



# Doing, Allowing, Gains, and Losses

Camilla Colombo<sup>1</sup>

Accepted: 22 October 2018 / Published online: 7 November 2018  
© The Author(s) 2018

## Abstract

This paper examines Kahneman and Tversky's standard explanation for preference reversal due to framing effects in the famous "Asian flu" case. It argues that, alongside with their "loss/no gain effect" account, an alternative interpretation, still consistent with the empirical data, amounts to a more reasonable psychological explanation for the preference reversal. Specifically, my hypothesis is that shifts in the baseline induce shifts in the agents' classification of the same action as "doing harm" rather than "allowing harm to occur", and that people are risk-seeking when it comes to avoid causing extra deaths—doing harm—and risk-averse when it comes to preventing more deaths—by the means of allowing other deaths to occur as a side effect. I then survey the two most influential concurrent accounts in the moral literature, with respect to the relation between the loss/no gain and the doing/allowing distinction: Horowitz's reductionist conclusion, which argues that the latter collapses into the former, and Kamm's rebuttal, which claims instead that the two distinctions can be pulled apart. I eventually explain why my interpretation differs from both these positions.

**Keywords** Doing and allowing · Prospect theory · Loss/no gain effect · Endowment effect

## 1 Introduction

Imagine you are faced with the following dilemma:

Your city is threatened by an "Asian flu" which is expected to kill 600 people, and you have to make a choice between these two alternative vaccination programs.

- If Program A is adopted, 200 people out of the 600 will be saved.
- If Program B is adopted, there is 2/3 probability that no-one will be saved, and 1/3 probability that all 600 people will be saved.

---

✉ Camilla Colombo  
C.F.Colombo@lse.ac.uk

<sup>1</sup> London School of Economics, London, UK

Which program would you choose?

Imagine now that, in the very same scenario, you could choose instead between C and D:

- If Program C is adopted, 400 people out of the 600 will die.
- If Program D is adopted, there is 1/3 probability that no-one will die, and 2/3 probability that 600 people will die.

What would be your decision this time?

If your choices were A and D respectively, your answers fit the experimental results famously discussed by Daniel Kahneman and Amos Tversky in their 1983 paper “Choice, Values, and Frames”.<sup>1</sup> The authors presented two groups of subjects with the same decision problem, but phrased in different forms; specifically, the first group had to choose between A and B and the second group between C and D. A and C, like B and D, are however extensionally equivalent, and describe the same vaccination program: “200 people will be saved and 400 will die” (A and C) and “there is 1/3 probability that 600 people will be saved and no-one will die and a 2/3 probability that no-one will be saved and 600 people will die” (B and D). Therefore, we could reasonably expect that the percentage of people opting for A and C (or for B and D) were similar in the first and the second group. Nonetheless, experimental findings showed that 72% of subjects in the first group chose Program A but, in the second group, 78% of subjects chose Program D.<sup>2</sup>

With the “Asian flu” case, together with other experimental findings, K&T brought attention for the first time to subjects’ conflicting choices when it comes to deciding over the same options described differently. K&T thus argue that, while making decisions, people are affected by how the decision problem is framed. More accurately, they suggest that, by setting a different reference point, two extensionally equivalent decision problems can trigger some amount to be perceived as either a potential gain or a potential loss (which is, in the “Asian flu” case, the amount of lives saved or lost). As people tend to show opposite risk attitudes with respect to gains and losses, they reverse their preferences depending on whether the outcomes are described as losses rather than gains from the reference point.

More technically, K&T show that value functions are usually convex for losses and concave for gains, and steeper for losses than for gains, and thus claim that shifts in the baseline can cause preference reversals. With respect to risk attitudes, this means that decision makers are more risk seeking when it comes to avoid a (perceived) loss with respect to the reference point, and more risk averse when it comes to pursue a (perceived) gain with respect to the reference point. The shift in the reference point can thus motivate different behaviours and result into preference reversal. Furthermore, K&T try to identify the psychological mechanism behind these risk attitudes as the “endowment effect” of the baseline.<sup>3</sup> By endowment effect, in psychology and behavioural economics, we mean the tendency to ascribe more value to things merely because we own them or, more generally, by the tendency to normativize the reference

<sup>1</sup> In *American Psychologist*, 39(4), pp. 341–350. From now on, “K&T” will refer to Kahneman and Tversky’s positions in this specific paper.

<sup>2</sup> In their experiment, K&T did not ask the same person to choose between A and B and C and D, but rather divided the subjects into two groups. The reader, therefore, can not be put in the same experimental setting designed by the authors, as she can not evaluate the two prospects independently. We could still reason about this case thinking whether we have a different perception of the health plans described, and whether our intuitions over the two decision problems differ.

<sup>3</sup> See, in addition to the 1983 paper, Kahneman et al. (1991).

point, i.e., assigning to the features of the situation described as the baseline a normative status.<sup>4</sup>

In this paper I first describe in detail how, according to K&T, the endowment effect of the baseline can explain the preference reversal in the “Asian flu” case (section 1). I then argue that, in the “Asian flu” case, K&T’s experimental results are consistent as well with the following interpretation: shifts in the baseline induce shifts in the agents’ classification of the same action as an instance of “doing harm” rather than “allowing harm to occur” (section 2). My claim is that people are risk seeking when it comes to avoid causing extra deaths—doing harm— and risk averse when it comes to preventing more deaths— by the means of allowing other deaths to occur. In short, I suggest, the same patterns in value functions call for a different psychological explanation, and agents are not only motivated by the perception of the same amount of lives as losses or gains with respect to the reference point. Alternatively, I argue that, in this example, the different risk attitudes could depend on the fact that people have different moral intuitions over the permissibility of doing harm and allowing the same harm to occur. As a result, when the reference point triggers different interpretation of the same action as “doing” rather than “allowing”, this, in turn, brings about different preferences over the two decision problems.

The relation between the doing/allowing distinction and the loss/no gain effect has been widely discussed in the moral literature. Sinnott-Armstrong (2008) and Horowitz (1998), for instance, argue that cases like the “Asian flu” show that the distinction between doing and allowing collapses into the loss/no gain distinction, and could thus be expressed in purely non-moral terms, as the result of different risk attitudes individuals exhibit with respect to losses and gains, together with agents’ sensitivity to framing. Arguably, as Kamm (2007) notices, this conclusion would threaten the significance and relevance of the doing/allowing distinction in moral theory. In order to challenge this position, she claims that a) the loss/no gain and doing/allowing distinctions do not coincide; and b) doing/allowing classifications are not subject to the same framing effects as loss/no gains perceptions are. In this paper, I do not address the entire doing/allowing debate, but I focus instead on the specific “Asian flu” example, trying to clarify my interpretation with respect to Kamm’s account (section 3). First, I argue that my hypothesis is consistent with Kamm’s claim a), and thus does not support the conclusion that the doing/allowing distinction can be only described in terms of losses and gains. On the other hand, I claim that Kamm’s analysis does not support b), and my interpretation acknowledges for the possibility that the doing/allowing distinction is subject to framing. Seemingly, the fact that both the doing/allowing and the loss/no gain distinctions are sensitive to framing effects, and induce similar risk attitudes, does not by itself exclude the autonomy and independent significance of the former. In section 4, I conclude with some remarks on how framing-dependence impacts the moral significance of the doing/allowing distinction.

### 1.1 K&T’s Interpretation: The Endowment Effect of the Baseline

In the introduction, I outline the two different descriptions of a choice problem agents are faced with in the “Asian flu” case, and report the experimental findings. This example is representative of a class of cases K&T use to theorize the notion of a *decision frame*. By a decision frame, the authors refer to “the decision- maker’s conception of the acts, outcomes, and contingencies associated with a particular choice”.<sup>5</sup> Roughly, a decision frame amounts to

<sup>4</sup> The endowment effect is defined as such by Horowitz (1998). Also, see Cushman et al. (2008).

<sup>5</sup> K&T, p. 455.

the way a choice problem is described to and perceived by the person who makes the choice. As “the frame that a decision-maker adopts is controlled partly by the formulation of the problem and partly by the norms, habits, and personal characteristics of the decision-maker”,<sup>6</sup> the same decision problem will often have several decision frames, even for the same decision maker (thus holding fixed the personal characteristics and changing only the formulation). In particular, they notice that, in very many cases, two decision frames do not involve different factual descriptions of the world, but rather assume a different reference point as the *baseline*.

By way of elaboration: according to standard expected utility theory, the change in a decision frame, or, more specifically, a change in the reference point, should be irrelevant to the choice of a course of action. Against this, K&T point out that decision makers do change their behaviour depending on the decision frame, as the “Asian flu” case clearly illustrates. This violates consistency norms rational decisions should supposedly conform to, as the two frames are, with respect to standard decision theory, two descriptions of situations that are identical in all relevant aspects.

Let’s now see in detail how reframing supposedly explains preferences reversal in the “Asian flu” case. K&T offer a specific interpretation of individuals’ behaviour, which has become the standard account in choice theory circles. Specifically, K&T argue that 1) the reference point matters and 2) people are more risk seeking when it comes to avoid sure losses from a given baseline, and more risk averse when it comes to pursue gains from a given reference point. In Kahneman and Tversky (1979) and in Kahneman et al. (1990), the difference in risk attitudes summarized in 2) is further examined and associated with a psychological mechanism known as *endowment effect*. With the support of further empirical evidence, the authors observe that agents prefer to avoid losses rather than to acquire equivalent gains, as they seem to value more an object (or an amount of a given currency) if they already “own” it, or they feel somehow attached or entitled to it; this, arguably, makes it worse for them to lose it with respect to the enjoyment they would experience in gaining the same object (or amount of currency).<sup>7</sup> As a result, people are less prone to take the risk of improving from the baseline and more prone to take the same risk to avoid getting worse compared to the baseline. Cases of endowment effect are ubiquitous, and are usually investigated in bargaining settings, by measuring the agents’ willingness to pay. One famous example of the psychological pull of the endowment effect, for instance, is the so called Knee example.<sup>8</sup> K&T polled subjects to ask an amount of money a) *in compensation* for not getting a lost knee back and b) *in exchange* for losing a knee. The results show that agents demand more money ex ante (b), i.e., when they are faced with the possibility of losing their knee, than they demand ex post (a), i.e., when they are told they don’t have a knee and could get it back. Apparently, this supports the conclusion that they value more the knee they already have with respect to the knee they would get back, even assuming that the new knee would be equal to the old one under all respects.

Let’s now take a closer look at how the endowment effect would work as an explanation of the “Asian flu” case. In the first decision problem (the choice between A and B), the use of the phrasing “saving” identifies the 200 lives as a gain, thus seemingly setting the adequate reference point at “all 600 people die”. With respect to the baseline “everyone dies”, choosing program A would amount to a sure gain from the reference point. Plan B, on the other hand, characterises a “bet”, as it involves evaluating a risky prospect. Specifically, with respect to the

<sup>6</sup> K&T, p. 455.

<sup>7</sup> Kahneman et al. (1990), p. 195.

<sup>8</sup> Kamm (2007) also compares the “Asian flu” case with the Knee case.

baseline “everyone dies”, Plan B could either deliver a bigger gain (all 600 people saved) or just make no progress at all from the baseline (all 600 die). When it comes to gains, K&T observe, decision makers tend to be risk averse, and, given the same expected lives saved of A and B, most opt for plan A, which guarantees a sure gain. In the second decision problem, the different framing of the decisions triggers a different evaluations of vaccination plans. Plan C, indeed, apparently presents the option of 400 people dying as a loss, as it uses the phrasing “die”; this description thus sets the baseline at “all 600 people live”. With respect to this reference point, Plan C involves therefore a sure loss. Plan D, again, amounts to a bet, where either losses with respect to the baseline are completely avoided (600 people saved) or a bigger loss could occur (all 600 people die). While C and D are expected-lives-saved equivalent, decision makers mostly opt for D, showing a preference reversal when compared to the choice between A and B.

To sum up, different framings select different reference points as the relevant baseline, namely “everyone dies” vs “everyone lives”, and this, in turn, induces a different perception of the options as losses rather than gains. Because of the endowment effect, agents would then tend to normativize the reference point, and to value more the same numbers of lives when they feel they already “own” them (or they feel they are already secured); therefore, people are supposedly risk seeking when it comes to avoid losing lives which are framed as losses with respect to the reference point “everyone lives”, and risk averse when it comes to save lives which are framed as gains from the reference point “everyone dies”. Consistently, they would tend to choose the course of action which involves a chance to completely avoid any loss (plan D over plan C), but are not as eager to take the same risk to save more lives (plan A over plan B).

This explanation, arguably, holds for all the cases where shifts in the baseline trigger a different perception of the same option as a loss from the reference point rather than a no-gain from a lower reference point. For this reason, the baseline sensitivity observed by K&T is often referred to as the “loss/no gain” effect.

## 1.2 The “Doing/Allowing Effect” of the Baseline

In this section, I argue that K&T’s empirical findings are consistent as well with a different interpretation.

The “Asian flu” case is in some respects different from other examples surveyed by Kahneman and Tversky, like the classical Knee case. In the latter, the knee is an object, or, more accurately, a possession, which agents are merely told to have or to be deprived of, without any further specification of how the loss is brought about. The former, on the other hand, involves a choice agents are asked to make, a choice which appeals to a moral sphere and has an impact over other people’s lives; the focus, arguably, is not just on the end states, like “knee” and “no knee”, but on how these end states are brought about by the agent’s decision. While the knee, or its absence, can thus be described as a loss or a gain from a given baseline, the number of lives saved or lost in the “Asian flu” example is also intrinsically related to the specific way the agent behaved and chose among options. At the end of section 1, when describing the endowment effect in the two decision problems, I explained the agents’ attitude towards the baseline as “owning a life”, or, slightly better “having a life secured”. I find this characterization, however, not persuasive as an adequate psychological explanation of the different intuitions we have over the permissibility of the health plans. I argue, therefore, that an additional element plausibly helps account for the observed choices in the “Asian flu” case, which acknowledges for the fact that the decision problems involve a moral choice.

Specifically, my hypothesis is that, in the moral dilemma set by the “Asian flu” case, agents’ assessment and appraisal of the options on the table depends on whether the latter are perceived as instances of *doing harm* rather than of *allowing a harm* to occur. In this paper, I do not endorse any particular interpretation of the doing/allowing distinction, and I do not take sides in the ongoing debate over the legitimacy and role of this dichotomy.<sup>9</sup> For the purposes of my analysis, it is sufficient to notice that, in many circumstances, agents share a strong and widespread intuition that doing a harm and allowing the same harm to occur amount to different moral conducts, and should be judged differently. In some pairs of cases, this distinction seems particularly salient: if I am jogging on the waterfront, I see a person drowning in the pond, and refrain to rescue her, I am allowing the death of this person by drowning. On the other hand, if I drown someone in the pond, I am actively doing a harm.<sup>10</sup> In this class of examples, the difference between the two behaviours is seemingly self-evident, as well as the moral value attached to the conducts, as drowning a person is for most people more wrong than a failure to rescue.

In moral theory, the persistency of this intuition has motivated several attempts to provide a consistent account of the doing/allowing distinction, and thus justify its employment in moral judgements. As Woollard (2015) puts it, “if there is no moral difference between doing and allowing, then morality must either be far more permissive than we generally suppose – permitting us to kill to protect our personal projects – or far more demanding – requiring constant sacrifice from us to save the lives of others”. To be sure, several objections to the feasibility of this project have been put forward, and no account seems completely successful in accounting for the moral relevance of the doing/allowing distinction.<sup>11</sup> For the limited scope of this section, however, the question of whether the distinction holds a legitimate moral significance can be set aside. My point here is that, at least in some decision contexts, the characterisation of a conduct as “doing” rather than “allowing” strikes most agent as evident; when this distinction is particularly salient, moreover, there is a strong agreement over the fact that doing harm is worse than allowing it to occur. This overwhelming intuition, arguably, must be accounted for even if we do not take the doing/allowing distinction as per se morally significant, but rather tracking other moral categorisations or principles. Whichever the case, the characterisation of a behaviour as an instance of doing rather than allowing triggers different moral responses, independently on how these responses are justified.

Having this said, my argument is that the “Asian flu” example belongs to the class of cases where the doing/allowing distinction is particularly salient, as, I will argue, the phrasing of the decision problem appeals to either doing or allowing characterisations. Agents’ appraisal and evaluation of different courses of actions, therefore, plausibly refers to this deeply-rooted intuition, and should be discussed and analysed accordingly.

My explanation for the preference reversal in the Asian flu case can be thus summarized as follows: different framings of the same decision problem can induce a different perception of the same option as an instance of “doing harm” rather than “doing good while allowing some harm as a side effect”; as agents are influenced by this classification, and regard doing harm as morally worse than allowing it, they are risk seeking when it comes to avoid doing harm (by causing a certain amount of deaths), and risk averse when it comes to doing more good, allowing harm as a side effect (by not saving the same amount of lives). This, eventually, leads to the preference reversal.

<sup>9</sup> For a survey of the main different positions in the doing/allowing debate, see Howard-Snyder (2007).

<sup>10</sup> For a discussion of this example, see Singer (1997).

<sup>11</sup> Famously, this position is upheld by Bennett in Bennett (1997).

I now test whether this hypothesis can successfully account for a different evaluation of the choice between A and B and the choice between C and D in the “Asian flu” case. For plan A and B, as already discussed, the appropriate baseline seems to amount to a situation where everyone dies; agents, as K&T argue, tend to take such baseline as the position they find themselves in. Endorsing plan A, therefore, with respect to this dreadful situation where all 600 people die, arguably appears from the agent’s perspective as actively saving 200 people, and *causing* the fact that they will live. On the other hand, the fact that the other 400 people will die reminds more of a side effect of the good action we are performing: in other words, it sounds more that we are allowing 400 people to die (for sure, a harm), while causing 200 people to live. Plan A, in short, may be interpreted as “doing good and allowing harm (for sure)”. As for plan B, agents must take a gamble: they could do even more good, eliminating completely the chance of even allowing harm, or they can end up allowing more deaths to occur. While equivalent with respect to expected harm (lives lost), most agents stick to plan A, as they are risk averse when it comes to doing good, while allowing harm to occur. This, I argue, happens because, in this specific choice problem, option A is framed as giving the agent the chance to do something good for sure, with respect to the reference point; the harm that will occur as a consequence of her choosing A is framed in a way that appeals to “allowing”, and thus to an occurrence of harm which is not as wrong as “doing”. On the other hand, B could give the chance of doing even more good, thus reducing the possibility of what is framed as allowing harm as a side-effect. I claim that agents are not eager to take this risk because this description frames the harm which will occur as an harm we would be merely “allowing”. Allowing harm, however, is not as objectionable as doing it, and this explains why agents are risk averse.

Turning to the choice between C and D, now the appropriate baseline is “everyone lives”. Plan C, a vaccination plan described as “killing” 400 individuals, appears in that respect as *causing* the death, for sure, of 400 individuals, thus actively harming them. In other words, I argue that the terms used for the description of Plan C are easily associated with a way of bringing about harm which is characterised as “doing”. D, again, describes a lottery where the agent could end up doing even more harm (600 people die), but also not harming anyone (600 people live). Plan D avoids, in this sense, “doing harm for sure”, which seems inevitable when Plan C is selected. Now, the framing of the decision problem suggests that the agents would do harm, which is more morally objectionable than merely allowing it. This difference explains why, in this second decision problems, agents are risk seekers, and more individuals take the chance to avoid doing harm.<sup>12</sup>

To sum up, while agents have of course interest in doing good and thus saving people, and do not want to harm, when the characterisation of the harm reminds more of “allowing” they are more risk averse, because this way of bringing about harm is less morally objectionable. On the other hand, when the description of the harm appeals to the idea of “doing harm”, agents are more risk seeking, as this seems a particularly wrong moral conduct.

---

<sup>12</sup> It could be interesting to discuss as well our risk attitudes in “doing good”. Arguably, they exhibit a symmetrical pattern with respect to our risk attitudes in “not harming”: we feel, indeed, that we have some kind of obligation to do good, but this is plausibly less stringent than the obligation not to do harm. We may, therefore, be risk averse when it comes to do (more) good. An interpretation which only relies on our attitudes in doing good, however, seems to perform worse in explaining both cases- the comparison between A and B and the comparison between C and D. When the baseline is “everyone lives”, indeed, it is not clear how we do good by choosing any of the two options. The distinction between doing and allowing, on the other hand, accommodates both examples, as it explains that we feel more obliged to avoid doing harm than to avoid allowing harm, because doing is worse than allowing.

This interpretation, I argue, is perfectly consistent with the experimental results and risk attitudes described by K&T, but implements this model with the observation that what seems to be really subject to framing is not the number of lives/deaths as losses or gains, but rather whether the same number of lives/deaths is perceived as the result of the agent's "doing harm" or "allowing harm". In this sense, baseline shifts change the perception of the way the agent brought about the same end-states, rather than the perception of the end-states as gains or losses. I call this psychological interpretation, which shares all the features and formal properties of the standard analysis, the "doing/allowing effect" of the baseline. I conclude that this kind of moral evaluations, rather than the endowment effect, can plausibly motivate the preference reversal, and that this hypothesis shares at least the same empirical support.

### 1.3 Loss/no Gain Vs Doing/Allowing

Since K&T's observations, moral philosophers have been concerned about the import of these experimental results on the status of moral judgements and ethical principles. Specifically, as moral classifications and principles are often tested against and inferred from people's intuitions in certain situations, the fact that our perception of cases and even choice behaviour is heavily dependent on framing effects seems to threaten the status and significance of the former. This reasoning is summarized by Walter Sinnott-Armstrong (2008) as the *master argument*,<sup>13</sup> and, in short, claims the following: if our moral intuitions are formed under unreliable circumstances, then they are not justified; as framing effects proves the unreliability of our evaluation of cases, moral intuitions can not be independently justified.

As discussed in section 2, the doing/allowing distinction, together with the moral principle "doing is worse than allowing", is mostly grounded in people's judgements and intuitions over some specific cases and scenarios. In this sense, all theoretical efforts to provide a justification for the different moral significance we attribute to "doing" and "allowing" actions are motivated by the fact that we persistently employ this distinction in real life. As Sinnott-Armstrong (2008) and Kamm (2007) argue, however, if our doing/allowing classifications were the mere result of framing effects, like the endowment effect and the baseline effect, then this whole justificatory project would become pointless.

The "Asian flu" case, among other examples, has been used by advocates of this destructive interpretation as a decisive test<sup>14</sup>: if "doing" and "allowing" attributions collapse into "losses" rather than "no-gains" framings, then K&T's analysis provides an explanation of the doing/allowing distinction in purely non-moral terms and, what is more, as the effect of psychological attitudes, idiosyncracies and reasoning biases. This position is upheld, in particular, by Tamara Horowitz, who argues that this example shows that what most people identify as "doing harm" equals to a loss from a given reference point, while what we perceive as "allowing harm" amounts to a "no gain" from another reference point. The shift in the baseline, which is a matter of phrasing, thus induces these different characterisations, which are nothing more than the effect of the change in the reference point and the framing of the same amount of lives as losses or no gains.

Before turning to Kamm's answer to Horowitz, I notice that my position is different from the one expressed above. While I agree that we may substitute the loss/no gain talk with the doing/allowing distinction, I do not conclude that the latter can be reduced to the former. On the other hand, I claim that the "Asian flu" case presents a moral aspect which can not be

<sup>13</sup> Sinnott-Armstrong (2008), pp. 48–56.

<sup>14</sup> This characterisation is upheld in Sinnott-Armstrong (2008) and Horowitz (1998).



adequately captured by the endowment effect of the baseline and by loss/no gain characterisations. While the risk attitudes remain the same, I claim that the characterisation of actions as instance of “doing” rather than “allowing” is performing the explanatory work.

Kamm famously criticises the thesis that the doing/allowing distinction collapses into the loss/no gain distinction, thus threatening the moral significance of the former.<sup>15</sup> Kamm’s argument can be summarized into two main claims: a) the doing/allowing distinction address a different aspect than the loss/no gain distinction does, and thus the former can not be reduced to the latter; b) the doing/allowing distinction, when properly characterised, is immune to framing effects.

When arguing in favour of a), Kamm starts with the observation that while losses and gains are descriptions referring to end-states, doing and allowing have to do with the way an agent brings about these end-states. As such, “doing” and “allowing” univocally identify the connection between the agent’s behaviour and the outcome. K&T’s and Horowitz’s analysis of the “Asian flu” case, in this sense, is misleading, as it does not account for this additional feature of the doing/allowing characterisation with respect to the loss/no gain description. Upon a closer look, Kamm argues, this example does not simply involve our different attitudes with respect to gains and losses, but rather our different intuitions over what she defines as *unprevented losses* and *denied gains*. In the first decision problem, indeed, agents are confronted with a “nearly death state”; the resulting lives saved with respect to the reference point are thus perceived as gains, but of a specific kind that Kamm defines as “maintaining what one is close to lose”.<sup>16</sup> To see this point, we can imagine that when opting for Program A, we arguably think that we are gaining 200 lives the Asian flu would instead take, if we do not act promptly. On the other hand, the 400 lives lost are a “no gain” which can be interpreted as a *denied gain*: i.e., with respect to the nearly death state, failing to save everyone is merely not maintaining what we are close to lose anyway if we do not act. On the other hand, the second decision problem sets the baseline as “everyone lives”. With respect to such situation, lives are framed as losses, Kamm defines as “losses which are happening independently of our intervention and which we fail to prevent”. To be sure, unprevented losses seems less morally objectionable than *imposed losses*; nonetheless, they are arguably more objectionable than denied gains. This distinction, she claims, explains people’s different responses to the two decision problems.

What this analysis shows is that in the “Asian flu” case the loss/no gain distinction does not fully explain our characterisations of the actions in the first decision problem as “allowing harm” and in the second decision problem as “doing harm”. The loss/no gain distinction, indeed, does not capture the crucial element of human decision-making, or the issue of “how the consequence is brought about”, which is extremely relevant in our doing/allowing attributions. All the experimental findings support, therefore, is that agents are sensitive to the agency-sensitive unprevented loss/denied gain effect, and perceive the same outcome differently depending on whether the outcome is perceived as the result of a conduct framed as the former rather than the latter.

I think that Kamm’s analysis is persuasive, and acknowledges for an issue I discussed in section 2: there is more to the “Asian flu” case than losing or gaining lives, as this choice appears as morally significant. In this sense, we may be interested in question of permissibility, responsibility or the like, which have to do with the way a consequence is brought about and the agent’s intentions (denying a gain rather than failing to prevent an harm). Therefore, my

<sup>15</sup> In this paper, I will not survey all the examples she employs as to argue her case, but I will only refer to the “Asian flu” case.

<sup>16</sup> Kamm (2007), p. 473.

hypothesis is perfectly consistent with claim a), as Kamm simply gives a specific and detailed account of how the shift in the baseline affects the perception and characterisation of behaviours, in a way that goes beyond mere gains and losses. Arguably, the doing/allowing distinction is thought of to conflate with the loss/no gain effect because of the fact that, in most cases, the two are difficult to disentangle, as doing harm often occur in cases where we suffer a loss, and allowing harm in cases where we do not achieve a gain. More accurately, as Kamm argues, it is quite difficult to design fully equalised cases, i.e., a pair of cases which is identical under all respect but are only distinguishable in the fact that one addresses the doing/allowing and the other the loss/no gain distinction. Such an example, however, would be required as to show that, when the loss/no gain effect is controlled for, the doing/allowing distinction disappears: this would prove, indeed, that there is nothing more to doing/allowing distinction than the loss/no gain effect. All examples used by the advocates of the reductionist thesis, however, seemingly share the same structure of the “Asian flu” case: a closer inspection reveals that there could be a further element in the deliberation, addressing the way the outcome is brought about.

So far, Kamm has successfully argued that the doing/allowing distinction does not collapse into the loss/no gain one, as the former does not only refer to end-states but also to how these end-states are brought about. We could think that, in her jargon, instances of imposed or even unprevented losses remind us more of “doing harm”, while instances of denied gains suggest a characterisation of a conduct as “allowing harm”. Reframing would then trigger perception of the same action as doing rather than allowing, explaining the preference reversal consistently with my hypothesis. Kamm, however, also argues that the doing/allowing distinction is not subject to framing effects as the loss/no gain distinction is, thus completely defeating the master argument.

To appreciate Kamm’s argument, it is useful to come back to the notion of fully equalised cases. Most counterexamples to the independence of the doing/allowing distinction, indeed, appear to be, as the “Asian flu” case, very confusing and misleading. According to Kamm, the phrasing, or the order in which events are presented, tends to hide the real or salient moral features of the case, thus inducing different moral judgements. Under a careful scrutiny, however, we should still recognize that one judgement is more adequate, or reliable, than the other, and thus “beat” framing effects. As for the “Asian flu” example, she argues, it is more reasonable, once we think more thoroughly about it, to conclude that any choice of vaccination plans would be an instance of allowing harm, rather than doing it. What is really causing the harm, here, is indeed something independent from our intentions and we can not possibly prevent: any bad consequence which would occur, therefore, should not be characterised as doing harm. What happens in the second decision problem, therefore, would be the result of a misleading reframing that fails to highlight this salient element of the choice at issue, i.e., that deaths are not brought about by our intervention. Once this aspect is made clear, arguably, the preference reversal will no longer be observed. The master argument would then be untenable, as we could maintain that some intuitions are more adequate and reliable than others, and these could serve a convincing account of the doing/allowing distinction.

I think Kamm’s conclusion here is not convincing, and that it fails to prove decisively that the doing/allowing distinction is immune to framing effects. First, it is worth noticing that Kamm is able to define both decision problems in the “Asian flu” case as instances of “allowing harm” because she relies on a prior definition of the distinction. Specifically, she refers to her “self-ownership” account, which states that a conduct classifies as “doing harm” if it is intruding or imposing over another individual’s body or possessions, while it is just an instance of “allowing” if this condition does not hold when bringing about the harm. In this sense, whichever health plan we choose, we

would not impinge on these people's self-ownership, and thus we would not be doing any harm. I believe, however, this argument amounts to a poor defence against the master argument, which assumes that adequate accounts of the doing/allowing distinction should be inferred from our intuitions about cases. If we already have a precise definition of the doing/allowing distinction, on the other hand, we would be adopting a different perspective on moral theorising, one which, in some sense, assigns a lesser role to our intuitions and insights. I will come back to this argument in section 4, discussing the significance of framing for the doing/allowing distinction.

Secondly, I am not sure whether our intuitions over the two decision problems can be dismissed as easily as Kamm claims. To be fair, as Horowitz's analysis seems to suggest, there could be at least disagreement over the characterisations of the conducts in the choice problem C vs D as "allowing". Other examples discussed by Kamm as to support this conclusion, like Schelling's Tax case, seems to show as well the same "instability" across individuals, insofar as there is a persistent disagreement over the correct evaluation of cases. In short, if my analysis were, if not conclusive, at least reasonable, we could argue that different classifications are seemingly fair and equally justifiable for different individuals. What is more, we might be still interested in an account of the doing/allowing distinction which explains the experimental findings, rather than one which solves the issue by claiming that one frame is more "right" than the other. In this sense, my hypothesis has the advantage of tracking how the reframing of the decision problems affects the perception of one behaviour as denying a gain/allowing harm rather than not preventing a loss/doing harm.

To sum up, my position is consistent with Kamm's claim a) that the doing/allowing distinction does not collapse into the loss/no gain distinction. We both agree that there is more in the "Asian flu" case than a different perception of outcomes as losses or no gains, and we both identify it with a concern for the way the outcome is brought about. Specifically, Kamm talks of denied gains and unprevented losses, a characterisation that is compatible with my psychological explanation for the different risk attitudes with respect to doing and allowing harm. I do not think, however, that Kamm succeeds in proving b), i.e., the fact that the doing/allowing distinction is not subject to framing. The fact that the doing/allowing (or unprevented loss/denied gain) distinction is different from the loss/no gain one, does not by itself rule out that the former could still be subject to framing. The latter hypothesis, I argue, is compatible with experimental findings and acknowledges for the instability of our intuitions and the persistency of preference reversal and interpersonal disagreement.

#### 1.4 Framing and the Master Argument

In this last section, I briefly discuss what my interpretation of the "Asian flu" case entails for the relevance of the doing/allowing distinction. As Kamm notices, we should be worried about framing effects insofar as they threaten the reliability of our intuitions over moral categories and principles, and, specifically, over the correct characterisation of actions as instances of doing rather than allowing. I have argued above that Kamm's answer to the master argument fails, as it assumes that we can somehow get rid of conflicting intuitions if we carefully reason upon cases and use the "correct" frame, the one which highlights the "right" moral features. My objection to Kamm is that this operation would only be possible if we already had an account of the doing/allowing distinction, which would not thus be constructed upon our intuitions about specific cases.

My different interpretation, on the other hand, i.e., that doing/allowing is subject to framing, leaves us vulnerable to the conclusion stated in the master argument, thus seriously threatening the significance of the doing/allowing distinction as a guide to our moral judgements. In this

sense, frame-dependency should pose some serious questions to the way we commonsensically think and reason from a moral perspective, as our moral reasoning would appear unreliable and not always to pick up on relevant features.

My intuition here is that there is no need to infer such a radical conclusion from limited examples like the “Asian flu” case. As Kamm persuasively argues, these counterexamples usually amount to very complex and difficult cases, where different effects and categorisations are mixed up in intricate ways, and, more importantly, the decision problem has no clear-cut solution. In both the choices between A and B and C and D, indeed, it is not self-evident that one option is morally required and the other impermissible, as the amount of interpersonal disagreement reported by experimental results shows. In these circumstances, I believe, agents try, throughout their deliberations, to “assimilate” the choice they face with more clear-cut examples and moral principles. In the “Asian flu” case, the phrasing and the use of specific terms may remind us and appeal to characterisations of “allowing harm” or “doing harm” respectively. For this reason, we pick up these cues and apply the principle “allowing harm is not as morally wrong/objectionable as doing harm” as to make a decision.

The kind of reasoning illustrated above, arguably, is not subject to frame-dependency when cases are well-defined and not as nuanced as the “Asian flu” examples. In most circumstances, therefore, framing should not affect the fact that we pick up the salient moral features, and will hence tend to agree upon which courses of action are permissible or required. To sum up, we might concede that our moral intuitions can generally be trusted, and can be used to build a coherent account of the doing/allowing distinction, while at the same time admitting that, for some of these “borderline” cases, framing effects could lead to different characterisations of the decision problem, that subsequently trigger different moral responses.

**Open Access** This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

## References

- Bennett J (1997) *The act itself*. Clarendon Press, Oxford
- Cushman F, Knobe J, Sinnott-Armstrong W (2008) In “moral appraisal affects doing/allowing judgements”. *Cognition* 108:281–289
- Horowitz T (1998) In “philosophical intuitions and psychological theory”. *Ethics* 108(2):367–385
- Frances Howard-Snyder (2007), *Doing harm vs. allowing harm*. *Stanford encyclopedia of philosophy*
- Kahneman D, Tversky A (1979) An analysis of decision under risk. *Econometrica* 47(2):263–291
- Kahneman D, Knetsch JL, Thaler RH (1990) Experimental tests of the endowment effect and the coase theorem. *J Polit Econ* 98(6):1325–1348
- Kahneman D, Knetsch JL, Thaler RH (1991) Anomalies: the endowment effect, loss aversion, and status quo bias. *J Econ Perspect* 5(1):193–206
- Kamm F (2007) Moral intuitions, cognitive psychology and the harming/not aiding distinction. In: *Intricate Ethics*. Oxford University Press, Oxford Chapter 14
- Peter Singer (1997), “The drowning child and the expanding circle”. In *New Internationalist*, April 17
- Sinnott-Armstrong W (2008) Framing moral intuitions. In: Sinnott-Armstrong W (ed) *Moral psychology: the cognitive science of morality: intuition and diversity*. MIT Press, Cambridge, MA, pp 107–144
- Wollard F (2015) *Doing and allowing harm*. Oxford University Press, Oxford, pp 17–20