THE KNOBE EFFECT WITH PROBABLE OUTCOMES AND AVAILABILITY HEURISTIC TRIGGERS

Tommaso OSTILLIO, Michal BUKAT

ABSTRACT: This paper contributes to the existing philosophical literature on the Knobe Effect (KE) in two main ways: first, this paper disconfirms the KE by showing that the latter does not hold in contexts with probable outcomes; second, this paper shows that KE is strongly sensitive to the availability heuristic bias. In particular, this paper presents two main findings from three empirical tests carried out between 2016 and 2018: the first finding concerns the fact that if the issuer of a decision with consequences on third parties is unlikely to be perceived as unfriendly, then KE is reduced or absent; the second finding regards instead the fact that if an action has two possible outcomes (one likely to obtain with strong intensity and one likely to obtain with less intensity), then KE does not obtain for decisions whose side-effects have limited consequences on third parties.

KEYWORDS: experimental philosophy, Knobe effect, cognitive bias, negative externality

The concept of intentionality has played and keeps playing a dominant role in contemporary epistemology, in contemporary philosophy of mind, in contemporary philosophy of action and in contemporary meta-ethics. This is because philosophers have struggled and still struggle with finding a definition of intentionality, which leads to long-term agreement among different schools of thought.

Historically speaking, the contemporary philosophical literature on intentionality has taken two main opposite directions: on the one hand, some philosophers find an association between intentionality and the reasons to act in a particular way;[1] on the other hand, some philosophers find instead an association between intentionality and the aboutness (i.e. the content) of mental states.[2]

[1] See Gertrud Elizabeth Margaret Anscombe, *Intention* (Cambridge: Harvard University Press, 1957); Donald Davidson, "Actions, Reasons, Causes," *The Journal of Philosophy* LX, 23 (1963): 685-700.

[2] See Daniel Clement Dennett, "Intentional Systems," *The Journal of Philosophy* 68, 4 (1981): 87-106; John Rogers Searle, *Intentionality: An Essay in the Philosophy of Mind* (Cambridge: Cambridge University Press, 1983), 1-36.

Besides, although the literature is exceptionally vast on both sides, no perfect argument to defend a particular definition of intentionality has been found on neither side.

At the same time, philosophers' overall troubles in defining intentionality have grown bigger since the so-called experimental philosophers have shown that there exists a discrepancy between the way philosophers understand intentionality and the way folks attribute intentionality to agents.

In this respect, Malle and Knobe investigate how folks attribute intentionally to agents empirically and find that, while philosophers usually relate intentionality to purpose or mental content, folks relate intentionality to possessing the right set of skills to carry out a given course of action.[3] That is, according to the folks surveyed by Malle and Knobe, an action is intentional if and only if an agent is able to carry out the course of action he or she intends to carry out.[4]

In the light of the findings of Malle and Knobe,[5] Knobe carries out another survey, which relates intentionality (understood as possessing the right skills to carry out the intended course of action) to the externality of actions.[6] In particular, Knobe constructs two vignettes where a fictitious character, Jake, is in need for money and gains the amount of money he needs either by participating in a rifle contest or by killing his old rich aunt.[7] Moreover, Knobe divides each vignette case in two sub-vignettes where two assumptions are dominant: either Jake is a skilled shooter or Jake is not a skilled shooter.[8]

In the first vignette, Jake participates in a rifle context where he is to shoot a bull in its eye from a big distance. If Jake succeeds at shooting the bull in its eye, he gets the money, whereas, if he does not, he gets no money. Yet Jake accomplishes his goal in both sub-vignette-cases regardless of whether he is a skilled shooter or not. QED, Knobe finds that when 37 random subjects are asked whether Jake acted intentionally or not, their general answer is that he acted intentionally in the first sub-vignette-case, but he did not do so in the second sub-vignette-case.[9] That is,

[3] Bertram F. Malle, and Joshua Knobe, "The folk concept of intentionality," *Journal of Experimental Social Psychology* 33 (1997): 101-121.

[4] Malle and Knobe, "The folk concept of intentionality."

[5] Malle and Knobe, "The folk concept of intentionality."

[6] Joshua Knobe, "Intentional Action in Folk Psychology: An Experimental Investigation," *Philosophical Psychology* 16, 2 (2003): 309-324

[7] Knobe, "Intentional Action in Folk Psychology."

[8] Knobe, "Intentional Action in Folk Psychology."

[9] Knobe, "Intentional Action in Folk Psychology."

Jake's accomplishment is intentional as far as he possesses the right set of skills to shoot the bull in its eyes from a great distance.

By contrast, in the second vignette, Jake gets the amount of money he needs if and only if he kills his old rich aunt, while she is at home, by shooting her through the window of the house in front of hers. As in the first vignette, Jake successfully accomplishes his goal in both sub-vignettes. Yet, when 37 random subjects are asked whether Jake acted intentionally or not, their general answer is that he acted intentionally regardless of whether Jake is a skilled shooter or not.

Thus, Knobe concludes that while it holds true that folks overall relate intentionality to the ability to accomplish a given intended goal, the gathered data show also the attribution of intentionality to agents is dependent on the externality of a given action. For folks consider Jake's murder of his old aunt as intentional in both sub-vignettes.[10]

On this basis, Knobe constructs two more vignettes, which put a stronger emphasis on the side-effects of an action. More specifically, the two vignettes recount the story of a firm's VP who wants to implement a business project aimed at increasing his firm's profits: in the first case, the business project is implemented successfully with a positive externality (i.e. its implementation helps the environment); in the second case, the side-effect of a success implementation is a negative externality (i.e. its implementation harms the environment).[11] QED, Knobe finds that when 78 random subjects are asked whether the VP caused both side-effects intentionally or not, their dominant answer is that he did so in the second case, but he did not do so in the first case.[12]

In the philosophical literature, the effect observed by Knobe[13] is usually referred to as the Knobe effect (i.e. folks' tendency to consider an action intentional if and only if it has negative side-effects) and, since the findings of Knobe[14] have been published, the Knobe effect (KE) has been the object of important debates in philosophy and in the social sciences. In fact, the findings of Knobe[15] have also gained a special place in the research programs of some researchers in business and economics because KE might explain how people perceive specific business or policy decisions (yet with some limitations).

---

[10] Knobe, "Intentional Action in Folk Psychology."

[11] Knobe, "Intentional Action and Side Effects in Ordinary Language," *Analysis* 63, 3 (2003): 190-94.

[12] Knobe, "Intentional Action and Side Effects."

[13] Knobe, "Intentional Action and Side Effects."

[14] Knobe, "Intentional Action and Side Effects."

[15] Knobe, "Intentional Action and Side Effects."

In this regard, Feltz *et al.* implement an experimental setting where a random sample of subjects undergoes a two-stage treatment: in the first stage, the surveyed subjects are asked to take actions with side-effects and then evaluate how intentional their actions are on a 5-points Likert scale; in the second stage, the surveyed subjects are asked to evaluate the intentionality of some actions carried out in some vignette case, which depict the events of the first stage, on a 5-points Likert scale.[16] Interestingly, Feltz *et al.* find that the surveyed subjects judge their actions in the first experimental stage as being less intentional than the actions depicted in the vignette cases of the second experimental stage.[17] That is, Feltz *et al.* find that a change from a first-person to a third-person perspective might affect how intentionality is evaluated and attributed to agents.[18]

On the other hand, Utikal and Fischbacher[19] object that the vignette cases of Knobe[20] do not properly consider the economic gains of the firm harming/helping the environment. Accordingly, Utikal and Fischbacher[21] translate the vignettes of Knobe[22] into a market-like setting with three scenarios where three players play respectively the role of the firm's VP (player 1), the role of the environment (player 2) and the role of an external judge (player 3) who can punish or reward player 1 depending on the outcomes of player 1's decisions. The experimental setting designed by Utikal and Fischbacher[23] is divided into two stages. The first stage X represents the default economic status of all the players and is divided in three sub-stage in the following way: in the first sub-scenario, a strong active player 1 affects a weak passive player 2; whereas, in the second sub-scenario, a weak (player 1 affects a strong passive player 2; and, in the sub-third scenario, a weak active player 1 affects a weak passive player 2. The second stage Y represents the final economic status Y of player 1 and player 2 after player 2 opted for one of the three following options: a bad outcome (harm); a good outcome (help); and a neutral outcome. Figure 1 (below) shows that, in each sub-scenario, the outcomes of player 1's decisions lead to different endowment reallocation. Eventually, after having observed what outcome obtains, player 3 can either reward player 1 (i.e.

[16] Adam Feltz, Maegan Harris, and Ashley Perez, "Perspective in intentional action attribution," *Philosophical Psychology* 25, 5 (2012): 673-687.

[17] Feltz*et al.*, "Perspective in intentional action attribution."

[18] Feltz*et al.*, "Perspective in intentional action attribution."

[19] Verena Utikal and Urs Fischbacher, "Attribution of externalities: an economic approach to the Knobe effect," *Economics and Philosophy* 30, 2 (2014): 215-240.

[20] Knobe, "Intentional Action and Side Effects."

[21] Utikal and Fischbacher, "Attribution of externalities."

[22] Knobe, "Intentional Action and Side Effects."

[23] Utikal and Fischbacher, "Attribution of externalities."

player 3 can subtract points from player 2 and reallocate them to player 1) or punish player 1 (i.e. player 3 can subtract points from player 1 and reallocate them to player 2). The latter option for player 3 represents the activation of KE.

		Default X		Endowment after change Y	
Setting		Player 1	Player 2	Player 1	Player 2
StrongActiveSmallHelp	Harm	50	50	60	30
	Help	50	20	60	30
	No side effect	-	-	60	30
WeakActiveSmallHelp	Harm	20	80	30	60
	Help	20	50	30	60
	No side effect	-	-	30	60
WeakActiveBigHelp	Harm	20	80	30	60
	Help	20	20	30	60
	No side effect	-	-	30	60

Figure 1 - Verena Utikal and Urs Fischbacher, "Attribution of externalities: an economic approach to the Knobe effect," 220.

In the light of the aforementioned premises, Utikal and Fischbacher[24] find that KE obtains only in the first scenario, while it reverses in the second and in the third scenario. That is, in the first scenario, player 1 is overall punished, whereas, in the second and in the third scenario, player 1 is overall rewarded by player 2 regardless of the option chosen by player 1. This is because, according to Utikal and Fischbacher,[25] Player 1 does not look unfriendly to Player 3 in the second and in the third scenario.

Most importantly, the findings of Utikal and Fischbacher[26] find some confirmation in an earlier study by Wible,[27] where 36 random subjects are asked to evaluate the following:

> *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 increase profits, and it will also help the environment.' The chairman of the board answered, 'Great! I care about*

[24] Utikal and Fischbacher, "Attribution of externalities."
[25] Utikal and Fischbacher, "Attribution of externalities."
[26] Utikal and Fischbacher, "Attribution of externalities."
[27] Andrew Wible, "Knobe, Side Effects, and the Morally Good Business," *Journal of Business Ethics* 85 (2009): 173–178.

> *helping the environment. I am happy that we can help the environment. I am happy that we can help the environment and make a profit at the same time. Let's start the new program.' They started the new program. Sure enough, the environment was helped.*[28]

Wible finds that 55% of the surveyed subjects says that the chairman acted intentionally. In other words, the fact that the intentions of the chairman were good and clearly stated impacts how intentionality is evaluated and attributed to agents.

Thus, considering the findings of Wible[29] and Utikal and Fischbacher,[30] there is room to argue that the availability heuristic bias[31] might nudge the activation of the Knobe effect in case like those described by Knobe.[32] In fact, the vignettes of Knobe[33] force the surveyed subjects to attribute intentionality to agents under uncertainty in presence of restrained data, which nudge stereotype-based judgements about the wrongdoings of greedy businessmen.

Furthermore, another objection to Knobe[34] might be that his vignettes represent cases where the telos of the events is given and taken for granted. That is, the intended outcomes entailed by the decision of the firm's VP are granted to obtain. Yet, when business projects are implemented, this is seldom the case because the unaccounted side-effects of a business decision might be more than executives can forecast alone.

Accordingly, in order not to fall into too speculative forms of argumentation about the vignette cases of Knobe,[35] this paper tests empirically whether the Knobe Effect is immune to the effects of the availability heuristic bias and whether the Knobe Effect obtains once the forecasted side-effects of an action are only probable. The next section presents the results of three survey-based experiments, which were carried out by the authors of this paper between 2016 and 2018.

Experiment 1

The first experiment took place in December 2016 within a different research project and involved two runs of testing: in the first run (Group 1), 40 master

[28] Wible, "Knobe, Side Effects, and the Morally Good Business," 174.
[29] Wible, "Knobe, Side Effects, and the Morally Good Business."
[30] Utikal and Fischbacher, "Attribution of externalities."
[31] See Amos Tversky and Daniel Kahneman, "Judgment under Uncertainty: Heuristics and Biases," *Science*, *New Series* 185, 4157 (Sep. 27, 1974): 1124-1131.
[32] Knobe, "Intentional Action and Side Effects."
[33] Knobe, "Intentional Action and Side Effects."
[34] Knobe, "Intentional Action and Side Effects."
[35] Knobe, "Intentional Action and Side Effects."

students of Finance at Kozminski University were asked to express their judgement on the vignette presented in Task 1 offline; in the second run (Group 2), 50 random individuals recruited on Amazon Mechanical Turk were asked to express their judgement on the vignette presented in Task 1 online.

The overall goal of the experiment was to test whether the surveyed subjects overall attribute intentionality to an action whose side-effects are only probable. More specifically, following the vignettes of Knobe,[36] we constructed a vignette where the outcomes of a business decision are double. That is, the latter decision leads to a bigger forecasted outcome that is likely to obtain with stronger intensity and a smaller forecasted outcome that is likely to obtain with less intensity.

On this basis, as shown below, Task 1 focused only on finding out whether KE activates only in the context of the smaller forecasted outcome that is likely to obtain with less intensity:

> **Task 1:** Assume that a hedge fund decides to finance a research project for the development of a new pain killer with €200M. Assume also that the project is carried out using dogs as test animals and that the dogs might either survive or die with some probability after the experiments is performed by researchers. In any case, the development of the pain killer generates returns that amount to 30% of the hedge fund's initial investment. You're asked to evaluate the following.
>
> CASE 1: The experiment is carried out successfully, the project generates returns that amount to 30% of the hedge fund's initial investment and the dogs used as test animals survive with probability with probability 0.75, i.e. few dogs die because of the side-effects of the experiment. Did the hedge fund cause the death of few of the dogs intentionally? Mark the option you choose.
>
> A) YES;
>
> B) NO.
>
> CASE 2: The experiment is carried out successfully, the project generates returns that amount to 30% of the hedge fund's initial investment and the dogs used as test animals die with probability with probability 0.75 because of the side-effects of the experiment, i.e. few dogs survive. Did the hedge fund cause the survival of few of the dogs intentionally? Mark the option you choose.
>
> A) YES;
>
> B) NO.

[36] Knobe, "Intentional Action and Side Effects."

RESULTS - CASE 1	Group 1 (N=40)	Group 2 (N=50)
YES	37.5%	40%
NO	62.5%	60%
Significance	$\chi^2 = 2.5\ (1)\ p = 0.114$	$\chi^2 = 2\ (1)\ p = 0.157$
RESULTS - CASE 2	Group 1 (N=40)	Group 2 (N=50)
YES	32.5%	20%
NO	67.5%	80%
Significance	$\chi^2 = 4.9\ (1)\ p = 0.027$	$\chi^2 = 9.68\ (1)\ p = 0.002$
RESULTS – COMBINED	CASE 1 (N=90)	CASE 2 (N=90)
YES	39%	30%
NO	61%	70%
Significance	$\chi^2 = 4.44\ (1)\ p = 0.035$	$\chi^2 = 14.4\ (1)\ p = 0.000$

Table 1 - Experiment 1: results

The results in Table 1 show that both Group 1 and 2 overall do not attribute intentionality to the hedge fund in CASE 1 and CASE 2. Yet the span between YES and NO is statistically significant only in CASE 2 for both Group 1 and 2. Hence, KE is not nullified.

KE is instead nullified when the results are combined. Therefore, there is room to argue that if a decision leads to a forecasted side-effect that is likely to obtain with less intensity, then there might be no attribution of intentionality on the issuer of that decision.

Experiment 2

After having presented the results of *Experiment 1* at some conferences and workshops, we received two main objections concerning our vignettes: first, the vignettes should have accounted also for the reverse case, i.e. for the case where the bigger side-effect obtains; second, the content of the vignettes is expressed in a very neutral language and nudges a biased evaluation under uncertainty. Both objections are addressed both by *Experiment 2* and *Experiment 3*.

More specifically, as shown below in Task 2.1, Task 2.2, Task 2.3, Task 2.4, *Experiment 2* provides a more explicit version of Task 1 including both the case where the big side-effect obtains and the case where the small side-effect obtains. Task 2.1, 2.2, 2.3, 2.4 are tested against the intuitions of 102 individuals randomly selected on Amazon Mechanical Turk.

> Task 2.1: A hedge fund decides to finance a research project for the development of a new painkiller with $500M. The researchers involved in the project use dogs and cats as test animals. In short, the researchers test the effectiveness of the

painkiller by causing some big harm to dogs and cats. Depending on how big a pain the researchers will inflict to dogs and cats, the test animals can either survive or die with some probability. Either ways, the hedge fund will turn a profit that amounts to 60% of the initial investment.

The experiment is carried out successfully. The hedge fund earns a profit of 60% on top of the initially invested capital. Yet dogs and cats survive with probability 0.75, i.e. few of them die and most of them survive. Did the hedge fund cause the survival of most of the test animals intentionally?

YES; NO.

Task 2.2: A hedge fund decides to finance a research project for the development of a new painkiller with $500M. The researchers involved in the project use dogs and cats as test animals. In short, the researchers test the effectiveness of the painkiller by causing some big harm to dogs and cats. Depending on how big a pain the researchers will inflict to dogs and cats, the test animals can either survive or die with some probability. Either ways, the hedge fund will turn a profit that amounts to 60% of the initial investment.

The experiment is carried out successfully. The hedge fund earns a profit of 60% on top of the initially invested capital. Yet dogs and cats survive with probability 0.75, i.e. few of them die and most of them survive. Did the hedge fund cause the death of few of the test animals intentionally?

YES; NO

Task 2.3: A hedge fund decides to finance a research project for the development of a new painkiller with $500M. The researchers involved in the project use dogs and cats as test animals. In short, the researchers test the effectiveness of the painkiller by causing some big harm to dogs and cats. Depending on how big a pain the researchers will inflict to dogs and cats, the test animals can either survive or die with some probability. Either ways, the hedge fund will turn a profit that amounts to 60% of the initial investment.

The experiment is carried out successfully. The hedge fund earns a profit of 60% on top of the initially invested capital. Yet dogs and cats die with probability 0.75, i.e. few of them survive and most of them die. Did the hedge fund cause the death of most of the test animals intentionally?

YES; NO.

Task 2.4: A hedge fund decides to finance a research project for the development of a new painkiller with $500M. The researchers involved in the project use dogs and cats as test animals. In short, the researchers test the effectiveness of the painkiller by causing some big harm to dogs and cats. Depending on how big a pain the researchers will inflict to dogs and cats, the test animals can either survive or die with some probability. Either ways, the hedge fund will turn a profit that amounts to 60% of the initial investment.

> The experiment is carried out successfully. The hedge fund earns a profit of 60% on top of the initially invested capital. Yet dogs and cats die with probability 0.75, i.e. few of them survive and most of them die. Did the hedge fund cause the survival of few of the test animals intentionally?
>
> YES; NO.

Answers (N=102)	Task 2.1	Task 2.2	Task 2.3	Task 2.4
YES	44%	59%	63%	42%
NO	56%	41%	37%	58%
Significance	$\chi^2 = 1.412\ (1)$ $p = 0.235$	$\chi^2 = 3.176\ (1)$ $p = 0.075$	$\chi^2 = 2.5\ (1)$ $p = 0.010$	$\chi^2 = 2.5\ (1)$ $p = 0.113$

Table 2 - Experiment 2: results

The results in Table 2 show that KE activates only in Task 2.3 because the span between YES and NO in Task 2.3 is the only statistically significant span. Indeed, while the YES are 59% in Task 2.2, there is no statistically significant span. Accordingly, there is room to argue that, regardless of the neutrality of language, KE activates only when a decision leads to a forecasted side-effect that is likely to obtain with stronger intensity. In this sense, the findings of Knobe[37] are correct.

Experiment 3

The last experiment was devised in order to account mainly for the objection of language neutrality, which is only partially addressed in Task 2.1, Task 2.2, Task 2.3 and Task 2.4.

Experiment 3 was carried out entirely online on Amazon Mechanical Turk where 69 randomly recruited individuals were asked to express their judgements concerning the following vignette cases: Task 3-6 attempt to nudge availability heuristic biases in the surveyed subjects; Task 7-8 replicate Task 1 by adding a few emotionally triggering words, e.g. investment bankers and puppies; Task 9-12 replicate the vignette of Task 2.1, 2.2, 2.3, 2.4 for a further test.

> **Task 3:** A crew of firefighters is called up to extinguish a blaze that has blasted in a building where 12 people live: 4 children, 5 women (3 of which are pregnant) and 3 men.
>
> Once the crew of firefighters reaches the building, the firefighters realize that the situation is pretty bad: the 4 children and the 3 pregnant women have remained trapped in the building. After having evaluated the gravity of the situation, the firefighters conclude that the chances of rescue success are 5%.

[37] Knobe, "Intentional Action and Side Effects."

Moreover, the firefighters know that they will get decorated and obtain a raise for bravery regardless of the outcomes of their action. Thus, the firefighters break into the building, but, given the situation, give up shortly after. However, they get decorated and obtain a raise for bravery.

According to you, did the firefighters intentionally act as they did just to get decorated and obtain a raise for bravery?

A) YES;

B) NO.

Task 4: A crew of firefighters is called up to extinguish a blaze that has blasted in a building where 12 people live: 4 children, 5 women (3 of which are pregnant) and 3 men.

Once the crew of firefighters reaches the building, the firefighters realize that the situation is pretty bad: the 4 children and the 3 pregnant women have remained trapped in the building. After having evaluated the gravity of the situation, the firefighters conclude that the chances of rescue success are 5%. Moreover, the firefighters know that the will get decorated and obtain a raise for bravery regardless of the outcomes of their action.

Nevertheless, against any rational forecast, the firefighters get into the building and manage to save the 4 children and the 3 pregnant women. Hence, they get decorated and obtain a raise for bravery.

According to you, did the firefighters intentionally act as they did just to get decorated and obtain a raise for bravery?

A) YES;

B) NO.

Task 5: An NGO operates in Africa where it provides locals with free vaccinations. In particular, the NGO raises funds with charity campaigns and then purchases vaccines from top pharmaceutical corporations.

According to the physicians working for the NGO, the last batch of vaccines is defective and potentially able to cause death. However, the board of the NGO does not want to ruin the good name of the NGO, which has always carried out valorous medical operations.

Thus, considered that a very bad epidemic is spreading in the countries where the NGO operates, the NGO's board decides to take the risk of handing out vaccinations to people because, in the worst case scenario, the NGO can lay the blame on its suppliers. As a result, all the people who were vaccinated survive and the name of the NGO is safe.

According to you, did the board of the NGO cause the survival of all the vaccinated people intentionally?

A) YES;

B) NO.

Task 6: An NGO operates in Africa where it provides locals with free vaccinations. In particular, the NGO raises funds with charity campaigns and then purchases vaccines from top pharmaceutical corporations. According to the physicians working for the NGO, the last batch of vaccines is defective and potentially able to cause death. However, the board of the NGO does not want to ruin the good name of the NGO, which has always carried out valorous medical operations.

Thus, considered that a very bad epidemic is spreading in the countries where the NGO operates, the NGO's board decides to take the risk of handing out vaccinations to people because, in the worst case scenario, the NGO can lay the blame on its suppliers. As a result, all the people who were vaccinated die. Yet the name of the NGO is safe because the press believes it's fault of the NGO's suppliers.
According to you, did the NGO cause the death of all the vaccinated people intentionally?

A) YES;

B) NO.

Task 7: A hedge fund run by investment bankers decides to run a project aimed at the development of a new shampoo. The fund invests $150M in a research project that is meant to generate returns up to 50% on top of the initial investment.

The fund purchases some puppies of tigers and panthers on the black market so that the researchers involved in the research project use those puppies as test animals. Importantly, the latter shall die with a 0.81 probability, i.e. only few of them survive.

The tests are carried out successfully, the project generates the expected returns and most of the puppies die as a result of the treatments.

Did the hedge fund cause the survival of few of the puppies intentionally?

A) YES;

B) NO.

Task 8: A hedge fund run by investment bankers decides to run a project aimed at the development of a new shampoo. The fund invests $150M in a research project that is meant to generate returns up to 50% on top of the initial investment.

The fund purchases some puppies of tigers and panthers on the black market so that the researchers involved in the research project use those puppies as test animals. Importantly, the latter shall survive with a 0.81 probability, i.e. few of them die.

The tests are carried out successfully, the project generates the expected returns and most of the puppies survive as a result of the treatments.

Did the hedge fund cause the death of few of the puppies intentionally?

A) YES;

B) NO.

Task 9: A hedge fund decides to finance a research project for the development of a new painkiller with $500M. The researchers involved in the project use dogs and cats as test animals. In short, the researchers test the effectiveness of the painkiller by causing some big harm to dogs and cats.

Depending on how big a pain the researchers will inflict to dogs and cats, the test animals can either survive or die with some probability. Either ways, the hedge fund will turn a profit that amounts to 60% of the initial investment.

The experiment is carried out successfully. The hedge fund earns a profit of 60% on top of the initially invested capital. Yet dogs and cats survive with probability 0.75, i.e. few of them die and most of them survive. Did the hedge fund cause the survival of most of the test animals intentionally?

A) YES;

B) NO.

Task 10: A hedge fund decides to finance a research project for the development of a new painkiller with $500M. The researchers involved in the project use dogs and cats as test animals. In short, the researchers test the effectiveness of the painkiller by causing some big harm to dogs and cats.

Depending on how big a pain the researchers will inflict to dogs and cats, the test animals can either survive or die with some probability. Either ways, the hedge fund will turn a profit that amounts to 60% of the initial investment.

The experiment is carried out successfully. The hedge fund earns a profit of 60% on top of the initially invested capital. Yet dogs and cats survive with probability 0.75, i.e. few of them die and most of them survive.

Did the hedge fund cause the death of few of the test animals intentionally?

A) YES;

B) NO.

Task 11: A hedge fund decides to finance a research project for the development of a new painkiller with $500M. The researchers involved in the project use dogs and cats as test animals. In short, the researchers test the effectiveness of the painkiller by causing some big harm to dogs and cats.

Depending on how big a pain the researchers will inflict to dogs and cats, the test animals can either survive or die with some probability. Either ways, the hedge

fund will turn a profit that amounts to 60% of the initial investment.

The experiment is carried out successfully. The hedge fund earns a profit of 60% on top of the initially invested capital. Yet dogs and cats die with probability 0.75, i.e. few of them survive and most of them die.

Did the hedge fund cause the death of most of the test animals intentionally?

A) YES;

B) NO.

Task 12: A hedge fund decides to finance a research project for the development of a new painkiller with $500M. The researchers involved in the project use dogs and cats as test animals. In short, the researchers test the effectiveness of the painkiller by causing some big harm to dogs and cats.

Depending on how big a pain the researchers will inflict to dogs and cats, the test animals can either survive or die with some probability. Either ways, the hedge fund will turn a profit that amounts to 60% of the initial investment.

The experiment is carried out successfully. The hedge fund earns a profit of 60% on top of the initially invested capital. Yet dogs and cats die with probability 0.75, i.e. few of them survive and most of them die.

Did the hedge fund cause the survival of few of the test animals intentionally?

A) YES;

B) NO.

Results	YES	NO	*Significance*
Task 3 (N=69)	51%	49%	$\chi^2 = 0.14\ (1)\ p = 0.904$
Task 4 (N=69)	32%	68%	$\chi^2 = 9.058\ (1)\ p = 0.003$
Task 5 (N=69)	42%	58%	$\chi^2 = 1.754\ (1)\ p = 0.185$
Task 6 (N=69)	61%	39%	$\chi^2 = 3.261\ (1)\ p = 0.071$
Task 7 (N=69)	25%	75%	$\chi^2 = 17.754\ (1)\ p = 0.000$
Task 8 (N=69)	64%	36%	$\chi^2 = 5.232\ (1)\ p = 0.022$
Task 9 (N=69)	35%	65%	$\chi^2 = 6.391\ (1)\ p = 0.011$
Task 10 (N=69)	57%	43%	$\chi^2 = 1.174\ (1)\ p = 0.279$
Task 11 (N=69)	70%	30%	$\chi^2 = 10.565\ (1)\ p = 0.001$
Task 12 (N=69)	36%	64%	$\chi^2 = 5.232\ (1)\ p = 0.022$

Table 3 - Experiment 3: results

In both Task 3-4 and Task 5-6, the Knobe Effect nullifies again as the results show that, in the harm-case, there is no dominant judgement due to the lack of

statistical significance. Most likely, the Knobe Effect is mitigated by the presence of both the probabilistic factor and the availability heuristic triggers in the thread of the vignette cases. Indeed, while the uncertainty factor is present, the firefighters and the NGO are unlikely to be perceived as unfriendly.

On the other hand, the same as in Task 3-6 is much evident in the reverse way. Indeed, in task 7-8 the hedge fund is run by investment bankers and there are no more dogs and cats, but puppies of panthers and tigers purchased on the black market. In this case, the Knobe Effect obtains regardless of the fact that the hedge fund takes a business decision with probable outcomes.

Eventually, once Task 2.1, 2.2, 2.3, 2.4 are repeated in Task 9-12, the experimental results of *Experiment 2* are confirmed. For, QED, the Knobe Effect obtains only for the side-effects that are likely to obtain with strong intensity.

Concluding Remarks

In the light of the results presented in the previous section, there is room to argue that the way folks perceive intentionality might be driven by some stereotypes concerning the agent who carries some action. In this sense, a firm's VP is likelier to look more unfriendly than an NGO who operates in underdeveloped countries or than a crew of firefighters. Moreover, it seems that if two outcomes (one big and one small) take place simultaneously, then ordinary folks judge the bigger outcome as more intentional than the smaller outcome. This is the case once the protagonists of the vignette take a decision with probable outcomes and different intensity. Eventually, the presence of triggering words (e.g. harm-help or similar) affects judgement. Thus, there is room to argue that the Knobe Effect is sensitive to framing and heuristic-related problems.[38,39]

[38] The first experiment of this study was partially supported by the National Science Centre of Poland (NCN) under the Grant [DEC-2015/17/D/HS6/02684] assigned to Lukasz Markiewicz, PhD (Kozminski University) who has kindly shared part of his funding to support this study. The second and the third experiment of this study were partially supported with the internal funds BST 2017 assigned to the Department of Modern Philosophy of the University of Warsaw. Hence, we would like to thank Lukasz Markiewicz, PhD (Kozminski University) and Prof. Marcin Poręba (University of Warsaw) for funding our work.

[39] We would like to thank the reviewers of this study for their time and professionalism. Last but not least, we are very much obliged towards Prof. Domenico Buccella (Kozminski University), Prof. Katarzyna Paprzycka (University of Warsaw), Konrad Werner, PhD (University of Warsaw), and Adrian Ziolkowski (University of Warsaw) for their insightful comments about the earlier versions of this study, which eventually made us understand where our research was heading. Thank you all for your help, time and support.