DECIDING WHETHER AND HOW MUCH TO HELP – ON JUSTIFICATION OF HELPING DECISIONS

Michał Kłusek

Katedra Filozofii Prawa i Etyki Prawniczej Wydziału Prawa
i Administracji Uniwersytetu Jagiellońskiego
ORCID ID: 0000-0003-3673-4332

Introduction

The growing experimental philosophy movement has been keenly interested in the intuitions behind our moral decisions and beliefs (e.g. Sytsma and Livengood 2015; Knobe and Nichols 2017). Much work has been done based on the psychological mechanisms underlying our decisions to hurt others (e.g. Greene 2016). The psychological basis for our helping intuitions, decisions and beliefs have received comparably little attention. This paper is a step in correcting this imbalance.

We make moral decisions every day, including decisions on whom and how much to help. On a given day, I might decide to give a sum of money to a homeless person or to donate to a charitable organisation. Philosophical reflection can help us assess whether our moral intuitions, decisions and beliefs are right or wrong. In other words, it can tell us what the moral truth is. It can also help us assess what we ought to believe is morally right – which moral intuitions, decisions and beliefs are justified. So much for philosophical reflection. What about psychology, neuroscience and cognitive science?

It is doubtful whether they can tell us much about which moral decisions are right. However, they can certainly help us to judge certain decisions as more or less justified. There are a several ways to do that. In this paper, I argue that moral intuitions and decisions are justified to

the extent that they respond to morally relevant factors. I also examine the extent to which our helping decisions and intuitions do that.

I will be considering, on the one hand, helping decisions regarding identified victims and statistical victims and, on the other – decisions to help and decisions on the size of help. Whereas the distinction between helping identified and statistical victims is clear, the difference between a decision to help and a decision on the size of help may seem unhelpful and artificial. After all, the two decisions seem to occur simultaneously and impossible to tell apart. One could plausibly claim that there is just one decision – to help person X in some way. However, they can be distinguished in a controlled environment of a psychological study.

Psychological research allows us to ask the following questions: Are our decisions more justified when helping identified victims or statistical ones? Are we more justified when we decide whom to help, or when we decide on the size of the help? The paper answers both.

Section 1 consists in a description of the relevant psychological research. Section 2 contains a short description of the positive and negative versions of 'the argument from What Our Intuitions Track', as it has been presented by Guy Kahane (2013; 2016). Section 3 combines the two and presents two claims: that our decisions to help statistical victims are more justified than our decisions to help identified ones, and that our decisions on whom to help are more justified than our decisions on the size of the help. Summary and discussion follow.

Psychological Research

The identifiability effect refers to a greater willingness to help identified people rather than merely statistical ones. Ever since the distinction was first introduced in the late 1960s (Schelling 1968), the psychological bias has been in the lab and in the field. For example, we have learned that people's donations to a child in-need are 60% larger when its name, age and photograph are provided (Kogut and Ritov 2005a), that we are more willing to share our money with others when given their identifying information (Bohnet and Frey 1999), and that donations to a charitable organization are larger if the participants are told the beneficiaries have been chosen in advance (Small and Loewenstein 2003).

There are good reasons to believe that the bias is driven by negative affect and emphatic pain. In one of the aforementioned studies, it

was shown that the greater generosity towards an identified victim has its basis in a stronger emotional reaction, and the participants reported a greater compassion towards identified victims (Kogut and Ritov 2005b). We also know that negative affect evoked by emphatic pain leads to greater donations and increases prosocial behaviours (Small and Verrochi 2009; Hein et al. 2010; Masten et al. 2011).

Based on those insights, a recent series of studies conducted by Szumowska et al. (2018, unpublished manuscript) aimed to examine whether the identifiability effect can be explained using the motivational readiness framework (Kruglanski et al. 2014). The results of the studies are worth a closer look, as they will serve the philosophical argument in Section 3.

In a nutshell, the motivational readiness theory explains willingness to act as determined by an interaction of two ingredients: *Want* and *Expectancy*.

$$MR = f(W, E)$$

Want refers to an individual's desire of some sort. *Expectancy* refers to the individual's subjective assessment of the probability that he will be able to satisfy the desire. *Want* is the main, crucial and indispensable driver of motivational readiness. There may exist some motivational readiness in the absence of *Expectancy*, but not in the absence of *Want*. *Expectancy* merely makes *Want* satisfaction seem more realistic, it serves an assisting role. Also, there is a minimal threshold level of both *Want* and *Expectancy* that needs to be reached for an individual to act.

The researchers run a series of studies in which they presented the participants with hypothetical cases and asked how much money they would be willing to donate to fund a medical procedure saving the life of a child (or of multiple children). Participants' affective response was measured, as well as the levels of their *Want* (desire to help) and *Expectancy* (perceived likelihood of success). They were also asked to decide whether they would wish to help the child and how much they would want to donate. Some of the presented scenarios involved identified victims and some involved statistical ones, which allowed the researchers to compare the participants' reactions. In those studies, identified victims were identified by names and were singular. The statistical victims were neither.

The results showed that the participants' experienced a stronger negative affective response when presented with cases involving identi-

fied victims than when presented with cases involving statistical ones. The negative affect predicted *Want* but had no effect on *Expectancy*. Both *Want* and *Expectancy* were significant predictors of the decision to donate to identified victims. When it comes to the decision on the magnitude of help, only *Want* was a significant predictor of the declared amount when both variables were included in the equation. Only *Expectancy* was a significant predictor of the decision to help statistical victims, *Want* had either a marginal effect or no effect at all.

Having established the psychological facts about helping decisions and intuitions, let us now sketch the philosophical argument we will use in combination with the empirical evidence to reach normatively valuable conclusions.

The Argument from What Our Intuitions Track

It is doubtful whether we can infer what we ought to do from psychological evidence without adding controversial metaethical premises. There are, however, ways to infer not what we ought to do, but what we ought to believe we ought to do. Our moral intuitions, judgements, beliefs and decisions stem from internal processes and mechanisms. Knowledge about the origins of those processes, what they are influenced by or responsive to, can affect our views about the justification of beliefs that stem from those processes.

Let us now take a closer look at the argument from what our intuitions track (Kahane 2013; 2016). It uses empirical evidence concerning what our intuitions are responsive to (or, in other words, what they track). If they respond to factors which are morally irrelevant, we can conclude that the beliefs are unjustified, in so far as they respond to the morally irrelevant factor.

Negative Argument from What Our Intuitions Track
P1. Empirical. Our intuitions that p respond to X.
P2. Moral. X is not a morally relevant factor.
Therefore
C. Normative. Our belief that p lacks justification.

We can also construct a positive mirror-image of the argument, where the fact that our intuitions respond to factors that are morally relevant gives us a *prima facie*, defeasible reason to hold a moral belief.

Positive Argument from What Our Intuitions Track

P1. Empirical. Our intuitions that *p* respond to *X*.

P2. Moral. Factor *X* is morally relevant, or at least does not seem morally irrelevant.

Therefore

C. Epistemic. We have *prima facie* reason to accept the parallel moral principle concerning *X*.

In this way, empirical evidence about what our intuitions respond to (or track) can lead to non-trivial normative conclusions about whether a belief (or decision stemming from the intuition) lacks justification or whether we have reasons to accept it.

Both versions of the argument require an additional moral premise concerning whether the X-factor is morally relevant. In each case this can be debatable. Our assessment of the moral relevance of certain factors can itself be responding to morally irrelevant factors, be shaped by epistemically irrelevant influences, be an outcome of an epistemically defective process etc. There are, however, many cases when our intuitions respond to factors that are clearly morally irrelevant. In Section 3 I will argue that our intuitions and decisions on whom to help are like that.

More importantly, it is not unlikely for our moral intuitions to respond to many factors, some of which are morally relevant, and some are not. Again – our helping decisions are like that. In such case, we can only conclude that our moral intuitions are unjustified to the extent that they respond to morally irrelevant factors and, *vice versa*, the more they respond to morally relevant factors, the stronger our *prima facie* reasons to accept them.

Finally, we need to distinguish between a moral intuition being responsive to morally irrelevant factors and being shaped by epistemically irrelevant influences. Morally irrelevant factors are present in the scenario. Epistemically irrelevant influences affect the agent when she is making a moral judgement. To quote Kahane (2016):

> "If the mere presence of an unpleasant smell in the room makes people judge that some hypothetical transgression should be punished more severely, this is an epistemically irrelevant influence. But if people judge a transgression more severely when the person committing the hypothetical transgression is described as smelly, that would (in most cases!) be a morally irrelevant factor."

Now that we have sketched the argument in a general form, let us return to psychological research on the identifiability effect. We will see the argument applied to helping decisions.

Justification of helping decisions

Both versions of the argument from Section 2 come with two premises. Firstly, that our intuitions that *p* respond to a factor *X*. Secondly, that the factor is or is not morally relevant. If the factor is irrelevant, our belief that *p* lacks justification in so far as it responds to factor *X*. If the factor is morally relevant, we have *prima facie* reasons to hold the corresponding belief.

We need to answer two questions before we use the arguments. Which of our intuitions respond to factor *X*? And – is factor *X* morally irrelevant?

As we have seen from the psychological research, our decisions to help are predicted by *Want* and *Expectancy*. The decisions on the size of help are predicted by *Want*, which in turn is predicted by the negative affect. In other words, our desire to help (*Want*) responds to the negative affect. The affect, in turn, responds to the identifiability of the victim (factor *X*), which in the aforementioned studies refers to the mere presence of identifying information about the victim – a name combined with singularity of the victim. Is the identifiability of the victim a morally relevant factor? I claim that it is not.

Some have argued that there are compelling reasons to give preference to identified victims. However, such arguments focus on a different aspect of identifiability. For example, for both Caspar Hare (2012) and Johann Frick (2015) the morally relevant factor is whether the victims are certain to suffer harm, or whether it is merely probable. Neither of them sees the mere knowledge of a person's name as a morally relevant factor. Mere identifiability is morally irrelevant. It surely is uncontroversial that whether or not we know a child's name ought not to influence whether or not we help them.

On the other hand, *Expectancy* is morally relevant. The subjective estimate of the likelihood of successful help is surely not morally irrelevant.

Having established that our desire to help (*Want*) responds to a morally irrelevant factor (identifiability), and that *Expectancy* does not, and is in itself is a morally relevant factor, we can use the argument from what our intuitions track to reach substantive conclusions.

1)   All else being equal, our decisions to help statistical victims are more justified than our decisions to help identified victims.

Whereas in the case of identified victim the main predictor of the decision to help is *Want*, in the case of statistical victims the main predictor is *Expectancy*. *Want* responds to negative affect, which responds to a morally irrelevant factor – identifiability of the victim. *Expectancy*, which simply refers to the subjective probability of successful help, does not track a morally irrelevant factor, and is itself morally relevant. We can conclude that, all else being equal, our decisions when helping statistical victims are more justified, as they respond to a clearly morally relevant factor and do not respond to a factor that is clearly morally irrelevant.

2)   All else being equal, in the case of identified victims, our decisions on whom to help are more justified than our decisions on the magnitude of help.

Both *Want* and *Expectancy* significantly predict the decision to help, but only *Want* is a significant predictor of the magnitude of help. We have established that *Want* tracks a morally irrelevant factor, and that *Expectancy* is itself a morally relevant factor. We can conclude that our decisions to help are more justified than our decisions concerning the magnitude of help. The former include a clearly morally relevant factor (*Expectancy*) whereas the latter do not. However, the fact that our decisions to help track a morally relevant factor does not mean they are justified. This is because they are also significantly predicted by *Want*, which responds to a morally irrelevant factor. What is more, *Want* is a stronger predictor of our decision to help than *Expectancy*. We can conclude that, all else being equal, our decisions to help are more justified than our decisions on the magnitude of help, but that does not mean they are justified in themselves.

Summary and discussion

In this paper I have argued that, based on psychological research, we have reasons to doubt our moral intuitions when helping identified victims, that our decisions on whom to help are more justified then our decisions on the size of help, and that we have *prima facie* reasons to trust our intuitions when helping statistical victims.

The paper also exemplifies the difficulties one faces when trying to reach normative conclusions concerning the justification of moral intuitions, decisions or beliefs. Our intuitions often respond to more than one factor, some of which are morally relevant, and some are not. We can assess the importance of different factors by measuring the strength of the responsiveness of the intuitions. This, however, forces us to ask additional questions about the comparative moral relevance of different factors. Finally, the fact that a moral intuition responds to a morally irrelevant factor does not mean that it responds only to that factor and that the factor is the most important one. I have presented psychological research supporting the argument that our desire to help (*Want*) responds to negative affect and is subject to the identifiability effect, which is clearly morally irrelevant. However, in each particular case our desire to help is likely shaped by many factors, the negative affect being just one of them. The affect itself may respond to many factors, the identifiability of the victim being just one of them.

Finally, I have shown that negative affect (and therefore our desire to help) responds to a morally irrelevant factor (the identifiability of a victim). One could also argue that responding to a negative affect is an epistemically defective process in itself, regardless of whether it tracks a morally relevant factor (Nichols 2014). This, however, is an altogether different matter.

References

Bohnet I., Frey B.S. (1999). The Sound of Silence in Prisoner's Dilemma and Dictator Games. *J Econ Behav Organ* 38(1): 43–57. https://doi.org/10.1016/S0167–2681(98)00121–8.

Frick J. (2015). Contractualism and Social Risk. *Philos Public Aff* 43(3): 175–223. https://doi.org/10.1111/papa.12058.

Green J. (2016). Solving the Trolley Problem. In: Sytsma J., Buckwalter W. (eds.), *A Companion to Experimental Philosophy*. Wiley-Blackwell, Hoboken, pp. 175–189. https://doi.org/10.1002/9781118661666.ch11.

Hare C. (2012). Obligations to Merely Statistical People. *J Philos* 109(5–6): 378–390. https://doi.org/10.5840/jphil20121095/615.

Hein G., Silani G., Preuschoff K., Batson C.D., Singer T. (2010). Neural Responses to Ingroup and Outgroup Members' Suffering Predict Individual Differences in Costly Helping. *Neuron* 68(1): 149–160. https://doi.org/10.1016/j.neuron.2010.09.003.

Kahane G. (2013). The Armchair and the Trolley. *Philos Stud* 162(2): 421–445. https://doi.org/10.1007/s11098–011–9775–5.

Kahane G. (2016). Is, Ought, and the Brain. Im: Liao S.M. (ed.), *Moral Brains: The Neuroscience of Morality.* Oxford University Press, Oxford, pp 281–311. https://doi.org/10.1093/acprof:oso/9780199357666.003.0013.

Knobe J., Nichols S. (2017). Experimental Philosophy. *The Stanford Encyclopedia of Philosophy* (Winter 2017 Edition). https://plato.stanford.edu/entries/experimental-philosophy/.

Kogut T., Ritov I. (2005a). The singularity effect of identified victims in separate and joint evaluations. *Organ Behav Hum Decis Process* 97: 106–116. https://doi.org/10.1016/j.obhdp.2005.02.003.

Kogut T., Ritov I. (2005b). The „identified victim" effect: an identified group, or just a single individual? *J Behav Decis* 18(3): 157–167. https://doi.org/10.1002/bdm.492.

Kruglanski A.W., Chernikova M., Rosenzweig E., Kopetz C. (2014). On motivational readiness. *Psychol Rev* 121(3): 367–388. https://doi.org/10.1037/a0037013.

Masten C.L., Morelli S.A., Eisenberger N.I. (2011). An fMRI investigation of empathy for 'social pain' and subsequent prosocial behavior. *NeuroImage* 55:381–388. https://doi.org/10.1016/j.neuroimage.2010.11.060.

Schelling T. (1968). The Life You Save May Be Your Own. In: Chase S.B. Jr. (ed.), *Problems in Public Expenditure Analysis.* Brookings Institution, Washington, pp 127–162.

Small D.A., Loewenstein G. (2003). Helping a Victim or Helping the Victim: Altrusim and Identifiability. *J Risk Uncertain* 26(1):5–16. https://doi.org/10.1023/A:1022299422219.

Small D.A., Verrochi N.M. (2009). The Face of Need: Facial Emotion Expression on Charity Advertisements. *J Mark* 46(6):777–787. https://doi.org/10.1509/jmkr.46.6.777.

Sytsma J., Livengood J. (2015). *The Theory and Practice of Experimental Philosophy.* Broadview Press, Peterborough.

Abstract

For a long time, experimental moral philosophers have been interested in mental processes underlying decisions to harm. Psychological mechanisms behind decisions to help received far less attention. This paper

makes use of psychological research on identifiability effect to assess how justified different kinds of helping decisions are. The effect can be explained using the motivational readiness theory, which states that willingness to act is determined by an interaction of two components: Want and Expectancy. The former refers to desire of some sort and is the main driver of motivational readiness. The latter refers to subjective probability of attaining the object of Want. The 'argument from what our intuitions track' claims that whether a given moral belief is justified is affected by whether it responds to morally relevant factors. Research has shown that, in the case of helping decisions, only Want, but not Expectancy is predicted by negative affect, which in turn is predicted by the identifiability of the victim. Both components predict the decision to help identified victims, but only Want predicts the decision on the size of help. However, only Expectancy is a significant predictor of the decision to help statistical victims. I use two versions of the argument from what our intuitions track to make two claims: First, that our decisions to help statistical individuals are more justified than our decisions to help identified individuals; and, second, that our decisions on the size of help are less justified than our decisions whether to help at all.

Keywords: helping decisions – identifiability effect – intuitions – motivational readiness

Czy i jak bardzo pomóc - o uzasadnieniu decyzji o pomocy

Abstrakt

Procesy psychologiczne towarzyszące decyzjom o tym, czy zadać komuś krzywdę, od dawna znajdują się w centrum zainteresowania eksperymentalnej filozofii moralnej. Mniej uwagi poświęca się decyzjom o pomocy. W tym artykule korzystam z badań psychologicznych nt. efektu ofiary zidentyfikowanej by ocenić w jakiej mierze uzasadnione są różne decyzje pomocy. Efekt ofiary zidentyfikowanej to prawidłowość psychologiczna w ramach której jesteśmy znacznie bardziej chętni pomagać ofiarom zidentyfikowanym (np. przy pomocy imienia lub zdjęcia) niż ofiarom niezidentyfikowanym (statystycznym). Efekt daje się wytłumaczyć przy pomocy teorii gotowości motywacyjnej (*motivational readiness theory*), która zakłada, że za naszą gotowość do działa-

nia odpowiada wzajemne oddziaływanie dwóch czynników: *Pragnienia* (*Want*) i *Oczekiwania* (*Expectancy*). Pierwszy z czynników w największej mierze odpowiada za gotowość motywacyjną pomagania. Drugi oznacza subiektywną ocenę prawdopodobieństwa, że przedmiot *Pragnienia* zostanie osiągnięty. Filozoficzny argument „z tego, za czym podążają intuicje" sprowadza się do konstatacji, że to, w jakiej mierze określone przekonania moralne są uzasadnione, zależy od tego, czy intuicje, na których bazują, odpowiadają na/podążają za relewantnymi moralnie czynnikami. Psychologiczne badania pokazują, że w przypadku decyzji o pomocy, *Pragnienie*, lecz nie *Oczekiwanie*, w dużej mierze daje się przewidzieć na podstawie negatywnego afektu. Intensywność negatywnego afektu zależy zaś od tego, czy ofiara, której chcemy pomóc, jest zidentyfikowana. Oba czynniki pozwalają przewidzieć decyzję o pomocy ofiarom zidentyfikowanym, ale tylko *Pragnienie* pozwala przewidzieć decyzję o tym, jak duża ma być pomoc. Jedynie *Oczekiwanie* pozwala przewidywać decyzje o pomocy ofiarom statystycznym. Używając dwóch wersji argumentu z tego, za czym podążają nasze intuicje, stawiam dwie tezy: po pierwsze, że nasze decyzje o pomocy ofiarom statystycznym są bardziej moralnie uzasadnione i po drugie, że nasze decyzje o wielkości pomocy są mniej moralnie uzasadnione niż decyzje o tym, czy w ogóle pomóc.

Słowa kluczowe: decyzje pomagania, efekt ofiary zidentyfikowanej, intuicje, gotowość motywacyjna

Nota o autorze

Michał Kłusek, doktorant, zainteresowania naukowe: eksperymentalna filozofia moralna, ekonomiczna analiza prawa, filozofia polityczna. Ostatnia publikacja: *The Identified Victim Effect – Utilitarian Analysis and Policy Recommendations*, „Principia" 65, s. 69–89. Doi: 10.4467/20843887PI.18.003.9886
Adres e-mail: Michal.klusek@doctoral.uj.edu.pl