Help). The Norm Violation Hypothesis may be upheld for the HARM condition, but it would be problematic in this hypothesis to identify a norm for the S2-Harm condition. This results from an analysis of the earlier (*Old*) and present (*New*) studies concerning this scenario which are significantly different statistically and asymmetrical at the same time.

However, the Normative Reason Hypothesis will also have trouble interpreting the results of these studies. If the normative reason is the duty to care about causing a side effect, which relates to the attribution of intentionality, then why does it appear in both conditions in stories S3, and in neither of them in story S2? It seems that neither of these two cases can be reconciled with the hypothesis proposed by Hindriks, Douven and Singmann (Hindriks et al. 2016).

Still other consequences arise from the results of the experiment carried out for the account proposed by Cova, Dupoux and Jacob (2012). Their account assumes three meanings for the term intentionality, of which one or the other are preferred depending on the situation. Therefore, when asking about the intentionality of causing a side effect, one should take into account the attitude of the agent to the expected effect and the (social) expectation regarding the attitude of the agent and possibly the skills of the agent. The account given by them does not work in the case of our experiment, because in the analyzed history there is not the necessary data

to judge on intentionality, and yet the respondents do it. Achieving a result similar to “Hard to say” would mean that their account using a changed meaning of intentionality is correct. However, this is not what occurs, because all the results are statistically significantly different from the expected response.

The results of our experiment support the hypothesis, however, that the issue of key importance for the attribution of intentionality in causing a side effect is that of the evaluative properties of the relationship occurring between the evaluative properties of the main effect and the evaluative properties of the side effect (Waleszczyński, Obidziński, and Rejewska 2019). This conclusion is based on the fact that when the structure of the story is the same, but the evaluative properties of the main effect and the side effect are modified, three different response patterns are observed for three different stories. Consequently, this leads to the conclusion that it is the (morally positive or negative) evaluative properties of the main and side effects that significantly affect the attribution of intentionality in causing a side effect.

When analyzing the significance of concealing the intentions, or more precisely, of the agent’s indifference to the expected side effect, two statistically significant changes and one tendency have been observed compared to the Waleszczyński, Obidziński and Rejewska (2019) study. One change and the tendency concern an increase in the attribution of intentionality in causing a side effect. The statistically significant change concerns the S3_Help scenario, and is characterized by a medium effect size. Its influence is reflected in the fact that responses in the S3_Help scenario and in the S3_Harm scenario are similar. This may mean that in a situation when the main effect (helping the environment) has a high positive value, and the side effect (company profit/loss) is not significant (has a low moral value) compared to the main effect, and we do not know the intentions of the agent with regard to the side effect, respondents tend to attribute intentionality in causing a side effect both in the situation of help and that of harm. The observed statistical tendency in the attribution of intentionality concerns the HARM condition in the S1 scenario. It shows a tendency to attribute intentionality in causing a negative effect, that is, harming the

environment (high moral value) when the main effect is the desire to increase company profits (low moral value).

The change most difficult to interpret is that in the S2_Harm condition, which is not only statistically significant, but which is also the only asymmetrical change in the experiment compared to the Waleszczyński, Obi-dziński, Rejewska study. Concealing the agent's attitude to causing a negative side effect (medium moral value) which is pneumonia, with the main effect (high moral value) of curing pancreatic cancer, indicates non-intentionality in causing it.

An analysis of the experiments performed so far supports the following tendency patterns in the attribution of intentionality in causing a side effect in view of the relationship between the main effect and the side effect. Table 4 presents a situation when the agent's intention/attitude to the side effect is unknown.

<i>L.p.</i>	<i>Main goal</i>		<i>Side effect</i>		<i>Causing a side effect</i>
1.	low-value	positive	high-value	positive negative	Unintentionally Intentionally
2.	high-value	positive	medium-value	positive negative	Unintentionally Unintentionally
3.	high-value	positive	low-value	positive negative	Intentionally Intentionally

Table 4. Tendency patterns in the attribution of intentionality in causing a side effect

The above Table shows a tendency in the attribution of intentionality based on the relationship between the main and the side effect. We can see that the asymmetry in responses appears only in relationship no. 1, found in Knobe's story. The symmetry and the discrepancy in results between the relationships in scenarios no. 2 and no. 3 does not allow us to conclude that there is an absolute moment of judging about the intentionality of causing a side effect. In other words, we cannot say that if we do not know the intentions or the attitude of the agent to the side effect, we cannot claim that such effect is always caused unintentionally.

7. Summary

Studies on the common understanding of intentionality show that it is a complex and multi-threaded problem which requires further in-depth studies. Particularly interesting and fruitful are studies on the so-called side-effect effect observed by Knobe. Experiments and analyses presented in this article were aimed at contributing new knowledge about the attribution of intentionality in causing a side effect, in particular the role played by intention. First, they have shown that the very originality of the story used by Knobe (2003) and the relationship between the evaluative properties of the main effect and the evaluative properties of the side effect results in an asymmetry of responses and contributes to the occurrence of the so-called side-effect effect. On the one hand, this means that the relationship between the main and the side effect significantly affects the so-called side-effect effect. On the other hand, it shows the role played by intention in the attribution of intentionality in causing a side effect. Second, using the story structure proposed by Knobe and concealing the agent's intention to the expected side effect is not in itself enough to obtain reproducibility of responses, that is, a predictable pattern in the common application of the notion of intentionality, which has been unequivocally shown by the discrepancy in the results of the studies concerning each scenario (S1, S2, S3). Considering earlier studies (Waleszczyński, Obidziński, and Rejewska 2019) and the results of the studies presented in this article, it may be concluded that a change in the evaluation of the relationship between the evaluative properties of the main and the side effect significantly affects the attribution of intentionality in causing the side effect.

The studies and analyses have shown that a significant impact on the occurrence of the Knobe effect has the story itself and the type of the main and the side effect, or to be more exact, the specific relationship between these effects. If the relationship between the two types of effect is significant for the attribution of intentionality in causing a side effect, then it must be a necessary condition for the evaluation of the intentionality in causing that side effect. Consequently, this means that there is a difference between the conditions of attributing intentionality in causing effect A as the main effect of action X, and the conditions of attributing intentionality in causing effect

B, which is a side effect of action X. In this situation, the question reappears about the role of moral factors and their possible impact on the occurrence of the Knobe effect. At this stage of research, this cannot be unequivocally established, as we cannot precisely determine what influences the significance of the relationship between the two effects for the attribution of intentionality in causing a side effect. Even if we identify the evaluative properties of an effect, we must still face the dispute as to whether they still have the nature of description, or whether it is already that of moral judgment. In that case, we would be dealing with a purely meta-ethical dispute.

The analyses performed so far concerning the occurrence of an asymmetry in the attribution of intentionality in causing a side effect have focused mainly on explaining the attribution of intentionality in the HARM condition. At the same time, it has been assumed that in the HELP condition, all conditions for the attribution of intentionality are met. Not enough attention has been paid, however, to the fact that standard accounts of intentionality define conditions for the occurrence of a single predicted and identified effect. This means that one action causes one effect A. In such case, it is enough to check the intention of causing a particular effect in order to attribute intentionality of causing it. However, in analyses concerning the attribution of intentionality in causing a side effect, a significant additional element has been omitted. The occurrence of a side effect is conditioned by the occurrence of the main effect. This means that one action causes two effects A and B, where effect B is a derivative of effect A. If a side effect occurs, at least two additional elements appear with respect to standard accounts of intentionality. These are: the occurrence of effect B, and the relationship between effects A and B, representing a significant change in the conditions of the analyzed situation.

We may ask whether observation of the side-effect effect materially affects the standard accounts of intentionality. The answer: no. The effect observed by Knobe concerns the attribution of intentionality in causing a side effect, and not the attribution of intentionality as such. Based on existing research, we may conclude that if one action causes one effect A, then in order to determine the intentionality of causing this effect, it is enough to check the intention of the agent with regard to this effect. When one action causes two effects A and B though, where B is a derivative of A,

then in order to attribute intentionality in causing effect B it is necessary to check the intention of causing effect B and at least one of the evaluative properties of the relationship between A and B. Now the problems would be, first, how to identify the evaluative properties of the relationship between a particular A and a particular B, and, secondly, how the evaluative properties influence the attribution of intentionality in causing a side effect.

References

- Adams, Frederick. 1986. "Intention and Intentional Action: The Simple View." *Mind & Language* 1: 281–301. <https://doi.org/10.1111/j.1468-0017.1986.tb00327.x>
- Adams, Frederick, and Annie Steadman. 2004a. "Intentional Action and Moral Considerations: Still Pragmatic." *Analysis* 64: 268–76. <https://doi.org/10.1111/j.0003-2638.2004.00496.x>
- Adams, Frederick, and Annie Steadman. 2004b. "Intentional Action in Ordinary Language: Core Concept or Pragmatic Understanding?" *Analysis* 64: 173–81. <https://doi.org/10.1093/analys/64.2.173>
- Cova, Florian, Emmanuel Dupoux, and Pierre Jacob. 2012. "On Doing Things Intentionally." *Mind & Language* 27: 378–409. <https://doi.org/10.1111/j.1468-0017.2012.01449.x>
- Dalbauer, Nikolaus, and Andreas Hergovich. 2013. "Is What is Worse More Likely? —The Probabilistic Explanation of the Epistemic Side-Effect Effect." *Review of Philosophy and Psychology* 4: 639–57. <https://doi.org/10.1007/s13164-013-0156-1>
- Głowala, Michał. 2013. "On Indirectly Intended Consequences of Action and Relinquishment. G.E.M. Anscombe and the Spanish Thomists of XVI/XVII Century." *Etyka* 46: 7–20.
- Guglielmo, Steve, and Bertram F. Malle. 2010. "Can Unintended Side Effects Be Intentional? Resolving a Controversy Over Intentionality and Morality." *Personality and Social Psychology Bulletin* 36: 1635–47. <https://doi.org/10.1177/0146167210386733>
- Hart, Herbert L. A. "1949. XI.—The Ascription of Responsibility and Rights." *Proceedings of the Aristotelian Society* 49: 171–94. <https://doi.org/10.1093/aristotelian/49.1.171>
- Hindriks, Frank. 2008. "Intentional Action and the Praise Blame Asymmetry." *Philosophical Quarterly* 58: 630–41. <https://doi.org/10.1111/j.1467-9213.2007.551.x>

- Hindriks, Frank. 2011. "Control, Intentional Action, and Moral Responsibility." *Philosophical Psychology* 24: 787–801.
<https://doi.org/10.1080/09515089.2011.562647>
- Hindriks, Frank. 2014. "Normativity in action: How to explain the knobe effect and its relatives." *Mind and Language* 29: 51–72.
<https://doi.org/10.1111/mila.12041>
- Hindriks, Frank, Igor Douven, and Henrik Singmann. 2016. "A New Angle on the Knobe Effect: Intentionality Correlates with Blame, not with Praise." *Mind and Language* 31: 204–20. <https://doi.org/10.1111/mila.12101>
- Holton, Richard. 2010. "Norms and the Knobe Effect." *Analysis* 70: 417–24.
<https://doi.org/10.1093/analys/anq037>
- Knobe, Joshua. 2003. "Intentional Action and Side Effects in Ordinary Language." *Analysis* 63: 190–94. <https://doi.org/10.1093/analys/63.3.190>
- Knobe, Joshua. 2004a. "Folk Psychology and Folk Morality: Response to Critics." *Journal of Theoretical and Philosophical Psychology* 24: 270–79.
<https://doi.org/10.1037/h0091248>
- Knobe, Joshua. 2004b. "Intention, Intentional Action and Moral Considerations." *Analysis* 64: 181–87. <https://doi.org/10.1093/analys/64.2.181>
- Knobe, Joshua. 2006. "The Concept of Intentional Action: A Case Study in the Uses of Folk Psychology." *Philosophical Studies* 130: 203–31.
<https://doi.org/10.1007/s11098-004-4510-0>
- Knobe, Joshua, and Arudra Burra. 2006. "The folk concepts of intention and intentional action: A cross-cultural study." *Journal of Cognition and Culture* 6: 113–32. <https://doi.org/10.1163/156853706776931222>
- Knobe, Joshua. 2016. "Experimental Philosophy Is Cognitive Science." In *A Companion to Experimental Philosophy*, edited by Justin Sytsma, and Wesley Buckwalter, 37–52. Wiley-Blackwell.
<https://doi.org/10.1002/9781118661666.ch3>
- Knobe, Joshua, and Shaun Nichols. 2008. "An Experimental Philosophy Manifesto." In *Experimental Philosophy*, edited by Joshua Knobe, and Shaun Nichols, 3–16. Oxford: Oxford University Press.
- Kuś, Katarzyna, and Bartosz Maćkiewicz. 2016. "Z rozmysłem, ale nie specjalnie. O językowej wrażliwości filozofii eksperymentalnej." *Filozofia Nauki* 24: 73–102.
- Leslie, Alan M., Joshua Knobe, and Adam Cohen. 2006. "Acting Intentionally and the Side-Effect Effect." *Psychological Science* 17: 421–27.
<https://doi.org/10.1111/j.1467-9280.2006.01722.x>
- Malle, Bertram F., and Joshua Knobe. 1997. "The folk concept of intentionality." *Journal of Experimental Social Psychology*, 33(2): 101–121.
<https://doi.org/10.1006/jesp.1996.1314>

- Malle, Bertram F., and Sarah E. Nelson. 2003. "Judging Mens Rea: The Tension Between Folk Concepts and Legal Concepts of Intentionality." *Behavioral Sciences and the Law*, 21(5): 563–80. <https://doi.org/10.1002/bsl.554>
- McCann, Hugh J. 1987. "Rationality and the Range of Intention." *Midwest Studies in Philosophy* 10: 191–211. <https://doi.org/10.1111/j.1475-4975.1987.tb00540.x>
- Mele, Alfred R. 1992. "Recent Work on Intentional Action." *American Philosophical Quarterly* 29: 199–217. <https://doi.org/10.2307/20014416>
- Mele, Alfred R., and Fiery Cushman. 2007. "Intentional action, folk judgments, and stories: Sorting things out." *Midwest Studies in Philosophy* 31: 184–201. <https://doi.org/10.1111/j.1475-4975.2007.00147.x>
- Mele, Alfred R., and Steven Sverdlik. 1996. "Intention, intentional action, and moral responsibility." *Philosophical Studies* 82: 265–87. <https://doi.org/10.1007/BF00355310>
- Nadelhoffer, Thomas. 2004a. "On Praise, Side Effects, and Folk Ascriptions of Intentionality." *Journal of Theoretical and Philosophical Psychology* 24: 196–213. <https://doi.org/10.1037/h0091241>
- Nadelhoffer, Thomas. 2004b. "The Butler Problem Revisited." *Analysis* 64(3): 277–84. <https://doi.org/10.1111/j.0003-2638.2004.00497.x>
- Nadelhoffer, Thomas. 2004c. "Blame, Badness, and Intentional Action: A Reply to Knobe and Mendlow." *Journal of Theoretical and Philosophical Psychology* 24: 259–69. <https://doi.org/10.1037/h0091247>
- Nadelhoffer, Thomas. 2005. "Skill, luck, control, and intentional action." *Philosophical Psychology* 18: 341–52. <https://doi.org/10.1080/09515080500177309>
- Nadelhoffer, Thomas. 2006a. "Bad Acts, Blameworthy Agents, and Intentional Actions: Some Problems for Juror Impartiality." *Philosophical Explorations* 9: 203–19. <https://doi.org/10.1080/13869790600641905>
- Nadelhoffer, Thomas. 2006b. "Desire, Foresight, Intentions, and Intentional Actions: Probing Folk Intuitions." *Journal of Cognition and Culture* 6: 133–57. <https://doi.org/10.1163/156853706776931259>
- Nadelhoffer, Thomas, and Eddy Nahmias (2007). "The Past and Future of Experimental Philosophy." *Philosophical Explorations* 10(2): 123–49. <https://doi.org/10.1080/13869790701305921>
- Nado, Jennifer. 2008. "Effects of Moral Cognition on Judgments of Intentionality." *British Journal for the Philosophy of Science* 59: 709–31. <https://doi.org/10.1093/bjps/axn035>
- Nakamura, Kuninori. 2018. "Harming is more intentional than helping because it is more probable: the underlying influence of probability on the Knobe effect." *Journal of Cognitive Psychology* 30: 129–37. <https://doi.org/10.1080/20445911.2017.1415345>

- Nichols, Shaun, and Joseph Ulatowski. 2007. "Intuitions and Individual Differences: The Knobe Effect Revisited." *Mind and Language* 22: 346–65.
<https://doi.org/10.1111/j.1468-0017.2007.00312.x>
- Obidziński, Michał, and Andrzej Waleszczyński. 2019. "The Significance of the Relationship between Main Effects and Side Effects for Understanding the Knobe Effect: Database.": <https://osf.io/ky3re>
- Obidziński, Michał, and Andrzej Waleszczyński. 2021. "The Problem of Intention and the Evaluative Properties of Effects in the Knobe Effect: Database.": <http://osf.io/8c5qk>
- Paprzycka, Katarzyna 2012. "Analityczna filozofia działania: problemy i stanowiska." In *Przewodnik po filozofii umysłu*, edited by Marcin Miłkowski and Robert Poczobut, 465–94. Krakow: Wydawnictwo WAM.
- Paprzycka, Katarzyna. 2015. "The Omissions Account of the Knobe Effect and the Asymmetry Challenge." *Mind and Language* 30: 550–71.
<https://doi.org/10.1111/mila.12090>
- Paprzycka, Katarzyna. 2016. "Poznań Studies in the Philosophy of the Sciences and the Humanities." In *The Sciences*, edited by Adrian Kuźniar and Joanna Odrowąż-Sypniewska, 107: 204–33. Leiden—Boston: Brill Rodopi.
- Piekarski, Michał 2017. "Efekt Knobe'a, normatywność i racje działania." *Filozofia Nauki* 97: 109–28. <https://www.fn.uw.edu.pl/index.php/fn/article/view/846>
- Shepherd, Joshua. 2012. "Action, Attitude, and the Knobe Effect: Another Asymmetry." *Review of Philosophy and Psychology* 3: 171–85.
<https://doi.org/10.1007/s13164-011-0079-7>
- Sneddon, Andrew 2006. *Action and Responsibility*. Dordrecht: Springer Netherlands. <https://doi.org/10.1007/1-4020-3982-4>
- Sripada, Chandra S. 2010. "The Deep Self Model and asymmetries in folk judgments about intentional action." *Philosophical Studies* 151: 159–76.
<https://doi.org/10.1007/s11098-009-9423-5>
- Sripada, Chandra S. 2012. "Mental state attributions and the side-effect effect." *Journal of Experimental Social Psychology* 48: 232–38.
<https://doi.org/10.1016/j.jesp.2011.07.008>
- Sripada, Chandra S., and Sara Konrath. 2011. "Telling More Than We Can Know About Intentional Action." *Mind & Language* 26: 353–80.
<https://doi.org/10.1111/j.1468-0017.2011.01421.x>
- Uttich, Kevin, and Tania Lombrozo. 2010. "Norms Inform Mental State Ascriptions: A Rational Explanation for the Side-Effect Effect." *Cognition* 116: 87–100. <https://doi.org/10.1016/j.cognition.2010.04.003>
- Waleszczyński, Andrzej, Michał Obidziński, and Julia Rejewska. 2018. "The Knobe Effect From the Perspective of Normative Orders." *Studia Humana* 7: 9–15.
<https://doi.org/10.2478/sh-2018-0019>

Waleszczyński, Andrzej, Michał Obidziński, and Julia Rejewska. 2019. “The Significance of the Relationship between Main Effects and Side Effects for Understanding the Knobe Effect.” *Organon F* 26: 228–48.

<https://doi.org/10.31577/orgf.2019.26203>

Wright, Jennifer C., and John Bengson. 2009. “Asymmetries in Judgments of Responsibility and Intentional Action.” *Mind & Language* 24: 24–50.

<https://doi.org/10.1111/j.1468-0017.2008.01352.x>

Zalla, Tiziana, and Marion Leboyer. 2011. “Judgment of Intentionality and Moral Evaluation in Individuals with High Functioning Autism.” *Review of Philosophy and Psychology* 2: 681–98. <https://doi.org/10.1007/s13164-011-0048-1>